

Making the Most of Policy Evaluations:

Overview and synthesis of evaluations of the
White Paper 'Our health, our care, our say'

Professor Chris Salisbury, University of Bristol.

Dr Kate Stewart, University of Bristol

Ms Ailsa Cameron, University of Bristol

Dr Stephen Peckham, London School Hygiene Tropical Medicine; SDO programme

Professor Mike Calnan, University of Kent

Dr Rachel Lart, University of Bristol

Dr Sarah Purdy, University of Bristol

Ms Helen Thorp, University of Bristol

Final Report
February 2010

Foreword

This report describes an overview and synthesis from findings of evaluations of initiatives promoted by the White Paper 'Our Health, our care, our say: a new direction for community services'. Through a series of case studies, lessons are drawn about how to improve the commissioning, conduct and use of evaluation of initiatives in health and social care.

We do not envisage that most readers will want to read the report in its entirety. Most people will find it more useful to read the executive summary and then to refer to specific chapters that are relevant to their interests.

In particular, chapter 4 consists of a series of structured summaries of the evidence obtained from evaluation of ten specific initiatives. This will hopefully provide a useful appraisal for readers interested in a particular initiative.

The final chapter discusses the overall findings from this project and includes specific recommendations and a bibliography of further resources.

Acknowledgements

The authors would like to acknowledge the support of the project advisory group, particularly Chris Ham, Jane Scott, Alan Glanz and Clare Croft-White.

We would also like to thank those people whose co-operation made this project possible by providing us with information or generously giving their time to talk to us about the initiatives and projects they were involved with.

Key messages

- There was a widespread commitment to the importance of the evaluation of initiatives promoted by the White Paper 'Our Health, our care, our say', as well as of health and social care initiatives more widely, in order to learn lessons about 'what works'. There was much less clarity about why evaluation is needed or how the findings will be used.
- There was some evidence that the initiatives are helping to achieve the goals of the White Paper 'Our health, our care, our say'. The evidence was stronger in relation to the feasibility and acceptability of the initiatives than it was in terms of costs or benefits.
- There was considerable variation between the different evaluations in terms of their scale, methods, funding and commissioning route, and this variation did not appear to have any strategic justification. Initiatives designed to support *better prevention and earlier intervention* were popular with users, but there was little evidence about health benefits and costs. Initiatives to give users *more choice and a stronger voice* were also popular but there was some evidence that impacts vary and those with the greatest needs were least likely to benefit. Many initiatives were designed to *improve access and tackle inequalities* and there was evidence of progress towards this goal. However, improved access can potentially increase rather than decrease inequalities, and lowering the threshold for care can reduce cost-effectiveness, again highlighting the need for more evidence about costs in relation to benefits. Several initiatives sought to provide *more support for people with long term needs*. Evaluation completed so far has been largely formative, but the two largest evaluations (IAPT and WSD) have not yet reported.
- There is considerable evaluation activity at both national and local levels, but much of this effort is not as productive as it could be because it fails to answer some of the most important questions raised by the goals set out in the White Paper.
- Most evaluations used formative methods, and provided useful information about implementation and how initiatives might work (formative evaluation). Many evaluations provided information about patient and staff perceptions or experience. Few provided robust evidence about whether initiatives achieve their objectives (summative evaluation) and fewer still provided evidence about the key issue of cost-effectiveness.
- Some evaluations were intended to assess the impact of initiatives, but did not use methods which could provide robust evidence about this. Some of the summative evaluations were limited by compromises in design or difficulties in collecting data. These were often related to not planning evaluation before the initiative was first implemented, insufficient resources or expecting findings in too short a time scale. Commissioning evaluations which cannot reach clear conclusions because of constraints such as these is a poor use of resources.
- All initiatives above a specified size should have a formal evaluation plan, which describes whether, why and how evaluation is to be conducted, and this plan should be reviewed by someone with research expertise.
- If possible, there should be a pilot phase in which an initiative is implemented in a limited number of sites, and decisions about future roll-out should not be made until the findings of evaluation are available.
- It is currently very difficult for policy makers or commissioners to know what evaluation is or has been conducted in relation to a topic of interest. Decision makers need an easier way of finding out about relevant ongoing evaluation and the findings of previous evaluations. There should be a publicly accessible register of evaluations, to increase awareness of the purpose and methods of current evaluations and to provide a record of their findings.

Executive summary

Background

The White Paper 'Our health, our care, our say' set out a new strategic direction for community care with four main goals:

- Better prevention and earlier intervention for improved health, independence and well-being
- More choice and a stronger voice for individuals and communities
- Tackling inequalities and improving access to services
- More support for people with long term needs.

The White Paper described a range of policy initiatives and included a commitment to evaluate these initiatives. It is important to learn lessons from these evaluations, both about the success of each individual initiative and also about the extent to which they are helping to achieve the broader policy goals described above.

The emphasis on evaluation which is apparent throughout the White Paper reflects an increasing recognition over the last fifteen years of the importance of using evidence to inform policy. This has become particularly important in the field of health care. However, previous research has demonstrated that the use of evidence in policy making is not straightforward. It is complicated by concerns about the relevance and quality of the evidence available, the value placed on different types of evidence, difficulties in ensuring that evidence reaches policy makers in a form they find useful, and also awareness that evidence is just one of many factors that influence policy.

The evaluations of initiatives arising from the White Paper were commissioned and conducted in a range of different ways. This raises questions about how and why these evaluations were commissioned, why they were conducted as they were, what they found and how these findings were used. A research project was commissioned by the Department of Health to study this process, to provide useful understanding about how best to commission and conduct evaluation of initiatives in health and social care in order to inform future policy.

Methods

The research drew on four sources of evidence:

1. In a mapping exercise, we sought to identify all evaluations of initiatives arising from the White Paper which were being conducted, or had been conducted, in England.
2. Based on the information collected in the mapping exercise, we conducted a survey to collect information about the context, methods and findings (or progress in the case of ongoing evaluations) of each evaluation. We used this survey to describe the range of evaluations being conducted, and the types of evidence that they would provide.

3. From respondents to the survey, we selected twenty-one evaluations for in-depth analysis using a case study approach. The case studies included centrally commissioned national evaluations for ten initiatives and a further eleven local evaluations. These case studies represented a wide range of evaluations in terms of initiative, methodological approach, size, funding arrangements and geographical location. Within each case study we sought to conduct qualitative interviews with those who led the service being evaluated, the principal investigator for the evaluation and a member of staff delivering the service being evaluated. Interviews explored what the initiative was intended to achieve, how and why the evaluation was undertaken, strengths and limitations of the evaluation approach chosen, factors which helped or hindered evaluation, and how the commissioning, conduct and dissemination of evaluation could be improved.
4. Interviews were held with key informants at the Department of Health, mainly civil servants responsible for delivery of the different initiatives, in order to understand their perspectives on the commissioning and conduct of evaluation and how the evidence arising from these evaluations would be used.

We synthesised the information from these four sources of data in three main ways:

First, we constructed structured summaries of the evidence available about the success of each of the white paper initiatives (summative evaluation) or ways in which these initiatives could be improved (formative evaluation).

Second, we drew lessons from across the different evaluations about the extent to which the initiatives are helping to meet the policy goals of better prevention, more choice, improved access and reduced inequalities, and more support for people with long term needs.

Third, we explored the findings in terms of how to improve the commissioning, conduct and use of evaluation of health and social care initiatives.

Findings

Types and strengths of evidence available from evaluation

It proved very difficult to identify all relevant evaluations of initiatives arising from the White Paper. Where we were able to identify evaluations, it was also difficult to obtain details of methods or results and therefore assess the strength of evidence available about initiatives. If we had difficulty obtaining this information, with the extra resources available to us as a research team, then commissioners of health or social care wanting to know whether or not there was evidence to justify commissioning a new service would find this almost impossible.

Most evaluations were intended to provide summative evidence about the effectiveness of initiatives and about half sought to provide formative information about how the initiative could be improved. However, in many cases the type of data collected would not allow conclusions to be reached about effectiveness. The majority of evaluations explored the process of care, patient experience, or staff views. About half the evaluations sought to provide evidence about the effectiveness of initiatives in terms of the impact on health or well-being. Fewer still (less than a third) sought to provide evidence about costs.

The most common research designs were based on qualitative methods (interviews or focus groups), and where quantitative methods were used only a minority of studies

used approaches which provide strong evidence by involving comparison with comparable control groups.

The quality of the evaluations was highly variable. Some of the national evaluations were ambitious, complex and combined multiple methods in carefully considered research designs. However, some evaluations did not have clearly specified aims and/or used methods which were not appropriate to achieve those aims.

Evidence about the specific initiatives

Readers interested in the evidence about a specific initiative are advised to read the relevant summary in the main report. This section very briefly summarises our interpretation of the most important findings. These conclusions may not necessarily correspond with the findings reported by those conducting the evaluations.

Summary of evidence from evaluation of White Paper initiatives

Care Closer To Home demonstration sites

Creative redesign of services demonstrates the 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 is underway with summative and formative elements. No findings yet available.

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 Prescription

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 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 are likely to require considerable resources and health or 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

Useful formative evaluation has been conducted to improve uptake and perceived usefulness of LifeCheck (separately for Baby LifeCheck, Teen LifeCheck and Mid-life Lifecheck). No evaluation yet about the costs or benefits of LifeCheck.

Partnerships for Older People Projects (POPP)

Have improved awareness of, access to and engagement with services amongst older people. This appears to have increased independence and quality of life, although this conclusion is based on limited data. POPP appears to be cost-effective, largely due to a reduction in emergency hospital admissions, but this conclusion is tentative since it is not based on direct comparison with controls not receiving POPP.

Self Referral to Physiotherapy

Popular with patients, 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 of the initiative: 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. Intended benefits were broad and success criteria not defined. No substantive evidence about the benefits for service users has been published so far.

Whole System Long Term Condition (LTC) Demonstration Sites

Ambitious evaluation is underway, using a randomised controlled trial combined with qualitative and economic methods. No findings have yet been published.

Progress towards White Paper goals

Several initiatives were designed to support **better prevention and earlier intervention** to improve health, independence and well-being. Examples include the LifeCheck programme and POPP. Evaluation suggests that these interventions have potential and are popular with users. Limited evidence from POPP suggests it may have important benefits in preventing hospital admissions – this needs replicating on a larger sample with a stronger research design. Prevention and early intervention programmes lower the threshold at which services become involved. This incurs costs which need to be justified in relation to their objective benefits. Robust research on the costs and benefits of these initiatives should therefore be a priority.

Individual Budgets, Information Prescriptions, and Self-Referral to Physiotherapy all reflect the policy goal of giving service users **more choice and a stronger voice**. Evaluations of these initiatives provided evidence of positive feedback from service users. There are important outstanding questions about which groups benefit most, as several evaluations suggested that impacts vary and evaluation of information prescriptions suggested that those with the greatest needs were least likely to benefit. More information is also needed about how people exercise choice, and the likely impact on services, workforce needs and health and social care budgets.

Many of the initiatives were designed to **improve access and tackle inequalities**. Several evaluations (e.g. Self-Referral to Physiotherapy, Care Closer to Home) demonstrated progress towards this goal. However, improved access can potentially increase rather than decrease inequalities. Attention needs to be given to improving access to effective care for those who need it most. Improving access can also increase costs by lowering the threshold for care and possibly increasing demand. This again highlights the need for more evidence about the costs in relation to the benefits of the initiatives.

The final goal was **more support for people with long term needs**. Several initiatives seek to address this goal, including Information Prescriptions, New Types of Workers, Improving Access to Psychological Therapies and Whole System LTC Demonstrator Sites. The evaluations which have reported so far have been largely formative and the latter two larger evaluations have not yet been completed. It is important to note that there are a wide range of other policy initiatives, not arising from the White Paper or included in this review, which are helping to improve care for people with long term needs.

Commissioning, conducting and using evaluation

In contrast to the common view that researchers and policy makers come from two communities with conflicting perspectives, the qualitative research based on interviews with key stakeholders identified a widespread commitment to the importance of evaluation of health and social care initiatives. There was also agreement about the characteristics of high quality and useful evaluation. However the rationale for why evaluation was important was much less clear and there was also less clarity about how evaluations would be used.

Both DH informants and principal investigators highlighted the importance of good communication between the 'customers' of evaluation in the DH and those with research experience at an early stage, well before the initiative is rolled out. This can be facilitated through involvement of research managers within the DH, or by

commissioning evaluation through one of the National Institute for Health Research programmes.

The evaluations conducted for different initiatives were highly variable in terms of scale, methods and funding, ranging from formative evaluation conducted by management consultants over three months to multimillion pound multicentre trials conducted over several years by a consortium of universities. There did not appear to be any strategic rationale for this variation, and it did not necessarily reflect either the resource implications of the initiative or the potential for health gain. A more consistent and considered strategic approach may lead to more effective use of the resources available for evaluation.

There were frequently tensions between the length of time it takes to conduct evaluation and the timescales for policy decisions. In some cases, by the time studies were commissioned the window of opportunity for robust and useful evaluation had passed. In addition, by the time evaluations were concluded the focus of policy had moved on and there was less interest in the findings, a problem exacerbated by rapid personnel turnover within the DH.

There were consistent problems in the way that evaluations were conducted. Although most evaluations stated an intention to determine whether or not an initiative achieved its objectives, few of them were designed to collect robust evidence about the most important objectives. Several initiatives had nebulous objectives, which limits any potential for summative evaluation. There was rarely a clearly articulated justification or theory to explain how and why an initiative would be expected to have particular benefits. Such an explanation would enable more informative evaluation about whether an initiative achieves its aims, the key determinants of success and why some initiatives are more effective than others.

Details of the methods and findings of evaluations were often not clearly described or in the public domain. In some cases this made it difficult to have confidence in the conclusions, as it is well recognised that without such clarity all research and evaluation is vulnerable to bias and selective reporting of results.

Several evaluations represented a considerable investment of resources, but did not use designs that would provide strong evidence about the most important questions. The most robust methods to determine the impact of initiatives on individual service users and the cost-effectiveness of initiatives are based on directly collected data from those receiving the initiative compared with those who do not receive the initiative. Ideally individuals or areas would be randomly allocated. Such evaluations are difficult and expensive, but may represent a better investment of time and effort than less rigorous methods which cannot provide clear answers. If this type of study is not possible, it may be more useful to conduct formative evaluation than to undertake an inconclusive summative evaluation.

There were tensions over the most appropriate time to conduct evaluation. If undertaken too early, the early adopter sites studied may be unrepresentative and the initiative will not have had time to 'bed-down' into normal operation. If undertaken too late, the findings will not influence policy decisions. If summative evaluation of impact is to be conducted, it should not be compromised by the need for quick results which may be misleading. Ideally the initiative should be tested in a limited number of sites and future roll-out should pause until the results of evaluation are available. However, if a decision to implement policy has already been made,

summative evaluation may not be worthwhile and formative evaluation to improve implementation may be more appropriate.

Conclusions

There was a commendable commitment to evaluate the initiatives promoted by the White Paper 'Our health, our care our say'. Several evaluations demonstrated progress towards the policy goals of the White Paper, although the evidence was stronger in terms of demonstrating the feasibility and acceptability of the initiatives than it was in terms of the costs and benefits.

There was considerable variation in the ways in which evaluations were carried out. Many were mainly formative and provided useful information about implementation. Some of the summative evaluations were limited by design compromises or difficulties in collecting data, which were often related to not planning evaluation before the initiative was first implemented and expecting findings in too short a time scale.

Because of the conflicting pressures on decision makers, particularly relating to the timescales within which decisions need to be made, this situation is unlikely to improve without organisational change which creates a requirement that evaluation is properly considered at an early stage in the decision making process. This would not mean that evaluation is always necessary or appropriate, but would mean that any evaluation which is undertaken is more likely to provide relevant and useful information.

If policy is genuinely to be based on evidence about 'what works', decision makers need an easy way of finding out what relevant evaluation is being conducted and much easier access to the findings of previous evaluations. In addition, greater transparency is needed about the methods used in evaluations so that it is possible to assess the strength of evidence that is provided and to have greater confidence in the interpretation and independence of findings.

These changes could be achieved by a system for registering evaluations similar to the registers which have been instituted for clinical trials, and by a formalised requirement that all new initiatives above a specified size should have an evaluation plan which includes information about whether, why and how evaluation is to be conducted.

In order to improve the usefulness of evaluation, it is important to provide greater clarity about the intended benefits of the initiative, to ensure that a means of evaluation is built into the policy making process from the outset, to have a clearly defined purpose for the evaluation at the outset ensuring that the methods chosen reflect that purpose, and finally to determine to take account of the findings in setting future policy.

Ten Recommendations

1) Before undertaking any evaluation, policy leads should address five key questions:

- Why do we need to undertake evaluation?
- Do we want to know whether an initiative is effective or cost-effective, or do we want to know how to improve implementation of an initiative?
- What would define success or failure of the initiative?
- What difference will the evaluation make to policy? (If evaluation suggests that the initiative does not achieve its objectives what will we do?)
- What mechanisms need to be in place to expand, halt or amend a policy in the light of evaluation?

The Department of Health (DH) should require answers to these questions to be recorded in a structured form, as part of an evaluation plan which should be mandatory as part of the internal approval process for any major initiative.

2) There should be a publicly available register of summative evaluations of health and social care initiatives which are publicly funded above a certain threshold (e.g. £50,000). The register would include key details of the methods of the evaluation and should be made publicly available online *before* the evaluation begins. This register could be co-ordinated by the National Institute for Health Research (NIHR).

3) The DH should give more attention to ensuring that policy leads are aware of existing guides to effective evaluation and able to take advantage of the training and professional development modules offered by the Government Social Research Unit.

4) Each NHS trust should have ready access to a source of help and advice with regard to evaluation design, data collection and analysis. The Government Social Research Unit should consider developing a web-based distance learning resource tailored to the needs of local NHS provider organisations and commissioners to help them conduct better quality and more useful evaluation.

5) The DH should always involve people with appropriate expertise in agreeing an appropriate design for evaluation. This may be achieved by involving the DH research and development directorate or by commissioning evaluation via one of the National Institute of Health Research (NIHR) research programmes.

6) Where possible, summative evaluation should be commissioned 'at arm's length' from the DH by organisations such as the NIHR Policy Research Programme or the Service Delivery and Organisation Programme.

7) If possible, there should be a pilot phase in which the initiative is implemented in a limited number of sites, and decisions by DH about future roll-out should not be made until the findings of evaluation are publicly available.

8) The DH should seek to maximise the independence of the team undertaking evaluation, and any potential conflicts of interest should be transparent.

9) The DH should make a commitment to publish all evaluation reports and to make them freely publicly available, irrespective of the findings

10) The DH needs to develop a mechanism for ensuring that evaluations follow their proposed plan (or changes are fully justified) and that the findings of evaluation are effectively disseminated to decision makers at local and national levels, given that the original policy

Contents

Executive summary	iv
1. Background	3
The NHS Context	3
The role of pilot programmes and evaluation	4
The impact of evidence on policy.....	7
How can the use of policy evaluations be improved?.....	9
Summary	10
Aims.....	10
2. Methods	11
Overview of data collection.....	11
Mapping exercise	11
Survey of evaluations	14
Case studies	17
Key informant interviews	20
3. Types and strength of evidence provided by evaluations of initiatives	23
4. Summary of evidence about each of the White Paper initiatives	29
Care Closer to Home	29
Improving Access to Psychological Therapies	35
Individual Budgets	42
Information Prescriptions	50
New Types of Workers	59
NHS LifeCheck	67
Partnership for Older People Projects (POPP)	74
Self-Referral to Physiotherapy	79
Social Enterprise Pathfinders	85
Whole System Long Term Condition Demonstrators.....	91
5. Evidence about the extent to which the White Paper initiatives are helping to meet policy goals	98
Better prevention and earlier intervention to improve health, independence and well-being	98
More choice and a stronger voice for individuals and communities.....	99
Tackling inequalities and improving access to services	101
More support for people with long term needs	102

Conclusion	103
6. Commissioning, conducting and using evaluation	105
Policy development	105
Strategy and planning	108
Engagement between policy makers, evaluators and sites: communication and co-ordination issues	113
The impact evaluation activities have on conduct and local practice in sites	115
The impact evaluation activities have on policy making	116
7. Discussion	119
Why do evaluation?	119
Commissioning evaluation	120
Conducting evaluation	121
Using evaluation	127
Postscript	130
Conclusion	131
Recommendations.....	132
8. Recommended guides to evaluation.....	135
9. References	137
10. Appendices.....	143
Appendix A Mapping responses according to contact type.....	144
Appendix B Survey Questionnaire (MS Word format)	145

1. Background

The NHS Context

The White Paper *'Our health, our care, our say: a new direction for community services'*⁴ was published in January 2006. It followed an earlier Green Paper on adult social care, and drew on a programme of consultation with the public about their priorities for the future of health and social care. The White Paper set out a new strategic direction with four main goals:

- Better prevention and earlier intervention for improved health, independence and well-being
- More choice and a stronger voice for individuals and communities
- Tackling inequalities and improving access to services
- More support for people with long term needs

The White Paper envisaged a strategic shift towards new forms of community services in order to achieve the above policy goals. It also described a range of policy initiatives and commitments that directly affect the relationship between community health and social care services, service users and carers. These are the means by which the policy goals would be achieved and provide a framework against which the implementation of the White Paper can be assessed (See Table 1-1).^{2;3}

Implementation of the initiatives described in Table 1-1 was supported by a number of changes in the 'architecture' of the NHS, including practice-based commissioning, payment by results, a framework for joint working between health and social care agencies, changes in local government and workforce reform. Therefore the changes brought about by 'Our health, our care, our say' cannot be considered in isolation. They sit alongside a range of other broader policies in health and social care services in England to widen access and choice, including increasing diversification of provision, reorganisation of commissioning, more integrated provision of health and social care and decentralising decision making.^{4;5}

It is important to note that the White Paper acted in three ways: (i) as an 'umbrella' under which to bring together a number of pre-existing initiatives within a coherent policy framework, (ii) as a 'catalyst' to promote initiatives which were already being developed by local pioneers, and (iii) by introducing new initiatives. Examples of the first category are the initiatives to improve access to general practice and to develop new work force roles during the decade before the publication of the White Paper. With regard to the second 'catalyst' role, the White Paper highlighted initiatives (for example, direct access to physiotherapy) which had been introduced in a small number of areas and were perceived to be promising ideas, and made a commitment to implement them more widely. Thirdly, the White Paper introduced entirely new commitments, mainly reflecting ideas prioritised by the public in the consultation exercise, such as NHS LifeChecks.

The key point is that all of these initiatives, whether established, in early development or entirely new, were viewed as reflecting the policy goals set out in 'Our health, our care, our say'.

Table 1-1 Specific initiatives described in White Paper in relation to policy goals

Better prevention and earlier intervention for improved health, independence and well-being

- NHS LifeCheck
- Improving Access to Psychological Therapies
- Partnerships for Older People Projects
- Bowel Screening Programme

More choice and a stronger voice for individuals and communities

- Individual Budgets and Direct Payments
- Patient Experience Survey
- End-of-Life Care Strategy

Tackling inequalities and improving access to services

- Shifting Care Closer to Home
- New Generation of Community Hospitals
- Urgent Care Framework
- Increasing Primary Care Provision
- Improving Access to Primary Care
- Self-Referral to Physiotherapy

More support for people with long term needs

- Expert Patients Programme
- Whole Systems Integrated Care
- Development of the Common Assessment Framework
- Information Prescriptions

Supporting changes (see below)

- Pilots of New Types of Workforce Roles
- Social Enterprise Pathfinders

The role of pilot programmes and evaluation

There was a recurring theme within 'Our health, our care, our say' of the need to evaluate the initiatives described. For example, this is reflected in the commitment to establish and evaluate 'demonstration sites' for the care closer to home initiative, 'pilot sites' for direct access to physiotherapy, and Social Enterprise 'Pathfinders'. This commitment to evaluation is expressed more strongly and consistently than in previous health White Papers.

There may have been a number of reasons for this emphasis on evaluation. Over the previous decade there had been increasing discussion about the need for evidence based policy throughout government.⁶⁻⁸ This was particularly relevant to health reform, given the importance of the 'evidence-based-medicine' movement which had a major influence on medical practice in the UK from the early 1990s onwards. Making a commitment to evaluation had become an essential part of the strategy to implement health policy and to defend new initiatives against potential opponents of changes within the NHS, as well as being important to inform policy. Several government reports have argued that systematic evaluation of the effectiveness of policies is an essential component of professional and modern policy-making.^{9;10} In 2004, the King's Fund report 'Finding Out What Works', informed by similar empirical evidence as this project relating to social programmes, suggested that a gap remained between the rhetoric of evidence based policy and practice.¹¹

The incoming Labour government in 1997 declared that 'what counts is what works' and this phrase was a key principle in the 1999 White Paper 'The New NHS'.¹² According to the Cabinet Office report *Professional Policy Making for the Twenty-First Century* (1999)¹⁰ the nine key features of modern policy making evaluation include using the best available evidence from a wide range of sources to inform policy, building systematic evaluation of early outcomes into the policy process, and learning lessons from experience of what works and what does not. The report then describes how evaluation should influence policy making (Box 1-1).

The justification for this use of evidence is simple – that it will lead to better and more effective policy. It will also lead to increased cost-efficiency in delivering services and increase public and professional confidence through greater transparency.

Box 1-1 Key features of modern policy making¹⁰

Systematic evaluation of the effectiveness of policy is built into the policy making process. Approaches to policy making that demonstrate a commitment to evaluation include:

- *Clearly defined purpose for the evaluation at the outset*
- *Success criteria defined*
- *Means of evaluation built into the policy making process from the outset*
- *Use of pilots to influence final outcomes*

The concept of conducting pilots before widespread implementation of policy became increasingly popular from 1997 onwards. The report 'Adding It Up',¹³ published by the Performance and Innovation Unit of the Cabinet Office called for more and better use of pilots to test the impact of policies before national roll-out. To help achieve this aim it recommended the establishment of a panel of enquiry to consider the future role of pilots. This enquiry was conducted by the Governments' Chief Social Researchers Office and the report was published in 2003.¹⁴ It supported the use of pilots before full-scale implementation of policy as the first stage of longer-term policy monitoring and evaluation.

The report made a number of key observations and recommendations, including the following:

- Pilots should be carried out in a genuine spirit of experimentation and are redundant if the policy is already set in stone.
- The benefits of a pilot will not be realised if the policy is implemented before the study has run its course.
- It is crucial that evaluation of pilots is independent if it is to be credible, and civil servants and ministers must distance themselves from decisions about methods and dissemination of findings.
- Methods must be rigorous but multiple methods should be considered to obtain a full picture of the impact of an initiative.
- A pilot that reveals a policy to be flawed or ineffective should be viewed as a success rather than a failure, having potentially helped to guard against the inclusion of embarrassing, expensive and preventable mistakes into new policy initiatives.
- Appropriate mechanisms should be in place to adapt or abandon a policy or its delivery mechanism in the light of a pilot's findings.

Several of the initiatives described in Table 1-1 carried with them a specific commitment within the White Paper to evaluation of pilot projects. However, this commitment to evaluation was taken forward in a number of ways. Some evaluations were directly commissioned by the Department of Health from independent organisations. Some, but not all, of these were commissioned by the Department of Health Research and Development Directorate. Some evaluations were conducted internally by the Department of Health with little or no involvement of other organisations, while others were commissioned by NHS research funding bodies such as the Service Delivery and Organisation (SDO) Research and Development Programme. Some evaluations were extensive and based on well recognised methods backed by detailed research protocols, while others were much more informal. Finally, in some cases it was clear from the outset that the evaluation would be independent and the results would be published, while in other cases there was no independence and no commitment to public dissemination of findings.

It is not immediately clear why this range of approaches to the evaluation of different initiatives came about, except that the largest (and therefore most expensive) evaluations were more likely to be commissioned via an external funding body. This raises questions about the reasons for these different approaches to commissioning evaluations, and the consequences in terms of how the findings from evaluation will be used.

It is important to recognise that in addition to the evaluations commissioned for the White Paper there is other relevant research, such as that commissioned by the SDO Programme on choice and self-care and by the Department of Health on Health Systems Reform Programme Evaluation. Further research has been instigated by other agencies and researchers themselves that may not be directly focused on the key policy questions raised by 'Our health, our care, our say' but is relevant (e.g. ESRC research on public service performance).

The impact of evidence on policy

Although the White Paper made a clear commitment to evaluation, there is debate about the use of this type of research evidence and its influence on policy.^{6;15-18} As Klein has observed,¹⁹ the implication that policy decisions should be based on research findings has not materialised despite the recognition of the need to 'get research into practice',^{7;20;21} also known as 'knowledge transfer'.^{22;23} Transferring knowledge from researchers who produce it, to the people who can use it (whether in policy making or in improving planning and management of services) is complex.

Research evidence is often difficult for policy makers to use because:

- it does not always address the questions decision makers need answering.
- it may not be timely as the timescales for undertaking high quality, reliable empirical studies are often longer than potential users can wait for the answers they need.
- research results may be expressed in such a way that it is difficult for potential users to pick up on the messages relevant for their circumstances.
- day-to-day organisational pressures and responding to other national targets are likely to be more influential on decision and policy makers than using formal research based evidence.

These issues are particularly salient for applied research, where expectations of impact on policy and practice will be higher, and where conventional methods of evaluation may prove inadequate.

Evaluation is one form of research, intended to determine whether and how an initiative 'works'. The former can be described as summative or impact evaluation, to *prove whether* an initiative works, and the latter as formative evaluation, aiming to *improve how* it works. Formative evaluation is generally seen as examining the early stages of development to help improve the initiative, and is related to process evaluation which is specifically designed to examine what happens during an initiative. Summative evaluation seeks to demonstrate whether the initiative has had the desired impact, although it is important to consider unintended as well as intended outcomes.

Research involves identifying and exploring a clearly defined question. Most traditional biomedical research, of which the randomised controlled trial is usually seen as the highest form, involves seeking to isolate and test the active intervention after excluding all possible sources of bias. Issues of local context are seen as nuisance 'confounding' variables which need to be balanced out across the intervention and comparison groups. However, in recent years there has been increasing recognition that in the real world social interventions act within complex social systems, and that context is crucial.⁸ This has led to increasing interest in theory driven models of evaluation, such as the 'realist' model, which argues that the key question is not 'what works' but 'what works for whom in what circumstances'.²⁴ These approaches aim to articulate and test the theory for how an initiative is intended to lead to specified benefits under certain circumstances. Evidence is then collected to populate this theoretical framework. Many different forms of evidence may be used to support, contradict or refine the theory.

The lack of impact of evidence on policy may partly reflect the fact that most evaluations of complex social interventions either involve an over-simplification of reality or are themselves highly complex and do not provide definitive answers about impact and effect. In addition it is unlikely that a single study will generate a clear answer about whether or not an initiative is effective, because the context (e.g. client group, other local services available) in which the evaluation was conducted will almost certainly be different from that in which the initiative is going to be implemented more widely, other policies will have been introduced since the evaluation was started which will modify the impact of the specific initiative which was the subject of evaluation, and the initiative itself will probably have changed because of lessons learned while the evaluation was underway and for other external reasons. The application of evidence from evaluation therefore always involves an element of interpretation.

For the reasons described above, decision-makers need to take account of a wide range of types of evidence in formulating policy.^{8;25} Evidence from evaluation is always likely to be just one factor that contributes to the complex mix of influences which ultimately determine policy. How this process occurs has been studied by a number of authors.^{8;18;21;25}

Weiss has described two key models of the use of research in policy: the 'problem-solving' and 'interactive' models.¹⁶ The problem-solving model proposes that once a policy problem is identified a solution is sought through research and then the information is transferred from the research arena to the policy arena. This rational model is widely assumed implicitly, but has been criticised as not reflecting what actually occurs. Real world policy decisions are often not clearly defined, decisions need to be made before evidence can be provided, and many factors affect decisions apart from evidence. Instead the policy/research interface is much more circuitous and complex. The interactive model proposes that research has a vague influence on policy but is just one of a number of other influences including policy-makers' own experience, politicians, the press, colleagues and service providers. Studies of the way in which policy makers use evidence support this interactive model rather than the problem-solving model.²⁵

In order for research evidence to have any kind of impact it has to move from the research arena to the policy arena. Dobrow distinguishes between the 'introduction' of evidence (how research is identified and introduced to the decision-making process), the 'interpretation' of evidence (recognition of the relevance, appropriateness and utility of the research) and the 'application' of evidence (how it is used to support or justify a decision).⁸ Contextual factors have a major influence on all of these stages, for example whether the evidence is in tune with prevailing ideology and the political acceptability of different decisions. Personal interaction between researchers and policy-makers are also very important.²¹

The importance of this personal contact was also highlighted in a study of how health policy makers perceive their use of evidence. Innvaer describes the apparent disconnection between researchers and decision makers as the 'two communities' thesis.²⁰ This thesis suggests the existence of two groups of people who lack the ability to take into account the realities or perspectives of one another. Social scientists tend to see themselves as rational, objective and open to new ideas and to see decision makers as action and interest orientated, indifferent to evidence and new ideas. Decision makers on the other hand see themselves as responsible, action-

orientated and pragmatic. They see social scientists as naïve, jargon ridden and irresponsible in relationship to practical realities. Absence of personal contact between policy makers and researchers and the associated mutual distrust were the most commonly mentioned barriers to the use of evidence.

It is important to note that the use of evaluation does not only depend on whether the information it provides reaches and is understood by policy-makers. Different policy-makers may be aware of the same evidence, but come to different decisions. Russell et al. criticise the 'naive rationalist' assumption that policy making is about finding and implementing the best available evidence about what works, but is instead 'the formal struggle over ideas and values'.²⁶ This approach recognises that there is no single 'right answer' in the messy world of policy making but only 'more-or-less good reasons to arrive at more-or-less plausible conclusions'.²⁶ The excellent guide to evaluation published by the US Kellogg Foundation highlights that the conduct and interpretation of evaluation is always a political and value-laden process in which many stakeholders have an interest and influence.²⁷

Freeman (2007) suggests that health policy making is characterised by 'bricolage', a process where the policy maker '... in contrast to the scientist or engineer, acquires and assembles tools and materials as he or she goes, keeping them until they might be used.'²⁸ Fundamentally, the development of policy is based upon values and the use of evidence is influenced by both ideological positions and what Sabatier has described as core beliefs and thus selective use of evidence is likely.²⁹ The importance of networks between researchers, policy makers and implementers has been shown to improve research use.

With all these provisos about the fact that evidence from evaluation is only one influence in the policy making process, it remains important to ensure that evaluations can contribute to this decision making process as effectively as possible. For this to happen it is essential that the appropriate form of evaluation is conducted in relation to the questions of interest, that the evaluation is well designed and carried out to high standards, and that the findings from evaluation reach policy makers in ways that are timely, understandable and relevant.

How can the use of policy evaluations be improved?

A number of authors have explored why evidence does not appear to have a strong influence on policy, and have made recommendations about how this situation can be improved. Specific findings include the following:

- To ensure that learning is incorporated into policy and practice there needs to be a facilitation process to bring researchers and policy makers together. Therefore it is essential to develop networks of researchers and policy makers and involve managers and policy-makers in an iterative synthesis process to ensure that the interpretation of the evaluation evidence from the studies is more relevant for policy and practice.^{15;20;30-32}
- Reviews of evidence need to highlight the relevance of information to policy-makers, giving greater attention to local contextual factors and the costs as well as benefits of an initiative.³²
- Researchers need to ensure that their research is perceived as timely and of high quality and show how their findings are relevant to current policy.²⁰

- Policy-makers need information presented to them in a way which allows rapid scanning for relevance and then graded entry to the information (for example a 1 page list of take home messages, with a 3 page executive summary and a 25 page report).³² These will highlight clear policy recommendations.²⁰
- Researchers and research funders need to have a more sophisticated understanding of how research influences policy.⁶ This may lead to iterative tendering for research involving dialogue between researchers and policy makers, involving those who need the research evidence more in the commissioning process, and a greater use of syntheses of results from multiple studies rather than single studies.

Summary

In summary, the White Paper 'Our health, our care, our say' committed the government to evaluation of a number of initiatives in health and social care. It is important to learn lessons from these evaluations, both about the success (or lack of success) of these initiatives, but also about the contribution they make to delivering the wider policy goals of the White Paper.

These evaluations have been commissioned and conducted in a range of different ways, which reflects a commitment to base policy on evidence about 'what works'. However, previous experience suggests that evaluation has only a limited and indirect influence on policy decisions about the wider roll-out of initiatives. These considerations lead to questions about how and why evaluations of the 'Our health, our care, our say' initiatives were commissioned, why they were conducted in different ways, what they found and how these findings were used in the policy making process.

Aims

The overall aims of this project are to:

- Synthesise findings from evaluation of specific White Paper initiatives to determine the extent to which individual initiatives have achieved their objectives (summative evaluation) and ways in which the initiative can be improved (formative evaluation).
- Synthesise findings from the evaluations of different initiatives to determine the extent to which the overall policy goals of the White Paper are being met.
- Increase understanding of the most effective ways of conducting evaluation of innovations in delivery of health care in order to generate findings which are useful to inform policy. These include the strengths and weaknesses of different approaches, learning lessons from successful evaluations and identifying barriers and facilitators to evaluation within health and social care.

The findings from this research will not only provide evidence to support decisions about the future of specific initiatives, and about progress with regard to policy goals, but will also provide wider learning about how to improve the process of commissioning and conducting evaluations of health and social care programmes.

2. Methods

Overview of data collection

The project design consisted of four interconnected elements. Although conceptualised as separate elements, as expounded below, in practice a degree of overlap exists between them. The elements were:

1. Identification of evaluations (the 'mapping' exercise): this sought to identify all studies that were evaluations of initiatives arising from the 'Our health, our care, our say' White Paper, or which related to the goals set out in that White Paper.
2. Survey of evaluations: to collect information about these evaluations in order to summarise and appraise this evidence.
3. Case studies: a range of large national and smaller local evaluations were selected from those identified in the survey and qualitative interviews were conducted with key individuals from these projects.
4. Interviews were held with key informants at the Department of Health about their perspectives on the evaluation of initiatives.

Mapping exercise

The mapping exercise was conducted by contacting a wide range of potential informants, and by undertaking internet based searches in order to identify all relevant evaluations of initiatives relating to the White Paper.

In discussion with the Department of Health we identified eleven initiatives which had arisen from the White Paper and which were subject to formal evaluation at a national level. These initiatives are listed in Box 2-1 below.

Box 2-1 List of White Paper initiatives subject to national evaluation

Care Closer to Home Demonstration Sites
Development of Common Assessment Framework
Improving Access to Psychological Therapies
Individual Budgets
Information Prescription
New Types of Worker Pilots
NHS LifeCheck
Partnerships for Older People
Self Referral to Physiotherapy
Social Enterprise Pathfinders
Whole System Long Term Condition (LTC) Demonstration Sites

It became clear that the evaluation of the Common Assessment Framework was unlikely to be commissioned until towards the completion of our overview, therefore this initiative was excluded, leaving ten national evaluations as the focus for our work. Furthermore, very little information was available about the evaluation of Early Years LifeCheck or Mid-life LifeCheck, therefore, in discussion with the Department of Health, we agreed to focus our assessment of evaluation of NHS LifeCheck on the Teen LifeCheck.

National evaluations of the above initiatives were instigated by the Department of Health, and progress of these was monitored by the DH in progress reports. Some of these national evaluations were commissioned directly by the DH, but others were commissioned by bodies such as the NIHR SDO Programme. However, it was recognised from the outset of this overview project that in addition to the national evaluations it was likely that other evaluations may have been undertaken (or were underway), which would provide relevant information about the success of the White Paper initiatives. These were likely to include smaller local evaluations conducted by universities or PCTs, and also possibly larger scale externally commissioned evaluation. The aim of this overview was to identify and synthesise any evidence we could find relating to evaluation of White Paper initiatives, whether or not it was commissioned as part of the programme of national evaluations. This was the purpose of the mapping exercise described below.

For the purposes of this mapping exercise, we defined 'evaluation' as any piece of work that:

- Included a written protocol or evaluation plan
- Included the collection and analysis of quantitative or qualitative data
- Measures progress towards pre-specified outcomes or benefits
- Was intended to result in a formal evaluation report

These four key criteria were based on an underlying conceptualisation of evaluation based on Patton's widely recognised definition of program evaluation:

*The systematic collection of information about the activities, characteristics, and outcomes of programs to make judgements about the program, improve program effectiveness, and/or inform decisions about future programming.*³³
p23

During the mapping phase of the study letters were sent to the following groups of people:

- Chief executives of all the PCTs
- Chief executives of all the SHAs
- Directors of adult social services
- Evaluation leads for the DH White Paper pilots
- The Department of Health policy leads for the White Paper pilots
- NHS Institute for Innovation and Improvement
- Improvement and Development Agency
- Integrated Service Improvement Programme
- NHS SDO R&D Programme

- Picker Europe
- Care Services Improvement Programme
- Improvement Foundation
- Health Foundation

This led to a total of 335 organisations or individuals being contacted: 151 PCTs, 143 Adult Social Services Departments, ten SHAs, nine funding bodies or other organisations, eleven DH policy leads and eleven principal investigators for DH led initiatives.

A letter was sent to these contacts outlining the eleven initiative areas we were interested in, and asking for information about any evaluation work they knew that was being conducted or funded that:

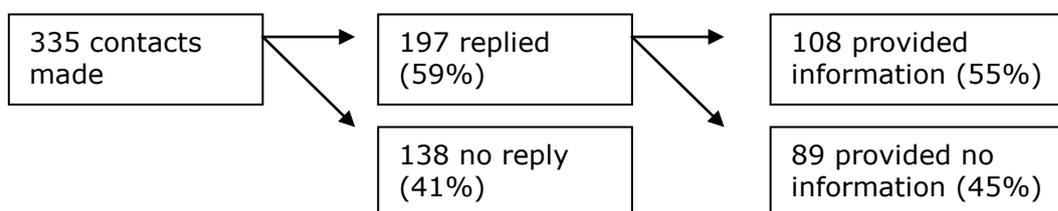
- Was intended to address the White Paper goals
- Broadly fulfilled the four characteristics of evaluation described above
- Related to one of the eleven initiative areas that were listed

The initial letter was followed up by a shorter reminder letter several weeks later. The chief executives of the PCTs, directors of adult social services, evaluation leads for the DH pilots and the DH policy leads were followed up for a final time by telephone if there was no response to the letters.

In addition to contacting the individuals listed above, we also sought to identify possible relevant evaluations in a number of other ways. These included searching the archived National Research Register (which provided data on studies until 2007) and the UKCRN Portfolio research database (which includes details of all research funded by the NIHR), as well as doing internet searches using Google, checking the Department of Health website and making personal contact with relevant researchers in the field.

The response rate to the letter sent to key individuals was 59%, with 108 (55%) providing information on potential evaluations. A full breakdown of responses according to contact type can be found in Appendix A, with a summary in Figure 2-1 below.

Figure 2-1 Summary of responses to mapping exercise



Some respondents provided information about several projects, and some projects were mentioned by more than one respondent. The mapping exercise yielded 188 contact names linked to potential evaluations. When we accounted for duplicate contacts, single contacts with multiple projects, projects clearly not fitting our criteria, and additional projects identified from internet based searches, we identified 141 potentially relevant evaluation projects.

The lead investigators for these projects were contacted with invitations to complete a questionnaire about their evaluation work, and eleven invitations (excluding Common Assessment Framework, but including two lead investigators identified for New Types of Worker) were sent to the lead investigators of the national evaluations (153 invitations total). In the vast majority of cases this invitation related to a single project, but in a small number of cases our information suggested the individual was involved in more than one project and they were invited to complete multiple responses.

It is important to note that these invitations to complete a questionnaire related to 'potential evaluations', meaning that we had been told that the individual concerned *may have been* conducting relevant evaluation. We recognised that in some cases this information would prove to be inaccurate once we made contact with the individual directly. Thus, there is a degree of overlap between the mapping and survey stages of the project.

Survey of evaluations

The mapping phase described above identified 'potential evaluations'. The purpose of the questionnaire survey was to confirm whether or not the research did indeed fulfil our criteria for an evaluation, and to obtain information about the methods used in each evaluation. In this way we sought to describe the volume of research being conducted into different White Paper initiatives, the types of research methods being used to evaluate initiatives, and to appraise the types and strength of evidence that would be available about the success of the different initiatives.

A survey questionnaire was designed to be administered to evaluation teams identified in the mapping exercise described above. The design of the questionnaire was informed by a number of sources, particularly the data extraction form developed by the Cochrane Effective Practice and Organisation of Care Group (EPOC).

This questionnaire included questions about:

- baseline details (title, evaluation team, management, funding)
- aims of the initiative (and how these relate to White Paper goals)
- context (patients/clients, area, local factors which may impact on evaluation)
- details of the initiative (services provided, number/type of people & staff involved)
- approach to evaluation (aims/objectives, design, methods, outcome measures, sample size, timetable)
- user involvement
- plans for dissemination
- progress and factors which facilitated or hindered the evaluation

Respondents were also asked to supply copies of their evaluation protocol, progress reports and final reports when available.

The survey was in the first instance administered online using the Bristol Online Surveys facility, which affords respondents convenience by allowing them to cut and paste information from existing documents, and facilitates swifter and more accurate

data entry by allowing responses to be downloaded directly into analysis software. The individual who was identified as leading the evaluation was contacted in writing, and the link to the online survey was then emailed within a few days of sending this letter. Reminder emails were initially sent with a further link to the online survey and then with a copy of the questionnaire in MSWord format for completion offline, to be returned by post or email. When there was no response the contact was followed up by telephone to try to establish whether the survey was relevant to the work being completed, and to encourage the project lead to provide us with more information, preferably via the survey but if necessary by simply sending us protocols or any other documents that would provide similar information about their methods to that requested in the questionnaire.

The responses to the questionnaire are summarised in Table 2-1. A copy of the questionnaire itself is provided as Appendix B.

A number of individuals responded to the questionnaire to indicate that the work that they were conducting did not fulfil our criteria for inclusion in the project and therefore they did not need to complete the questionnaire. These individuals should be viewed not as declining to respond to the questionnaire, but rather as having provided further clarification to the mapping exercise.

After excluding people who responded to say that the questionnaire was not applicable because they were not conducting an evaluation, the overall response rate was 32% (36/112). However it is quite likely that a number of the non-respondents did not reply because the contact details we had been given were inaccurate or because the questionnaire was not relevant to them, so it is uncertain what proportion of all relevant evaluation work was identified by this exercise. We were able to obtain responses from most of the national evaluations which were the main focus of our project. The exceptions were as follows. It became clear during the mapping exercise that evaluation had not yet been commissioned for the Common Assessment Framework. With regard to Self-Referral to Physiotherapy, there was no clearly identifiable principal investigator, but a report on the data collected by pilot sites was subsequently completed by a researcher brought in on a consultancy basis and we were able to appraise the evaluation from this report. The Whole System Demonstrator project was very complicated and could not be summarised within our questionnaire. However the research team sent us their application for ethical approval and some presentations about the research and we were able to understand the methods used from these documents.

Table 2-1 Questionnaire responses

	a Questionnaire invitations	b Not conducting an evaluation, questionnaire not completed	c National Evaluation questionnaire returns (incl. partially completed)	d Non DH questionnaire returns (incl. partially completed)
Care Closer to Home	30	8	1	2
Common Assessment Framework	6	4	0	0
Improving Access to Psychological Therapies	17	5	1	4
Individual Budgets	17	2	1	2
Information Prescription	9	4	1	2
New Types of Workers	9	0	2	1
NHS LifeCheck	4	3	1	0
POPP	17	1	1	5
Self Referral to Physiotherapy	8	3	0	3
Social Enterprise Pathfinders	7	2	1	0
Whole System for LTC	26	7	0	1
Other	3	2		7
Totals	153	41	9	27

Note: The mapping exercise had identified some projects as relating to an initiative area which the respondent in the questionnaire subsequently indicated related to a different initiative. Columns a & b are categorised according to information from the mapping exercise, and column d is categorised according to questionnaire response. Thus, appropriate caution should be exercised reading across the table rows.

Assessment of strength of evidence

The information provided in the mapping and survey stages of the project was subjected to a process of critical appraisal. The purpose of this was to summarise the overall body of evidence which would be available in relation to each initiative area, rather than to carry out a critique of the work of individual research teams. In this way we intended to identify consistent themes in terms of the types of evaluation, and the strengths and weaknesses of the research, being conducted in relation to White Paper initiatives.

A proforma was developed to record an initial summary of each of the thirty-five projects (eight national DH evaluations who responded and twenty-seven local evaluations) identified through the mapping exercise. In completing the summary about each evaluation, the research team made use of all information available, including that provided by investigators in their questionnaire responses along with any information available from protocols, interim or final reports and any other documents or publications. The proforma was designed to collect in a structured form:

- The type of evidence provided by the project (that is, the types of outcomes or domains that the evaluation provided evidence about, such as the process of care, patient experience, effectiveness, cost-effectiveness)
- Whether the project fulfilled our original criteria to be considered an evaluation
- Whether the evaluation should be characterised as formative, summative, or action research
- The methodological approach(es) used within the project
- Key indicators of the reliability, independence, and generalisability of the methods used
- Comments about the strengths and weaknesses of the project

After identifying the methodological approach for each evaluation, the reviewers used the relevant guidelines developed by the Clinical Appraisal Skills Programme (CASP) ³⁴ to summarise in more detail the strengths and weaknesses of each project. A range of appraisal guideline tools was considered, and the CASP guidelines were chosen due to their simplicity, their appropriateness for a range of project types, and their flexibility of use given the variation in available information for each project. Appraisal of the projects was shared between the research team, with each appraisal reviewed by a second team member and all overseen by the principal investigator.

This process of critical appraisal was undertaken for all thirty-five projects for which we had obtained some information. In some cases it became clear that although the respondents had described their work as being an evaluation, it did not fulfil our criteria for evaluation. For example, some studies did not have any written plan or protocol. Others simply involved collection of monitoring data without any articulation of the intended benefits of the initiative or whether the data demonstrated or described progress towards any objectives. In other cases, very little information about study methods was available, and it was not possible to conduct meaningful critical appraisal. However, an inclusive approach was taken and proformas were completed for all thirty-five projects as far as the available information allowed.

Case studies

A case study approach was used in order to explore the process of evaluation and ways in which this could be improved. Case study designs combine different appropriate methods to explore relevant contextual features in a way that also allows emerging issues to be incorporated into the study.^{35;36} Case study designs are often associated with small sample sizes, raising concerns about generalisability due to the possible lack of typicality within the sample.^{35;37} However, validity in case study research rests not on typicality, but on cogent theoretical reasoning from the

collected data,³⁷ with the employment of agreed protocols in team based approaches helping to reduce associated risks of researcher bias.³⁵

Our case study design included the reviewing of the documentary evidence and survey data collected in the stages discussed above, together with a series of qualitatively informed semi-structured interviews, which formed the central part of the case study work. Qualitative approaches are ideally suited to accessing areas of enquiry that may not be amenable to quantitative methods, or where exploratory depth, insight and understanding of a specific phenomenon are sought.³⁸ Semi-structured interviews identify key areas to be discussed consistently with all participants, and allows scope for the inclusion and exploration of issues not previously anticipated by the researcher.³⁹

At the outset we envisaged including as case studies all of the multi-site independent evaluations and ten to fifteen local evaluations. As planned, all ten of the commissioned national evaluations were included as case studies. In addition, a process of selection was undertaken to identify a sample of local evaluations to act as case studies, selected to represent each of the initiative areas, and an appropriate variety of methodological approaches, project sizes, local features and geographical locations. This purposive strategy allowed us to ensure that an artefactual emphasis on larger scale projects resulting from the mapping and survey stages was balanced in the case study stage. An initial selection of eleven local projects was made by the research team and data were collected between September 2008 and June 2009. These represented eight of the initiative areas: no potential local evaluations were available for selection for NHS LifeCheck, Social Enterprise Pathfinders or the Common Assessment Framework. However there were several local evaluations of Partnership for Older People Pilots using different approaches, so four of these were selected as case studies. This provided an interesting variation of methodological approach in relation to the evaluation of one initiative.

Each case study sought to include a face to face or telephone interview with the principal investigator ('PI') for the evaluation, a key stakeholder in driving the innovation (the 'service leader'), and an NHS member of staff involved in the initiative. PIs were contacted by letter initially and then followed up with a reminder email and then a telephone call if necessary. We asked PIs to help identify an individual who we could interview and who fitted the role of service lead for the evaluation they were conducting. We then contacted the service lead by letter or email, depending on what details had been provided. The service lead was asked to provide contact details of a member of staff who had worked on the intervention being evaluated. Often the PI was able to provide information on both these individuals and they were therefore contacted concurrently. Each individual was contacted twice in writing; either by email or letter and then followed up with at least two telephone calls. A number of individuals gave consent to be interviewed but then found that they were unable to arrange a date within the time frame of the project. There were also a number of situations in which after speaking to the PI or service lead it was felt that it was not appropriate to contact anyone further about the project, or that there were no individuals felt to fulfil this category.

Table 2-2 Summary of characteristics of selected local case studies, by initiative type

	Self-Referral to Physiotherapy	Whole System for Long Term Conditions	Individual Budgets	Care Closer to Home	Improving Access to Psychological Therapies	POPP				New Types of Worker	Information Prescription
						i	ii	iii	iv		
Mixed methods	✓					✓	✓	✓	✓	✓	✓
Mainly quantitative									✓		
Mainly qualitative		✓	✓	✓							
Action research element								✓			
Larger budget					✓						
Mid budget			✓			✓		✓		✓	
No specific funding	✓										
Rural						✓			✓		
Urban				✓	✓			✓			
Northern regions		✓			✓				✓		
Southern regions	✓									✓	
Mixed socioeconomics							✓				
High elderly						✓	✓		✓		
Ethnic mix								✓			
Vulnerable/ disabled adults					✓						✓

Interviews with the PI covered why the evaluation was being undertaken, choice of evaluation design, data collection and outcome measures, strengths and limitations of the approach chosen, factors which helped and hindered the evaluation and ways in which commissioning, conduct and dissemination of evaluation could be strengthened.

Interviews with the service leader covered issues of local context, what the intervention was intended to achieve (outcome), how this relates to White Paper goals, and the proposed 'mechanism' (what the innovation consists of, and how this is intended to achieve the aims). Service leaders were asked their views about evaluation and how the results of evaluation would be used.

Interviews with staff covered the relevance of the evaluation and the impact of evaluation on their work.

A total of fifty-nine interviews (thirty-four face to face, fifteen telephone) in twenty-one case studies were completed, as summarised in Table 2-3 and Table 2-4:

Table 2-3 Summary of case study interviews by person group

Principal Investigators	22
Service Leads	15
Staff	12

Table 2-4 Summary of case study and key informant interviews by area

Care Closer to Home	5
Common Assessment Framework	1
Improving Access to Psychological Therapies	8
Individual Budgets	9
Information Prescription	9
New Types of Worker	6
NHS LifeCheck	3
POPP	6
Self Referral to Physiotherapy	3
Social Enterprise Pathfinders	6
Whole Systems for LTC	3
N/A	2

Key informant interviews

Interviews were conducted with policy leads and other key informants at the Department of Health, in order to gather information across the full range of initiative areas related to the White Paper. We were unable to complete interviews with key informants for two initiative areas: NHS LifeCheck and POPP.

These key informants at the DH were conceived as the 'customers of evaluation' at a national level, in that they were generally the people who had commissioned the evaluation and/or who were in a position to make use of the information arising from evaluation. Interviews focused on the parameters that determine how and why evaluations are commissioned, difficulties experienced in obtaining evidence from evaluations which is useful for policy making, examples of successful evaluations and factors which characterise them, and ways in which evaluations of health service initiatives can more effectively meet their needs. All of these questions were addressed in relation to the specific initiative for which the informants were responsible.

The completed interviews lasted between 40-80 minutes, and were fully transcribed for analysis with an appropriate computer assisted qualitative data analysis software package. A total of thirteen key informant interviews were conducted.

During the course of the project it became apparent that it would be useful to re-interview these key informants towards the end of the project (generally about a year after they were originally interviewed), despite this having not been part of the original research design. Having asked them in the original interviews about what they hoped to gain from evaluation of their initiative, we wanted to find out whether the evaluation had provided or was providing the type of information that they had hoped for. In addition we wanted to discuss with them their experience of working on an initiative that was being evaluated, and where evaluation findings had become available, how they had used these findings.

We attempted to identify and contact key informants for all ten of the national evaluations. However, adding this additional data collection stage had only limited success, and we were only able to carry out follow-up interviews with two key informants. In two cases we were unable to identify any key informant with specific responsibility for the initiative and its evaluation, in five cases we identified a probable contact but they did not reply despite three contacts from us by email and/or telephone, and we were therefore unable to arrange an interview before the end of this project. In one case, a scheduled interview had to be cancelled and could not be rescheduled before the end of the project.

Analysis of interview data

All interviews were digitally recorded, transcribed verbatim and analysed using a constructive process of thematic content analysis, which involved the audio reviewing, and reading and re-reading of transcripts in order to identify and develop themes and theories identifiable from the data. Analysis of the audio files and transcripts was managed using the software package QSR NVivo8 (QSR Victoria, Australia). Primary responsibility for the data analysis was undertaken by one member of the team (KS), with analyses validated using a process that echoes 'inter-rater reliability', whereby other members of the research team reviewed the transcripts independently to develop the thematic framework, and then subsequently contributed to its ongoing development. The involvement of additional experienced qualitative researchers can help to guard against the potential for lone researcher bias and help to provide additional insights into theme and theory development.⁴⁰ Qualitative analysis is an ongoing, iterative process. Early data collection and analysis informs sharpened subsequent data gathering and analysis, which in turn informs a further refining of ideas. Thus in qualitative data collection, unlike quantitative designs, insights from early interviews should and do refine subsequent interview schedules.⁴¹

Accordingly, the following analysis stages were undertaken:

1. The identification of initial thematic areas from the research aims and interview schedules
2. Initial familiarisation with interview data to inform the identification of further thematic areas, and the iterative development of later data collection
3. Coding of transcripts and development of more detailed thematic codes
4. Reflective discussion of emerging coding, further re-interrogation, refinement and application of coding categories
5. Identification of relationships between coding categories
6. Development of key themes identified by these categories and relationships

Synthesis

Using the four elements of data collection described above we obtained a range of quantitative and qualitative data about the evaluations of White Paper initiatives. We then synthesised this information in three main ways.

First, we created structured abstracts to summarise the evidence which had been obtained from evaluation of each of the ten White Paper initiatives. Where evaluations had not yet been completed we sought to describe the type and strength of evidence that would be available in due course, assuming that the evaluations continued as planned. Within each of these structured abstracts we described the background to the initiative (pre-existing evidence or policies which had informed the development of the initiative), the evaluation of the initiative, the findings from evaluation, and the strengths and limitations of the evidence arising from evaluation. These structured abstracts were based on information obtained from all possible sources, including protocols, responses to our questionnaire survey, and interviews with principal investigators. The abstracts focused on the national evaluations but also briefly described other local evaluation work we identified.

Second, we synthesised the evidence from across the different initiatives about the extent to which these initiatives were promoting the four main policy goals of the White Paper. This information came largely from the structured abstracts described above. For each policy goal we highlighted specific examples of achievements (i.e. where there was concrete evidence of an initiative supporting the policy) and also examples of challenges (where evidence suggested the policy was not being achieved) and unanswered questions (priorities for further research and evaluation).

Third, we synthesised the information coming from both the quantitative survey data and the qualitative interviews to explore lessons about how evaluation is commissioned, undertaken and used within health and social care. Methods for undertaking this type of synthesis are not well established, although increasingly discussed.^{42;43} We conducted the synthesis through repeated reading of the qualitative and quantitative findings, and discussion of the key themes arising in meetings of the research team. This team included members from a variety of disciplinary backgrounds and with different areas of methodological expertise. The purpose of this analysis was to increase understanding of the most effective ways of conducting evaluation in order to generate findings which are likely to be useful to inform policy.

Finally, we used the findings from this synthesis to make recommendations and provide guidance to those who commission and conduct evaluation of innovations in health and social care.

3. Types and strength of evidence provided by evaluations of initiatives

This brief chapter summarises the findings from our questionnaire survey of initiatives identified in the mapping exercise.

It is interesting to summarise the range of organisations conducting the different national evaluations, and the different ways in which the evaluations were commissioned. This is shown in Table 3-1 below. It can be seen that most were commissioned directly by the Department of Health, although our interviews (reported later) suggested that was only sometimes with the support of the Policy Research Programme. Just over half the evaluations were led by academics from universities, while three were commissioned from consultancies.

Table 3-1 Organisations commissioning and conducting national evaluations

Evaluation	Commissioned by:	Conducted by:
Care Closer to Home	DH Policy Research Programme (PRP) with NPCRD	University – Manchester
Increasing Access to Psychological Therapies	DH PRP with SDO	University – Sheffield
Individual Budgets	DH PRP	University – York
Information Prescription	DH	Consultancy – OPM
New Types of Worker	Skills for Care and additional funding from DH	Universities – Oxford & Sheffield Hallam
NHS LifeCheck	DH	Teen LifeCheck: University – Institute of Education. Early Years & MidLife LifeCheck: Consultancies (People Partnership, Bunnyfoot, Fatherhood Institute)
POPP	DH PRP	University – Warwick
Self-Referral to Physiotherapy	DH	Analysed retrospectively by NHS Quality Improvement Scotland
Social Enterprise Pathfinders	DH PRP (Social Enterprise Unit)	Consultancy - Newchurch
Whole System Demonstrators	DH PRP	University – University College London

As described in the preceding chapter, we originally obtained questionnaire responses describing thirty-five evaluations. After including further evaluations about which we were able to obtain documentary evidence such as protocols or reports (n=3), we were able to appraise thirty-eight evaluations.

These evaluations were categorised according to the type of domains or outcomes that were studied within the evaluation. As can be seen from Table 3-2, the most common type of outcome studied was patient experience or satisfaction. More evaluations examined processes of care than examined outcomes of care. Less than a third of the evaluations collected data about costs.

Table 3-2 Domains of evidence provided by evaluations

What domains will this evaluation provide evidence about?	Yes	No
Effectiveness (the impact of the initiative on patients' health or well-being)	14 (45)	17 (55)
Patient/client views, experience or satisfaction	24 (77)	7 (23)
Staff views, experience or satisfaction	17 (55)	14 (45)
Commissioners'/ managers' views, experience or satisfaction	6 (19)	25 (81)
Process of care (e.g. numbers of people receiving intervention, number of referrals)	21 (68)	10 (32)
Costs	9 (29)	22 (71)
Cost-effectiveness (costs in relation to outcomes)	6 (19)	25 (81)
Organisational structure or function	9 (29)	22 (71)
Not clear or none of the above	7 (18)	31 (82)

We then considered whether the study fulfilled our original criteria for an evaluation, bearing in mind that we included all projects in this analysis which were considered to be an evaluation by the research teams conducting them. These criteria were chosen to represent minimum criteria for any meaningful evaluation.

In a substantial minority of cases it was difficult to answer these questions because the information provided (mainly with regard to the local 'evaluations') offered only limited or vague details of the methods which were being followed. It is notable that a substantial minority of the evaluations did not appear to have a written protocol or plan. In some cases the descriptions of evaluations provided almost no details of the methods used to collect data. Many of the descriptions of evaluations did not specify the outcomes or benefits that were intended to flow from the initiative and which were being assessed.

Table 3-3 Whether evaluations fulfilled our criteria for evaluation

Criteria for evaluation	Yes*	No*	Not clear
Is the initiative intended to address White Paper goals?	36 (100)	0	2 (5)
Is the evaluation based on a written protocol or evaluation plan?	22 (67)	11 (33)	5 (13)
Does it include the collection and analysis of quantitative or qualitative data?	31 (97)	1 (3)	6 (16)
Does it measure progress towards pre-specified outcomes or benefits?	26 (87)	4 (13)	8 (21)
Is the project intended to result in a formal evaluation report?	28 (90)	3 (10)	7 (18)
Does this study fulfil our original definition of an evaluation? (tick yes if all of the above are yes)	20 (67)	10 (33)	8 (21)

* Percentages are of those in which it was possible to state whether or not the evaluation fulfilled the criteria, excluding 'not clear'.

We also sought to categorise each evaluation into broad categories to describe whether the evaluation, according to its objectives, was intended to be summative, formative, or action research. It was interesting to note that the majority of evaluations were worded to suggest that the intention was to make a summative judgement of the effectiveness or worth of an initiative, but in many cases the design of study or the data to be collected would not provide this type of information.

Table 3-4 Whether evaluations were intended to be summative or formative

Is the study intended to be:	Yes	No
Summative (to reach a judgement about the effectiveness or worth of an initiative)	22 (79)	6 (21)
Formative (designed to find ways to improve the initiative)	13 (46)	15 (55)
Action research (designed to solve a specific problem)	3 (11)	25 (89)
Not clear, or none of the above	10 (26)	28 (74)

We also attempted to categorise the evaluations according to the overall research design. It is interesting to note that the majority of evaluations were based on mainly qualitative methods. This illustrates the discrepancy between the aims of the studies, summarised above, and the methods used, as qualitative methods are not an appropriate way to assess effectiveness or value, although they may help to explain how and why participants value an initiative.

Table 3-5 Research designs used in evaluations

Which of the following descriptions best summarises the research design	Yes	No
Randomised controlled trial	2 (7)	26 (93)
Quantitative controlled before and after study	5 (18)	23 (82)
Quantitative before and after study without controls	4 (14)	24 (86)
Quantitative observational study (after only, without controls)	10 (36)	18 (64)
Interrupted time series design	0	28 (100)
Qualitative, based on interviews or focus groups	19 (68)	9 (32)
Action research	2 (7)	26 (93)
Case study approach	6 (21)	22 (79)
Economic analysis	6 (21)	22 (79)
Documentary analysis	1 (4)	27 (96)
Rapid appraisal	0	28 (100)
Not clear, or none of the above	10 (26)	28 (74)

Finally, we identified a range of characteristics which were key questions within the questionnaire sent to PIs and which were intended to reflect minimum standards for the conduct of an evaluation. We examined the extent to which the studies met these standards. Again, the studies were often poorly described and specified so it was not always possible to answer these questions from the information provided. However, it is clear that many of these minimum standards are not being met, in particular some evaluations did not have a written analysis plan, and in many cases those conducting the evaluation were not fully independent.

Table 3-6 Key methodological characteristics

Other characteristics	Yes	No	Not clear
Does the evaluation have explicit aims or objectives?	29 (91)	3 (9)	6 (16)
Is the methodological approach appropriate to the aims and objectives?	22 (85)	4 (15)	12 (32)
Is there a written analysis plan?	15 (60)	10 (40)	13 (34)
Does the study involve triangulation of data from more than one source?	22 (76)	7 (24)	9 (24)
Is there (or going to be) a formal final evaluation report?	28 (90)	3 (10)	7 (18)
Is the evaluation intended for national, public dissemination?	25 (83)	5 (17)	8 (21)
Is the evaluation independent of any body commissioning, promoting, planning or providing the initiative?	10 (32)	21 (68)	7 (18)
Does the study relate to a single site?	12 (44)	15 (56)	11 (29)
Does the study relate to multiple sites?	15 (56)	12 (44)	11 (29)
Is the study likely to be generalisable to a range of sites?	22 (81)	5 (19)	11 (29)
Is the study likely to be useful to decision makers?	24 (89)	3 (11)	11 (29)

Summary

The purpose of this exercise was to determine the types of approaches to evaluation which are being conducted both locally and nationally with regard to White Paper initiatives, and to identify common themes in terms of the strengths and weaknesses of these evaluations.

The exercise proved to be less useful than we had hoped. Some of the respondents to our questionnaire provided relatively little information about their methods. It became clear that many of those conducting local evaluation had little or no experience of research or evaluation. Our questions asked about concepts such as summative and formative evaluation, or about types of study design. Although we explained each of these terms in standard English throughout the questionnaire, and piloted the questionnaire, it appears that some of the respondents found the questions hard to understand or to apply to their own project. This is probably because it became clear that some of the studies were not evaluations, but were simple audits or patient surveys without any clear objectives, therefore many of our questions would not have been entirely applicable. Some of these small local audits and surveys were not intended for a wide audience, and this may be why the processes of conducting the study had not been formalised in any way. However, even these small local studies would probably have been improved if they had been based on a written plan, since this would have increased clarity of purpose and reliability of data collection, and they would have been more likely to be helpful to local decision makers if the findings had been written up into a report.

The overall impression was that a number of NHS organisations were collecting data because they were interested in the performance of their local service. However several of the evaluations had ambitious aims which could not possibly be achieved using the methods they had chosen to approach the evaluation. In addition there was often a lack of clarity about how data would be collected, analysed or used. The overall effect was that much of this effort could have been more productive if local evaluation teams had better access to (or had taken advantage of) advice about basic study design.

It is important to bear in mind that our findings probably reflect only a minority of the work being conducted within the NHS which is considered to be 'evaluation'. It is likely that those people who completed our questionnaire were conducting more formalised evaluation than those who did not reply, and our results include the national evaluations of White Paper initiatives.

4. Summary of evidence about each of the White Paper initiatives

Care Closer to Home

Background

The White Paper identified a need to move from health and social care that is provided in institutions such as hospitals and residential care homes to a situation where care is provided in more local and convenient community settings.

Key policy drivers for the Care Closer to Home programme were:

- Public expectation of greater independence, more choice and control⁴⁴
- Changing technology that allows the patient pathway to be planned so that specialist skills are integrated within it^{45;46}
- The ageing population which is an economic driver to focus on prevention and supporting individuals in the community rather than in institutions⁴⁷
- A need for the NHS to focus on delivering better care with better value for money⁴⁸

A number of national and international models were seen to be exemplars of the Care Closer to Home approach. These included the Kaiser Permanente health care system in the USA and the reintroduction of the concept of polyclinics in Germany, where specialist care is provided outside hospitals and multidisciplinary teams provide coordinated care. In Kaiser Permanente lower utilisation of acute bed days is achieved through integration of care, active management of patients, the use of intermediate care, self care and medical leadership.

The exemplar services promoted by the White Paper in the NHS included community based anticoagulation nurses, intermediate care services which provide some specialist care in primary care settings (e.g. musculoskeletal services) and direct access to audiology for diagnostic, monitoring and treatment services. For example, in a service offering direct access audiology, a 90% reduction in referrals to ENT was achieved with children being treated in six weeks as opposed to six months.¹ The views of the public were also considered; 54% of participants in the 'Your health, your care, your say' consultation were in favour of providing more care in community settings.

Patient choice and practice-based commissioning were seen as key tools for delivering the Care Closer to Home agenda. Pathway redesign and review of the place of delivery of care were central to delivering this programme. The role of health care professionals with a special interest (PwSI), including general practitioners with a special interest (GPwSI), was also seen as crucial. The programme reflects a goal of shifting care from the acute sector to the community, including to a new generation of modern NHS community hospitals offering diagnostics, day care surgery and outpatient facilities. There was also a desire for greater integration between health and social care organisations through initiatives such as health living centres and also more widely, for example with local authorities to ensure public transport is available locally.

In order to reinforce the shifts to community care and increased preventative care, from 2008 PCTs were scrutinised annually for achievements against these goals in terms of both spending and achievement of targets.

Aims

The aims of the initiative were to reduce demand on acute hospital services, while also improving patient choice and convenience. These aims related to the wider policy goals which are outlined above. In the White Paper the emphasis is placed on improving choice and meeting patient expectations. However, a further underlying theme of the initiative appears to be driven by the need to shift care from the acute sector because of the ageing population, pressure on services and economic drivers to reduce costs. Drivers of this shift included patient choice, the tariff, practice based commissioning (PBC) and PCT commissioning. This shift was also intended to increase commitment to spending on preventative services. The UK historically spends less than 2% of total health spend on prevention and public health, compared with 5% in the Netherlands.¹

The benefits anticipated were that the new models of care would be more effective than those previously offered. The demonstration sites established by the Care Closer to Home initiative would provide PCTs and GP practices with evidence that would enable them to commit to fundamental service redesign and to the development of more local models for delivering care. The demonstration projects were also intended to inform development of: care pathways (for example for use in national framework contracts); a tariff based on best practice delivered in a community as opposed to acute setting; performance measures for PCTs and future multi-skilled workforce requirements.

The initiative

The White Paper committed to funding demonstration sites for the Care Closer to Home initiative. The aim of these pilots was to ensure that clinicians, PCTs and local authorities worked together to ensure that these sites provided transferred care and did not create demand for new services. Five demonstration sites were identified with Closer to Home services in each of six specialties: ENT, orthopaedics, dermatology, urology, gynaecology and general surgery. These sites were already in operation at the time of selection for the evaluation and were spread across England. The sites adopted various approaches including consultant led clinics provided in community settings, surgery led by GPwSIs, nurse led services, dedicated telephone follow up systems and home chemotherapy. Evaluation of the Care Closer to Home Demonstration Sites took place between September 2006 and May 2007.

The evaluation

The Department of Health funded an overall programme to evaluate and report results on a consistent basis across all the demonstration sites. This report was commissioned from the National Primary Care Research and Development Centre at the University of Manchester.⁴⁹

Aims and Objectives

In thirty Closer to Home sites covering six specialities, the aim was to demonstrate the range of ways in which secondary care services might be moved to community settings.

Objectives were to:

1. Describe the structure and organisation of Closer to Home services
2. Identify factors that facilitated or impeded their set-up
3. Assess the impact of the services on access
4. Assess the impact of the services on quality of care
5. Assess the impact of the services on costs

Design

This was an observational study which included:

1. Face to face interviews with the service manager and other stakeholders at each demonstration site to describe the set-up of each service, its operation, and the perceived impact that it was having.
2. Telephone interviews with key stakeholders (practitioners, commissioners) in twelve of the demonstration sites (two per speciality).
3. A postal questionnaire survey to 1500 patients who had attended demonstration sites services to ascertain their views of service access, quality and coordination of care.
4. An economic evaluation in six sites (one per speciality).
5. A postal survey of 'control' patients attending conventional outpatient clinics in the specialties included in the Closer to Home sites (but selecting areas where only conventional services were being provided for those specialties). However, data are only available from six control sites. These did not include all specialties in the Closer to Home services.

Analysis

Qualitative interviews were audio recorded, transcribed and thematic analysis was undertaken. Analysis of quantitative patient survey data involved regression modelling to examine the relationship of study site (demonstration vs. control) to patients' views of service access, quality and coordination after controlling for potential confounding variables such as age and sex. Economic analysis was based on anonymous patient-level data to quantify activity levels, estimating capacity and utilisation by clinical role, and cost per case. Cost data provided by each site were used as a basis for estimating the cost per consultation and cost per patient. National tariff prices for 2006/07 and reference cost data for 2005/06 were used as a basis for estimating what the activity of Closer to Home sites would have cost if undertaken by secondary care providers. The evaluation used a previously described analytical framework to classify the different types of shifted services from hospitals to the community.⁴⁹

Evaluation team's findings and conclusions

1. Challenges in service design

Three main challenges were found in designing new services:

- Deciding what services to move Closer to Home
- Finding and equipping a suitable venue

- Changing health professional roles

2. Factors facilitating implementation

The factors said to facilitate service implementation included:

- Learning gained from similar services elsewhere
- Local service champion(s) and continuity of leadership
- Positive prior working relationships among key stakeholders
- Prior stakeholder consultation

3. Impact of the new services on access, quality and cost

- Services established to reduce waiting lists were generally successful.
- Patient safety was the central consideration in services where care was delegated from consultants to PwSIs
- Patients reported positive evaluations of the quality of care they received
- Closer to Home services were often apparently cheaper than services provided in hospital outpatient clinics

4. Workforce implications

- Existing nurse guidelines may be insufficient to address their new roles in Closer to Home services
- Better PwSI training programmes
- Need for community based specialist services to become accredited as locations for the training of junior doctors

5. Impact on the wider health economy

- Impact of expansion of Closer to Home services on acute hospitals is uncertain
- Rates of referral for specialist care might increase
- Direct access for patients to specialist services without GP referral could have a significant impact on overall demand

Conclusion

Closer to Home services demonstrate that there is considerable potential to move care into the community and so improve access and convenience for patients. However, important issues of quality, safety, cost and staff training need to be considered as community based services are expanded.

Comments on the evaluation

The national evaluation of the Care Closer to Home employed mixed methods to address the varied and complex initiatives included in the programme. The participants, although based on convenience samples in most cases, were representative of the target groups and sample sizes and follow up of greater than 40% of patients on average appears to have been adequate. The risk of bias in the quantitative data was minimised although the risk of confounding is harder to assess as the comparison group for the patient survey was not drawn from the same population or health economy. In the economic analysis, the additional assessment of opportunity costs and a sensitivity analysis would have been informative.

Our interpretation of the findings

The findings are limited by the lack of a formal comparison group, which places constraints on the validity of the findings. Based on the available information, developments were often driven by the enthusiasm of local clinicians and the availability of a suitable venue, as much as by healthcare needs. Not surprisingly, sites that had local champions and that consulted widely with stakeholders found it easiest to establish services. Some sites creatively redesigned services to break down traditional barriers between primary and secondary care. However, it is worth noting that many of the projects evaluated as part of the Care Closer to Home were in place prior to the publication of the White Paper.

The aim of improving patient satisfaction appears to have been achieved. Patients found the new local services more convenient, they experienced shorter waiting times, and they were happy with the quality of care they received. However, no objective measures of quality, outcome, or competency were available. This is concerning, as care is being transferred from one type of practitioner to another and from centralised units to smaller peripheral centres. Both hospital specialists and some of the community practitioners expressed concern about this matter. In particular, some nurses were worried about their lack of training for the new responsibilities they had been given. Training needs and accreditation criteria have now been defined for general practitioners and pharmacists, but for nurses this is an ongoing problem that needs to be resolved, and robust arrangements to audit quality and outcomes are essential.⁵⁰

The cost-effectiveness analysis is limited by the lack of assessment of opportunity costs and a sensitivity analysis. In addition, the implications for the wider NHS are not fully explored. Under payment by results, hospitals are paid using a fixed national tariff, which is based on an estimate of the average cost of providing care within broad categories such as general surgery outpatients.⁵¹ A standard tariff is used to encourage providers to focus on quality and quantity of referrals rather than price. However, many of the demonstration sites provided care for simple procedures at well below the tariff. This does not mean that these demonstration sites necessarily represent better value, as it is important to distinguish between the cost of providing a service and the price hospitals have to charge. Diverting low cost cases, on which hospitals make a profit, while leaving them with the complex and expensive cases, on which they make a loss, is unsustainable.

In addition, most demonstration sites were designed to increase capacity so that waiting lists could be cut; this represents an additional cost. Savings can be made only by disinvesting in hospitals, but if the marginal cost of providing low complexity care in hospitals is less than the cost of establishing new services in the community this may not ultimately be good value. Previous studies have shown that care in the community is generally more costly than hospital based care.^{52;53} The increased capacity, accessibility, and popularity of closer to home services are likely to lead to an increase in demand, particularly if (as in some cases) these services provide direct access for patients without referral from a general practitioner. Therefore, this policy could actually increase total costs to the NHS.

Finally, a tension exists between promoting patient choice and providing value for money. The demonstration sites seemed to be designed to increase choice for commissioners rather than for patients, because they often involved triage of patients referred for secondary care. Some patients prefer to attend hospitals, and it is unclear

whether commissioners allow this choice to be exercised if the price of hospital care is higher.⁵⁴

Overall, the evaluation appears to support the notion that there is potential to move care into the community, with the caveats stated above and by the evaluation team. However, most of the demonstration sites were of small scale, and the evaluation provides limited evidence about the costs and benefits of the policy. It highlights the need for careful attention to implementation, costs, quality, and training as the policy is rolled out more widely.

Other relevant evaluation work

Two other local evaluations of Care Closer to Home were identified. The first of these was a research report resulting from an evaluation undertaken towards a higher degree. This addressed the implementation of an intermediate care service at a single site and used a qualitative approach with the intention of providing information for local stakeholders, particularly commissioners. This evaluation identified a number of issues relating to the organisation of the new service and the difficulties associated with delivering the commissioning goals within the constraints of the local health care economy. The second potential project identified did not fulfil our criteria for an evaluation.

Conclusions about the evidence that the initiative achieved its objective and delivered policy goals

The national evaluation described contains limited information that the Care Closer to Home Demonstration Sites were making progress towards the goals of the White Paper. In terms of the aim to reduce demand on acute hospital services, Care Closer to Home had an impact by reducing waiting times for patients but the impact on acute hospital services was uncertain. Further aims were to improve patient choice and convenience. No conclusions can be drawn on patient choice but patient feedback was positive in terms of the quality of care received. In terms of the underlying theme of the initiative to shift care from the acute sector because of pressure on secondary care services and economic drivers to reduce costs, the care delivered by Care Closer to Home Demonstration Sites appeared to be less costly than the hospital outpatient equivalent. However, this conclusion should be interpreted with caution for the reasons discussed above. In addition, there was a concern that demand may be increased by the new services.

It was expected that the new models of care would be more effective than those previously offered. However, the lack of evidence available so far about the effect of Care Closer to Home on patient outcomes is an important omission. It was also hoped that the success of these demonstration projects would encourage local service redesign in other areas. The national and one local evaluation included are not supportive that this expectation will be met. There is no evidence that the initiative resulted in increased commitment to spending on preventative services.

In summary, the evaluations suggest that there is potential to move care into the community and there are some exemplars of this. However, there are concerns about workforce, organisational and economic impacts of Care Closer to Home as a wider initiative.

Improving Access to Psychological Therapies

Background

The Improving Access to Psychological Therapies (IAPT) programme was launched in May 2006 and sought to deliver on the Government's 2005 manifesto commitment to improve access to psychological therapies. This was in response to a recognition of the imbalance between the high prevalence of common mental health disorders such as depression and anxiety and the limited availability of effective help. The aim was to provide better access to a choice of evidence-based psychological therapies for people with depression and anxiety disorders.

A key precipitant for the IAPT programme was an influential report published by Professor Richard Layard, an economist at the LSE and life peer in the House of Lords.^{55;56} Layard's report highlighted that as well as having an impact on individuals and their families, common mental disorders also have an impact on society particularly because of the impact on employment.

Mental health disorders are extremely common and disabling. Approximately 16% of adults report suffering from depression, anxiety, phobias or related conditions at any one time. The management of these disorders accounts for a quarter of all consultations in primary care and expenditure of over £400 million on prescriptions for anti-depressants.

Layard estimated that the cost to society of depression and anxiety was approximately £25 billion, made up of lost production, health service costs and carers' time, representing some 2% of GDP. The cost to the exchequer of benefits resulting from these disorders was approximately £10 billion. Mental health problems were the primary reason for 40% of people receiving incapacity benefit (IB), and a secondary reason for a further 10%. Fewer than 10% of these people were in contact with specialist mental health services. As many as 90% of those with mental health problems who are unemployed would like to work and work has been shown to have a beneficial effect on mental health, while unemployment has been found to adversely affect mental health.

Layard also pointed out that there are a range of treatments for common mental health disorders with good evidence of effectiveness, including antidepressants for moderate and severe depression, and structured psychotherapies such as cognitive behavioural therapy (CBT). NICE guidance on the management of depression^{57;58} and anxiety⁵⁹ in primary care recommends that patients with mild depression should be initially offered guided self-help, including problem-solving, brief CBT or counselling and more formal therapies (e.g. courses of CBT over six to nine months) should be offered to people with depression who have not responded to other treatment. These NICE recommendations endorsed the earlier recommendations of the National Service Framework (NSF) for mental health which emphasised the importance of providing accessible, high quality services offering effective evidence-based treatment.⁶⁰ A review of progress on the NSF highlighted the need to improve access to services, increase availability of psychological services, and to place more priority on patient choice.⁶¹

However, the stepped care model advocated by NICE, which included greater use of psychological therapies, could not be implemented without a significant increase in the number of trained therapists and better models of organisation. Layard argued

that 10,000 extra therapists were needed to provide people with the choice of psychological therapy, and that this expansion should be under-pinned by a network of treatment centres led by specialists. These would provide supervision, training and support on a hub and spoke model. Treatment centres would provide rapid access to people who need them, backed by waiting time targets. Although this would involve a considerable investment of resources, Layard's thesis was that this cost would be recouped by helping people come off IB. He highlighted evidence from the pilots of the Pathways to Work programme (which included provision of CBT and problem solving approaches) showing a 50% increase in the numbers leaving IB, mostly into employment.⁶²

Layard's report, which was endorsed by a range of influential NHS and academic signatories, set out a clear 'call to action', insisting that most people with mental illness should be offered the choice of psychological therapies.⁵⁶

Aims

As summarised by the White Paper progress report 'Our health, our care, our say: making it happen'³ the purpose of the national and regional demonstration sites in the IAPT programme was to 'collect evidence of effectiveness of providing increased access to psychological services, assessing:

- clinical effectiveness and well-being
- employment status and economic impact
- provision of service user choice
- service user satisfaction

In the case of the IAPT initiative, the mechanism by which these benefits should be achieved was very clear, thanks to the hypothesis set out in Layard's report. This hypothesis can be summarised as follows. Mental health disorders are very common, and are a very major cause of people being unemployed and claiming benefit. This has adverse consequences for the economy as well as for the individuals concerned and their families. There is good evidence that psychological therapies are effective, but provision is insufficient to meet the need. If access to services was considerably expanded then more people would receive effective treatment, which would enable them to recover, come off benefit and return to work. The economic benefits would more than pay for the investment in the expansion in mental health services. The evaluation aims described are therefore appropriate in assessing the domains which are key steps in the hypothetical mechanism of action for the IAPT programme.

The initiative

The Improving Access to Psychological Therapies programme originally comprised:

- Two national demonstration sites, jointly sponsored by the Department of Health and the Department for Work and Pensions.
- A national network of twenty-three local improvement programmes led by the Care Services Improvement Partnership (CSIP) Regional Development Centres and supported by the National Institute for Mental Health in England (NIMHE). These programmes sought to improve access to psychological therapies by re-organising existing services without the investment of significant extra resources.

- A partnership with the Primary Care Collaborative to improve management of common mental disorders within primary care and to support self-care.

The two demonstration sites were established in Doncaster and Newham. These developed different approaches to expanding the provision of psychological therapies, making it possible to learn lessons from different approaches in different settings.

Doncaster delivers high volume, low intensity treatments as part of an integrated stepped care model in which trained non-medical care co-ordinators have a pivotal role. They work with GPs, care managers in the Pathways to Work programme, and support employers. Newham provides enhanced services in primary care settings, a treatment centre for people with moderate anxiety or depression and employment support. A key challenge in Newham was ensuring that services were accessible for the high number of people from black and minority ethnic groups in the area.

In 2007, the IAPT programme was extended via an additional eleven Pathfinder sites across England. The aim of these sites was, according to a CSIP report published from the Department of Health,⁶³ 'to provide additional evidence and experiential learning to support the case, including affordability, for continuing the roll-out of psychological therapy services for people with anxiety and depression.' The report went on to state that 'Clear evidence will be required to demonstrate that psychological therapies can deliver real and significant benefit in terms of health and well-being gains to patients, healthcare efficiencies and resource savings in the NHS'. The IAPT programme website claims that this extra investment of £173m over three years will allow 900,000 more people to access treatment, with half of them moving to recovery and 25,000 fewer on sick pay and benefits, by 2010/11.⁶⁴

A particular expectation for the new Pathfinder sites was that they should enhance social inclusion by targeting one or more of a range of particular client groups, such as older people or those from black or ethnic minority groups. The Pathfinder groups were required to collect routine data to provide evidence of performance in four domains: health and well-being; choice and access; inclusion (including employment); patient experience.

The evaluation

Aims and objectives

The national evaluation of the two demonstration sites had three objectives:

1. To discover whether providing improved access to psychological therapies at two national pilot sites is cost effective when compared to providing the existing level of psychological services at comparable sites.
2. To identify the organisational implications of implementing improved access to psychological therapies within an existing mental health system.
3. To find out how acceptable patients find the extended service delivered at the pilot sites, and what their experience of the service is.

Design

The evaluation was designed to test one central hypothesis, namely that access to primary care CBT psychological therapies via newly configured service delivery models implemented at two demonstration sites will be cost-effective when compared with existing organisation of services drawn from comparator sites. The study utilises

and integrates both quantitative and qualitative methods and comprises three key strands:

1. evaluation of the comparative costs and outcomes of the demonstration sites and control sites
2. evaluation of the organisational system impacts
3. study of the acceptability and patient experience of the demonstration services

The first strand, the costs and outcomes study, was a prospective controlled comparison. Within each of the two pilot sites (Newham and Doncaster) a representative sample of the GP practices making referrals to the pilot service were identified. Within these practices, a cohort of patients who met the criteria for referral to psychological treatments for anxiety or depression after the start date was identified. Outcomes for these patients were compared with outcomes for patients from comparator sites where there is no IAPT pilot service. The comparator cohorts were recruited from GP practices in each of four other Primary Care Trust (PCT) sites, two matched to Doncaster and two to Newham on a range of factors such as levels of deprivation and ethnic diversity. The same criteria were used to select patients in the cohort at both demonstration sites and all four control sites i.e. patients who were recognised by GPs as suffering anxiety and/or depression, who had already been offered two primary care interventions or one intervention if employment, housing or physical health were at risk. The aim was to recruit 360 participants at each demonstration site and 180 at each comparator site. Data were to be collected via postal questionnaire at baseline and after three and twelve months follow-up. These questionnaires included a comprehensive battery of validated outcome measures relating to mental health, quality of life and satisfaction with services received.

In addition to this prospective cohort study, a retrospective comparison was made of diagnoses, treatments and referrals for anxiety or depression amongst practices participating in the study in any of the six PCTs. These data were obtained using routine data collected in GP record systems, and related to the period between 2004 and 2008 (before and after the introduction of the IAPT services). The outcomes of the cohorts were also compared ('benchmarked') to existing large datasets using some of the same outcome measures, to place the activities and outcomes in the demonstration sites into the context of routine primary care and employment support practice.

Strand 2, the systems impacts study, was a descriptive study of the nature of the services in each demonstration site and an account of lessons learned during their implementation. Key informants from the commissioners, managers, referrers, practitioners, service user groups, etc. who were involved in setting up the service were interviewed. In addition, routine data gathered by the pilot services on their whole caseload (not just the cohort members) were inspected to collect data about waiting times, drop-out rates, length of treatment, etc.

Strand 3, the patient experience strand, had both qualitative and quantitative elements. Data about satisfaction were obtained from people participating in strand 1 through the questionnaires at three and twelve months follow-up. Selected members of the cohort, chosen to represent a range of participants, were also asked to take part in a face-to-face interview to explore their experiences and perceptions of the services they received.

Evaluation team's findings and conclusions

At the time of writing, no findings have been reported from this evaluation.

Comments on the evaluation

This appeared to be a well-designed evaluation, conducted by an experienced team, which was planned in response to a commissioning process managed by the NHS SDO R&D Programme. However, there will be inevitable and important limitations on the extent to which it can provide reliable evidence about the costs and benefits of the IAPT demonstration sites. In particular, the demonstration sites were placed in areas which had particular needs and socio-demographic characteristics, making it difficult for the evaluation team to identify suitable comparator populations. This also makes it difficult to generalise the findings with different population characteristics. In addition, the plan to obtain data about outcomes from a large sample of patients with depression and anxiety via postal questionnaires is likely to be challenging. Similarly, it is well recognised that it is very difficult to extract reliable routine data from GP computer systems especially when there has been no long term commitment to consistency in data collection and coding.

These problems appear to have arisen because the timeline for commissioning the evaluation was inconsistent with the very rapid set-up of the demonstration sites themselves. By the time the evaluation was commissioned the demonstration sites had already been set up, and there was little opportunity for the evaluation team to influence the data about activity that were routinely collected. The concept of IAPT programme demonstration sites could have provided an excellent opportunity for robust evaluation which would have provided good evidence to support or refute the Layard hypothesis, for example by randomising practices to have access to the new services with other practices in the same area acting as controls. In addition a larger evaluation budget would have been needed to achieve high rates of patient follow-up. The postal questionnaire method planned for this evaluation is likely to lead to low follow-up rates leading to concerns about response bias. In the light of the resources invested nationally in the IAPT programme (more than £173M), and the potential benefits for the national economy if improved access to psychological therapies does indeed help many people recover from depression and anxiety and come off benefits and back into work, a greater investment in evaluation at an earlier stage would have been justifiable. This opportunity has now probably been missed, meaning that the evidence about the costs and benefits provided by the IAPT demonstration sites will be much more uncertain.

The rapid set-up of the demonstration sites appears to have been driven by the desire to produce evidence of benefit in time for the comprehensive spending review of 2007. The language used in the policy documents from 2006-7 (for example the CISP document on commissioning local IAPT services⁶³ described on page 37) strongly suggest that the intention was to find evidence to support the Layard hypothesis, in order to justify greater investment and expansion of the programme, rather than to test whether the hypothesis was correct.

Our interpretation of the findings

No findings are yet available.

Other relevant evaluation work

As well as the evaluation commissioned by the NHS SDO R&D programme, Clark, Layard and Smithies conducted a rapid audit of the impact of the demonstration sites in their first year of operation, based on data about the process of care, the number and flow of referrals and some before and after data about individuals contacting the services.⁶⁵ This rapid evaluation demonstrated that (i) high numbers of people were accessing the demonstration sites, (ii) there were good rates of recovery in those patients treated (iii) there was an increase in rates of employment amongst those accessing the sites, but no decrease in the number of people claiming benefits. There were no data from comparable concurrent control patients, which makes it difficult to draw clear conclusions from these findings.

The authors compare their findings with benchmarks from other studies to show that recovery rates amongst patients contacting the demonstration sites are higher than natural recovery rates. However this approach has important limitations. The analysis of the IAPT is not based on the principle of 'intention to treat' (i.e. including all eligible patients in the analysis). It is not surprising that there are good recovery rates in those patients who stay in treatment and it is important to note that a considerable proportion of patients dropped out of treatment. Patients who contact the IAPT demonstration sites are not necessarily comparable with those included in other studies used as benchmarks.

A report describing the experiences of the eleven Pathfinder sites has also been published.⁶⁶ This report uses a more qualitative approach to describe emerging patterns about the apparent impact of these Pathfinders, lessons about improving data collection for evaluation and recommendations for those commissioning services in future. Key findings were that most patients are referred by GPs (rather than self-referred), that patient experience was generally very positive, and that patient outcomes were broadly in keeping with studies of other interventions. The report provides a clear summary of the data obtained from the sites, with useful recommendations about priorities for improving data collection. It is important to note however that completion rates of some outcome measures were low. In particular, caution is needed because it is well recognised that observational studies such as this often suggest benefits that are not supported when subsequent evaluations with more rigorous methods and incorporating randomised control groups are conducted.

Finally, we identified four projects being conducted locally in PCTs, which were described as evaluations of IAPT initiatives. Two of these fulfilled our criteria as an evaluation, and one was studied in more detail as a case study. This evaluation comprised interviews with patients about their experience of using the IAPT service, interviews with stakeholders about the implementation of the service, and monitoring data about the extent to which the service met its performance targets e.g. in relation to waiting times.

Conclusions about the evidence that the initiative has achieved its objectives and delivered policy goals

There is considerable attention being given within the IAPT programme to the collection of comprehensive data about the activities of sites, and the experiences of patients who access the IAPT services. It is clear that the new IAPT programme has improved access to psychological services and is helping to meet a previously unmet demand for therapy. Patients who use the services have generally positive

experiences of them, and those who stay in treatment have recovery rates which are comparable with those expected from effective interventions. All of these are very positive findings.

Improving access to psychological therapies is fully consistent with the White Paper goals to improve access to and choice of health services, to provide more care outside hospital, and to support people with long term needs.

The Layard report made much stronger claims about the economic impact of an expansion in the availability of psychological therapies, based on the hypothesis that this would lead to people coming off benefits and into work, which would more than pay for the increased investment in therapy services. Little robust evidence is currently available to support or refute this hypothesis, and it is unfortunate that the opportunity has not been taken to add to the empirical evidence relating to the Layard hypothesis.

Individual Budgets

Background

Cash payments for social care users have formed part of government policy in England for some time. The Community Care Act⁶⁷ was passed in 1996 and since 1998, local authorities have been able to offer direct payments instead of services in kind to disabled and older adults. Direct payments were introduced in response to demands by younger disabled people for greater control and choice over their support arrangements and tended to be used to employ a personal assistant who can provide help in flexible way.⁶⁸ Evidence about the impact of cash payments on outcomes is in short supply although there is a body of research mainly employing qualitative methods which suggests that those who receive direct payments prefer receiving personal assistance in this way.⁶⁸⁻⁷² However, the evidence also suggests⁷³ that overall take-up of direct payments had been low and varied between different care service user groups and local authorities suggesting major barriers to uptake and implementation.⁷⁴

The emphasis in government policy on consumerism and personalisation⁷⁵ encouraging users of social care services to exercise more control and choice was manifest in the decision to pilot individual budgets (IBs)⁷⁶ which built on both the direct payments scheme and the In Control scheme developed for people with learning disabilities. IBs were first proposed in the 2005 Cabinet Office report 'Improving the Lifechances of Disabled People'. The commitment to pilot IBs was repeated in the 2005 Green Paper on adult social care⁷⁷ and subsequently taken forward in the White Paper 'Our health, our care, our say'.¹ IBs appeared to address the White Paper's key themes of providing more support for people with long term needs, better prevention and early intervention for improved health, independence and well-being as well providing stronger voice and more choice.¹

IBs were shown to be at the forefront of government policy on social care as reflected in both recent political statements from the previous minister for social care⁷⁸ and policy documents^{79;80} and now there are also plans for them to be piloted in health care,^{81;82} although possibly using a different model.⁸³⁻⁸⁵

IBs can be distinguished from the approach used in Direct Payments in a number of ways.⁸⁶ Firstly, needs identified following assessment are translated into a specific budget expressed in monetary terms rather than hours per week. Secondly, people have the choice of using services directly from a local authority or making their own arrangements or a mixture of both. Thirdly, IBs include funding streams that cover housing support, adaptations and equipment for disabled people and financial support to enable disabled people to work. Finally, IBs can be used to purchase a variety of support options.

Aims

Individual Budgets were believed to be an effective method for giving people more control over the support they that receive and tailoring this support to an individual's needs – not just to what is available. The intention was to build on the best features of the direct payments scheme whilst finding a way to overcome some of the barriers to uptake and implementation.¹

The initiative

IBs were introduced in 2003 in the Control model but the IBs in the trial had several new features such as multiple funding streams so IBs in this form were being evaluated for the first time. This was a cross government project involving the Department of Health, the Department for Work and Pensions, the Office for Disability Issues and the Department for Communities and Local Government. The pilot project tested the inclusion of a range of income streams: local authority-funded adult social care, Supporting People, Access to Work, Independent Living Fund and Disabled Facilities Grants and included all the main adult groups that social care is provided for:

- Older people
- People with physical and/or sensory disabilities
- People with learning disabilities
- People with mental health problems

IBs were piloted in thirteen local authorities in England: Barnsley, Barking and Dagenham, Bath & NE Somerset, Coventry, Essex, Gateshead, Lincolnshire, Leicester, Norfolk, Manchester, Oldham, West Sussex and Kensington & Chelsea.

The evaluation

Aims and objectives

To identify whether individual budgets offer a better way of supporting older people and adults with social care needs compared to conventional methods of funding, commissioning and service delivery; and the relative merits of different individual budget models for different groups of users.

The specific objectives were to:

1. Describe the processes of implementing individual budgets, including identifying those factors that facilitate and inhibit implementation and the potential implications for sustainability and roll out.
2. Assess the effectiveness and cost-effectiveness of different models of implementing individual budgets, for different groups of service users, compared to standard funding, commissioning and service delivery arrangements.
3. Assess the experiences of individual budgets for different groups of service users and carers.
4. Assess the wider impacts of individual budgets on social care and other services.

Design

This evaluation was carried out by IBSEN^{75;87} which was a consortium of three DH-funded research units across five universities drawing on a range of different disciplines and methodological expertise.

The evaluation adopted a mixed methods design, the core of which was a randomised controlled trial (RCT) examining the costs, outcomes and cost-effectiveness of IBs compared to conventional methods of service delivery. Unusually, individuals were

randomised and then approached to take part in the study. 2,521 individuals were randomised across the thirteen pilot sites of which 1,594 (63%) agreed to take part in the study. Not all of these people remained in the study and in all 959 service users were interviewed six months after they were allocated to either the IB group (510) or to standard services (449).

The primary outcome measures used were the general health questionnaire (twelve item) and the Adult Social Care Outcomes Toolkit (ASCOT) supplemented by secondary outcome measures which included satisfaction measures and self reported impact of IBs on daily life. The distribution of the sample across 'primary' user groups was:

- 34 per cent were working-age physically disabled or sensorily impaired people.
- 28 per cent were older people.
- 25 per cent had learning disabilities.
- 14 per cent used working-age mental health services.

In-depth interviews were conducted with a subsample of 130 people about two months after they had been offered an IB to explore their first experiences of the new processes and what benefits and drawbacks they anticipated.

Interviews were also conducted with lead officers responsible for IB implementation, and with other staff responsible for commissioning, resource allocation, support planning and brokerage, service provision, and different funding streams.

Representatives of user and carer organisations in some sites were interviewed and each interview examined implementation processes and experiences. Front-line staff (care co-ordinators) and first-tier managers were also interviewed about the impact of IBs on their workloads, job satisfaction, training needs, and adult safeguarding, as well as collecting staff activity data.

In a separate study not directly commissioned by the Department of Health, funded from the resources of two of the three DH-funded units that carried out the main evaluation, structured interviews were conducted with up to 200 carers from the sample of service users who took part in the IBSEN evaluation. Half were carers who looked after an Individual Budget holder and half were carers of people using conventional services. This enabled comparison of the health and well-being of carers between the two types of support arrangements. The impact of IBs on carers was further explored through semi-structured interviews with a small group of carers approximately six months after the service user had been awarded an Individual Budget. The interviews focused on the carers' expectations of any change, how these were realised, any difficulties that arose and how they were overcome. Semi-structured telephone interviews were conducted with Carers Lead Officers in the Individual Budgets Pilot sites.

Evaluation team's findings and conclusions

Outcomes: When pooling data across the sample as a whole, the findings showed that the IB group were significantly more likely to report feeling in control of their daily lives, the support they accessed and how it was delivered. No significant differences between the IB and comparison groups were found in the other outcome domains, although the tendencies in the data generally suggested that the IB group

was enjoying slightly better outcomes. There were variations in outcomes by user group. The most positive outcomes in overall well-being were found in the mental health group with the least benefit from IBs being observed in older people.

Almost half of those who accepted the offer of an IB who were interviewed for the qualitative study described how their aspirations had changed as a result, in terms of living a fuller life, being 'less of a burden' on their families, and having greater control and independence. These informants were also more likely than other IB holders to report satisfaction with the support planning process and financial arrangements. IBs were typically used to purchase personal care, assistance with domestic chores, and social, leisure and educational activities.

Costs: There was very little difference between the cost of support received by the comparison group and the cost for IB holders. Over the full sample, IBs funded an average of about £280 of support per week compared with an estimated mean weekly cost of about £300 for support packages for people receiving standard mainstream services. This difference was not statistically significant, but it is likely from this evidence that IBs could be cost-neutral, if the significant costs of implementation were disregarded.

Integration of funding: IBs rarely included resources from different funding streams due to major legal and accountability barriers.

Cost effectiveness: For people who use mental health services, IBs appeared to be more cost effective than standard arrangements on both the social care and psychological well-being outcome measures. Similarly, for younger physically disabled people, there appeared to be a small cost-effectiveness advantage for IB over standard support arrangements with respect to both the social care and psychological well-being outcomes. For people with learning disabilities, IBs were found to be cost-effective with respect to social care, but this advantage is only visible when the data covered only people who had support plans in place. Standard care arrangements appeared to be slightly more cost-effective than IBs with respect to psychological well-being. There was no evidence of a cost-effectiveness difference between IBs and standard support arrangements for older people in terms of social care outcomes. Standard support arrangements were marginally more cost-effective than IBs with respect to psychological well-being.

One conclusion from these analyses is that IBs have the potential to be more cost effective than standard care and support arrangements but only for certain groups of clients.

Workforce implications: Major shifts in working culture, roles and responsibilities were reported for in-house staff. Some saw these shifts – and IBs more generally – as a reinvigoration of social work values, while others saw them as eroding social work skills. On the positive side, the opportunity to work with people to identify a wider range of goals and aspirations, and to develop support plans to achieve them, were specifically linked to traditional social work roles. But particular concerns were that the care management process was being fragmented, that professional skills and qualifications would not be thought to be needed to undertake tasks related to IBs, and that crisis work, safeguarding and high-end complex casework would come to dominate social work practice.

A widespread view among staff was that more and better training was needed in order to implement IBs. Analysis of data from care co-ordinators diaries showed that

those staff with some IB holders on their caseloads spent significantly more time than others on a number of activities: completing assessments with service users, assessing carers needs, support.

Key concerns for staff included determining the legitimate boundaries of social care expenditure within a support plan; and managing the potential financial and other risks sometimes involved with achieving desired outcomes while at the same time being responsible for safeguarding vulnerable adults.

Impact on carers: The separate study described above investigated the impact of IBs on carers in terms of assessment, support planning, costs and outcomes. When carers of people with IBs were compared with carers of people using conventional services, IBs were significantly associated with positive impacts on carers' reported quality of life and, when other factors were taken into account, with social care outcomes.⁷³ Positive outcomes for carers partly reflected their being more able to engage in activities of their choice.

There was no statistically significant difference between the costs of IBs and conventional services, nor in the time spent or opportunity costs of the help given by carers in either group, although the direction of effect suggested higher opportunity costs to carers in the IBs group.

The thirteen pilot sites varied in how help from carers was treated in service users' IB assessments. Especially early in implementation, carers' own needs risked being overlooked. Where carers' needs were included in IB assessments, this did not always comply with current legislation giving carers rights to assessments of their own needs. Sites also varied in how help from carers was treated in calculating the monetary value of service users' IBs. Typically IB amounts were lower if help from carers was taken into account. Only a minority of carers received any payment from service users' IBs. This was always far lower than the value of the help they actually gave. Local authority officers had very mixed views about paying carers from service users' IBs.

Our interpretation of the findings

The IBSEN study, using a mixed method approach, was the most extensive and rigorous evaluation of self-directed support to date in England. The general message which emerged from the RCT provided ammunition for those who were either enthusiastic or sceptical about IBs in that it showed IBs had some, if limited benefits, and only for some groups such as those using mental health services and to a lesser extent for those people with physical and/or sensory impairment. IBs were more cost effective than standard care for these two groups but there was little evidence of any benefits for older people, whose lower levels of allocated resources may not allow access to a wide range of choice and an extensive range of services, once pressing needs for personal care have been met. The process evaluation showed the implementation of IBs had significant implications for the staff and the organisations involved. Moreover, there were major barriers to integrating funding streams which was one of the key aims of IBs. When carers of people with IBs were compared with carers of people using conventional services, IBs were significantly associated with positive impacts on carers' reported quality of life.

The strength of the evaluation appears to lie in the design of the study and the use of a randomised controlled trial, supplemented by the more detailed qualitative

investigations of the processes and perspectives of a wide range of users and stakeholders. The process evaluation made a major contribution to recognising the challenges all local authorities would likely face in implementing IBs. RCTs have not hitherto been often used in social care partly because of feasibility as they are difficult to implement and partly because of the appropriateness of using experimental or quasi-experimental designs on their own for evaluating complex, social interventions.⁸⁸

In many ways this very ambitious study illustrates some of the difficulties of using an RCT design in the evaluation of complex policy interventions. Although it represents a huge amount of work, and was presumably expensive to undertake, there were many complications which limit the usefulness of the findings. Many of these limitations are recognised and discussed by the authors.⁸⁷

First, it is unconventional to randomise individuals before approaching them to participate, although this is sometimes necessary when whole organisations or areas are randomised ('cluster randomisation') which is not the case here. The disadvantage is that different sorts of people are likely to agree to take part in each comparison group, depending on what sort of care they would like. Indeed the consent rate in this study was different for those recruited to the IB and the comparison group, and although the evaluators maintain the two groups remained comparable, this introduced potential bias. This approach was a result of local pilot authorities' ethical concerns about gaining consent before randomisation, who felt that they could not inform users and promote the idea of IBs, and then to refuse them the IBs because of the randomisation outcome.

Second, there are problems of generalisability. This is often the case when evaluations are done on sites which are early pioneers for an intervention. In this example, the thirteen sites which took part in the pilots had particular advantages. Senior officers and most care co-ordinators and team managers expressed enthusiasm for the principles of IBs, they were backed up by a dedicated national implementation team and they may have been motivated by the spotlight of a high profile national pilot and evaluation. These features may not be present in wider rollout, or at least not on the same scale. In addition, each pilot site had a much higher uptake of direct payments than the national average prior to the study. This raises a further problem, in that 26% of the people in the study (in the comparison arm as well as the IB arm) had already received direct payments. This would considerably dilute the potential of the study to detect any differences between the IB and comparison groups.

Third, the evaluation timescale set by the Department of Health, compounded by implementation delays, permitted only a six-month interval for users to receive their IB and commission their own services, whereas a longer follow-up period would have allowed more time to develop a more accurate picture of experiences and impacts. Some of those offered IBs in the 'IB' arm did not take them up and others did not receive them before outcomes were collected at six months. Only 45% of the people in the IB sample were actually in receipt of the IB at the time of the six month interview and of these around a half had held the IB for under a month. Thus, the restricted timescale severely limited the strength of the evidence which emerged from this large scale study. In addition around a quarter of the interviews carried out after six months were carried out with someone other than the service user (usually a

family member) which suggests that interpretation of the analysis of the outcome data should be treated with caution.

Fourth, and related to the above two problems, the pilot sites were evaluated while they were still at an early stage of implementing the IB initiative, so the evaluation may not represent 'steady state' implementation. And there were a number of policy changes and external pressures on the sites during the period of the evaluation, so the evaluators were trying to take a 'snap-shot' of something which was in reality a moving picture.

Finally, there are a number of other limitations to the RCT aspects of the study. Few details are provided to justify the size of the sample included, which is important in understanding the power of the study to detect meaningful differences. There is no statement *a priori* of which outcomes are considered primary or most important. Related to this is the problem of multiple outcomes. Data were collected about a wide range of outcomes. For most of these outcomes, no significant differences were detected between the groups, but the few differences which did reach statistical significance are highlighted in the results. But if many outcomes are examined (and in this case each outcome was also analysed for four groups of people) then a number of 'statistically significant' differences are likely to arise by chance. This may explain the inconsistent findings of the effect of IBs for different groups of people.

Other relevant evaluation work

An innovative action research evaluation was carried out in one of the pilot sites⁸⁹ commissioned by the local authority who wanted a more detailed picture of the local experiences of IBs. It was carried out by a team from Coventry University⁹⁰ and involved the use of a range of different qualitative methods (including group artwork, film projects, posters, Dictaphones to record individual thoughts) to elicit and construct stories from users (n=30) about their views and experiences of IBs and the extent to which IBs have been effective in empowering them to make decisions for themselves. The findings showed that IBs had been a very positive development for family roles, independence and promoting person-centred approaches. However, the evaluation identified a range of barriers to the use of IBs which included: the varied levels of knowledge and awareness of IBs; the difficulties in users obtaining IBs; lack of supply of good care staff; having the right skills to become an IB holder; the need for supply to be more responsive and flexible and the need for professionals to recognise that their role is changing. The findings from this local evaluation were disseminated widely in an imaginative and accessible form.⁸⁹

Conclusions about the evidence that the initiative achieved its objectives and delivered policy goals

The main aims of IBs are listed again below, with comments about the evidence so far available:

Improved outcomes/cost effectiveness: There is some evidence that IBs may have a positive impact on social care and psychological outcomes and can be cost effective but only for some groups of service user. Previous research suggest cash payments are attractive to younger, disabled adults and the evidence from this study suggests it could be extended to mental health users although the latter were the smallest of the user groups and diverse in terms of problems. The variable impact of IBs on different user groups needs to be explored in future research.

Improved patient choice/control: There is some, if limited evidence, that IBs increased feelings of control amongst some groups of user. IBs were typically used to purchase personal care, assistance with domestic chores, and social, leisure and educational activities although this did not indicate a significant increase in the exercise of choice. Positive outcomes for carers partly reflected being more able to engage in activities of their choice.

Costs: There was very little difference between the cost of support received by the comparison group and the cost for IB holders. However the evaluation identified a range of other factors likely to affect the overall and long term costs of implementing IBs on a national scale for all adult social care users.

Workload for social care workers/managers: Major shifts in working culture, roles and responsibilities were reported for in-house staff and there was a need for more and better training. Key concerns for staff included determining the legitimate boundaries of social care expenditure within a support plan; and managing the potential financial and other risks sometimes involved with achieving desired outcomes while at the same time being responsible for safeguarding vulnerable adults.

Integration of funding: Little evidence that IBs included resources from different funding streams due to major legal and accountability barriers.

White Paper goals: In terms of the White Paper goals, it is consistent with ensuring that some user groups have more control and have greater independence and a better quality of life. However further evidence is needed before it can be concluded that IBs are associated with major improvements in long term social care and psychological outcomes. It must also be remembered that this initiative, or different models of it, formed part of government policy on social care some time before the White Paper was published.

Information Prescriptions

Background

The White Paper 'Our health, our care, our say' contained an explicit commitment to introduce Information Prescriptions. However, this can be seen as one particular form of information giving, a broader theme in health and social care for some time. The asymmetry of information between the health care professional and the client/patient has long been seen as an important characteristic of that relationship, and one which would need to be addressed as part of creating a more patient-centred approach to delivery of health and social care, or supporting moves to self-management of long term conditions.

A number of policy documents in recent years have emphasised the importance of access to information for users of services. For example, *Building on the best: choice, responsiveness and equity* reported that 'almost 90% of respondents to the Choice Consultation Survey stated that they needed more information in order to make decisions and choices about their treatment or care' and set as a policy goal 'to ensure people have the right **information**, at the right time, with the support they need to use it.'⁹¹

Better information, better choices, better health developed these ideas and introduced the concept of 'information prescription' as part of a process of routinely assessing individuals' information needs and signposting them to directly relevant information about their condition and the options open to them, within consultations.⁹²

In the consultation exercise 'Your health, your care, your say', a majority (61%) of participants in the core sample said that being given more information about health conditions would make a big difference, but also recognised the challenges and costs of providing information effectively.⁹³ In the sample of people with long term conditions, 65% said they wanted more information about their condition and about services, though this rose to 84% - 89% when asked about information on specific services (to support independent living, and on benefits, for example) (ibid:22). Finally, the Darzi Review has explicitly linked information needs and provision to the promotion of choice and personal control for users of services,⁹⁴ and the NHS Constitution makes the provision of appropriate information to support choice a right and refers specifically to the introduction of Information Prescriptions.⁹⁵

Referring directly to the success of other initiatives which have used the model of 'prescribing' ('exercise on prescription', for example)⁹⁶ the White Paper made a commitment that all people with long term health and social care needs and their carers would receive an 'information prescription' by 2008 (ibid:114). However, the White Paper contained no further detail as to how information prescriptions would work, referring instead to the development work being undertaken to pilot the idea.

Aims

The provision of Information Prescriptions is intended to enable users and carers to better manage their own care through the provision of information. That information should be delivered in a timely, personalised and accessible fashion. In terms of the White Paper goals, they are intended to address the following: more support for people with long term needs; tackling inequalities and improving access to services; more local and convenient care; more choice and a stronger voice for individuals and

communities; and better prevention and early intervention for improved health, independence and well-being.

The initiative

The term 'Information Prescription' covers a range of practices for giving individual service users access to better information about their condition, about the treatment options available to them and the pros and cons of different options, about the support networks they can access, and about ways in which they can manage their own health and well-being. The distinguishing feature of an Information Prescription should be the degree of individualisation involved – the term 'prescription' is intended to convey that it is a personalised recommendation and not just a standard set of information. The corollary of 'prescribing' is 'dispensing'; the process by which the service user is given access to the information. It can take the form of written material, but can also be provided in other media; DVDs or podcasts, for example. Beyond this somewhat abstract description, the concept is open to a range of interpretations and forms.

In order to trial the introduction of Information Prescriptions, the DH recruited twenty pilot sites nationally (i.e. within England). These pilots were established in January 2007, with the aim of having introduced Information Prescriptions by January 2008. A feature of the initiative was the diversity represented by the pilots in terms of settings and conditions addressed. They covered health and social care settings including primary care, secondary care, acute healthcare and community mental healthcare, and addressed a range of long term conditions. Of the twenty, seven addressed mental health conditions, but within this there was a mix of age ranges and types of conditions. The next largest group of pilots (five) addressed mixtures of long term conditions such as diabetes, asthma, arthritis, HIV; most of these pilots were based in or led by primary care or social care with high levels of involvement of voluntary sector partners, and they again covered a range of age groups. Four pilots provided prescriptions to people with cancer of various kinds, again covering primary and acute settings. Two pilots addressed sensory impairments, and two addressed a single long term condition; cystic fibrosis and Parkinson's disease respectively.

The diversity of setting, target condition and intended age group of recipient was intentional as the initiative was intended to 'test and provide evidence on the effectiveness and impact of information prescription on users, carers, professionals, and organisations, including the voluntary sector'.⁹⁶ Much of the work of the pilot sites was developmental; there was no single pattern for what information prescriptions were or how they should be delivered; each site created a locally agreed and owned model for delivery. Key elements of this were the type and quality of information to be included, the means by which it should be offered to the recipient and by whom ('prescribing'), and then how the recipient would in practice access and receive the information ('dispensing').

There was always the intention that the development of Information Prescriptions at local level should be supported by, and linked to, the development of a range of ways to provide health information nationally, through the NHS Direct and NHS Choices websites, for example. However, by the end of the pilots, there was concern in some sites that a great deal of the effort that had gone into identifying, checking and making available some kinds of information had then been made redundant by the development of NHS Choices. This particularly applied to those sites which addressed

a range of conditions and tried to make information available in non-specialist settings and to a wider group of people ('light touch prescribing' – see below).

The evaluation

Aims and objectives

The aim of the evaluation was to assess the overall effectiveness of the pilot programme, and to gather learning about what works well in respect of the Information Prescription process, from the different approaches being adopted in each of the twenty pilots. More specifically, the evaluation was designed to help inform the four main goals of the pilot programme:

- To shape the practical design and delivery of Information Prescriptions nationally, including how this will be supported nationally at the locality level
- To provide evidence on the effectiveness and impact of Information Prescriptions on the public, professionals, and organisations alike
- To contribute to successful national implementation of Information Prescriptions by 2008 to people with a long term condition.
- To inform the policy direction, ensuring that the implementation of Information Prescriptions is integrated with other major policy drivers

Design

The evaluation was carried out by a consortium comprising the Office for Public Management, (OPM), GfK-NOP and the University of York. The evaluation was informed, at least initially, by a theoretical model drawing on 'theory of change' methodology.⁹⁶ This methodology is referred to in the final report as entailing the development of 'a coherent conceptual understanding of how an initiative operates and intends to bring about changes to outcomes for services, users and carers' (p10). It is clear that this theoretical framework informed the initial stages of the evaluation, being used in an initial, scoping phase of the project to identify an 'impact model' setting out the main activities, products and outcomes associated with the development of Information Prescriptions. The framework also involved identifying the key 'change mechanisms'; the assumptions about how and why Information Prescriptions might make a difference to services and users and carers. This model informed decisions about the structure of the evaluation, and the design of research instruments. However, it is not explicitly referred to further in the report and is not used in the final analysis of data.

The evaluation was a mixed methods design, with the following main elements:⁹⁶

1. Qualitative research: two waves of qualitative fieldwork were carried out. Each wave comprised a focus group and set of in-depth semi-structured interviews with key stakeholders from each of the twenty pilot sites. Over 150 staff took part in this element of the study.
2. Survey of patients/users, carers and professionals involved in Information Prescriptions: Two surveys of patients/users, carers and professionals involved in Information Prescriptions were conducted. The first of these surveys included respondents from twelve of the twenty sites. The survey included 74 service users who had received information prescriptions (nearly all from cancer services), and a further 94 who would have received them had

their site started issuing them (none from cancer services). It was originally intended to compare these two groups, but as they were not sufficiently comparable on a range of characteristics, the non-recipient group was instead used as a comparison group for survey responders from the same sites in the second survey. While these sites were not yet formally offering Information Prescriptions, staff said, and users confirmed, that information was offered to most users, and the term 'Information Prescription' was not used in any of the survey instruments with users or carers. The second survey was carried out in all twenty sites and included 299 service users (a 36% response rate), 164 carers (total number not clear, but an estimated response rate similar to that for users) and 243 professionals (49% response rate).

3. Activity data collection: Pilot sites were asked to make monthly pro forma returns to the evaluation team giving information about: the numbers of Information Prescriptions issued and by whom; the number of 'items' on a prescription and where and by whom the dispensing process was carried out and; estimates of staff time involved and any other identifiable expenses. By the end of the evaluation, 70 monthly returns had been made in total, including data from each of the pilot sites. Pilots were also asked to complete a 'stocktaking form' about previous information giving activity. The information on this form was discussed and added to during site visits by the evaluation team. The purpose of this strand of work was principally to estimate the potential cost of national Information Prescription roll-out, and to conduct an approximate comparison of the costs of different theoretical models of Information Prescription delivery. This was done by combining the data from the pilot sites with nationally available data on prevalence of long term conditions and associated service usage rates.
4. Literature search: The University of York Centre for Reviews and Dissemination conducted a literature search on research evidence pertaining to Information Prescriptions and the impact of information on patients. The results of this are presented as a narrative review, with discussion of: themes from the general literature on health information; findings of specific studies of Information Prescriptions and; identification of where additional research is needed.
5. Review of Information Technology: The Consortium conducted a review of the IT implications involved in the development of Information Prescription systems; this ran in parallel with the main Information Prescription evaluation. This review examined the use of IT in pilot sites, exploring stakeholder views on the current systems and the potential uses of IT to support information prescription in future.

As well as the data gathering described above, the evaluation team carried out the following learning and support activities:

1. Action learning: The Consortium organised five action learning events during the course of the evaluation. Each event was designed to enable people from the pilots to work together to identify common challenges and develop joint solutions. Following each event, OPM produced a report summarising the learning from the day. The emerging themes from these sessions fed back into the ongoing evaluation work.

2. Website: A project website was created to provide a forum for information exchange and learning. All those involved in the project – pilot site representatives, the consortium organisations and DH teams – were given access to a secure section of the site, inaccessible by the public. This restricted area included a discussion forum and information and tools to support pilots.

Evaluation team's findings and conclusions

The headline findings were reported⁹⁶ in terms of process issues and impact. The evaluation also provided estimates of the local and national resource implications of mainstream implementation of Information Prescriptions, based on the costs and activity levels of the pilots.

Process issues. The evaluation report divides the findings on the implementation process into stages: preparation for, development of, and delivery of Information Prescriptions. The preparation stage involved for most pilots a process of definition and of agreeing 'principles' on which to provide Information Prescriptions, and the sourcing and collating of information. Definitions of Information Prescriptions varied; in particular there was concern in some sites about the term 'prescription' being too medicalised. Development involved a process of engaging local users, carers and professionals with Information Prescriptions; quality assuring information; developing IT systems and training staff in delivery of Information Prescriptions. The delivery stage demonstrated the diversity of the pilots, with different approaches reflecting: the different settings of the pilots; the characteristics of the local user and carer population; local geography and the structure of local services; and the extent of local professional involvement. The evaluation team describe three broad models for delivery reflecting different levels of 'depth' of prescribing, and mechanisms for dispensing. These models were:

- *Model One: Light touch prescribing and self dispensing*
In this model the prescribing process is quick and generally consisted of a 'tick box' template or similar, pointing the user towards information they can access themselves via the internet, libraries or voluntary sector organisations. The model was used mostly in primary care, for people with conditions that were relatively stable and did not require extensive input from secondary care.
- *Model Two: In-depth prescribing and linked dispensing*
In this model prescribing is usually done in a pre-arranged consultation session as part of a care pathway – at referral or review, for example. The prescription is a tailored list of resources, which is then taken to a designated dispenser. Dispensers include a range of providers – community information points and NHS Direct, for example.
- *Model Three: In-depth prescribing with information centre*
This model involves prescribing by a specialist professional – a consultant, a care manager or a specialist nurse, for example – in a structured and tailored way as part of a care pathway. It may involve a two stage prescribing process, with initial and more detailed assessments. Dispensing is then done within a specialist health information centre. This model is predominantly used in secondary care, for conditions which are high risk or highly complex.

Impact on users. In the second survey high numbers of users overall (73%) reported having more confidence in asking questions about their condition. Half of all users

(52%) who had received information said it had improved their care, while two thirds (66%) said they now felt more in control of what was happening with their condition. However, detailed analysis of these figures suggested consistent differences between groups of users, with those who were in poorer general health, those living in disadvantaged areas, people under 65 and those where the Information Prescription was delivered through a 'light touch' model less likely to report these favourable outcomes.

The data from the second survey of users was compared with those from users in the first survey who had not been issued with Information Prescriptions. While there was a drop in the proportion of users who said that professionals did not discuss information with them (40% in the first survey: 12% in the second), and that any information given was not easy to understand (31%: 11%), the same proportion (90%) said that overall information given was useful.

Impact on carers. Very high numbers (89%) of those carers who had received or seen an Information Prescription found it useful. However, a third (35%) overall and nearly half (44%) of those in 'light touch' sites did not know about the Information Prescription.

Impact on professionals. Two thirds of professionals (66%) were satisfied with how Information Prescriptions were implemented in their site, with only a few (7%) actively dissatisfied. Over half (57%) of professionals said that Information Prescriptions compared favourably with the ways in which information had been given in their site previously. However, a third (36%) felt it was about the same and a small number (5%) felt it was worse. Professionals reported making an assessment before offering an Information Prescription according to whether they thought the user and carer would be able to make good use of the information. Overall one in four professionals reported offering everybody an Information Prescription, though this was much higher in mental health where half of professionals did so.

Resource implications of mainstream implementation. The collection of baseline activity data showed great variation between the pilot sites, and in some cases between different parts of the same site, in the extent, sophistication of delivery and uptake by users, of information provision prior to the pilot initiative. The report provides estimates of demand, and resource implications of national implementation of each of the three models, and identifies key steps to be taken at local and national level to implement each. If Model One ('Light Touch') was rolled out in primary care, it is estimated that 65% of consultations for long term conditions would result in the take up of an Information Prescription. For a general practice with a list of 10,000 patients, this would mean 5575 such prescriptions a year for patients with the long term conditions covered by the Quality Outcomes Framework (QOF) registers (and over 8,000 if a wider range of long term conditions is used). This would also generate 1115 (1673) contacts with local and national voluntary organisations, 613 (920) contacts with Benefits Offices and 279 (419) contacts with local libraries, though it is stressed that these are all estimates. It is estimated that for a small to medium sized PCT, the implementation of this model of Information Prescription would require a full-time senior coordinator, plus a junior support post.

The resource implications of the other two models are much harder to gauge as it proved very difficult to identify the time spent by professionals on the prescribing element: estimates ranged from 10 – 15 minutes for relatively straightforward

conditions to over an hour for more complex ones though this could include 'dispensing' where this was done within the same consultation. The resource implications in terms of staff time, in particular for Model Three, where senior clinicians would be heavily involved, were high but no actual figures were given in the report. Where dispensing was done by designated other agencies, this would need to be resourced either through commissioning or some form of payment by results.

Literature review

The literature review provides some useful background to the evaluation and identifies how little is known about the operation of this highly personalised and individualised way of providing health information.

The Information Technology review

This review extended beyond the scope of the main evaluation, drawing on the pilots' experiences in implementing Information Prescriptions as one source of data. The report of the review was published as an appendix to the final report. The review identifies the multiple stages of Information Prescriptions and comments on appropriate technology and issues to be resolved at each stage. The review identifies issues reflected elsewhere in the evaluation: the time consuming nature of using electronically generated Information Prescriptions; issues of quality assurance of information; accessing and storing appropriate information (in particular how to reconcile the need for standard, national, accredited banks of information and quite detailed local information); the complexities of recording prescriptions, to be kept with individual records; inequities of access to information technology.

Conclusions/recommendations

The report makes eleven key recommendations on how Information Prescriptions should be implemented. These are to do with stakeholder involvement in planning, the content and accreditation of information directories, the importance of templates and 'structured scripts' in delivery, the need for personalisation through a diversity of delivery mechanisms and formats, and issues for local partnerships in planning and delivering Information Prescriptions. The recommendations also include the need to ensure that Information Prescriptions are accessible and useful to disadvantaged users and carers.

Comments on the evaluation

This was a complex evaluation, combining developmental and learning activities with more traditional evaluative methods. As indicated by its stated aims and objectives, the intention of the evaluation was not to provide evidence for a decision about whether or not Information Prescriptions should be implemented; this had already been determined by a prior policy decision announced in the White Paper.

Several of the objectives of the evaluation appear to have been fulfilled in that it provided information about, and evidence on, the process of developing Information Prescriptions, and captured the different models adopted by the pilot sites. The formative aspect of the evaluation was strong; the action learning events and the role the evaluation team played in bringing the pilots together enabling them to learn from each other was important to the development of the pilots and in making the lessons learnt in terms of process available more widely. The Information Prescriptions website, and the online resources available from it, provide a wealth of information

and tools for other sites wanting to develop Information Prescriptions. The report identifies very clearly the steps that would need to be taken both locally and nationally to implement Information Prescriptions. The DH responses to the report's recommendations largely refer back to the online resource⁹⁷ created as part of the evaluation, demonstrating the significance of the formative role the evaluation has had. Where the evaluation is weakest is in terms of looking at the effectiveness and impact of Information Prescriptions.

It would be fair to say that the diversity of models, and the extent to which Information Prescriptions were developed individually within the pilots, had two major impacts on the evaluation of the initiative and the kind of evidence it was possible for the evaluation to provide. The first is that the development stage took longer than anticipated and many pilots issued far fewer Information Prescriptions in the evaluation period than had been expected. This reduced the number of respondents to the surveys of users and carers. The second is the difficulty of interpreting evidence from so many disparate sites and contexts in a way that makes it useful, in particular for service delivery past an initial stage of development.

In terms of the usefulness of the survey data, there is likely to be a bias towards positive evaluation of Information Prescriptions because of the selection process professionals used in deciding to whom to offer information, and because of the response rate. It is striking that similar proportions of those who did not get Information Prescriptions in the first survey, and those who did so in the second, thought that any information they did receive was useful. However, it is likely that even though these users were not given a formal Information Prescription at the time of the first survey, the services involved in the sites were ones where information giving was valued and which were preparing for Information Prescriptions.

All the findings on activity, and inferred resource implications, are estimates. It was difficult to get good data on time and resources expended on the process. It was also very hard to identify what sites had been doing before, and therefore what difference the Information Prescription pilot actually made, in quantifiable terms. The report assumes that take up would decrease over time as people became better informed, however, it is also possible that people would come to expect a higher level of information giving in consultations. In particular the aim of personalising information suggests that an individual's information needs would be reviewed on a regular basis – Information Prescription in this case would not be a one-off process.

Our interpretation of the findings

This evaluation appears to comprise a number of disparate elements of data collection, with some elements being much stronger than others. The evaluation had a strong formative function in helping those implementing Information Prescriptions to learn from each other. The qualitative work also provided a useful typology of the different approaches to implementing Information Prescriptions that developed. However, the collection of data on outcomes was hampered by the smaller numbers of respondents than intended and the extent to which aggregating results across the pilots makes it difficult to interpret in the light of different models for delivery. Some analysis by broad category of model ('Light touch', etc.) and by different groups of recipient was carried out. But this analysis raises rather than answers questions that would be important in terms of the White Paper goal of tackling inequalities.

Other relevant evaluation work

We also looked at the development of Information Prescriptions in a site not included in the national initiative. Experience here echoed that in the pilots – that the process of developing a model and quality assuring information was extremely time consuming. This site had, at the time of our fieldwork, not reached a stage of issuing any Information Prescriptions, and so was unable to provide any evidence on delivery or outcomes.

Conclusions about the evidence that the initiative achieved its objectives and delivered policy goals

The initiative was intended to enable people with long term conditions to better manage their own care through the provision of timely, personalised and accessible information. The evidence is that setting up systems to make that information available is complex and potentially resource intensive.

In terms of the White Paper goals, Information Prescriptions are clearly a means of providing more support for people with long term needs. What is not clear, because of the lack of data on outcomes, is whether they are a cost effective means of doing so and whether they do in fact lead to better prevention and early intervention for improved health, independence and well-being, and more choice and a stronger voice for individuals and communities. There are also issues raised by the evaluation about whether they are a means of tackling inequalities and improving access to services.

New Types of Workers

Background

The White Paper 'Our Health, our care, our say' aims to realign the health and social care system. The White Paper states that one of the keys to closer integration of services between health and social care services is the development of an integrated workforce ensuring that services are designed around the needs of people who use services and supported by 'common education frameworks, information systems, career frameworks and rewards' (8.35). Building on existing work being carried out by the NHS Workforce Change programme and the Skills for Care New Types of Worker (NToW) pilot, the White Paper suggested that the development of integrated services and new health and social care multi-skilled teams would need to be based on the development of common national competencies and occupational standards (8.38), creating career pathways across health and social care (8.39). These competencies and standards would enable staff to work in integrated settings using common tools and processes. The White Paper confirmed that Skills for Care and Skills for Health would lead this work (8.40).

The Department of Health's interest in the development of new professional roles dates to the early 1990s when a number of projects, including the Exploring New Roles in Practice project, were commissioned under the Policy Research Practice Programme's Human Resources and Effectiveness initiative.⁹⁸ In recent years interest in developing new ways of working for health and social care workers has become an integral part of the government's modernisation agenda reflecting the wider aim that services should be developed in ways that meet the needs of those who use them rather than to fit traditional professional or organisational boundaries. The NHS Plan for example, identified the need to develop a new approach to the provision of services which would recognise and enhance the skills of the workforce enabling staff to make a more effective contribution to health care.⁹⁹ The establishment of the Changing Workforce Programme added further momentum to this strategy. The programme supported the development of roles that were outside traditional professional or organisational boundaries, within health as well as social care services.¹⁰⁰

The difficulty of evaluating the impact of new roles and new ways of working is well documented.¹⁰¹ Whilst existing research evidence suggests that developing new ways of working can help to support specific government initiatives such as the introduction of the New Deal for Junior Doctors^{102;103} the evidence does not, as yet, demonstrate improvements in patient outcomes.¹⁰¹

Aims

Stage 1 of the NToW programme ran from 2003-2006. The programme was funded by the Department of Health to meet the following targets:

- 'Identify and trial additional innovative types of working that are central to the White Paper vision of a service-user commissioned service.'
- 'Embed sustainable new types of working in the sector through informing the development of career pathways, performance management, NOS [National Occupational Standards] and qualifications.'

- 'Ensure that new types of working in social care make a full contribution to community regeneration.'

The NToW programme aimed to:

- Give guidance and examples of how the social care workforce can be redesigned.
- Demonstrate the workforce changes required to implement new service models.
- Show what those changes will mean for all aspects of workforce development.
- Be an exemplar of best practice in people and carer engagement.

The Stage 1 Evaluation Report (2007) noted that the NToW programme did not have 'tightly' drawn objectives because, it was argued, these could have restricted the development of new and innovative approaches to work. The report also comments that the work took place against a backdrop of rapid policy development, including the publication of 'Our health, our care, our say' and as a consequence the programme had to 'reframe' itself to ensure that it reflected these developments.

The objectives of the pilots also reflected the priorities of Skills for Care (SfC), these were to develop:¹⁰⁴

- person-centred services (for example developing person centred planning or assessment tools)
- voices of people who use services and of carers (improving service user engagement in the development and delivery of services)
- independence for people who use services and for carers (developing independent living in the community)
- prevention/early intervention (improving service response often across organisational boundaries)
- service access for people who use services and for carers (improving access to services by those who were marginalised by mainstream services)
- workforce capacity
- skilled/competent workforce (pilots with a training component including those that sought cultural change as a means to embed new practice)
- joined-up services

Each of the twenty-eight pilots had one of these objectives as their key objective and at least one other objective, with a maximum of six objectives.

The report notes a lack of precision with regards the SfC programme objectives arguing that there was a lack of clarity about whether they were looking at new work roles or new ways of working. As a consequence the researchers point to a lack of shared understanding about the aims of Stage 1 of the programme.

The Stage 1 evaluation report identifies the following national policy objectives that are reflected in the objectives of the pilot sites (however little indication of how the work of specific pilots contributed to these objectives is provided):

- Person centred services (voice/dignity/well-being/independence)
- Prevention/early intervention

- Access
- Support for long term needs
- Workforce capacity building

The evaluation report also identifies the following national policy objectives that were reflected in the work that pilots were engaged in (however little detail to illustrate how these objectives were met is provided):

- More care at home
- Better joined up services at the local level
- Innovation
- Allowing different providers

The initiative

Pilot sites were established in twenty-eight locations. The report does not provide detailed information about individual pilot interventions however it does note that most involved the development of a new role and that individual pilots reflected one of the priorities of Skills for Care identified above.

Of the twenty-eight pilots, four were based in the North, two in the East, eleven in the South, eleven in the South West and two were based in organisations providing a nationwide service. No indication is provided about how pilots were chosen. The report notes that the pilots were concentrated in the south and that services for older people were underrepresented. Six pilots were funded for three years, eleven for two years and eleven for one year.

The evaluation

Aims and Objectives

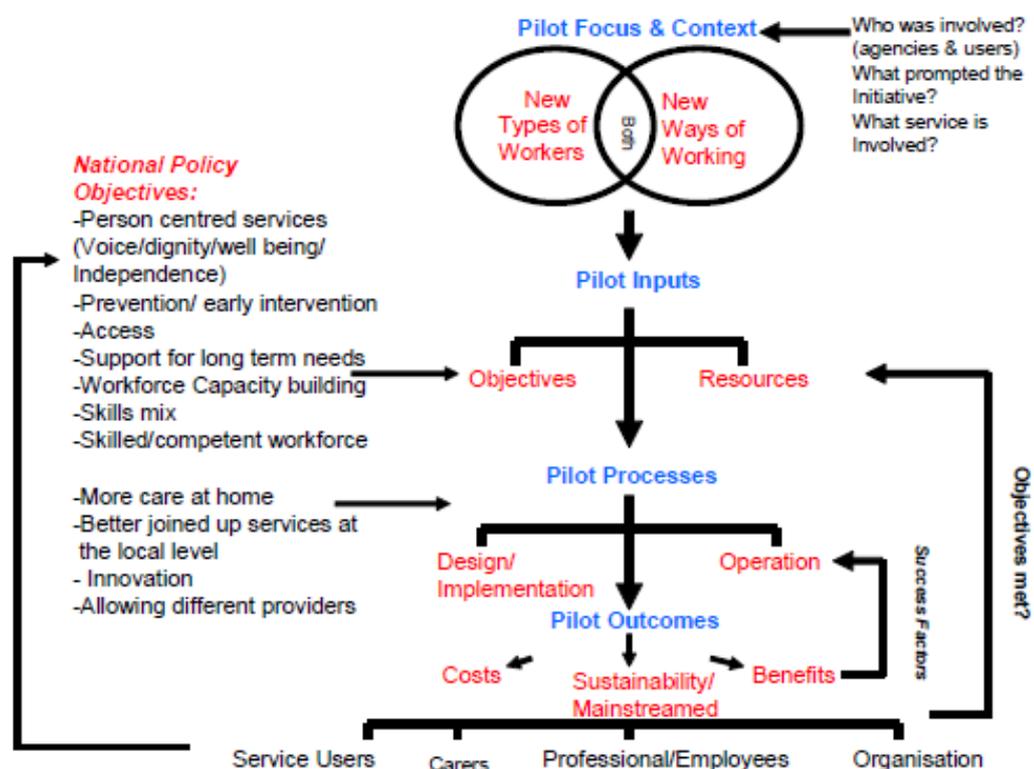
The NToW objectives for the Stage 1 evaluation were to support pilots in order that the programme could contribute to:

- The development of service provision in a 'person-centred' way
- The generation of new work roles
- The creation of new (inter) organisational relationships
- The formulation of management systems to support these relationships
- The stimulation of learning at different levels

Design

The design of the evaluation was based on a systems model with 4 main elements: context (who was involved, why initiative was set up, what the service is), inputs (objectives and resources), process (design/implementation, operation) and outputs (costs, sustainability, benefits from a range of perspectives: service users; employees, carers and the organisation). The design is presented in diagrammatic form below.

Figure 1: Evaluation model¹⁰⁴



The evaluation was completed in three months, from October to December 2006 and incorporated the following methods:

- Documentary analysis of pilot proposals and final reports.
- Telephone interviews with twenty-five of twenty-eight pilot managers.
- Follow-up work at eleven pilot sites, this work included interviews with workers and people using services, focus groups with people using services; occasional observation for example of steering group meetings. Follow-up sites were chosen to reflect user group and region.

A total of 51 interviews were completed during the evaluation. No detail of the analysis strategy or the process by which research ethics review was secured is provided.

Evaluation team's findings and conclusions

New roles were developed in twenty-three of the twenty-eight pilots. The roles were characterised in terms of their purpose: specialist role (breaks away from general occupation to focus on a task and/or user groups); person-based role (a role which is performed by a person who uses the service); co-ordinating role (organises activities involving different parties); boundary spanning role (crosses traditional organisational/client jurisdictions) and their design: a re-labelled role (new title but no significant change in content); a re-packaged role (combines established tasks in a new way) and a re-created role (newly generated tasks and responsibilities).

The evaluation highlighted the importance of process in the development of new roles, for example in managing the new project and managing the different stakeholders.

All pilots were required to carry out a local evaluation but had the discretion to decide how to do so. Six pilots commissioned independent evaluations and the remainder carried out self evaluations. The report records minimal information about the local evaluations recording only whether it included a survey, interviews or 'other' methods. The report does not consistently record the number of participants taking part in local evaluations.

Three categories of outcomes were identified: organisational; employee, and outcomes for people who use services.

1. Organisational outcomes. The report notes that these were difficult to identify and measure particularly given the diffuse objectives and time limited nature of the pilots. Several pilots attempted to undertake a financial cost benefit analysis but there were difficulties apportioning financial values to the costs and benefits associated with these developments. The measure used to capture organisational outcomes was whether or not the role/work continued after the pilot funding ended. Of twenty-four new roles/ways of working supported by the programme these continued in full at nine pilots sites and in part at ten sites. The new roles/ways of working were terminated at three sites. Tentative suggestions are given as to why some roles/ways of working were sustained, these included whether the institutional culture was receptive to new roles and or new ways of working. Discontinued/vulnerable new roles were typically based in the voluntary or private sector; predominantly smaller voluntary organisations.

2. Employee outcomes. Recruitment difficulties were experienced at seven of the twenty-eight pilots; these difficulties were thought to be associated with the short-term nature of the posts and shortage of suitably qualified personnel particularly in sensory impairment services. The report also suggests that there was a lack of people with appropriate skills for those pilots developing new roles or combining tasks in innovative ways and allied to this that the unique nature of the roles raised uncertainty for potential applicants who therefore didn't apply. Retention difficulties were not noted; indeed some pilots reported that the new ways of working had improved retention. The need to provide suitable training was identified as an important issue at several pilots, for example to familiarise workers with the relevant legislative frameworks. Career uncertainty was identified amongst some workers. The importance of developing appropriate line management arrangements was noted and led to difficulties occurring for those roles that crossed organisational boundaries. The development of roles staffed by people who use services themselves raised specific issues about employment status, workload expectations and the impact of 'work' on an individual's rights to benefits. Although it is difficult to gauge the strength of these findings the themes emerging in relation to employee outcomes mirror those identified in the broader new role literature.

3. Outcomes for people who use services. The report notes that one of the main aims of the NToW programme was to improve service provision from a user perspective. Given the diverse nature of the pilots the evaluation notes the difficulty of establishing criteria to assess outcomes for people who use services. At pilots where the new roles worked directly with service users the report notes some

'impressive outcomes' from the local evaluation. For example 65.5% of people using one scheme who returned a questionnaire reported 'improved outcomes'. However, given the lack of detail provided about the local evaluations it is difficult to gauge the strength of this finding. No outcomes for people who use services are presented for those pilots where the new roles/ways of working had only an indirect impact on service users, for example where they worked as training co-ordinators.

Our interpretation of the findings

The NToW programme evaluation was commissioned and carried out approximately two years after the programme began. The evaluation design appears to have been constrained by the requirements to complete the work within three months. The programme evaluation notes the lack of 'tightly' drawn objectives against which individual pilots could be monitored. Without clear and measurable objectives for the programme it is difficult to interpret the strength of the findings presented.

The evaluation design for the national programme evaluation is poorly described. No rationale is provided to explain why the pilot sites were chosen and no definitive indication of sample size is provided for the follow-up work at eleven pilot sites. It is unclear how the data were analysed.

The evaluation report provides a *post hoc* rationale for why the programme fits into the aims of the White Paper.

Other relevant evaluation work

One of the pilots identified in the NToW report, an 'In Reach Model of Care in LA Care Homes' was independently evaluated by a team based at the University of the West of England.¹⁰⁵ The two year evaluation was part funded by the Joseph Rowntree Foundation and the Office of the Deputy Prime Minister. The evaluation used a mixed method design combining interviews with stakeholders, home managers and care staff as well as questionnaires, focus groups, a meeting with residents and an audit of activity. The study was originally designed to include five homes, one of which was to act as a control, however due to the closure of three of these homes data collection was restricted. The inclusion of one non-intervention home meant some comparative data were collected. However data collection across the evaluation was not 'stable'. The evaluation team collected a wide spectrum of data related to the activity of the in-reach team including: number of residents referred to the team, reasons for referral, main primary and secondary diagnoses for referrals, outcome of referrals, type of interventions, length of stay in the in-reach service, prevention of hospital admissions, estimated cost saving due to prevented hospital admissions, early hospital discharge and estimated cost saving, nursing home transfers prevented and estimated cost saving, GP visits prevented and estimated cost saving and previously undetected illnesses and resident assessment.

The results of the evaluation suggested that: residents of the homes appreciated the positive impact the in-reach team had on the quality of care provided; the development of the new style of working raised communication difficulties; the development of the new role had resource implications i.e. increased salaries for those working in the new roles and increased workload for those in allied roles. However the evaluation suggested that these costs could be off set against saving to the PCT in relation to reductions in the need for hospitalisation of residents, facilitation of early discharge and early detection and treatment of residents' illnesses.

Conclusions about the evidence that the initiative achieved its objectives and delivered policy goals

The pilots were commissioned against the following Department of Health targets:

- 'Identify and trial additional innovative types of working that are central to the White Paper vision of a service-user commissioned service.'
- 'Embed sustainable new types of working in the sector through informing the development of career pathways, performance management, National Occupational Standards and qualifications.'
- 'Ensure that new types of working in social care make a full contribution to community regeneration.'

Whilst the pilots did trial innovative types of working it is not clear how innovation was conceptualised and little information is provided to make any judgement as to whether the pilots increased our understanding of the impact of new roles/ways of working on career pathways, performance management, National Occupational Standards or qualifications. The evaluation did not provide any information about how new types of working can contribute to community regeneration.

The conclusion of the report notes the limited nature of the NToW programme reporting that 'across 25 of the 28 pilots, fewer than one hundred employees have undertaken the new roles.' The conclusions also note the limitations of the evaluation arguing that such innovations in workforce developments need to develop appropriate outcome measures. With these limitations in mind it is difficult to see this evaluation as providing anything more than an impressionistic view of these developments.

The programme does not provide explicit evidence to address the points noted in the White Paper, i.e. contribute to the development of integrated workforce planning ensuring that services are designed around the needs of people who use services and supported by common education and career frameworks (8.35), or provides evidence about the development of common national competencies and occupational standards (8.38), creating career pathways across health and social care (8.39).

The evaluation report cites a number of different sets of objectives including: NToW programme aims and objectives, Skills for Care New Types of Worker programme objectives and boarder objectives drawn from Topps England. These different objectives serve to confuse the aims of the evaluation.

Addendum: Other relevant evaluation work

(At the time of writing, this section is based on pre-publication findings not for public release).

In November 2008 Skills for Care issued a further document,¹⁰⁶ which summarised the evaluation of the initial phase of the programme as well as presenting findings from a mapping exercise and an evaluation of six continuation sites. The methodology for the mapping exercise and evaluation are not described in detail and the report notes that the evidence base was limited. The mapping exercise highlighted the poor alignment of social care learning and National Vocational Qualifications as one of the constraints of the programme overall. The evaluation echoed previous findings with regards the need to base new roles on written agreements, protocols and joint training and the importance of involving human resources departments in such developments. The report also noted how structural issues, such as the use of short

term contracts and the lack of career pathways can undermine the development of new roles.

In March 2009 Skills for Care issued the evaluation of phase 3 of the New Types of Worker Programme which constituted thirty sites funded from the NToW fund.¹⁰⁷ These sites were commissioned in 2007-08 and were selected on the basis that they offered the potential to explore issues related to the careers of those working at the sites. Brief descriptions of each site are provided in the report. The methodology for phase 3 was not described in detail; it included a combination of face to face interviews, telephone interviews and email correspondence, however it is not clear how many interviews were conducted at each of the sites or who these interviews were with. The report notes a number of factors that constrained the evaluation; these included the disparate start-up dates and end dates, delays in getting started and the disparity between the evaluation cycle and the funding of the sites.

The report specifically looks at the evidence in relation to how the sites contribute to the policy objectives detailed in 'Our health, our care, our say', in particular the personalisation of services. All of the sites supported people to live independently and confirmed the importance of staff and services working flexibly in order to meet the needs of individuals using services as well as the importance of multi-agency working. The report noted the lack of career paths for personal assistants who do not have professional qualifications and suggests that human resources departments need to consider the wider implications of personalisation on the careers of those working in this sector.

Given the limitations identified by the authors of both reports and the lack of detail about the methodology, including the lack of explicit outcomes for phase 2 and 3, it is difficult to take anything more than an impressionistic message from these reports. However the conclusions do support themes identified in the wider new roles literature.

NHS LifeCheck

Background

In the White Paper it was proposed that a 'LifeCheck' service would be developed to help people assess their own risk of ill health and this would be available at critical points in an individual's life. The commitment in the White Paper was to develop and evaluate the service with roll-out taking place after this. LifeCheck was described as a personalised service with two components: Individuals complete the initial assessment themselves and then they are offered advice about what can be done to maintain and improve their health. Initial development was proposed to be around the following life stages: the first year of life, the teenage years and those around the age of fifty. These subsequently become known as Early Years LifeCheck, Teen LifeCheck and Mid-life LifeCheck. Early LifeCheck has recently been renamed 'Baby LifeCheck'.

The idea of NHS LifeChecks came from the public consultation of 'Your health, your care, your say'. A regular health 'MOT' was seen as a way of improving health and preventing disease. This regular health MOT was proposed to be a check up to include things such as a blood pressure check, diabetes test, and height and weight measurement. Regular screening was suggested as a way of identifying those at risk of major health conditions. The options of screening being carried out either across the whole population or targeted at groups who do not often use services (e.g. young men) were discussed. At the Citizens' Summit, the health MOT received the most votes as an option in a range of possible initiatives. More than three out of four people (77%) chose it as one of their top three priorities.

Although the concept of regular health checks is clearly popular with the public, interviews with stakeholders subsequently revealed concerns about the idea. There were concerns about the possible effects they may have on the quality of existing services, whether they are likely to be cost-effective, and whether potential adverse consequences might outweigh any potential benefits. It is well recognised that screening programmes of this type have disadvantages as well as the obvious possible advantage of early detection of disease. Disadvantages include generating anxiety amongst participants, false positive results leading to unnecessary, costly and possibly risky further investigations, diversion of funds from other health services with greater potential for health gain, and exacerbation of health inequalities (as screening programmes tend to be accessed most by those at least risk).

Stakeholders consulted felt that there is strong clinical evidence that regular physical checkups are not a good use of resources. For example, screening for cardiovascular risk factors was introduced within the 1990 GP contract, and this led to two large randomised controlled trials, Oxcheck and the British Family Heart Study.^{108;109} A review of the evidence from these studies concluded that with the information available it was not possible to judge whether the interventions were cost effective.¹¹⁰ Mant also raised concerns about whether there are enough resources in general practice to provide adequate intervention and follow up after screening in order to reduce risk.¹¹¹ He warns against these issues leading to a move away from preventative medicine and suggests that a population approach may be a more effective way of reducing the burden of disease.¹¹¹

It was due to these concerns about checkups that the idea of an online self-completion survey called LifeCheck was developed.¹¹²

Development and evaluation of the three different NHS LifeCheck stages have progressed at different rates since the publication of the White Paper, as described below.

Aims

NHS LifeCheck was proposed to 'empower people to take charge of their own health and 'well-being', addressing the White Paper goals of better prevention and earlier intervention for improved health, independence and well-being. LifeCheck is designed to support people in making changes that will help them to improve their health and therefore is proposed to have a role in helping the NHS place more emphasis on preventative care. LifeCheck was planned to focus on risk factors such as diet, physical activity, smoking, alcohol, stress and risky sexual behaviour. The LifeChecks also have a role in tackling inequalities by improving access to services as well as improving choice for individuals.

Baby LifeCheck is aimed at parents and carers of babies between five to eight months. It is an online tool which aims to provide advice and information to enable parents to keep their baby healthy and make changes to improve their child's long term health.

Teen LifeCheck is aimed at eleven to fourteen year olds as a self-assessment tool to help young people assess their health and well-being. The aim of Teen LifeCheck is to provide information and advice for healthy lifestyles and to provide information about sources of further advice or support available in order to make changes. Teen LifeCheck was also planned to reach the most vulnerable groups

The Mid-life LifeCheck is aimed at those aged forty-five to sixty. This online tool focuses on smoking, healthy eating, alcohol use, physical activity and emotional well-being and provides feedback, helping people to plan lifestyle changes and providing information and support based on the answers provided.

The initiative

The Baby LifeCheck pilot was launched on the 7th February 2008 at twenty-three Sure Start children's centres. Trained facilitators supported the parents in navigating the LifeCheck.

The Teen LifeCheck was piloted in the four teen health demonstration sites of Bolton, Hackney, Northumberland and Portsmouth. Teen LifeCheck was launched in these sites in late February to early March 2007 with much of the promotion work being conducted between May and August 2007. The demonstration sites were asked to place particular emphasis on reaching the most vulnerable groups.

The pilot Teen LifeCheck consisted of less than thirty questions asking teenagers about their attitudes and behaviour around several topics such as smoking, drug taking, sex, school work, friendship, dealing with problems and emotions. Most questions had four possible responses to choose from. Each page contained links to other organisations' web pages which dealt in more detail with the particular topic concerned. On completion some text is generated that gives advice about dealing with particular issues and making healthy choices. This text varies depending upon the

answers given. There is also a list of links to local or national organisations that may help the individual with particular areas of concern that arose from their responses.

The NHS Mid-life LifeCheck is still being developed. Pilot programmes were due to be introduced in a range of settings in spring 2009. The settings included faith groups, workplace, acute hospital trusts, GP and health trainers, black and ethnic minority groups, the homeless and learning difficulties. Pilots were announced for Bristol, South of Tyne and Wear, Birmingham, London Borough of Camden, Tower Hamlets PCT and Guy's and St Thomas' Hospital London. Since the White Paper was originally published, the introduction of vascular checks has been announced.¹¹³ Development of Mid-life LifeCheck is being aligned with this policy.

The evaluation

At the time of undertaking this review, we were unable to obtain information about the evaluation of either Baby or Mid-life LifeChecks. Therefore in discussion with the Department of Health we focused our review on the evaluation of Teen LifeCheck. Subsequently, the Department of Health have published a summary of evidence about developing all three LifeChecks and this has been included more briefly in this appraisal.¹¹⁴

Baby LifeCheck

The Department of Health commissioned several pieces of work as an evaluation of Baby LifeCheck:

- Bunnyfoot, a behavioural research consultancy aimed to determine how well the self-completion tool performed by identifying problems that could affect its efficiency and recommending improvements to the tool itself.
- The People Partnership explored key stakeholders' views of how best their agencies or organisations could maximise their support of the LifeCheck. They also explored the acceptability of the tool to parents who had completed the LifeCheck and professionals involved in the pilot. In addition to this The People Partnership were asked for feedback on the idea for a LifeCheck which would focus on the parents' lifestyle and risk factors as opposed to their baby's.
- The Fatherhood Institute advised on the effectiveness of the tool in engaging fathers.

Design

Bunnyfoot sought the views of two experts on the usability of technology, tested it with twelve members of the public and considered how it could be used by people with disabilities.

The People Partnership carried out short *ad hoc* interviews and focus groups with parents who had used the LifeCheck during the pilot. They also conducted in-depth interviews with children centre staff.

Few details are provided of how the Fatherhood Institute provided input to the evaluation.

Evaluation team's findings and conclusions

Improvements to the tool suggested by Bunnyfoot were mostly incorporated into the pilot version of the LifeCheck that was launched in February 2008.

The People Partnership found that overall reactions to the LifeCheck were very positive. Centre staff suggested a number of ways and times to encourage parents to make use of the tool. It was thought to have the potential to trigger behaviour change. Amongst parents the most positive reactions came from younger less experienced parents, those from deprived backgrounds and those lacking support networks. This group included recent immigrants and fathers.

Overall the Baby LifeCheck was found to be appealing, empowering and informative. It was also thought to have a good balance between helping the 'worried well' and ensuring referral where appropriate. However it was advised that information accompanying the Early Years LifeCheck should place more emphasis on the goal setting aspect of the tool. Changes to the tool were also suggested in order to make results and feedback pages more accessible. Suggestions were made for the marketing and promotion of the tool when rolled out nationally which included the need to provide guidance for agencies about the roles and responsibilities for those involved in promoting use of the LifeCheck.

The fatherhood institute suggested separate tracks for mothers and fathers to make the tool more inclusive, and to ensure that the tool was marketed in a way that would appeal to men and addressed their needs.

Teen LifeCheck

Aims and objectives

The evaluation of Teen LifeCheck pilot was conducted in the four pilot areas in order to get feedback about and make changes to the Teen LifeCheck. This evaluation work was conducted by the Social Science Research Unit over an eight month period until September 2007.

Issues explored by the evaluation included:

- The level of awareness of Teen LifeCheck amongst all young people aged eleven to fourteen as well as amongst the target groups
- The level of completion among all young people and the target groups
- How acceptable young people find the Teen LifeCheck
- How satisfied young people (from different target groups) are with the Teen LifeCheck
- The impact of Teen LifeCheck in relation to knowledge/attitudes/behaviour (help seeking behaviour)
- How professionals working with young people perceive the Teen LifeCheck (including concept, content and design) and its likely impact on services
- What professionals and young people believe are the most effective methods of engaging young people in the Teen LifeCheck

Design

Data were collected from young people, parents and professional stakeholders.

A survey of eleven to fourteen year olds was undertaken to establish levels of awareness and completion of the Teen LifeCheck. It was completed by 2982 young people in secondary schools and pupil referral units in all four areas.

Data generated by the Teen LifeCheck website in the four areas covered by the pilot were collected between 6th February and 7th August 2007. This included data generated by people using the website as well as responses from a short evaluation page that was added to the Teen LifeCheck.

Focus groups were conducted with 156 young people aged nine to eighteen to explore the impact of and their satisfaction with the Teen LifeCheck. Although the Teen LifeCheck was targeted at eleven to fourteen year olds, focus groups were also conducted with nine to ten year olds and sixteen year olds in order to obtain views on extending the age ranges in both directions. To consider the views of lesbian, gay, bisexual or transgender young people a focus group was conducted with seventeen to eighteen year olds as it had been problematic accessing younger people to discuss issues in relation to sexual orientation. Focus groups were targeted at various groups of young people in order to obtain views from a range of participants and contexts (e.g. mainstream schools, pupil referral units, youth clubs, travellers and children with special needs).

Focus groups were also conducted with parents to consider their satisfaction with and the acceptability of the Teen LifeChecks. Fourteen mothers and one father took part in four focus groups in three of the four geographical areas covered by this pilot.

Semi-structured interviews with professionals and other local and national stakeholders were conducted to investigate staff perceptions and opinions of Teen LifeCheck. These included the four local Teen LifeCheck co-ordinators, people working as managers or directly with young people in health, educational or youth work settings and four national stakeholders.

Analysis

Survey data were analysed using cross sectional correlation analysis. The DH provided analysis and user rates from the online information collected when the Teen LifeCheck was completed. Thematic analysis techniques were used to identify the key themes arising from the focus groups and interviews.

Evaluation team's findings and conclusions

Young people's survey and focus groups: Levels of awareness were lower than hoped for but highest amongst the vulnerable groups, which suggested that targeted promotion had been effective. Eleven percent of those who responded to the survey had completed the Teen LifeCheck. When the Teen LifeCheck was explained in advance full completion was more likely. Teen LifeCheck was reported to encourage young people to consider various positive health behaviours but did not encourage them to use local services. The majority were positive about Teen LifeCheck as a concept and most were happy with the range of subjects covered. There was some dissatisfaction with the content and design and suggestions were made as to how these could be improved.

Parent's focus groups: Teen LifeCheck was thought to be a useful resource for both parents and young people but there was a minority with serious concerns about the quality, suitability and usefulness of the checks.

Professional stakeholder interviews: It was felt that significant improvement was needed and suggestions focused more on the content than the design. Professionals were generally positive about the concept.

Findings from young people and professionals: The Teen LifeCheck was viewed as having the potential to impact on knowledge, attitudes and behaviour if improvements were made. The assurance of confidentiality was an important issue for young people and professionals. School was thought to be a suitable location for promoting Teen LifeCheck and raising awareness but there was a concern about whether this setting would provide adequate privacy for completion.

Mid-Life LifeCheck

The Mid-Life LifeCheck was developed following input from expert advisors and then initial usability testing was undertaken by Bunnyfoot. The People Partnership also conducted sixty interviews with potential users to explore their views of the tool. They were generally positive about the concept and it appeared to have potential to help behaviour change, but it needed to be embedded within the other services available. The Mid-Life LifeCheck is therefore being redesigned. A series of pilot sites are planned to further refine and develop the tool using focus groups and interviews with stakeholder and potential user groups. The recent evidence summary report states that: 'A series of pilots planned for early 2009 will look not only at the effectiveness and appropriateness of the tool, but also at how the Department of Health can encourage take-up by building partnerships with community organisations.'

Comments on the evaluation

The Department of Health website refers to a rigorous independent evaluation of the LifeCheck.¹¹⁵ The development of the Baby LifeCheck appears to have been based on a thorough process, taking account of the views of experts, professional groups and potential users. This is a good example of formative evaluation, but does not provide summative evidence of the benefits of the LifeCheck.

Unfortunately due to the tight timescales for the evaluation some of the research on Teen LifeCheck was conducted before much of the promotion of this check had been done. Of those surveyed 31% had heard of Teen LifeCheck and although there were young people who had looked at Teen LifeCheck or answered a few questions, only 11% of respondents had completed it and were able to comment on their experiences of this.

Our interpretation of the findings

The evaluation of both Baby and Teen LifeCheck has been most useful in terms of suggestions for improving the resource and as an opportunity to explore how best to promote these checks and engage potential users. The evaluation so far provides limited evidence about the impact of NHS LifeCheck on attitudes and behaviour, or subsequent service use.

Conclusions about the evidence that the initiative achieved its objectives and delivered policy goals

The main aim of the LifeCheck initiative was to 'empower people to take charge of their own health and well-being'. This aim is too general to form a basis for evaluation, but it can be understood in terms of people having information which enables them to make choices about improving their health by reducing risky health behaviours. In addition there was an aim to reduce health inequalities by improving access to care by the most vulnerable groups.

The evaluation so far conducted has been largely formative and intended to improve the LifeChecks themselves. In this sense it has been useful. However robust evidence is lacking about the most important questions. These include:

- The extent to which health checks are undertaken by people in the target age groups
- Whether the health checks are used by those in high risk, vulnerable or marginalised groups, thus achieving the aim of reducing inequalities in access to care
- Whether those people completing health checks do actually make changes to their behaviour which improve their health
- Any impacts of the checks on patient empowerment, informed decision making and levels of anxiety
- The costs as well as the benefits of the NHS LifeCheck initiative, taking account of all costs (e.g. of subsequent investigations, consultations and other use of health service resources)
- Changes in health outcomes (e.g. heart attacks, strokes) amongst populations given access to health checks in comparison with similar populations without such access (possibly through comparison between areas, assuming that not all areas introduce NHS LifeChecks at the same time)

It is too early to expect evidence about these questions at present, but more detailed evaluation should be planned now and undertaken in due course.

In terms of the White Paper goals, there is good reason to believe that the LifeCheck initiative will improve access to services and choice for individuals. It has the potential to reduce inequalities in access. Evidence of the type described above will be needed to determine whether LifeChecks lead to 'better prevention and earlier intervention for improved health, independence and well-being'.

Partnership for Older People Projects (POPP)

NB This summary was based on reports available up to September 2008 only. The final report was published in January 2010.

Background

The White Paper made a commitment to provide support for people with long term conditions to enable them to manage their conditions themselves with appropriate help from health and social care services, and to develop assistive technologies to support people in their own homes.

Five key policy drivers have been identified for the Partnership for Older People Programme:

1. Public Service Agreement targets¹¹⁶ relating to long term conditions and the support of vulnerable older people.
2. The National Service Framework for Older People,¹¹⁷ a ten-year programme setting out the standards of care for people over the age of fifty-five, which requires health and local authorities to develop new ways of working and new ways of providing services.
3. Department for Work and Pensions LinkAge Plus pilots, which provide older people with access to more integrated services.¹¹⁸
4. The 2004 Public Health White Paper 'Choosing Health' which identified six key priority areas including tackling health inequalities.⁴⁴
5. The NHS and Social Care Model for long term conditions⁴⁶ which set out a model designed to improve the care of people with long term conditions

The National Service Framework for Older People was published in March 2001, setting national standards to improve services for older people at home, in residential care and hospitals. The 2004 Spending Review provided ring-fenced funding for Councils with Social Services Responsibilities to establish innovative pilot projects in partnership with PCTs, the voluntary, community and independent sector. The pilots were aimed at large-scale system reform across health and care services to deliver improved outcomes for older people through greater investment in prevention. This is in direct support of the key aims of the White Paper 'Our Health, our care, our say' and delivery of the two Public Service Agreement (PSA) targets with regard to long term conditions and supporting vulnerable older people. These were firstly, relating to long term conditions, to improve health outcomes by offering a personalised care plan for vulnerable people most at risk and to reduce emergency bed days, and secondly supporting vulnerable older people by increasing the proportion of older people being supported to live in their own home.

The £60m Partnerships for Older People Projects (POPP) grant was announced in March 2005, to encourage councils in England with their NHS, local government, voluntary and community sector partners to devise innovative approaches to establishing sustainable arrangements for supporting older people in active and healthy living.

Aims

The overall aim of the POPP is to improve the health, well-being and independence of older people. A total of twenty-nine local authority-led partnerships including health and third sector partners were funded by the Department of Health to deliver and locally evaluate innovative schemes for older people. The projects' aim was to:

- Provide person centred and integrated responses for older people
- Encourage investment in approaches that promote health, well-being and independence for older people, and
- Prevent or delay the need for higher intensity or institutionalised care

The initiative

The POPP programme has two rounds of pilot sites: nineteen Round 1 sites were established in May 2006, developing 381 projects, and ten Round 2 pilot sites followed in May 2007, initiating 89 projects, giving a total of 470 projects. The Round 1 pilot sites were required to conclude their piloting work by 31st March 2009. At the end of March 2008, with the exception of one site (which had completed as a pilot and had moved fully into mainstreaming), the remaining eighteen sites were continuing to run for up to a further twelve months during which period they will incrementally move into mainstreaming.

Common aims of the pilot projects included a shift to prevention and the reduction of emergency admissions and/or bed days for older people, meeting the specific needs of socially excluded older people, those with or at risk of mental health problems and black and ethnic minority groups. Most (71%) of the projects focused on 'Universal Services', such as handyman schemes, gardening, shopping, leisure and signposting services, designed to support older people in maintaining independent lives in their own homes and to improve their general well-being, with 14% of projects providing 'Additional Support', such as medicines management, telecare services, falls services, holistic assessments and mentoring services to support older people 'at risk' of admission to hospital. 8% of the projects provided 'Specialist Support', such as Community Rapid Response, Hospital at Home and Intensive Support Teams, to help older people at serious risk of imminent hospital admission. The final 7% of projects did not provide services directly to older people, but were designed to support projects through staff training, capacity building in the voluntary sector and needs mapping.

The evaluation

Design

The National Evaluation is taking a case study approach as developed by Robert Yin.³⁵ This approach was chosen because it takes full account of the importance of local context. The evaluation incorporates both quantitative and qualitative methods. The case study approach is said to offer a robust framework that allows evidence from very different paradigms to be brought together and synthesised.

The evaluation involves three phases of working:

- **Phase 1 includes a literature review and explores** activity across all of the twenty-nine sites. This involves analysis of key documents including project initiation documents, minutes of relevant meetings, strategy

documents such as local area agreements and documentation exploring key, for example, partnership structures. This phase also includes the collection and analysis of data within a core, or minimum dataset. Finally within this exploratory stage, a key informant questionnaire has been administered.

- **Phase 2** is an **explanatory** phase, using initial data to try to explain what is happening within the sites. For example, why is it that some of the sites appear to be improving the quality of life and/or well-being of their clients or perhaps providing real progress toward the PSA indicators, whilst others may be less successful in achieving their targets? Five sites were selected to begin to build an explanatory model for why one partnership approach, for example, may be more effective than another, based on the key informant interviews, interviews with older people and focus groups.
- **Phase 3** involves a Framework Development, bringing together the theoretical and empirical work to develop a model of partnerships/financial organisation/preventative models that can be transferred to other care groups.

Evaluation team's findings and conclusions

Interim Reports have been completed in October 2007 and October 2008, with a Final Report due in October 2009.

The Key Messages from the reports as at October 2008¹¹⁹ were:

- 99,988 individuals had received, or were receiving, a service within the POPP programme
- POPP pilot sites have a demonstrable effect on reducing hospital emergency bed-day use when compared with non-POPP sites
- The POPP projects are having an effect on how users perceive their quality of life as a whole. Following the project, users report they see their quality of life as improved
- Users also reported that their health-related quality of life improved in five key domains, (mobility, washing/dressing, usual activities, pain and anxiety), following their involvement in the POPP projects
- An analysis of those sites where data are currently available (eleven out of twenty-nine sites) appears to demonstrate the cost-effectiveness of POPP projects
- The POPP programmes also appear to be associated with a wider culture change within their localities, with a greater recognition of the importance of including early intervention and preventative services focused toward well-being
- POPP partnerships across the health and social care economy seem to have strengthened and accelerated developments around joint commissioning. In particular, there has been recognition of the value of involving voluntary and community organisations in service planning and delivery
- Involvement of older people within the POPP sites appears to be focused on the delivery of services; almost half the staff in the projects across the POPP programme are older volunteers

- Only fifteen (4%) of the total 470 projects across the POPP programme have indicated that they do not intend to sustain their service after the end of DH funding

Comments on the evaluation

The case study method approach uses both quantitative and qualitative data, but is primarily based on a qualitative rather than quantitative paradigm of knowledge. As such it is good for obtaining rich data from multiple sources, and for generating explanatory frameworks to explain how apparent benefits from a new model of organisation might occur. The evaluation has used this approach to generate useful information about the range of activities undertaken within sites, partnership working between organisations, and the involvement of older people themselves.

The POPP interim report makes relatively strong headline claims about the effectiveness and cost-effectiveness of the projects, but these conclusions should be tempered by awareness of the limitations of the study design and data available (as recognised by the report's authors within the report itself). Data about effectiveness is based on just 551 individuals who completed questionnaires before and after their involvement in POPP. No information is given about the response rate or representativeness of these individuals and it is not clear that adjustment is made for clustering of individuals with the small number (n=11) of sites contributing data. The lack of data from concurrent controls means that it is impossible to know if any changes observed over time are due to the intervention or to other factors which changed over time.

Data about costs are based on routinely collected data and involve comparison with non-POPP sites. No information is provided about how closely these control sites match the POPP sites. It is well recognised that routine data are often incomplete and of uncertain reliability. Some data about the costs of service use came from the survey of 551 individuals, with the same limitations as discussed above.

These concerns mean that based on information currently in the public domain, conclusions about cost-effectiveness should be very tentative. Despite this concern the magnitude of the effects observed means that there is good reason to think that the POPP may well prove to be cost-effective.

By using a case study approach rather than an RCT, the POPP national evaluation team acknowledge that attributing effects to the programme is problematic. Proponents of the case study approach would argue that triangulation of the qualitative findings with the quantitative helps to provide evidence about why observed changes have occurred, but it is arguable that this type of evidence is still closer to hypothesis generation than hypothesis testing. It is only possible to approximate what appears to have been prevented by the projects, due to the nebulous and incalculable nature of prevention as a measure.

Other relevant evaluation work

The local projects included in our overview were all also participating in the national POPP pilots or LinkAge projects. These projects included additional research activity beyond the evaluation activities associated with the national evaluation, in order to address specific local needs. Most of these involved only internal reporting of these additional elements, but where information was available to us the findings suggest

that projects did increase supported living at home, reduce hospital admissions, increase quality of life and increase community activities.

Conclusions about the evidence that the initiative achieved its objectives and delivered policy goals

In advance of the forthcoming final report of the national evaluation, the evidence from the POPP evaluations strongly indicate that the programmes have improved awareness of, access to and engagement with an improved range of services among older people. Improvements in health related quality of life, and reduced emergency hospital bed day use are also suggested by the evaluation, although the strength of evidence in regard to these outcomes is weaker for the reasons discussed above. The long term persistence of these improvements and the sustainability of projects also remain unclear from the available evidence, although the fact that most sites intend to continue the service after the pilot phase suggests that they are perceived locally to be successful.¹¹⁹ Overall, the evidence from the evaluations suggests that the White Paper goal of supporting people with long term conditions in their homes has been successfully addressed by these projects in these pilot sites.

Self-Referral to Physiotherapy

Background

The concept of self-referral (without referral from a GP) for NHS services gained momentum during the 1990s. The White Paper noted that although self-referral to allied health professionals was not a high priority raised by patients in the listening exercises, there were good examples from around the country of how this approach had increased the accessibility of services. An example highlighted in the White Paper was the PhysioDirect service in Huntingdon, in which patients can contact a physiotherapist by telephone and then be seen in person if necessary.

Providing timely access to physiotherapy has long been a problem in the NHS, so physiotherapy is a good example of the need for simpler, quicker and more convenient routes to care. The concept of self-referral to physiotherapy began to gather momentum in the 1990s, reflecting many other initiatives to promote a more patient-focused NHS.¹²⁰⁻¹²³ In addition, there has been a strong push from the profession, represented by the Chartered Society of Physiotherapists, to allow patient self-referral. To some extent this reflects a long term trend of allied health professional groups seeking to strengthen their right and ability to manage their own patients, without referral from doctors.

An international perspective shows that models of care based on direct access by patients have been developed since the 1970s in several developed countries, often couched in the rhetoric of promoting professional identity. In Australia, for example, a seminal publication in 1977¹²⁴ pioneered the 'professional advance' of direct access in physiotherapy. In the United States, since the early 1980s individual states have been 'fighting for the right of patients to see physical therapists without having to obtain a physician referral'.¹²⁵

In the late 1990s, self-referral to physiotherapy began to be implemented and studied in the UK. A pioneering study found that patients used a self-referral service appropriately and that GPs wanted the service to continue.¹²⁶ Fears of being inundated with referrals and missing important diagnoses were not realised. A series of observational studies by Holdsworth et al. have expanded the evidence base, including an observational study of self-referral by patients from twenty-nine general practices throughout Scotland.¹²⁷⁻¹³⁰ These studies have suggested that self-referral is feasible, acceptable to users and providers, and leads to reductions in GP workload. Referral rates differ according to geographical setting, with rural areas experiencing the highest rates of referral.¹³⁰ Less than 20% of areas experienced an increase in referrals following the introduction of a self-referral system, but it is notable that most of the areas studied had a much higher level of provision at baseline (53/1000) than is the norm in England (GP referral rate 23/1000).¹³¹ It is also worth noting that the Scottish studies only studied the impact on demand for physiotherapy over a short period of time. Only about 22% of patients did in fact refer themselves,¹³⁰ with most patients being referred by GPs as usual, so it is difficult to assess the long term impact on demand as patients become more aware of the possibility of self-referral.

Holdsworth et al. also showed that, compared with patients referred by GPs, those who referred themselves tended to have had their condition for less time, and had been absent from work in lower proportions and for less time,^{127;128} suggesting that self-referral may increase access for people with less severe problems. If self-referral

does increase demand on physiotherapists this is beneficial if it represents meeting previously unmet need, but may not be cost-effective if it involves diverting resources to treating people who are less in need.

In 2004 the Chartered Society for physiotherapists called for self-referral for all primary care patients by 2007, and leadership by senior physiotherapists to deliver the changes.¹³² The intended benefits are described below:

Aims

The aims of self-referral can be identified from the CSP briefing document¹³² as well as in the White Paper itself.

The main intended benefits are:

- Improved clinical outcomes, particularly with a faster return to work
- Reductions in waiting times for patients
- Improved patient choice
- Reduced costs
- Reduced workload for GPs and other help-professionals.

In terms of wider NHS goals, this is consistent with the goals of improving access to services, increasing patient choice and putting patients in control of how their care is delivered, improving public health and improving the quality and consistency of health care.

The ways in which these benefits would be achieved (the 'mechanism') are also clearly articulated in the CSP briefing document. Self-referral would mean that patients did not have to go through the step of visiting the GP acting in their role as gate-keeper. Since musculoskeletal problems are one of the most common reasons for consulting a GP, and physiotherapy is the treatment of choice for many of these problems, this would have a substantial impact on GP workload, reduce costs, reduce one step in the access chain for patients and enable them to see a more appropriate health professional more rapidly, therefore improving patient outcomes.

The initiative

The DH worked with the CSP to identify six pilot sites, which would test the impact of self-referral to physiotherapy in England. Three of these sites were in or around London, two were in the Southwest of England and one was in Birmingham. From December 2006 these sites would accept patients who self-referred for physiotherapy. Patients had to complete a self-referral form, and the remainder of the care pathway was not dissimilar from previous systems – there was no PhysioDirect type telephone assessment or email access. Self-referral was introduced in addition to patients being referred by GPs in the usual way. All referrals were prioritised and treated in the same way, irrespective of the source of referral.

The evaluation

Aims and objectives

The aim was to evaluate the impact of introducing self-referral to musculoskeletal physiotherapy,¹³³ and in particular to:

- Seek the views of patients, GPs and physiotherapists

- Establish whether significant numbers of users of private physiotherapy services would return to the NHS
- Identify changes in waiting times
- Understand differences in accessing self-referral according to ethnic background
- Gauge the difference that self-referral made to GP workload
- Build on rather than duplicate existing work

Design

Patients at the six pilot sites were able to refer themselves, as described above. From December 2006 patients were allowed to self-refer. Data were collected about referrals (self-referrals, GP referrals, and 'GP-suggested' referrals) over a one year period from December 2006 to December 2007. The sites received no extra funding apart from initial costs of advertising and relevant paperwork. Data were collected as follows:

- Historical data about demand, activity levels and waiting times
- Anonymised data about 2835 patients. This data included details at baseline about socio-demographic status, severity of condition, duration or symptoms, employment status, use of alternative providers. At discharge, data were only collected about reason for discharge, patient perception of condition severity and total number of contacts
- Feedback forms sent to GPs and physiotherapists

Data were subsequently analysed by Dr Lesley Holdsworth, who had conducted the original research in Scotland. Analysis was mainly descriptive, but some comparisons are made between the groups (self-referred, GP-referred, GP-suggested). Data from one site were omitted because of concerns about data reliability.

Evaluation team's findings and conclusions

The headline findings reported¹³³ were that the service led to:

- High-levels of user satisfaction
- A more responsive service
- Patients were empowered to self-care
- Lower levels of work absence
- No increase in demand for services
- No evidence that black and ethnic minority groups used self-referral less than white groups
- No return to the NHS by patients traditionally seen by private physiotherapists
- Well accepted by physiotherapists and GPs
- Lower NHS costs

Comments on the evaluation

It is important to note that this is a simple observational comparison of patients who chose to self-refer or had a referral made by or suggested by their GP. It can provide evidence about differences between the types of people who chose different routes of referral. But in the absence of any control group it does not provide any evidence about whether offering self-referral leads to the benefits intended for the service e.g. improved patient satisfaction, reduced costs, shorter waiting times, faster return to work or improvements in other outcomes. To make this type of comparison it would be necessary to compare patients who were self-referred or GP-referred in terms of their outcomes after they accessed physiotherapy, after adjusting for any differences between the patients at baseline (and it is evident that people who use self-referral have different characteristics from those referred by GPs).

The design of the evaluation means that it could not address many of its aims (aims 2, 3, or 5). Nor could it test whether the intended benefits of self-referral to physiotherapy are achieved (e.g. reductions in waiting times or costs). This would require a trial comparing patients in areas which do or do not offer self-referral (ideally but not necessarily randomised).

The findings from the evaluation appear to have been over-stated in the evaluation report in some cases. For example, it is claimed that self-referral led to less work absence. However, no data about work absence was collected at follow-up, so it is not possible to demonstrate an impact of self-referral. The data show that those people who chose to self-refer had had their symptoms for a shorter period and had been off work for less time *before* they accessed physiotherapy compared with those who were referred by GPs. This suggests that the group who self-referred were less severe and/or accessed physiotherapy at an earlier point in the illness episode, but it does not provide any evidence that this resulted in them getting back to work any more quickly. Similarly, the headline findings claim that self-referral was associated with reduced costs but no robust evidence is provided in the report to demonstrate this. Instead, reference is made to the earlier Scottish research, which has the same limitations as this evaluation since it was an observational study with no control group.

Our interpretation of the findings

This evaluation suggests that self-referral is appreciated by patients (although no details are given about response rates to the patient survey, or comparison with patients who were GP-referred), and is also popular with GPs and physiotherapists.

Comparison with historical, routinely collected data shows no evidence of an increase in demand following self-referral. It is important to note that routinely collected data are often of doubtful reliability and may not be comparable between areas or services.

There is a suggestion that self-referral may be accessed differently by different ethnic groups and this requires further investigation (even though there were only small numbers of BME patients in the study, the association between referral type and ethnicity was of borderline statistical significance ($p=0.055$)).

There is no robust evidence from this evaluation about the impact of self-referral on the important outcomes of patient clinical outcome, return to work, waiting times, GP workload or NHS costs.

Other relevant evaluation work

Three local evaluations of self-referral were identified in this overview. These were conducted by physiotherapy services within PCTs. One of these was examined in detail as a case study. All three local evaluations were based on an audit of performance against standards and included patient feedback using questionnaires. Two of the three local evaluations described ambitious aims, such as to determine whether self-referral reduces GP workload or increases demand on the physiotherapy service, and involved collecting considerable amounts of data. However the design of the evaluations, which included no control group or before-and-after data, meant that it was not possible to answer these questions. It was encouraging to observe the willingness of local services to assess their activities against performance standards, but there was a lack of clarity about how the data collected would answer the evaluation questions.

Conclusions about the evidence that the initiative achieved its objectives and delivered policy goals

The main aims of self-referral are listed again below, with comments about the evidence so far available:

Improved clinical outcomes, particularly with a faster return to work. There is currently no evidence about the impact of self-referral on clinical outcomes or return to work. This is an important priority for further research.

Reductions in waiting times for patients. It might be predicted that self-referral would lead to increased demand which would increase rather than reduce waiting times. Assessing the impact of service change on waiting times is difficult, because so many other factors and concurrent initiatives affect waiting times and it is difficult to attribute changes to any one factor. Overall, the evidence from the White Paper self-referral pilots and the Scottish data suggest that, in itself, self-referral does not increase demand on physiotherapy services compared with GP referral when the supply of physiotherapy services is sufficient to offer services to 50/1000 patients per year. However this level of supply is much higher than currently provided in some areas of England, so self-referral may expose the service to increased demand due to previously unmet need.

Improved patient choice. It is self-evident that self-referral achieves this aim, so evaluation may not be necessary. However, it would be worthwhile to consider whether self-referral increases choice for all groups of patients e.g. those with language or literacy difficulties.

Reduced costs. Although patients who self-refer have lower GP consultation costs and other NHS costs than those who are referred by GPs, these patients also have less severe illness than those referred by GPs.¹²⁹ In the absence of a self-referral option it is possible that these patients would not have consulted either a GP or a physiotherapist. Therefore there is an *a priori* reason to think that self-referral may increase rather than decrease costs, both within physiotherapy services and in the wider NHS. There is at present no robust evidence that self-referral is associated with reduced or increased costs.

Reduced workload for GPs and other help-professionals. For the same reason as the above, there is no robust evidence about this issue.

White Paper goals. In terms of the White Paper goals, it is evident that self-referral improves patient choice in how they access services and it also is consistent with ensuring that patients are in control and services are responsive to patients' wishes. However further evidence is needed before it can be concluded that self-referral is associated with improvements in the quality of care or health outcomes.

Social Enterprise Pathfinders

Background

Social enterprises can be defined as organisations using business methods to address social problems although social enterprise is normally explained as a rational and functional solution to public sector funding and philanthropic resource constraints.¹³⁴ In addition social enterprise is also said to represent a strategically better option for organisations to fulfil their pro-social mission.^{134;135} However, social enterprises in the UK are not a new organisational form nor do they reflect a new approach. There is a long tradition of co-operative business organisations and voluntary sector agencies providing health and welfare services.¹³⁶ In particular since the 1970s there has been an increasing involvement of voluntary organisations and charities in the delivery of social care funded by local authorities with further growth in both health and social care since 1990 and the commitment in the Community Care sector to encourage voluntary and private involvement in the health and social care sectors. This can be seen as part of the growing shift in the last twenty years to a mixed economy of care and a concern that the delivery of services by State bureaucracies was inefficient and unable to adequately respond to the needs of individuals. Voluntary, and increasingly, private agencies were seen as more responsive and innovative. These broad ideas about the nature of welfare services and the role of the State are important as they provide the context for the development of alternative approaches to the organisation and delivery of health and social care services and changing structure of health services in England. Social Enterprise organisations are seen as being more innovative and responsive than traditional NHS and other public service organisations.

Since being elected in 1997 the Labour Government has sought to promote social enterprise as a mechanism for service delivery and community regeneration over the past decade. In 1998 the Government established the New Deal for Communities (NDC) to promote regeneration through partnerships of local organisations including business, voluntary bodies and the private sector.¹³⁷ Then in 1999 the DTI established the Phoenix Fund allocating £29 million between 2001 and 2004 as 'part of the Government's strategy for using entrepreneurship to tackle social exclusion'¹³⁸ followed in 2002 by a joint DTI/ODPM Adventure Capital Fund to invest in 'independent community-based organisations'.¹³⁹ There is now some enthusiasm in policy circles to promote the ideals of social enterprise as a 'third way' to deliver public services to deprived communities where the public and private sectors have failed to do so.¹⁴⁰

The Labour Government has also launched the 'Social Enterprise Coalition' and created a 'Social Enterprise Unit' to improve the knowledge of social enterprises and, above all, to promote social enterprises throughout the country. Within the same framework, the Department of Trade and Industry¹⁴¹ has defined social enterprise (and also the 'Community Interest Company') organisations driven primarily by social objectives and achieving sustainability through trading but stresses that 'A social enterprise is, first and foremost, a business. That means it is engaged in some form of trading, but it trades primarily to support a social purpose'.¹⁴¹

While most activity in health and welfare has focused on social care mainly funded by local government there has been a long tradition of agencies providing health services in areas such as palliative care, rehabilitation and home care for example. Before the publication of 'Our health, our care, our say' there had been a growing interest in

government in developing new forms of community and social ownership organisations for the delivery of public services involving either the transfer of public services from the state or funding new organisations to develop new welfare services.¹⁴² The inclusion of proposals for social enterprises in 'Our health, our care, our say' was not, therefore, unsurprising and late in 2006 the Department of Health announced a fund for supporting social enterprise organisations and invited proposals. From this twenty-six 'pathfinder projects' were selected.

Box 4-1 Examples of Social Enterprise Pathfinders

SCA Healthcare is an industrial and provident society. It is proposing providing a range of community services including a long term condition resource centre and support to carers and relatives from a community hospital in Southampton.

The Bridge is a proposal to set up a social enterprise to deliver alcohol and substance misuse programmes in London. The service will include residential accommodation, therapy programmes, modular treatment centres, and a back to work programme.

Leicester Homeless Primary Care Service is proposing providing primary medical services to vulnerable patients. Its primary health care centre will be co-located with a 42 bed night shelter and a YMCA drop in centre.

The Open Door is proposing providing a wide range of primary care services for vulnerable people in the Grimsby area, in addition to support into employment, gardening and music therapies, exercise and cooking skills.

DCP is a third sector organisation that will set up a social enterprise in Newcastle to deliver services to people living with dementia, and their carers.

Based in Middlesbrough, the Developing Partners Project aims to develop and provide user led training for health workers, and user led research and evaluation of health services.

To support the development of the social enterprise pathfinders the Department of Health established the Social Enterprise Unit in 2007 to lead the Social Enterprise Pathfinder Programme (SEPP). The programme is supporting twenty-six social enterprise pathfinder projects to deliver health and social care services with total funding of £9m over two years. The Department also set up the Social Enterprise Investment Fund (SEIF) which provides an additional source of funding of £100 million to further support and encourage social enterprises to enter the market.¹⁴³ This is subject to a separate evaluation to the evaluation of SEPP. The Darzi review confirmed the Government's support for expansion of the role of social enterprises in the provision of health and social care and further expansion of social enterprises and community interest companies remains a key policy objective of the Department of Health and Government. However, a recent survey of Primary Care Trusts examining the extent to which they had separated their commissioning and providing functions found that most PCTs stated preferences to become foundation trusts with only 12% stating a preference to develop social enterprise organisations for their provider functions.

Aims

The aims of the SEPP as outlined in 'Our health, our care, our say' were essentially two fold – to provide increased capacity and choice and address inequalities. A key aim was to expand community health services' capacity and responsiveness to patients by '...encouraging or allowing new providers, including social enterprises or commercial companies, to offer services to registered patients alongside traditional general practice. Increased capacity – and contestability – will allow people to choose services that offer more convenient opening times, tailored specialist services or co-location with other relevant services' (para. 3.11). The White Paper also highlighted current inequalities in service and the need to develop service in areas where services were lacking or by being more responsive to needs than traditional services. The aim was to '...ensure that both new and existing providers are allowed to provide services in underserved areas. Social enterprises, the voluntary sector and independent sector providers will all make valuable contributions in the longstanding challenge of addressing inequalities. The voluntary and community sectors often have strengths and experience in delivering services to those people who are not well served by traditional services'(para 3.43).

The initiative

The Social Enterprise Pathfinders cover a wide range of community health and social activities. The DH advertised in October 2006 for organisations to become social enterprises – these could be new organisations or existing agencies wanting to develop their activities. A key goal was not specifically to establish the social enterprises but for the DH to develop its own role as a social investor supporting the development of new agencies. This is not a new role for government and local authorities and other central government departments have been involved in similar activities for many years. The development of the social enterprise programme was seen as part of a broader initiative in the Department of Health on social enterprise and developing new forms of health and social care providers. The programme was partly in response to commissioners asking for evidence about the benefits of social enterprise organisations. In a sense the pathfinders were in many ways a 'research cohort' with or without the evaluation. The experience of the programme has informed Department thinking on social enterprises and the role of the enterprise investment fund. Two key areas were of interest to the Department of Health. The first was information about the way the organisations worked and what they did and to gain information about the market and market behaviours. The Department ran a selection process for the pathfinder programme. The incentive for organisations to apply was that they could get additional funding from a £1,000,000 start-up fund.

The Department held an open tender process for the programme evaluation. Newchurch was the successful applicant. The policy team has worked closely with the evaluation team and they hold regular meetings.

The evaluation

Aims and objectives

The aim of the national evaluation project was to inform future policy towards social enterprises in health and in practice in the NHS and in social enterprises themselves. The Department was keen to find out the advantages of social enterprises, what they could deliver, what kinds of changes they brought about in local health economies

and for service users and how they affect health care. The research project aimed to evaluate success in meeting the Social Enterprise Programme goals and test assumptions about the ability of third sector organisations to play an important role in delivering health and social care services. It was intended to have a national evaluation right from the very beginning and the policy group sought advice from the Department's Policy Research Programme staff. The policy team were also very keen to get feedback on their own role.

Design

The research involves both a formative evaluation to engage the pathfinders in the design of the research and in sharing learning through workshops and a summative evaluation employing a multi-level framework of evaluation criteria at programme level, local health economy level and the individual social enterprise level. The approach involves examining system level effects and by examining actions against objectives.

Methods

The evaluation used a mixed methods approach including interviews, workshops and focus groups and quantitative methods including analysis of survey data and unit costs. The SEPP evaluation is based on data collection from all twenty-six (subsequently twenty-five) pathfinder social enterprises with in depth case studies of a sub sample of six pathfinder projects. The first stage of the research was to develop the evaluation framework. The framework has a number of broad dimensions – context, organisational capability, contribution (of the SEPPs), impact on local health and social economies, quality (choice, access, risk management, user views) and workforce. There was also an attempt to examine cross cutting issues (such as sustainability, transformation). The research involved an initial telephone survey of all pathfinders followed by case study research and then another survey of all pathfinders.

Evaluation team's findings and conclusions

(At the time of writing, these are pre-publication findings not for public release)

The final report of the evaluation provides details related to the evaluation framework criteria. This provides a good overview of the organisation and activities of the pathfinder projects. The report highlights the generally supportive attitudes of NHS organisations and the positive views of the pathfinder projects themselves about the benefits of social enterprise. However, there were concerns about whether expected incomes had been achieved and also all but four SEPPs were delayed in setting up due to the complexity of what they were trying to achieve. The evaluation highlighted the need for stronger commissioner support for setting up social enterprises and the then subsequent letting of contracts to them. The evaluation also highlighted the need for additional resources to support the programme. These were considered as key to the success of SEPPs. The evaluation concludes that as twenty-four of the original social enterprise pathfinders were in existence at the end of the evaluation and that sixteen were set up as social enterprises and twelve were trading that this made the programme worthwhile. In particular the evaluation team highlight factors such as gaining local contracts and national recognition as success factors and proof that the social enterprise model is 'fit for purpose'.

Comments on the evaluation

Over the evaluation period, difficulties establishing pathfinder sites resulted in shifts in the evaluation, and consequently affected the degree to which the original evaluation framework remained relevant. This, together with the broad nature of the evaluation framework makes it unclear how the framework has been applied except in a descriptive way. The evaluation categories have been used as a way of presenting organisational data but it is less clear how they informed the framework for analysis. The discussion is not linked to any concepts or organisational model or framework for analysis. The definition of performance or success is not identified in an explicit way. There is a tendency towards generalisation rather than constructive analysis.

Our interpretation of the findings

It is difficult from the report to link methods, findings and analysis, therefore making it difficult to attribute particular findings to any specific organisational model or specific context. The use of qualitative data is limited, with the report summarising key points without reference to the type or form of organisation of SEPP. This problematises how confidently statements can be made about the formation of social enterprises and what they have achieved. As the evaluators highlight, the start-up process of pathfinder sites took considerably longer than anticipated, undermining the capacity of the evaluation to meet their evidence goals. In turn, it is difficult to set apparent successes against the original research objectives.

Other relevant evaluation work

No other reports were identified and no further studies in progress were found. However, the national evaluation report refers to the fact that some SEPPs were being evaluated locally. No information was found about these studies.

Conclusions about the evidence that the initiative achieved its objectives and delivered policy goals

The final report gives assurance to the Department that social enterprises are successful and a suitable organisational model for developing further provider services. While there must have been a substantial amount of data collected during the evaluation it is not clear from the final report what data was collected or how it was analysed. Key success indicators in the final report appear to be about whether the organisations were still in existence and whether they had either achieved social enterprise status or had received contracts for services. There is little analysis of organisational models or reflection on whether the investment put into social enterprises has produced additional benefits. Lack of clarity about evaluation objectives is apparent given the three fold areas of analysis relating to whether social enterprises work, whether voluntary and third sector agencies can contribute effectively to health and social care services and providing guidance on developing social enterprises. The evaluation does provide some information about the need for better support and resourcing but does not specifically relate this to whether the social enterprises have brought additional benefits. Clearly in the time available any summative assessment of benefit to service users is difficult – especially as many pilots experienced substantial delays in their development. There is already substantial evidence and experience that third sector organisations can deliver a wide variety of health and social care services pre-dating the social enterprise model. In fact some of the pilots already operated in the voluntary sector prior to joining the

pathfinder programme. Also local pilots had different criteria regarding how they might be evaluated in terms of focusing on local context and service developments.

We have noted that the organisation evaluating the social enterprises is also involved in supporting NHS organisations to develop social enterprises. It had been providing consultancy services to local PCTs advising them to develop social enterprise organisational models (e.g. Kingston PCT). The evaluator's internal governance procedures guard against such potential conflicts of interest introducing bias, and it is important to ensure that where evaluators are in this position such measures are in place.

Whole System Long Term Condition Demonstrators

Background

'Our health, our care, our say' makes a commitment to provide people with good quality social care and NHS services in the communities where they live. This commitment is detailed in chapter 5 of the White Paper, which addresses supporting people with long term needs. This chapter on ongoing care and support includes discussion on:

- empowering those with long term needs to do more to care for themselves, including better access to information and care plans
- investment in training and development of skills for staff who care for people with ongoing needs
- new supports for informal carers including a helpline, short-term respite and training
- collaboration between health and social care to create multidisciplinary networks to support those people with the most complex needs

Previously, the Government introduced a model designed to help improve the care of people with long term conditions in Supporting People with Long Term Conditions, published in January 2005.⁴⁶ This followed the setting of a national Public Service Agreement (PSA) target for improving outcomes for people with long term conditions by offering a personalised care plan for vulnerable people most at risk, and reducing emergency bed days by five per cent by 2008 through improved care in primary and community settings.

Within 'Our Health, our care our say', a number of 'Whole System Demonstrators' were promised to explore the possibilities of integrated health and social care working, supported by advanced assistive technologies such as telehealth and telecare. The demonstrators aimed to offer a better understanding of the level of benefit associated with such developments, as well as help fast track future change by addressing the key implementation barriers and providing solutions for the wider NHS and Social Care.

The development of the initiative was informed by evidence of the potential benefits of telehealth and telecare from both smaller scale and pilot projects in the UK as well as the integrated health systems provided by the Veterans' Health Administration in the USA, which provides such systems for 7.6 million veterans and has been shown to both reduce hospital admissions and lower costs.¹⁴⁴ In the UK, the evidence base relating to the benefits of telehealth and telecare was felt to be insufficient, but did point to emerging or possible benefits for improved quality of life and clinical outcomes for certain conditions, in the provision of information, and in cost benefits relating to hospital admissions.^{145;146} There is good evidence that telehealth is feasible and can improve health behaviour, but a lack of evidence about mechanisms of action, clinical outcomes, cost-effectiveness, patient satisfaction, impact on service utilisation, and acceptability.^{147;148} A need was identified for robust randomised control evidence pertinent to the UK experience.¹⁴⁵

The Department of Health strategy unit identified that there was good use of telehealth in America and only some small pilots in the UK in addition to investment in

some areas of telecare. It was decided that the environment was sufficiently distinct in the UK to necessitate the establishment of Whole System Demonstrators to examine the usefulness of telehealth and telecare systems more widely.

Aims

The initiative aims to examine the impact of whole-systems redesign and the introduction of telecare and telehealth on:

- Levels of hospitalisation and patterns of service use
- Costs of health, social care and other services, and wider societal costs
- Cost-effectiveness
- Quality of life
- Clinical outcome

And also to:

- Examine the potential mechanisms by which telehealth and telecare may impact on patient and carer outcomes
- Determine the organisational factors that facilitate or impede the adoption of telecare and telehealth
- Identify patient factors that predict the impact and use of telecare and telehealth
- Examine patient, carer and professional experience of the implementation of telehealth and telecare
- To explore patient/service user, carers and professionals' attitudes to and engagement with advanced assistive technologies and their impact on caring relationships
- To determine whether telecare and telehealth increase patient/carers autonomy and empowerment

The initiative

Whole Systems Demonstrator sites were set up to test the benefit of integrated health and social care systems, supported by assistive technologies. Individuals were assessed for appropriate assistive technologies available from the Telecare National Framework which supported individuals at home to maintain their independence, according to their individual circumstances.

Three PCT sites with an established record of joint health and social care working (Newham, Kent and Cornwall) were designated as Whole System Demonstrator sites and implemented integrated care and other innovations as part of a 'whole systems redesign'. 'Whole systems redesign' was supported through advanced assistive technology (based on telecare and telehealth) for patients with long term conditions and complex health and social care needs. Within these sites, the initiative was targeted at the following groups:

1. Over eighteens with social care needs
2. Individuals with health care needs at risk of future or current hospital admission due to:
 - COPD
 - Heart failure
 - Type 2 diabetes
3. Those with both health and social care needs as defined above

The programme was launched in December 2006, with demonstrator selection completed in April 2007 for commencement in June 2007 and completion planned by March 2010.

The evaluation

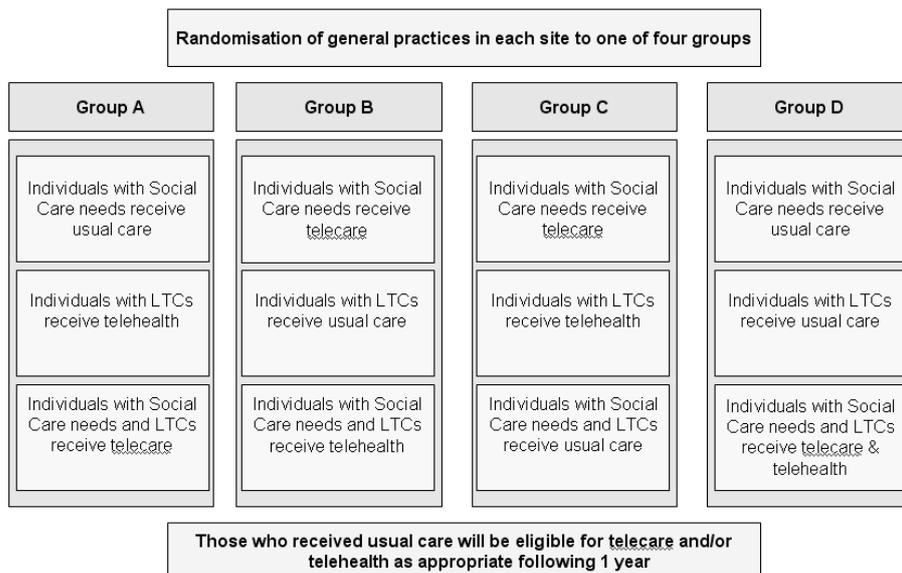
Aims and objectives

The Whole System Demonstrators national evaluation is a large cluster randomised controlled trial, believed to be the largest randomised controlled trial of telehealth or telecare undertaken in the world to date. Over a two year period, the plan was to recruit and follow up 2,000 service users in each of three demonstrator sites in a real time, pragmatic study to assess the impact of advanced assistive technologies in addition to usual care.

Design

GP practices are the unit of randomisation, and as originally planned each participating practice would be allocated to one of four groups (Figure 4-1):

Figure 4-1 Pragmatic Cluster Randomised Control Trial with Embedded Factorial Component: original design



The evaluation is a programme categorised into five themes, commissioned as one programme of work to evaluate the whole system integration of health and social care, and the delivery of telecare and telehealth within that comprehensive approach to care.

There are five major research questions corresponding to these themes:

Theme 1: Impact on service use and associated costs for the NHS and social services (return on investment). Does the introduction of telehealth and/or telecare result in reduction in service utilisation and costs of care?

Methods for Theme 1: A comparison of costs of NHS and social care amongst the 6,000 service users. The evaluation will test whether these technologies reduce demands on health and social care services, to see if reductions in cost outweigh the cost of delivering telehealth and telecare.

Theme 2: Participant-reported outcomes and clinical effectiveness. Does the introduction of telehealth and/or telecare result in improvements in quality of life and clinical outcomes in people with one or more of the following long term conditions: Chronic Obstructive Pulmonary Disease (COPD), diabetes and heart failure?

Theme 3: Cost and cost-effectiveness. What is the cost-effectiveness of the introduction of telehealth and/or telecare?

Methods for Themes 2 and 3: A comparison of clinical and patient outcomes amongst a proportion of the service users using patient questionnaires addressing severity of symptoms, participant-reported quality of life, psychological well-being, reductions in informal carer burden and social isolation, increases in self-care behaviours, self-efficacy and empowerment as well as attitudes to and levels of use of telehealth/telecare.

In addition, a questionnaire survey was conducted on a subset of informal carers (n=660 in the plan), measuring outcomes of particular relevance to them (such as burden). This research also asks about the extent to which patients use health and social care services to determine whether assistive technologies are cost-effective and offer good value for money. At the beginning of the study, the questionnaires were completed in the participants' homes with the support of a researcher, followed by postal questionnaires at three months and twelve months, with telephone and face to face support offered where necessary.

Theme 4: Qualitative study. What are the service users, informal carers and health and social care professionals' experiences of telehealth/telecare?

Methods for Theme 4: A small number of participants/informal carers (planned n=45) were recruited for interview, to cover a number of aspects of their current health and social situation, experience of health services and attitudes to and use of the technologies. Three interviews are to be conducted over a one year period. Twenty of these interviewees will also be asked to complete written diary records and twenty will be asked to allow a researcher to observe a visit from a relevant health or social care professional. A further fifteen to thirty individuals who decide that they do not want telehealth or telecare technologies installed will be interviewed about their reasons for refusal. Health and social care professionals (n=75) will be interviewed about their views about these services.

Theme 5: Qualitative study of organisational issues. What organisational factors assist in the successful introduction of telehealth and telecare?

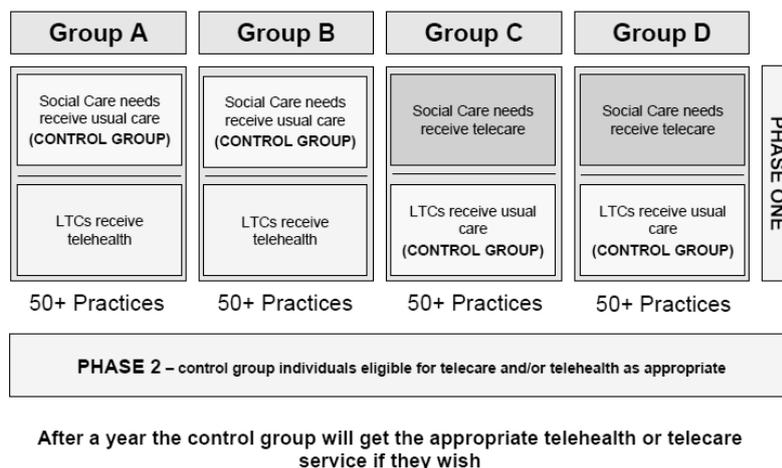
Methods for Theme 5: Interviews with forty-five key managers and commissioners in health and social services to assess their views on telecare and telehealth and the implications of the rollout of these technologies on the organisation of care, repeated at three points in time over the two-year duration of the study; observations of some strategic implementation meetings to provide contextual background information concerning the organisations' implementation plans, and help inform the qualitative interviewing.

Progress and changes to design

This evaluation involves large-scale, formal evaluation using methods which are generally thought to provide the strongest evidence of effectiveness within a biomedical research paradigm, as it reduces scope for bias due to differences between people who receive the intervention and those who do not.

In order to succeed in its aims, a study of this nature needs a large sample, which can make recruitment challenging. The evaluation has experienced problems in recruiting the planned numbers. Challenging time schedules saw some sites rolling out the initiative before the evaluation was ready, affecting randomisation. Additionally, fewer of the target population than expected had both health and social care needs, meaning very few individuals were eligible for both telehealth and telecare. Consequently the group receiving both telehealth and telecare would not have been statistically significant and was dropped from the design.

Figure 4-2 Revised RCT design



The national evaluation is due to present its final report in March 2010. There have been no interim findings published so far.

Comments on the evaluation

This is a very large and very ambitious evaluation. It is probably one of the largest and most comprehensive evaluations of a health or social care initiative ever attempted. The whole programme appears to have been planned following the best principles of research, and (unlike some of the other evaluations reviewed in this

chapter) careful consideration appears to have been given to key issues such as defining inclusion and exclusion criteria for participants, specifying primary and secondary outcomes, and describing the approach to analysis. The research team includes people with a high level of expertise in the field.

Despite these advantages, or perhaps because of the ambitious nature of the evaluation, the study has experienced problems and the research team have had to modify the design. It remains to be seen whether or what to what extent the team are able to complete their research in line with the modified plan.

The difficulties experienced illustrate some tensions in the use of this type of 'gold-standard' approach to evaluation, particularly the incompatibility between the length of the time it takes to plan and begin research and the pressures to implement services before the evaluation is ready. In addition, identifying and recruiting both sites (in this case GP practices) and patients to participate in randomised controlled trials is very challenging. If low rates of participation are achieved this may threaten the validity of findings since those recruited may not be representative of the wider population.

If the research team are able to complete the study as planned, this type of evaluation design, which uses mixed quantitative and qualitative methods to gain a full understanding of whether and how telehealth and telecare can benefit patients, will provide much more robust and rounded findings than the other evaluations described. Widespread roll-out of telehealth and telecare has not yet occurred so the evaluation has the potential to genuinely inform policy decisions. However these benefits will have required the investment of considerable resources in evaluation.

Other relevant evaluation work

Through our mapping exercise we identified two other evaluations being conducted in local areas and relevant to the WSD programme:

1. A local evaluation of a community alarm-based telecare service, providing risk management equipment to people in their own homes, within the Bury Metropolitan Borough Council area. The aim is to allow those at risk to remain as independent and safe as possible within their own homes, for as long as possible, in order to reduce long term admission to hospital or residential care. Drawing on routinely collected data, the aim of the evaluation is to assess the effectiveness of the telecare equipment to the end user, but also to demonstrate significant cost savings across health and social care with Bury, in order to secure long term funding for the initiative. The findings are intended for local dissemination only and are not more widely available.
2. A local action research project, looking into the workforce development to meet the needs of population with long term conditions in Tower Hamlets PCT. The project aims to address workforce aspects of long term conditions in the context of whole care pathway redesign; emphasise shifting care and avoiding hospital admission through developing self-management skills and capacity in primary care and community services to manage the bulk of LTC care; reconfigure multi-agency community workforce roles around the integrated care pathway for patients with LTCs; address the plurality of provision through joint health and social care planning and working and third sector involvement; and develop a community wide (GP, social services and acute

trust) clinical information system based on EMISweb to support identification and management of LTC patients.

Neither of the respondents to our survey who described the above evaluations provided any information about methods to be used, beyond saying that they were based on routinely collected data.

Conclusions about the evidence that the initiative achieved its objectives and delivered policy goals

No findings are yet available.

5. Evidence about the extent to which the White Paper initiatives are helping to meet policy goals

As discussed in chapter 1, the various initiatives discussed in this report were promoted by the White Paper 'Our health, our care, our say' because they were seen to reflect the policy goals of:

- prevention and earlier intervention to improve health, independence and well-being
- more choice and a stronger voice for patients
- tackling inequalities and improving access
- more support for long term conditions

This chapter seeks to synthesise key findings from the case studies of the various evaluations of these initiatives. In this way, it summarises the evidence about how much has been achieved through these initiatives in meeting the policy goals and also highlights challenges to the policies and outstanding questions where further evidence is needed.

Better prevention and earlier intervention to improve health, independence and well-being

Achievements

Several initiatives were clearly designed to support this strategic aim. Examples of progress towards the aim include the following:

- The NHS LifeCheck Programme provides people with an opportunity to assess their own health and well-being in order to help them make changes to improve their health. The evaluation of the Teen LifeCheck found that most people were positive about the concept. Both young people and health professionals were positive about the potential of LifeChecks to impact on knowledge, attitudes and behaviour.
- The Self Referral to Physiotherapy initiative has enabled patients to receive care earlier for musculo-skeletal problems. Levels of patient satisfaction have been high.
- The Improving Access to Psychological Therapies (IAPT) programme is a further example of how developments in services have allowed people easier and earlier access to healthcare. The evaluation has not reported yet but will demonstrate the extent to which this has been achieved.
- Both physiotherapy and psychological therapy for mental health problems represent forms of treatment which have long been subject to long waiting lists under the NHS. These initiatives represent good examples of the way in which the NHS is introducing new ways of working in order to ensure earlier intervention.

- The POPP programme has improved how older people perceive their quality of life, and appears to have led to a reduction in emergency hospital admissions. This is associated with a wider culture change with stronger partnership working between health and social care organisations around prevention.

Challenges and outstanding questions

Cost-effectiveness

Providing earlier intervention can increase health service costs. Delays in care are a form of rationing in that some people get better before they need care and other people do not seek care because they know that it will not be available for a long time. The increased costs of earlier intervention are worthwhile if they lead to health benefits. It is notable that none of the evaluations of LifeChecks, the IATP Programme or Self Referral to Physiotherapy provide good evidence of either health benefit or costs. The evidence about cost-effectiveness from POPP is indirect, based on limited data and a range of assumptions.

It is important to assess the costs as well as the benefits of these programmes so that their cost effectiveness can be compared with other competing demands for NHS resources.

Take-up

Although LifeChecks were introduced largely because of popular demand, evaluation suggests that take up so far has been limited, at least for Teen LifeCheck. In addition, the research suggests that further work needs to be done with regard to the content and design of the LifeCheck. Evaluation so far has been most useful in terms of suggesting ways in which to improve the LifeCheck. Subsequent evaluation needs to be conducted to test the effectiveness of different ways of delivering LifeChecks and different approaches to improving uptake. Similarly, with regard to direct access to physiotherapy, although this is improving people's access to care, the majority of people accessing physiotherapy still do so via a GP referral, even in areas where self-referral is available. More attention needs to be given to exploring the reasons for this, and ways of increasing awareness of the service.

More choice and a stronger voice for individuals and communities

Achievements

- The Individual Budgets initiative is clearly designed to allow individuals more choice and control over how their services are provided. The evaluation generally suggests that this initiative has been a success. People receiving Individual Budgets are more likely to feel in control of their lives, how they access support and how the care is delivered. The evaluation also described how people who accepted the offer of an Individual Budget changed their aspirations in terms of living a fuller life or being less of a burden on their families and having greater independence. They were more satisfied with the support planning process. Benefits were also evident for carers in terms of improved quality of life and social care outcomes.

- The evaluation of Information Prescriptions also demonstrated that this initiative has helped patients. Those people who were given information on prescriptions gained more confidence in asking questions about their care. Carers also found these prescriptions very useful.
- Although the evidence from evaluation is less clear, several other White Paper initiatives are also clearly designed to provide patients with greater choice in how service are delivered. The Self Referral to Physiotherapy initiative allows people to make their own decisions about when and how to access care. Both the IAPT programme and the social enterprise pathfinders provide people with a wider range of possible providers for their care when in the past options were very limited.

Challenges and outstanding questions

Which groups benefit?

With regard to Individual Budgets, the evaluation suggested that the benefits were greater for some groups of people (e.g. those with mental health problems) than for others (e.g. the elderly). Further research needs to explore which other groups of patients are most likely to benefit. This evaluation also raised a number of other important questions for further research. More attention needs to be given to how to value and pay for the input of carers in relation to individual budgets. Individual Budgets had a big impact on staff, leading to concerns that they may fragment the care process. They have important implications for the training needs of staff, as patients may choose to purchase very different types of care from those previously provided. Finally, although the findings from the evaluation of Individual Budgets were generally positive there were various concerns about the robustness of the evidence generated by the evaluation for the reasons described on pages 46 - 48. Some of this work therefore needs further investigation and replication.

The evaluation of Information Prescriptions also suggested that the benefits varied for different types of people (e.g. those in poor health or living in disadvantaged areas were less likely to benefit). Further research is needed into which groups of people are most likely to benefit and why this might be the case.

Trading time against benefits

The evaluation also demonstrated that Information Prescriptions led to more discussion with patients, which is likely to be beneficial but also takes time. Further research is needed into how much additional time this process takes and whether this investment is compensated by, for example, reductions in subsequent consultations or improvements in patient outcomes.

How do people exercise choice?

With regard to Social Enterprise Pathfinders it was clear that this initiative led to a wider range of alternative providers of care. Further research should investigate how patients make choices between alternative providers, what factors they value in terms of different models of this provision, and whether these different models are more or less costly and lead to better or worse health outcomes.

Tackling inequalities and improving access to services

Achievements

There is probably more evidence about the success of White Paper initiatives with regard to improving access to care than there is with regard to any of the other policy goals.

- The Self Referral to Physiotherapy scheme clearly improved access to care and was popular with patients but did not appear to lead to any overall increase in demand for physiotherapy.
- The Care Closer to Home initiative was also popular with patients, who appreciated the availability of more local services. It also appeared to lead to reductions in waiting times for care. In addition, when asked to rate how easy it was to get to a clinic, over 70% of patients rated their experience excellent or good in Care Closer to Home sites compared with 54% in control hospital clinics. This improved access was achieved while also reducing costs both for PCTs as commissioners and for patients.
- Social Enterprise Pathfinders provided an increase in capacity which was intended to improve access to care and in particular to reducing inequalities of access by providing people with more choice in how they obtain services.
- The IAPT programme also involved a considerable increase in the capacity of the NHS to deliver care. Although findings from the evaluation are not yet available it is likely that this has had a significant impact on improving access to care for mental health problems.
- Finally, as previously noted, the NHS Teen LifeCheck Programme achieved a high level of awareness. In particular there was a high level of awareness in the most vulnerable groups, so the targeted approach to promoting LifeChecks appears to have been effective. With regard to the Early Years LifeCheck, the most positive responses came from the younger and less experienced parents as well as those from more deprived backgrounds and those lacking support networks.

Challenges and outstanding questions

Does improving access reduce or increase inequalities?

Improving access could reduce inequalities if certain groups of patients were disadvantaged in accessing services in the past and if the new services were particularly accessible for these disadvantaged groups. However, improving the capacity of services can potentially increase inequalities if they increase access most for those with the least needs. This could be the case if, for example, new services require people to have access to the internet or to travel to a destination which is not easily accessible by public transport. Further research needs to give more attention to the differential impact of these initiatives on access for care for different groups, particularly in terms of ethnicity, age and whether or not people have long term conditions.

Lack of control groups

It is unfortunate that the evidence from several evaluations (e.g. Self Referral to Physiotherapy, Care Closer to Home, IAPT) was weakened by the lack of appropriate concurrent control groups. Although some evidence from comparison groups was obtained for the Care Closer to Home and IAPT evaluations, there were important limitations on the comparisons available for reasons discussed in chapter 4. These evaluations could appropriately have been conducted in the context of a randomised controlled trial, with individuals or areas randomly allocated to one form of service provision or another, and this would have provided much stronger evidence about the benefits of these initiatives. Although the difficulties of conducting RCTs of these initiatives are not to be under-estimated, these difficulties are not insurmountable. The benefits of having stronger evidence of effectiveness would probably have outweighed the costs, given the total national investment in these initiatives.

Improving access to effective or ineffective care?

Improving access to services is not an end in itself. Improved access is only of benefit if it improves access to effective services provided by appropriately trained staff. There was a concern in some evaluations (e.g. Care Closer to Home) that patients had been given greater access to less high quality care or less appropriately trained staff. Further research is needed to examine the benefits and costs of providing care of different types and in different settings, e.g. further high quality studies comparing consultant care versus that provided by practitioners with special interests.

Social Enterprise Pathfinders provide another example. They were introduced as a way of increasing the capacity of the care system in order to improve access and choice. The evaluation discussed in this report demonstrated that it was possible to create social enterprise providers but did not provide any robust evidence about their costs and benefits. Although there is evidence from other sources that enterprise models can make a positive contribution to healthcare, further robust evaluation is needed in the NHS context.

Lowering the threshold for care

Improving access lowers the threshold for obtaining care. This could potentially decrease cost effectiveness if access is improved for patients with fewer health needs who have less to gain from health or social care interventions. Further evidence is required to explore the thresholds at which care is beneficial. This particularly applies to the Self Referral to Physiotherapy initiative and the IAPT programme. It may be important to have efficient mechanisms to triage patients, given the high prevalence of both minor musculoskeletal conditions and low level mental health problems, which may not necessarily benefit from early NHS intervention.

More support for people with long term needs

Achievements

Initiatives which help to provide support for people with long term conditions include Information Prescriptions, New Types of Workers and the IAPT programme.

- The evaluation of the Information Prescription initiative demonstrated that these prescriptions were helpful to both patients and carers. The evaluation of

the New Types of Workers initiative gave examples of how workers operating within new roles were able to support people with long term conditions to enable them to live at home.

- The IAPT programme also provides better access to treatment for people with long term mental health needs who were previously often unable to access effective services. The evidence so far available about the IAPT programme shows that people who use the newly available services have positive experiences of them and recovery rates which are comparable with those which would be anticipated from better access to effective interventions. These findings are based on the early rapid audit conducted by Clarke et al. Further information about costs and benefits will be available in due course from the national evaluation.

The most relevant initiative to this goal is the Whole System Demonstrator programme, but information from this evaluation is not yet available.

Challenges and outstanding questions

Appropriate ways to disseminate information

The evaluation of Information Prescriptions demonstrated that although providing better information is an important aim, it is a difficult one to achieve. Providing appropriate information was resource intensive and delivery platforms change rapidly (e.g. some providers of information felt their work was overtaken by the development of the NHS Choices website).

Demonstrating the benefits of greater information

It was also difficult to demonstrate benefits from the great availability of information. This should be a priority for further research. Suitable topics would be the impact of further information on patients' health related risk behaviours, use of health services and health outcomes.

The evaluation of initiatives relating to this policy goal was largely formative. For example, one of the initiatives most relevant to this policy goal is the New Types of Worker initiative, and further evaluation is needed to provide more rigorous evidence about the impact of New Types of Workers compared with traditional forms of service delivery.

Conclusion

In summary, 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 'Our health, our care, our say'. The evidence is strongest in relation to improving access to care and less strong in relation to better prevention and greater support for long term conditions.

In general the evaluations have been largely formative, describing the processes being undertaken within each initiative and ways in which the initiative might be improved. Few evaluations have sought to demonstrate the impact of initiatives in terms of improved health outcomes. Where evaluation of health outcomes has been

attempted, this has often not made 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. There were 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.

Much of the evidence from the evaluations described patient satisfaction or experiences or staff perceptions of benefit. Patient satisfaction is an extremely important outcome to be considered and patient support for these initiatives may well be a good justification for them. However, patient experience should also be considered alongside assessment of health benefits and costs. As previously discussed, evidence about these is extremely limited. Only about half of the evaluations have considered costs, and none have expressed costs in relation to outcomes (e.g. in terms of Quality Adjusted Life Years) in the way that is now conventional with evaluation of other health technologies. It is arguable 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 which might have offered greater benefits. This approach is now well recognised and encouraged by government in relation to the appraisal by NICE of new drugs and technologies. It needs to be more widely applied to the introduction of other new initiatives within health care.

The quality and depth of evaluation conducted for the different initiatives was highly variable, ranging from large and expensive multi-centre randomised controlled trials to small retrospective analyses of routinely collected data. It is not clear how and why this variation came about, but it appears to have arisen from the different origins and histories of the different initiatives, with little sense of a co-ordinated approach to the evaluation of 'Our health, our care, our say'. This does not mean that a co-ordinated approach was necessarily required, but it does raise questions about why detailed evaluation was commissioned for some initiatives but not for others.

6. Commissioning, conducting and using evaluation

The interview data collected from the sources described in chapter 2 were analysed with the primary purpose of increasing our understanding of the most effective ways of conducting evaluations in order to inform policy. The interrogation of the data included, but was not restricted to, exploring the perceived strengths and weaknesses of different approaches among the different participant groups, establishing what these groups identified as the useful and successful aspects of evaluations, and identifying the different barriers and facilitators they perceived from their distinct yet associated roles. Exploring these issues using this integrated qualitative approach clarifies how the process of commissioning and conducting evaluation could be improved to make it more useful to these groups.

This analysis identified five key areas relating to the place and use of evaluation in policy making:

1. The commitment to evaluation as an integral part of the policy making process
2. Issues relating to the strategies for and planning of the evaluation work
3. Engagement between policy makers, evaluators and sites regarding activities relating to ongoing evaluation work
4. The impact evaluation activities have on conduct and local practice in sites
5. The impact evaluation activities have on policy making

The data suggests that optimising the usefulness of evaluation can be achieved by a thorough and connected approach to these areas. We conclude each section by indicating a number of key questions that should be considered at each of these stages.

The themes that have been identified are discussed in the remainder of this chapter, using extracts from the data to illustrate these emergent themes and issues. Extracts have been chosen that concisely convey the point raised, presented in order to illustrate the patterns identified by systematic analysis of the interviews, and no bias should be inferred from them. Extracts are identified by an anonymous unique alphanumeric identifier, which indicates only whether the speaker was a respondent from our Department of Health policy lead and key informant interviews (DH00), an evaluation investigator (PI00), a service lead from an initiative site (SL00), or a member of staff from an initiative site (Staff 00). Any other information that may identify individuals has been removed from the extracts.

Policy development

Commitment to evaluation

The first key area in the strength of a good evaluation strategy is the place evaluation has in policy development. Overwhelmingly, the parties we spoke to expressed a commitment to evaluation, and saw it as an integral part of the development and delivery of initiatives.

In terms of policy aspects now they're seen as an important and essential component. How anyone will approach development of it that will be part and parcel of what a policy maker will want to actually take into account. DH51

If we don't do the evaluation we won't do the programme. So I think that cleared the air and made sure everyone understood that if we can't evaluate it we are not doing it. DH59

This widespread commitment ostensibly provides an excellent foundation for capitalising on evaluation work, however the data indicates that there are some areas that would help consolidate this position. It was apparent in the accounts that this expressed commitment to evaluation can at times be barely more than an acknowledgement that evaluation should always be an element in policy development, without a fuller articulation of the ways in which such evidence is useful. The accounts lacked clear articulations of *why* evaluations should be an integral part of the development and delivery of initiatives. We have found that despite the commitment to evaluate, the issue of what evaluation is for is more problematic, and an area that could be strengthened.

We need it because it is the right thing to do and people will say 'Why haven't you evaluated it?' DH52

The commitment to evaluate, clear in the White Paper and articulated clearly in the interview data, is not accompanied by a clear rationale for the evaluation processes that follow on from that commitment.

Characteristics of Good Evaluations

We also found a shared view of the core characteristics of good evaluation, consistent among policy makers, evaluators and service delivery personnel. These characteristics related to design issues, characteristics of evaluation teams and results:

i. Design issues

- A design which is appropriate to the task and anticipated audience(s)
- A study with clear, comprehensive, understandable aims

A good evaluation is based on a really good question. Clarity about what you want to know DH52

ii. Evaluator Characteristics

- Commitment to the aims of the initiative
- The use of appropriate expertise in the evaluation team
- Good communication between policy leads and evaluators

I've always felt is that it is incredibly important that if you want major change as a manager you have to involve people in the project and I think you have to let them be part of the data coverage and know the results and I think it would have been hugely helpful if the research team had worked, at least some of us to help us to do, even some action research, it would have engaged and helped people understand how doing an evaluation in an organisation of this, how it can help? SL17

iii. Results

- Clear and anecdotally meaningful conclusions which are easy to understand
- A pragmatic timetable of reporting and feedback
- Delivering on time and to specification

Personally I am keen on evaluation because how do you actually capture the learning if you're not going to actually stand back and reflect what you're doing and try and catch that in some systematic way? How do you learn from it? So, I would find it difficult to argue a case for not evaluating to be honest. DH58

While these criteria did not favour any particular research design or discipline per se, some respondents felt that certain types of research design were more suited to satisfying criteria relating to clarity of aims and conclusions. While for some this favoured more quantitative designs such as for the production of robust findings, to others qualitative work were felt to produce more meaningful and useful results. The interviews gave us some examples of how this variation was perceived: the accessibility and plausibility of narratives offered by qualitative data were seen as being especially useful for providing illustrative vignettes; cost related data was seen as being particularly useful to the Treasury; the collection of feedback and routine data was seen as being of potential benefit to providing formative support to sites.

I don't like paper based anything me, so anything that involves ticking boxes, I personally get a bit wobbly about, it's the tension isn't it between how you capture soft evidence and statistical quality versus quantitative stuff. Staff 42

I think understanding what type of evidence you need as an outcome. So different audiences have different requirements. And if you understand what those audiences need to be able to convince them of the worth of something then you can start by deciding what type of evaluation you need because in this sense we had a clinical audience we thought it was important after quite a lot of that discussion that we would go for a randomized control trial because there's quite a lot of sort of overheads with doing that sort of trial. DH59

Discord among the different respondent groups we interviewed arose not from any broad disagreement about what makes a good evaluation, but rather from different parties' perceived abilities to contribute to ensuring these characteristics would be met from their perspective. These areas of discord between the different parties involved in the evaluations most usually relate to the issues discussed in the 'Strategy and planning' section below. This suggests that in order to achieve a genuinely informed commitment to evaluation as opposed to its simple acceptance as a usual part of policy development, a strategic approach to the elements raised below is necessary.

Key questions for policy makers:

Why is evaluation important in this programme of work?

What evidence will move the policy area forward, and how can this best be gathered?

For whose use is the evidence intended?

Strategy and planning

Where evaluation was deemed most useful by the respondents in our study, there has been early involvement of appropriate expertise to match a good evaluation design to the initiative. Our data point to a number of key challenges in achieving this, which when engaged with effectively, can result in the establishment of a good evaluation design, conducted by an appropriate team, which will deliver answers to well defined goals. These challenges related to how expertise is identified and engaged, the early involvement of evaluation teams, the extent of evaluation activities within participating sites, and the turnover of personnel.

Expertise

Our data point to the importance of having either appropriate in-house research expertise on the initiative team, or access to relevant research expertise to inform the strategic process of planning the evaluation. The evaluation teams in our study were chosen or composed in a number of ways. These can be characterised as being expertise led processes, research question led processes and open approaches.

Expertise led

For some initiative areas, the policy lead had a clear idea of the type of expertise they required, which was often but not necessarily associated with a clearly defined research question for the evaluation. This approach was seen as a way of ensuring quality and robustness in the execution of the evaluation:

It wasn't open to tender. We were invited to bid for it on the basis that there was a very considerable body of social research expertise in [the team's] funded units. PI4

One of the lessons we've learned from other evaluations ... is that we need to get research and evaluation expertise in right at the start, before the demonstrator sites are set up. Actually in setting out what we want them to bid on, we want to make sure that covers the issues and mechanisms in a way that can be evaluated properly. So that at the moment we're seeking and negotiating around some expertise to actually give us some input into that research and evaluation thing, into our thinking. DH51

Research question led

Other initiative areas began with looser notions of the expertise required for the evaluation, but a clear idea of the research question or specification. Here, appropriate expertise is drawn in based on research needs. Whilst this was viewed as a way of ensuring the most appropriate design and expertise for the perceived research needs of the evaluation, it can also require those commissioning evaluation to have a broader expertise in different research approaches to be able to determine the best way forward:

We had an outline [for the evaluation] of what we wanted from a spec' ... then I suppose they picked up our outline specification and turned it in to something more robust, which was consistent with their process again, putting out to tender and all the steps that they need to go through. DH53

I think very much the decision was made by Department of Health as far as possible to replicate the work we had done [previously], it seemed sensible then to try to add to the evidence base. PI1

Open approaches

Finally, a small number of evaluations adopted an approach in which both specific research questions and appropriate expertise were expected to emerge through the tendering process. This approach was more characteristic of evaluations of a smaller scale, or those consisting of a larger overall collection of smaller scale projects. While this allows for a degree of flexibility and responsiveness, it does risk resulting in incomplete or unsatisfactory overall evidence bases:

We weren't clear of the outcomes, so they [sites] could never be sure what they were measuring or evaluating, and I think that's where we the task and finish group let the pilots down. That's why the evaluation is so sort of, so sort of hit and miss for some. It's strong for some. DH57

Team structures

The composition of the evaluation teams included in our study ranged from small local projects driven by a single individual to large, complex consortia undertaking multi-phased programmes of work. Variation in team structure can be accounted for by the following:

Preferred organisation type

The evaluations we studied were conducted by a variety of academic and non-academic organisations. Non-academic organisations were viewed favourably for their speed, perceived non-critical approach to policy and their adaptability to changing requirements. Academic institutions were viewed favourably for their methodological and disciplinary expertise, status, and perceived robustness of method.

The consultant companies will come in and we are their clients. We could be whoever. We could be murderers from rotten row and they wouldn't judge us. Whereas I do sometimes find academic institutions quite judgemental. DH56

I think [the Department of Health] very much wanted an academic contribution and [colleague] and I fulfil that role. PI46

Available resources

While the large national evaluations were usually adequately resourced to meet the research requirements of both policy leads and evaluation teams, the availability of resources was still a factor that directed how the evaluation team was composed. For the smaller local projects we also included in the study, resource availability was a much more significant impediment to putting together research teams to conduct an evaluation:

Quite honestly I think that if you had much more resources or whatever you could do a few aspects a bit more in depth. PI1

They gave us extra money to approve more sessional interviewers and they cracked the whip to make sure so that names and addresses were coming back to us, that sort of thing. PI5

Timetable for reporting requirements

The initiative areas differed in how soon they required evaluation teams to report findings that would feed into ongoing policy development or initiative implementation. For almost all policy leads it was perceived that a balance needs to be struck between the need for robust evidence with a clear conclusion which was often seen by them to be associated with longer evaluation designs, and projects which could provide useful information on a shorter timescale but with less perceived certainty. This tension can produce differences with evaluators who may conversely see these requirements as risking running counter to methodological robustness, and also with service delivery personnel whose preference for swift, contextually appropriate, accurate, pragmatic feedback may also feel at odds with reporting structures.

All the way through there's been if you like a disjunction between the formal evaluation, or what you might call the independent evaluation, and what the people on the ground actually needed, which was a kind of quick and dirty kind of way. PI13

One approach to trying to strike a satisfactory balance was through multi-faced evaluation structures where parallel work by different teams could increase both the volume of evidence collected and the speed with which findings were available.

From a point of rigour, then I suppose the external longer term evaluations have certainly got more longitudinal data and so they will stand up to scrutiny from an academic audience which is extremely important and influential, its problem is that it takes more time. So, sometimes perhaps it's the short internal evaluations that have more influence in terms of change of direction of travel because that is the only data that you have got in the short term. DH 53

Properly designed, swiftly conducted research can produce robust findings, and by no means does a lengthy project guarantee valid, reliable results, yet our data suggest that concerns along these lines endure. Timescale considerations do need to be addressed at the research planning stage in order to minimise these concerns that best possible designs are compromised by squeezing them into desired timescales rather than planned robustly within them.

Early involvement

All the groups we interviewed indicated the need for establishing a working relationship with evaluators at an early stage in the implementation of an initiative. Delays in commencing an evaluation, especially if compounded by a lack of research expertise in the early stages, may mean that the opportunity for evaluators to contribute useful insights to the structure of the programme of work may be lost. Both policy makers and evaluators felt there were occasions where such an early involvement of evaluators could have pre-empted difficulties which later arose.

Indeed, we were working fairly hard on it for six months before the funding actually started but that did mean that the evaluation process could feed into the implementation process and there were certain aspects of implementing [the initiative] that were driven by the demands of the evaluation. So actually, [name] highlights that as a real advantage the evaluation team were right in there at the beginning almost before the pilot sites had been selected. PI4

It would be good if the process of starting the evaluations off could be streamlined, because it takes so long to actually do them anyway, why does it take a year or more to kick them off in the beginning? DH53

Extent of evaluation activities within sites

We learned of the importance of establishing early and clearly with sites what the nature of committing to evaluation activities would be, and also of taking on board the impact that evaluation activities have on the day to day activities of staff at the sites. Tensions arising from additional workloads associated with new systems or data collection, or lack of feedback meaningful to staff activities can risk damaging relationships with staff at sites.

A couple of them [local sites] have grumbled since because of the time commitment, because the busy people try to do other things, but there hasn't been, I haven't picked up any fundamental misunderstanding of what we are trying to achieve with the evaluation and indeed quite a few of them are also involved in separate local evaluations as well. DH58

I think all of us had different work streams as part of the overall project so we were all thinking how best to capture data with customers, there was a lot of formal data gathering as well which involved filling in forms from the local evaluators and nationally from [the national evaluators] and whatever. I have to say I know that we have to be evaluated and I do understand that but the disadvantage of that was it turned people off a little bit... That created a bit of bureaucracy which I'm sure you can appreciate so in a way we lost sight in some ways of what it was all about and why we were doing it. Staff 42

It was a condition of their participation that they were part of an evaluation, and they needed to sign up to that, that was all part of the award letters that went to each of the participating authorities. The extent to which the authorities actually were aware of what that would entail I think only became apparent once the evaluation had begun. DH54

By clarifying what the evaluation activities will involve, policy makers are better able to assess what are realistic tasks and targets for sites to achieve, and staff at the sites are less inclined to see any extra workload from evaluation activities as a problematic drain on their time. Rather, when well consulted, they are able to contribute to and appreciate the learning opportunities of the evaluation work.

The evaluation, in its nature, has significantly affected what the sites are doing. And the sites have been involved from very early on as well and there's been a healthy dialogue. It's quite a challenging dialogue I think at sometimes between the sites and evaluators because we want to make sure that everything we do contributes towards this evaluation because otherwise we're spending money for no real benefit...I think we've had to push in two ways, we've had to push the evaluators to go faster than they have perhaps wanted to go, and we've had to make sure the sites don't just rush into things without fully understanding what the requirements of the evaluation are. DH59

Personnel turnover

In many of the initiative areas, we identified a high turnover of personnel, particularly at the Department of Health, and this appears to be usual. Where personnel turnover is anticipated as likely, it becomes more important to formally retain clear information about strategic rationales and decisions, which in many cases were lost or became vague when personnel changed. Often our Department of Health respondents were unclear about why decisions relating to choice of sites or evaluation design were made by predecessors, and therefore felt less well able to draw lessons from the work. From the interview data, it was not possible to build a comprehensive picture on detailed commissioning processes across all the initiative areas, partly due to this issue of personnel turnover.

Interviewer: Could you tell me a little bit about your involvement in the process of the commissioning of the evaluation teams or partners who were doing the evaluation?

Respondent: Well again that's probably quite, I mean, because I wasn't actually there. I can sort of give you a sort of second hand, though probably not as detailed as the people who were involved at the time.
DH54

Very often people with a historical knowledge of what went on before have either moved on or moved out, so that's lost. DH57

Most commonly this issue of personnel turnover related to roles within the Department of Health. Those that are most apparent in our data relate to the policy lead role, though a number of respondents also made reference to changes in ministerial personnel having consequences for the focus or impact of ongoing evaluation work.

They sat on it for ages. It went into the DoH in March and didn't come out until October, which co-incidentally was two weeks after [minister] was moved portfolios, so part of me thinks he was sitting on it because he didn't like what it said. SL43

The other thing is you get different ministers who are into different things at different times, so. [A] really wasn't interested in the evidence base at all, [B] was much more interested in evidence bases and that's anything from where's it been tried before... and of course now you've got [C], who just takes a completely different view to [D] for example.
DH56

The production of detailed research protocols in commissioning and tendering processes meant that personnel turnover was much more easily accommodated within research teams, though issues did occasionally arise:

Some aspects of the data were incomplete and I flagged this up initially and was told to continue but then managed to find a bit more data that had been left behind somewhere that hadn't been entered on the database. PI1

Key questions for policy makers:

Who can best take forward the evaluation goals, and what might they already be able to offer?

Is there sufficient understanding between the sites, evaluators and Department of Health regarding what the evaluation activities will both involve and yield.

How clearly chartered are the decisions and rationales behind the evaluation strategy?

Engagement between policy makers, evaluators and sites: communication and co-ordination issues

We explored the structures and channels of communication that enable all involved parties (policy making, service delivery, evaluation) to communicate their activities and issues with each other as the work progresses. The complexity, size, and design of the initiative and evaluation will require different approaches to best ensure the programme of work remains well co-ordinated. Some of the teams we met have found that the deployment of a clearly defined co-ordinator role has been a successful approach. This may take the form of an individual who pulls together evaluation activities of different, often geographically disparate, teams, or who acts as a conduit between evaluators and the Department of Health.

PI co-ordinator

While for many teams, principal investigators combined research activity work with communication and liaison activities, in some of the teams we studied the principal investigator had little or no involvement in the conduct of the research but instead took a co-ordinating or project management role.

I sit on the board at [research consultancy company] and am involved personally in a number of evaluations as the project manager and project director. PI20

Co-ordinators within multi-team design

Where evaluations were being conducted by teams from multiple organisations, key individuals had specific remit to co-ordinate between teams and with customers.

I think the sort of two key people who hold it together are [name] who's got that co-ordination role, that's his role and responsibility in the evaluation. He isn't undertaking in any of the research on the sites. He is sort of co-ordinating the efforts of all the universities together, and [name] who's the principal investigator and hence is responsible for the ethics submission and therefore needed to get all the universities to co-ordinate their activities into the ethics and also to come be present at the multi site ethics committee and meeting as well. DH59

We have also set up a ... project advisory group. Which again is multidisciplinary and that sits as an advisory group for the evaluation so that the evaluation team will take things to them for guidance and discussion and has that mechanism resulted in changes to how the evaluation is conducted from how it is envisaged at the start. DH58

Communicators between evaluators and DH

In this third co-ordination model, an individual with no direct involvement with the conduct of the evaluation work undertook a liaison role to communicate between the evaluators and the Department of Health. This approach was seen by those involved to add a constructive impartiality to relationships between evaluators, sites and the Department of Health.

Respondent: In the olden days, you may remember the Department used to have people called Research Liaison Managers. I think he did that job.

Interviewer: So he was employed by the DH?

Respondent: He was employed by DH and it was very much a sort of, yes, a source if you like for research liaison role. He was very useful, there were times when he would be very robust in defending the research, even against policy customers and he was also very good the other way as well. He used to say to use very gently: I think you're going to have to do that another way. I think you're going to have to take this seriously. So he was brilliant. Very important role. PI4

Communication mechanisms

In addition to the communication roles of individuals, others have discussed with us, and highlighted the usefulness of, planned activities such as workshops, learning networks and internal reporting as useful mechanisms for different elements of the work to remain connected and share learning. Our research does not indicate any one approach to be superior, but that seeking mechanisms appropriate to the circumstances are effective and valued by all involved parties.

The participants... they have workshops with the evaluation team, so the evaluation team can reflect back to them in terms of the base line – what they've found so far. For the [sites] that will probably be already a learning experience from all the different stages of the development we will be looking at the outputs of the evaluation at each stage and using that to inform our work programme going forward. DH58

We're going to establish a learning network that will share the lessons learnt that we get from the sites relatively early, and it will also pick up what barriers other people are facing because we recognise that these sites do have quite a lot of additional resource... we will hold master classes, we will have some national events, we will set up a website for them, we will set up discussion forums for them...So by that we think it's sort of a two way benefit really. We can share the benefits from the sites to make sure that people don't repeat things that don't necessarily work, or pick up on things that do work straight away. But we will also learn from other people who are trying to do this, on the ground without any external support and without perhaps the focus that a big national programme. That's our way of helping to inform any policies that might develop out of this. DH59

These mechanisms are useful in balancing the potentially conflicting needs (discussed in 'Team Structures' above) relating to the timetables and requirements for reporting.

Key questions for policy makers:

What are the mechanisms for sites, evaluators and policy makers to communicate with each other about the ongoing work?

What will this learning be used for?

The impact evaluation activities have on conduct and local practice in sites

Our data indicates that there is a dynamic relationship between the ongoing evaluation and the initiative: each changes as the work progresses. The data point to an artefactual impact of being evaluated on the initiative sites, where the process of being evaluated changes that which is being studied. Evaluation impacts are not unidirectional: they feed learning 'up' to decision makers, but equally importantly influence the day to day work of those involved in the initiative.

The whole thing is ongoing definitely; it does keep people and the team conscious of what we should be doing. We do a lot anyway, but we tend to forget to recognise it, formally record what we are doing and I think that's been another aspect of it that we are actually now recording the work that we have already been doing. Staff 60

Dissemination

In the accounts we collected, evaluations were most often deemed useful when lessons were being learned as the work progresses, particularly regarding local practice in the sites. While it is clear from the data that this is often achieved, the accounts also suggest that there can be a lack of strategic ambition to achieve this goal, and its achievement at times appears opportunistic rather than planned. Most talked about uses of evaluation, in terms of the impact evaluations had on sites, were to facilitate or refine the implementation of initiatives.

Interviewer: So what impact are you hoping that would have?

Respondent: That other places would be less frightened because I've helped two or three sites, you know people now contact me ... because why should people have to try and invent, re-invent the wheel if we have already set it up and also I chat to them about the for and the against and how to set, go forward with it with their worries and everything else. PI39

I think it's probably the same across all programmes. I think that the difficulty we had with this is that it's a relatively new sector, a new market and so therefore everyone's learning a bit as they go along both the sites and the evaluators, about what they might do and how they might do it. DH59

The mechanisms for ongoing communication described above provide a medium for the dissemination of findings during the course of the initiative; however the longevity of these mechanisms rests on the finite nature of the evaluation process, problematising the continued dissemination and use of findings at the end of this time. How successfully this is achieved appears from the data to be another area that can be determined by propitious circumstances rather than strategic approaches.

Our interviews illustrated the wide range of audiences to whom the findings of evaluations are disseminated.

We're working quite hard on it at the moment in terms of dissemination, we've had two conference presentations this week, to

groups of sort of practitioners, local authority managers, local authority councillors, I think and between us will probably over a period of three of four months will probably have clocked up fifteen or twenty dissemination events to sort of local conferences and indeed national conferences. PI4

The extent to which staff respondents at initiative sites were familiar with and discerning of evaluation findings and used them in their ongoing work was an encouraging feature of the interview data.

Interviewer: Have you looked at the National Evaluation?

Respondent: I've read bits of it and it's not made good reading (laughs)...

[later]

Interviewer: You know about the evaluation, you know about the local and the, in terms of your work, you know we talk about evidence based policy and practice, do you use it?

Respondent: Evidence?

Interviewer: Yeh

Respondent: Loads Staff 42

The evidence also suggests that usefulness of evidence has to strike a balance between being locally meaningful to those specific contexts in which the work was conducted with being useful in contexts where initiatives were yet to be implemented.

Interviewer: Did you find the evaluation useful?

Respondent: It didn't tell me anything I didn't already know. I think for Local Authorities that hadn't already done it, it was probably quite useful. SL2

The impact evaluation activities have on policy making

The data suggest a range of ways in which initiatives and their evaluations have an impact on their relevant policy fields. While evaluations were most often deemed useful when lessons were being learned as the work progresses, these lessons were not confined to activities at the sites, but also related to being able to link the learning to both the policy origins and the development of future or wider policy. Again, the data suggest great strength and potential for this process of ongoing learning, yet there appeared to be a lack of strategic ambition to achieve this goal. Evaluations were chiefly discussed in terms of how they facilitate or refine the implementation of initiatives, as discussed above, rather than their place in the policy making process.

There's recognition that it can't be used in a kind of summative way. We'll wait for some evaluations and then decide what to do. Because you have got to carry on before. You learn from it and tweak. DH56

Dissemination

Procedures for formally reporting findings to the Department of Health in ways that are agreed are robustly adhered to. However, confidence in how reported findings will be used once delivered is less strong, with many evaluators having little sense of how their work will be applied, or if it will at all.

Interviewer: Do you think the researchers had any impact on local policy and practice or national policy and practice?

Respondent: The research that we did?

Interviewer: Mmm.

Respondent: Umm. I honestly don't know the answer to that. None that I am aware of to be quite honest (laughs). I mean the report was received by the steering group, as I'm aware. PI45

It contained a lot of information about where the programme fell on its face, basically. I suppose that's something, I mean they will publish it in a much edited form, because it's too excruciating otherwise, but what was very clear was that scaling up from a pilot site is incredibly difficult. PI46

Influences on impact

The interviews explored respondents' perceptions of the factors that influence the nature and extent of the impact that the evaluations they were involved in would have. Those involved in the collection, analysis and reporting of the evaluation evidence felt largely removed from this aspect of the evaluation, and are unclear in what way, if at all, their work is being taken forward.

I didn't see any concerted effort at the time or to address it and in terms of our impact I think we have had very little impact... In neither case do they seem to be paying any attention to the evidence from researchers to what they are doing. PI6

We certainly have very strong evidence that ANY evidence is being ignored by policy makers, and the twin spectres of finance and politics, and strategy are driving the programme. PI15

Timeliness

The timing of the availability of evaluation findings appeared from the data to be an important consideration in the impact of evaluations. Policy debates and decision making processes occur separately to the schedule of any relevant evidence production, and consequently if research reports in a coincident window of time to relevant debates the evidence can be expected to enjoy more attention.

I think we need another kind of commissioning process for the short cycle research that's very much fit for a policy purpose but then the three year evaluation ...They didn't think ahead about what we will need to know in four years from the point at which they were starting to think about commissioning the tender. PI13

In practice, there is a perception that policy debates usually move on before evidence from evaluations is able to usefully inform the debate.

Political interest

Interest from ministers in particular schemes was discussed by our respondents as driving initiatives in two ways. Firstly, ministerial interest in a particular service area at the start of the programme of work is seen to drive it at the start of and during the initiative:

[Minister], when we'd done all the briefings and speeches and things, stood up and out of the blue delivered this speech ... and I can

remember my manager sort of coming to me and saying right now, we've got to really run with this and look at what it actually means. DH57

Secondly, later in the process, when initiatives are established and learning emerging, by harnessing ministerial interest in the outcomes of the work:

One of the things we've arranged actually is for the [evaluation] team to come and present to ministers directly, so there's a meeting at the end of this month where the team will come along and tell them what they've found, which I think is quite powerful. Because they probably wouldn't believe us (laughs) as officials. You know what I mean? They'd probably think that we'd hidden something away. DH54

While feeling excluded from such influences, evaluators do perceive that the influence of their work is dependent upon such factors rather than its intrinsic worth or robustness.

I'm convinced that if political imperative had turned so that, if you like, it was advocating that this was not a good thing, that data, that report would have said that, would have used the same data, to have made that argument. PI1

Policy context

Respondents discussed securing the continuation of an initiative, or aspects of it. Some of the initiative areas were a product of this process from an earlier initiative, and as such had a degree of momentum to continue further, while others lacked this pedigree, and those supporting the initiative faced establishing the idea as something that could be developed further, often by widening to other policy contexts.

I think that as a policy team one of our tricks is going to be looking at the evaluation then looking across the departmental strategies to see actually what is the evaluation telling us in terms of how [this initiative] can support these other priorities and trying to make sure were weaving the outputs of an evaluation into a way that's meaningful across the wider health and social care reforms. DH58

Data suggests that continued momentum of an idea is important. Evaluation may be carried forward, as above, but if not immediately carried forward it is not habitually returned to as part of an evidence base informing future ideas. If not used immediately, our study suggests that evaluation evidence dies, rather than hibernates.

Strength of evidence

Counter to the acceptance of the principle that certain research designs were not better able to answer research questions, the data does indicate that respondents feel that certain kinds of evidence (however produced) may be more persuasive than others.

My sense is that numbers matter and money matters. If you can demonstrate cost effectiveness, cost benefit, ... that's important, clinical cost effectiveness really, if you can do that then you have got a case and implementability as well. If you can demonstrate that people can do it and they're up for it, that's important. DH52

Key questions:

How and when will learning from the evaluation be used?

7. Discussion

The purpose of this concluding chapter is to discuss the wider lessons from this overview of evaluations of White Paper initiatives with regard to how to improve the process of commissioning and conducting evaluations of health and social care programmes. This chapter will draw on the findings from both the structured review of evaluations in chapter 4 and the stakeholder interviews in chapter 6.

Why do evaluation?

As discussed in chapter 1, researchers and policy makers are frequently described as coming from two distinct communities, each unable to understand the other's point of view.²⁰ To some extent our findings challenged this belief, in that there was widespread agreement about several key issues. There was a clear shared commitment to the importance of evaluation. Both researchers and policy makers also recognised and shared a sense of frustration over the difficulties of generating high quality evidence from evaluation soon enough for it to influence policy.

However, there was less clarity and agreement about *why* evaluation should be conducted. Providers and some policy leads tended to want to know how implementation of an initiative could be improved, whereas other policy leads were more interested in whether or not an initiative achieved its objectives. In several interviews policy strongly emphasised the importance of evaluation, but in some cases this appeared to be largely rhetorical. 'Doing evaluation' was a given, or an end in itself, and being seen to have evaluated was perceived as being more important than any findings that resulted.

Both in interviews and in policy documents there was discussion of the need to 'learn lessons' from evaluation, but there was little clarity about what lessons needed to be learnt or a sense that policy makers were waiting to learn lessons from evaluation before making decisions. Although it was usually clear where the driving force for conducting an evaluation had come from initially, in many cases by the time the evaluation was completed it was not clear who was waiting for or needed the answers generated.

There were several reasons for this. Firstly, most of the civil servants who had been responsible for commissioning evaluations of White Paper initiatives had changed their jobs before the findings were obtained, and there appeared to be little 'organisational memory' to provide a sense of continuity. In many cases the civil servants had been involved in commissioning evaluation as part of getting the White Paper initiative started, but once it was underway there was no subsequent individual who was particularly interested in the findings. Second, during the course of the evaluation of initiatives arising from the White Paper 'Our health, our care, our say' other initiatives were introduced (specifically related to the NHS next Stage Review, published in 2008) which were higher priorities for staff within the DH, meaning that evaluation findings from earlier initiatives were of less interest. Thirdly, in some cases, commissioning evaluation appeared to be less about obtaining evidence in order to inform policy, but more about defusing any possible criticism of a controversial new initiative. Therefore the priority was simply to ensure that an

evaluation was being undertaken, and the nature of and findings from this evaluation were of less interest.

As discussed in chapter 1, a number of reports were produced within government from 1997 onwards which aimed to promote the systematic use of evidence to inform modern and professional policy making. The Government Social Research Unit (part of the Cabinet Office) was set up in October 2002 to provide strategic leadership to social research across government. Its 'Policy Hub' website¹⁴⁹ provides links to a number of excellent sources of guidance about the use of evaluation to inform policy making.

It was notable that these sources of information were not discussed by any of the DH policy leads interviewed for this project. It appears that the 'evidence informed policy making' movement has succeeded in creating a sense of expectation within the DH that evaluation should be undertaken, but that there is a lack of awareness about the resources available to help decision makers to consider *how* and *why* evaluation should be done.

These observations lead to several recommendations. Before commissioning an evaluation, it is important to consider whether evaluation is needed, and if so, why? Those making decisions about commissioning evaluation should ensure they are well informed about general principles of evaluation, such as the advantages and disadvantages of different approaches, perhaps making use of the 'Policy Hub' website mentioned above. Policy makers should consider who will make use of the information generated by the evaluation (given that specific individuals may change roles over time), whether any findings from the review will in fact make any difference to policy decisions, and if so what mechanisms need to be in place to expand, halt or amend a policy in the light of evaluation.

Commissioning evaluation

One notable finding from the summaries of evaluations reported in chapter 4 was the wide variety of approaches to evaluation conducted. There was variation in terms of means of commissioning, approach to evaluation, methodological rigour, size, budget and timescale. The reasons for these different approaches appeared to relate to timing, the public profile of the initiative, and the interests of different government departments (for example the need to provide evidence to convince the Treasury appears to have been influential in the approach to the evaluation of both the IAPT initiative and the IB initiative).

The different commissioning routes were associated with different approaches to evaluation. For example, at one extreme, evaluation of LifeChecks was commissioned directly from consultancies and was formative, short term and small in scale, while at the other extreme the evaluation of the Whole System Demonstrator programme was commissioned by the Policy Research Programme and involved an extremely complex study based on a factorial randomised trial conducted by a consortium of universities. There does not appear to be any consistent approach to considering what information is needed from evaluation of any particular initiative, and how best to commission the evaluation in order to ensure that this information was obtained effectively. The extent to which the different evaluations arising from the White Paper were co-ordinated was also limited, although the progress report documents,^{3;2;150} and indeed the overview project reported here, are an attempt to draw together findings

retrospectively. It would, however, have been advantageous to have greater co-ordination of evaluations before they were commissioned.

Another notable finding from the overview of the White Paper evaluations is how many of them experienced difficulties in conducting their studies as planned. Several evaluations failed to recruit as many participants as anticipated, experienced serious delays, obtained poor response rates or were unable to obtain other necessary data. Although on the face of it this may suggest poor planning by researchers, since this is a recurring finding it also suggests inadequacies in the commissioning process. Many of the evaluations had extremely ambitious aims and objectives, in some cases probably reflecting tender briefing documents which also set out similarly ambitious requirements, and in hindsight it is not surprising that they have proved to be unachievable. It is tempting for researchers to 'over-promise' what they can achieve in order to win contracts for work, just as it is tempting for commissioners to accept ambitious proposals rather than ones which are more cautious and limited. Commissioners of evaluation are responsible for robustly scrutinising the plans of research teams, asking them to justify that their methods are achievable.

One recurring theme from the interviews with both DH informants and principal investigators was the importance of good communication between the 'customers' of evaluation in the DH and potential 'providers' of evaluation at an early stage. This can be facilitated in various ways. Some interviewees highlighted the benefits of having someone within the DH who had research expertise who could liaise between researchers and policy makers in order to ensure that the evaluation commissioned was appropriate, of high quality, and achievable. A similar function is achieved when an evaluation is commissioned on behalf of DH by the NIHR SDO Programme or PRP programme. However, where evaluations were commissioned directly by policy leads, who may have had little or no research experience, there may be more risk that the evaluation was not appropriate to the need.

Conducting evaluation

The appraisal of evaluations reported in chapter 3 and the structured summaries of evaluations reported in chapter 4 highlighted a number of consistent problems with the conduct of evaluations of White Paper initiatives. In several cases the evaluations did not appear to address the most important aims of the initiative. For example, a key justification for the Improving Access to Psychological Therapies initiative was the Layard hypothesis that improved access to therapy would have economic benefits, but the evaluation was not designed to provide robust evidence about these economic effects. This is not a criticism of the evaluation team, but a reflection on wider problems with evaluation which will be discussed further below. Furthermore, in some cases (e.g. New Types of Worker) the initiative itself did not have clearly specified objectives, making it difficult to design meaningful evaluation.

As well as problems matching the evaluation to the initiative, in other cases (e.g. evaluation of Self Referral to Physiotherapy) the design of the evaluation was not appropriate to answer the research questions. This was particularly noted during the critical appraisal of many of the smaller local evaluations identified in the mapping exercise. In some cases, much effort had been expended on collecting large amounts of data of uncertain quality, without any written protocol, clearly specified research questions or clear plan for analysis. This reflects an admirable enthusiasm to evaluate

services but is wasteful of time and resources and suggests a lack of experience or expertise.

With regard to the national evaluations, in some cases there was a marked lack of clarity over key issues which determine the robustness of evaluation. These issues included:

- how pilot sites were chosen
- the extent to which pilot sites were representative
- justification for the choice of research methods
- how participants were selected and recruited
- details of the main outcomes assessed
- how these outcomes were measured
- how and when data were collected
- justification for the sample size
- the approach taken to analysis of data
- ethical considerations and whether ethical committee approval was obtained

The above considerations are fundamental to sound evaluation, but were sometimes not reported. The absence of information about these issues undermines the confidence that outside observers can have in the findings.

As previously discussed, there was a wide variation in the scale and methodological approach used within different national evaluations. Choosing the most appropriate methodological approach is in theory dictated by the aims and objectives of the evaluation, but in practice can be a complex decision. The 'best' method may not be feasible, and judgement needs to be made about the type of evidence that is obtainable, proportionate and likely to be useful.

In particular we noted ambivalence about the use of randomised controlled trials (RCTs), and this ambivalence affected all stakeholders including staff working within initiatives, service leaders, researchers and policy makers. On the one hand, there was a perception that within a medical environment the RCT sits at the pinnacle of the hierarchy of evidence. There was recognition that a well conducted RCT can provide the best evidence about effectiveness and other quantitative outcomes because of its ability to reduce bias and confounding variables. On the other hand, various objections to RCTs were made by interviewees or were evident in reports. Some interviewees questioned the feasibility, appropriateness or acceptability of randomisation, and there were concerns that RCTs may not provide valid findings unless they fully reflected the real world context.

These issues are illustrated by the RCTs being conducted within evaluations of White Paper initiatives. The RCT of Individual Budgets provided much stronger and more detailed evidence about costs and effects than the other evaluations described in this report. On the other hand, the need for tightly controlled and pre-specified methods did not sit well with the realities of service delivery. About half of the people in the 'Individual Budgets' arm of the trial had not received their budgets by the end of follow-up and some people in the 'control' arm also received Individual Budgets. Although the other RCT (Whole System Demonstrators) has not yet been completed, the evaluators appear to have experienced major problems in conducting the study as

originally planned because of difficulties recruiting patients and the challenges of conflicting pressures and timescales affecting service providers. This meant that some sites implemented the new services before the evaluation was ready to proceed. The WSD evaluation team has therefore had to modify the design of their study during the evaluation.

A determination to conduct RCTs in order to evaluate services can therefore lead to large and complex research studies which can be very difficult to manage and deliver. Such studies may also be very expensive and possibly disproportionately so in comparison with the cost of the service itself. In addition the constraints of an RCT can influence the provision of services so that the service is not representative of how it would operate outside a trial.

However, if careful account is taken of the above problems, RCTs on service innovations can be conducted and can provide the most robust estimates of the effects of initiatives. There are many examples of RCTs which have been successfully conducted to evaluate initiatives in health or social care. Deciding to use an alternative design, such as a non-randomised observational comparison between areas with or without access to an initiative, avoids some of the problems associated with RCTs but at the significant risk of introducing biased findings and reaching less reliable conclusions.

It was notable that several of the evaluations reviewed for this project did not include any comparison group at all (e.g. evaluations of Self Referral to Physiotherapy, Care Closer to Home, New Types of Workers, Social Enterprise Pathfinders, LifeCheck). Such descriptive evaluations may be useful for formative evaluation but cannot provide anything other than impressionistic evidence about the effectiveness of initiatives.

These considerations demonstrate that all approaches to evaluation have different advantages and disadvantages, and that any evaluation has costs as well as benefits. It is important to make clear judgements about what questions the evaluation should seek to answer, and how reliable and accurate the answers need to be. There were several examples amongst the evaluations reviewed for this project where money was used ineffectively in collecting data that would not answer the most important questions. It may be better to invest considerable resources on conducting a robust multicentre RCT that provides clear answers about the cost-effectiveness of an initiative, rather than spending smaller sums of money on quantitative studies which are never likely to produce clear or reliable results. For example, we argued in chapter 4 that the IAPT initiative represents a topic that could have been suitable for a RCT designed to assess cost-effectiveness. The total budget committed to the national evaluation of IAPT was £450,000, which would not be sufficient to conduct a multi-centre RCT. However, given the investment in the national roll-out of the IAPT programme (£173M between 2008 and 2011), and the even greater potential economic benefits if the programme was successful, a much larger sum for evaluation could have been justified in order to determine whether or not the policy was effective.

As discussed in chapter 4, the most important outcomes for a summative evaluation are patient or user benefits in terms of health or care outcomes in relation to the costs of the intervention. Although the importance of cost effectiveness was recognised both in interviews with policy leads and evaluators, and also in the

descriptions of aims of initiatives, few of the evaluations of White Paper initiatives examined either costs or benefits in a rigorous way (see Table 7-1). Instead, most of the evaluations using quantitative methods collected data about processes of service delivery. This provides much less robust evidence for benefit of an initiative, and it relies on an assumption that improved processes of care will lead to benefits in terms of outcomes for patients. The extent to which this assumption is justified probably varies for different initiatives.

Table 7-1 National evaluations assessing costs and benefits of initiatives

Initiative evaluation	Robust evidence* on:	
	Health or care outcomes	NHS costs
Care Closer to Home Demonstration Sites	x	limited
Improving Access to Psychological Therapies	(✓) [†]	(✓) [†]
Individual Budgets	✓	✓
Information Prescription	Limited	x
New Types Of Workers Pilots	V. limited	x
NHS LifeCheck	x	x
Partnerships for Older People	Limited	Limited
Self Referral to Physiotherapy	x	x
Social Enterprise Pathfinders	x	x
Whole System Long Term Condition Demonstration Sites	(✓)	(✓)

* 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.

() Planned, but results not yet available.

[†]Control group may not be comparable.

Conducting a robust summative evaluation requires a commitment at the outset to both allocate considerable resources for evaluation, and also to plan evaluation from the outset as part of the implementation of the initiative. Attempting to conduct a summative evaluation which uses a weak design, or is too small or unrepresentative to generate generalisable information, or which cannot recruit patients is wasteful and investing money in formative evaluation may be more productive. Several evaluations reviewed for this project fell into a middle ground – neither providing useful formative

evaluation to help services improve, nor providing reliable summative information about the success or failure of an initiative.

Some of the most useful aspects of the evaluations we reviewed were the formative elements. Examples include the evaluations of Information Prescriptions and some aspects of the evaluation of NHS LifeChecks, and the formative components of the evaluation of Care Closer to Home and Individual Budgets. These formative evaluations generally relied on qualitative methods and analysis of routinely collected data about processes. If funding for evaluation is limited, it may be better to undertake a well conducted and useful formative evaluation than to attempt an inconclusive and uninformative summative evaluation.

At the outset of this synthesis and overview we planned to analyse the evidence for each initiative using a realistic evaluation framework. In this way we would explore whether the evaluations provided evidence to support the underlying theory about why each initiative should have the intended benefits in certain contexts. However, we were not able to apply this framework for a number of reasons. First, most of the initiatives and evaluations were not based on any clearly articulated theory which was being tested (i.e. a justification that if specified processes were followed then the initiative would have certain specified effects for particular target groups of people or in a particular context). In some cases the theoretical justification and the intended outcomes were clear (e.g. the IAPT initiative) but in other cases the intervention itself was not clearly defined (e.g. New Types of Workers) or the intended benefits were also vague (e.g. Social Enterprise Pathfinders). Some of the problems of the evaluations can be traced to this lack of clarity. A few of the evaluations appeared to be more about theory generation (i.e. introducing an intervention and then assessing who it seems to help and in what way) rather than testing whether an intervention achieves specified aims and objectives. If the nature of the initiative, how it is intended to work and the proposed benefits had been more tightly specified at the outset this would have helped to ensure that the evaluation collected the appropriate information about processes and outcomes.

Interviews with service leaders and staff working in services which were subject to evaluation highlighted the reciprocal and dynamic relationship between the initiative and the evaluation. Conducting evaluation can mean that the service has to be run in a particular way, for example to make it possible to recruit patients before they receive any treatment. Studies that were based on data about the process of care often required staff to collect extra data, sometimes in a different format from usual. This can generate considerable extra work for staff, who may not understand how these data will be used. As well as creating extra burdens for staff and sometimes for patients, this extra bureaucracy can mean the service being studied runs less efficiently than it would during normal conditions.

As well as the evaluation having an impact on the service, the reverse can also be the case. Several evaluations had to respond to changes in the service while it was being studied. Examples include the evaluations of Whole System Demonstrators and of Individual Budgets. In addition, there is often a requirement within evaluations to feed back interim findings to the services involved as soon as possible in order to learn early lessons and improve the initiative. However this may lead to changes in the way in which the service is provided, and it is very difficult to provide generalisable lessons from an evaluation of an initiative which is constantly changing. The implication of these observations is that it is important to describe carefully the

ways in which a service is provided at the time it is evaluated and to seek to conduct the evaluation over a short term to avoid the effects of change over time.

This raises the issue of what is the most appropriate point in the lifecycle of an initiative to conduct evaluation. Evaluation is conducted in order to improve implementation of an initiative or to measure its success. This argues for conducting evaluations as early as possible so that lessons can be learnt from a small number of pilot sites before widespread roll-out of the initiative. This was the model followed in several evaluations reviewed for this report (e.g. Self Referral for Physiotherapy, IAPT, Care Closer to Home). However, pilot sites are often chosen for particular reasons, such as their enthusiasm to participate, and the experience of early pioneers under these circumstances may not provide a good guide to the future success of the scheme in more representative sites. Also, new initiatives often need time to 'bed-down' before they operate normally.

This has several implications. First, that every effort should be made to recruit pilot sites which are representative of the range of settings in which the initiative will be implemented. Second, that full details are given of how pilot sites were selected and their characteristics. Third, that evaluation should not be conducted until sites have overcome early operational difficulties and are operating normally. Several of the evaluations reviewed for this report failed to fulfil these requirements and there was a marked lack of information about how pilot sites were chosen.

However, if evaluation is not conducted at an early stage then the findings come too late to have an influence on policy decisions. The tension between the timescales needed to make policy decisions and to conduct research was a recurring theme in interviews and evaluation reports. It is clear that the length of time it takes to conduct high quality research is not compatible with the policy making process. The combined effect of the requirements of research ethics and research governance frameworks, the need to plan evaluation, to recruit sites and participants, and to collect data which may mean following up participants for extended periods, means that research often takes several years to generate results, and it is not realistic to expect policy decisions to be delayed this long.

Although this problem is well recognised it is less clear how to address it. Some of the evaluations reported here faced major constraints because of this problem. In an attempt to provide evidence quickly in order to inform policy some evaluations were undertaken against unrealistically tight timetables (e.g. New Types of Worker, Individual Budgets, LifeCheck). The consequence was to severely limit the reliability and usefulness of the findings. In other cases, the pressure to implement the initiative quickly meant that it was planned and implemented well in advance of the evaluation (e.g. New Types of Worker, IAPT, Common Assessment Framework), which meant that opportunities for robust evaluation using pre and post data or an experimental design were lost.

Resolving the different requirements of research and policy is not easy, but several recommendations can be made. If there is a pressing need to implement policy, it may be inappropriate to commission evaluation at all, unless there is a commitment to take note of the findings. Formative evaluation, aiming to improve implementation of the policy may be a better investment than summative evaluation.

If summative evaluation is to be conducted, it is better not to compromise it by requiring the evaluators to generate very quick results. These may be misleading for

the reasons alluded to previously: the first pilot sites may be unrepresentative, they may not operate normally when first established, and outcomes for participants may not become apparent until they have been followed up for a reasonable length of time.

If possible, the initiative should be rolled out to a limited number of sites and then further roll-out should pause until the results of summative evaluation are available. If this is not possible, and evaluation will not influence roll-out of the policy, it is better to make this clear at the beginning. The emphasis of the evaluation may then be different, as the priority will be to learn lessons which might be used to modify the initiative over time (for example which groups of people are most likely to benefit and to whom the initiative should be targeted). In addition the evaluation might seek to learn broader lessons (for example about the relationship between improved access and inequalities in use of service) that can inform other related initiatives, even if they are not decisive for the future of the initiative being evaluated.

Using evaluation

The 'naïve rationalist' position, as described in chapter 1, would suggest that evaluation is undertaken in order to learn lessons from early experience of an initiative, and that these lessons will therefore be used to inform policy. This position assumes that lessons from the evaluation reach the people who will make the policy decisions, that the information provided by evaluation is useful and relevant, and that evidence from evaluation is an important factor in the decisions which need to be made. This overview has provided data which helps to inform each of these steps.

The importance of effective dissemination of research findings was well recognised by several of the teams conducting evaluations reviewed here, and in some cases detailed plans had been made to ensure findings were well publicised. The DH itself has produced attractive and interesting summaries of evidence from some evaluations (e.g. LifeCheck, Self Referral to Physiotherapy). Some teams have publicised the type of evaluation that they are undertaking during the course of the work, so that decision makers are aware of when findings will be released and the type of information that will become available. However, in most cases if a decision maker wanted to know what evaluation had been or was in the process of being undertaken on a particular initiative, it would be almost impossible for them to find out. Our experiences in the mapping exercise suggested that whether at local (PCT), regional (SHA) or national (DH) level, there was a lack of strategic awareness of what evaluation was being or had been conducted. Without this awareness there is great potential for duplication of effort and missed opportunities for useful evaluation, and it is difficult for commissioners, particularly at PCT level, to learn from each other's experience.

In terms of the usefulness of evaluation, we have already highlighted some of the limitations of the evidence that will be provided by many of the evaluations we reviewed. In our interviews, policy makers and service leads were aware of some of these problems, and several interviewees expressed doubts about whether evaluation would tell them things that they did not already know. As previously highlighted in the King's Fund report 'Finding Out What Works'¹¹, there are tensions between the needs of different parties involved in evaluations (politicians, civil servants, researchers and practitioners), and clarity is needed about what these are and the

problems that arise from them in order to maximise the usefulness of evaluation. This problem can only be overcome by greater clarity about the most important questions that need to be addressed by evaluation, and the use of high quality methods to answer these questions. But this review also highlighted several issues in relation to the way in which evidence from evaluation is used to inform policy.

The Cabinet Office reports discussed in chapter 1 described the key characteristics of how evaluation should be used within modern policy making.^{13;14} It is interesting to reflect on the extent to which the evaluations reviewed for this report demonstrate these characteristics.

- *Pilots should be carried out in a genuine spirit of experimentation and are redundant if the policy is already set in stone.* As demonstrated by the quotations in chapter 6, several of the evaluators expressed scepticism about whether their findings would have any impact on policy, particularly if they suggested that the initiative was not as successful as hoped. They based their scepticism on their own previous experience of conducting evaluation and on their observations about policy decisions. There was an implication in at least one of the tender documents for evaluation that the main purpose was to find evidence in favour of a policy rather than to conduct a balanced assessment of an initiative and whether or not it achieved its objectives. Similarly, some of the DH summary reports about the results of evaluation appeared to be designed to show evidence to support an initiative, and in doing so tended to overstate the findings from the data.
- *The benefits of a pilot will not be realised if the policy is implemented before the study has run its course.* In most cases the policy had already been implemented before the study had run its course.
- *It is crucial that evaluation of pilots is independent if it is to be credible.* Although many of the evaluators described themselves as independent, it appeared to us that this could be questioned in some cases. Some of those involved in evaluation had also been involved in promoting or developing the initiative in question. This is not altogether surprising or necessarily inappropriate, as these individuals are often also those who are most knowledgeable about the subject area. But it is difficult for proponents for an idea to be completely objective about it, and in our view any potential conflicts of interest should be transparent so that readers can take these interests into account when reading evaluation reports.
- *Methods must be rigorous but multiple methods should be considered to obtain a full picture of the impact of an initiative.* As previously noted, the rigour of the methods used in different studies was highly variable. However, several of the larger evaluations did make good use of multiple methods.
- *A pilot that reveals a policy to be flawed or ineffective should be viewed as a success rather than a failure, having potentially helped to guard against the inclusion of embarrassing, expensive and preventable mistakes into new policy initiatives.* As described above, this message does not appear to have been entirely accepted. At least one of the evaluators interviewed felt that their findings suggested that the policy was flawed, but that these findings were ignored. Some of the policy leads suggested that evaluations which did not give the desired results were a problem, rather than an opportunity to

disinvest in an ineffective policy. However, a bigger problem was that several evaluations were unable to provide clear evidence about whether the initiative was a success or failure since success criteria were not clearly defined and the evaluations did not collect evidence about health or social care outcomes.

- *Appropriate mechanisms should be in place to adapt or abandon a policy or its delivery mechanism in the light of a pilot's findings.* Our review was conducted too soon to be able to tell what impact the evaluations will have on policy. There was evidence from some of the evaluation reports that delivery mechanisms had been adapted in the light of evaluation findings, but there did not appear to be clear mechanisms to adapt or abandon policy in the light of evaluation.

Some of the above points risk falling into the trap of assuming that evaluation findings can ever be more than one consideration in the policy making process. It is important to recognise that there will always be a tension between the researchers who conduct evaluations and those who use the information within the policy process. The essence of research is to subject ideas to critical scrutiny, and researchers can only ever disprove theories, not prove them. However policy makers have to promote and implement ideas, sometimes against resistance, and criticism is likely to make this more difficult. Some of the individuals interviewed recognised the different roles and needs of researchers and policy makers. In general, we identified a healthy willingness of policy leads to engage with evaluators and to subject their policies to scrutiny, even if the findings sometimes caused them difficulty. When evaluation was commissioned from university based academics, there appeared to be well established mechanisms to allow free and open publication of results subject to the DH having an opportunity to see reports prior to publication so that they could prepare their response. This mechanism generally worked well (although evaluators often have to be reminded to send advance notice of their draft publications). However, when evaluation was commissioned from consultancies the conduct and the findings of evaluations were much less transparent, and some of the reports cited in DH documents in support of initiatives did not appear to be publicly available.

Some of the literature on the role of evaluation in policy making discussed chapter 1 emphasised the important role of personal contact between evaluators and policy makers. Our findings supported this. A recurring theme in the interviews was the benefits of discussions during the process of planning the evaluation, continued close contact during the conduct of the evaluation, and the importance of influential networks in disseminating the results. Our findings also tended to support previous literature which suggests that the role of evidence in policy making is not linear, planned and staged, but is rather negotiated, influenced by personal relationships and subject to the random play of outside events. This may partly explain the finding that some of the evaluations we reviewed had a much higher public profile than others.

Our findings also supported the concept of 'bricolage' introduced by Freeman²⁸ – that evidence is collected in a piece-meal fashion and then used selectively to support particular arguments. If findings were not in tune with a current policy they might be used to support another policy now or in the future. This was consistent with the observation that even though the findings from summative evaluation may be published too late to influence the roll-out of a particular initiative, they may be useful in terms of decisions about future related initiatives.

Postscript

As this report was being finalised, the Government Office for Science (GO-Science) published a Science Review of the Department of Health.¹⁵¹ This was based on collecting evidence from a consultation process with external stakeholders, interviews with DH staff and peer reviews and case studies of projects. The review highlighted many examples of good practice in the way that DH commissions, manages and uses science, which other government departments could usefully copy. Many of the specific findings and recommendations from GO-Science review echo the findings of this report. In particular:

- The need for more transparency in the research procurement process
- The need to maintain and expand the transparency of research findings and the evidence which underpins policy making.
- The benefits of having an academic scientific advisor to provide expert advice at all stages of planning, monitoring and evaluation, as used successfully by the PRP programme
- The benefits of initiatives (e.g. master classes and training programmes) for policy makers and scientists to develop a shared understanding to encourage better policy making
- The need for DH to have a mechanism for ensuring that all relevant scientific advice is considered in policy
- More could be done to ensure that stakeholders are aware of the evidence base and its role in policy decision making. The DH should consider the effectiveness of its systems to ensure that there is access to the evidence, including historical data.
- The need to ensure that methods and resources for evaluation and feedback are built into programmes and projects from the outset
- A need for more consistency in the use of pilots and their evaluation before roll-out of policies
- The need for a better internal knowledge management system and evidence log in policy areas to maintain and share existing evidence.

Conclusion

There was a commendable commitment to evaluate many of the initiatives promoted by the White Paper 'Our health, our care our say'. There was considerable variation in the ways in which these evaluations were carried out. In general, the formative evaluations were probably more useful to policy makers in the short term. Some of the summative evaluations were limited by major design compromises or difficulties in collecting data, which were often related to not taking account of the evaluation when the initiative was first implemented, commissioning the evaluation once the initiative had already started, and expecting findings in too short a time scale. Despite these problems, some of the evaluations were well planned and highly ambitious, and the fact they were undertaken at all is an achievement.

It is pertinent to recap the key features of evaluation cited in 'Professional Policy Making for the 21st Century'¹⁰ as important in modern policy making. There should be a clearly defined purpose for the evaluation at the outset, defined success criteria, a means of evaluation built into the policy making process from the outset, and a commitment to allow the pilots to influence final outcomes. Almost all of the evaluations reviewed for this report would have been more effective and useful if these principles had been followed more closely when evaluation was first considered. These therefore remain excellent principles to guide future evaluation of initiatives in health and social care.

Recommendations

There is no shortage of good advice and resources available to policy makers and service managers who wish to evaluate new initiatives. However advice alone has not proved sufficiently effective in changing practice. Experience in other fields suggests that rapid and sustained change is more likely to occur if systems are changed which facilitate a change in behaviour.

Over the last decade, a number of initiatives within health research have been introduced in response to deficiencies of much medical research of relevance to the NHS. Researchers who wish to ultimately publish their findings are now required to register and publish key features of their study before they begin the work in a publicly available register of clinical trials. There is increasing pressure for pharmaceutical companies to do the same for drug trials. These registers, which provide a form of pre-study disclosure, help to avoid problems of selective reporting of findings. In addition, many journals require authors to adhere to guidelines such as CONSORT¹⁵² (for RCTs), TREND^{152;153} (for non-randomised public health interventions) and STROBE^{152;154} (for epidemiological studies such as cross-sectional designs). Use of reporting guidelines can lead to improved accuracy and transparency of publications; it facilitates appraisal for research quality and relevance, and can improve the efficiency of literature searches and ability to find information. The National Knowledge Service (NKS),¹⁵⁵ established to develop a strategic approach to the management of the £150M annual NHS expenditure on knowledge and information services, works with journal editors, research funding bodies and other collaborators with a mutual interest in improving the quality of research publications and of research itself. We believe it would be timely to follow a similar process in relation to the conduct and reporting of evaluations of health and social care initiatives relating to NHS policy.

The following recommendations are intended to improve the quality of evaluations and the confidence which can be placed in the reported findings:

1. Before undertaking any evaluation, policy leads should complete an **evaluation plan** which addresses the following five key questions:
 - Why do we need to undertake evaluation?
 - Do we want to know whether an initiative is effective or cost-effective, or do we want to know how to improve implementation of an initiative?
 - What would define success or failure of the initiative?
 - What difference will the evaluation make to policy? (If evaluation suggests that the initiative does not achieve its objectives what will we do?)
 - What mechanisms need to be in place to expand, halt or amend a policy in the light of evaluation?

Policy leads, in discussion with evaluators, should set out explicitly the theoretical basis for the policy. In particular, this theory should specify (i) how the initiative is intended to work, (ii) in what contexts and/or for which groups of people is it most likely to be beneficial, and (iii) what are the intended beneficial outcomes.

It may be helpful for policy leads to commit the answers to these questions in a written evaluation plan and to discuss them with the research and development directorate of the DH. This would help policy leads to improve the clarity of purpose of the proposed evaluation and would increase the likelihood that it would generate useful findings. The DH should require such a written evaluation plan as part of the internal approval process for any major initiative.

2. There should be a publicly available **register** of summative evaluations of health and social care initiatives which are publicly funded above a certain threshold (e.g. £50,000). This should be completed and made publicly available online *before* the evaluation begins, with a dated record of any changes to the plans. This approach is already standard practice for randomised controlled trials in registers such as controlled-trials.com, and a similar approach would be beneficial for evaluations using other methods. A suitable register might be appropriately co-ordinated by NIHR. The register should contain fields such as:

- Title
- Brief details of the initiative to be evaluated
- Keywords for health or social care conditions (e.g. mental health problems)
- Keywords for type of initiative (e.g. telehealth)
- Main aims of initiative
- Main aims of evaluation
- Study design (e.g. RCT, controlled before and after study, observational comparison)
- Primary outcome measures
- Secondary outcome measures
- Methods used to collect data
- Length of follow-up (between receiving intervention and outcome data being collected)
- Participants (inclusion and exclusion criteria)
- How participants will be identified and sampled
- Settings in which participants will be recruited, number of sites
- How participants are allocated to or included in comparison groups (if randomised, how? If not randomised, what determines which group they are in?)
- Proposed sample size and justification
- Approach to analysis
- Funding body for the evaluation
- Has ethical approval been obtained?
- Start date and end date

3. More attention needs to be given to ensuring that policy leads are aware of and encouraged to use existing **guides** to effective evaluation (see bibliography), and to

take advantage of the training and professional development modules in Policy Evaluation and Analysis offered by the Government Social Research Unit.

4. Many local NHS Trusts will wish to evaluate their local services and this should be encouraged. Each Trust (including PCTs, hospital trusts, and foundation trusts) should have ready access to a source of **help and advice** with regard to evaluation design, data collection and analysis. In addition they need to be made more aware of the help available at a national level, for example from the resources listed in the bibliography to this report. The Government Social Research Unit should consider developing a web-based distance learning resource tailored to the needs of local NHS provider organisations and commissioners to help them conduct better quality and more useful evaluation. This could be based on their existing sources of guidance. Alternatively the DH could commission such a resource from another organisation.

5. The DH should always **involve people with appropriate expertise** in agreeing an appropriate design for evaluation. This may be achieved by involving the DH research and development directorate or by commissioning evaluation via one of the NIHR research programmes.

6. Where possible, summative evaluation should be commissioned '**at arms length**' from the DH by organisations such as the NIHR Policy Research Programme or the Service Delivery and Organisation Programme.

7. If possible, there should be a **pilot phase** in which the initiative is implemented in a limited number of sites, and decisions about future roll-out should not be made until the findings of evaluation are publicly available.

8. The DH should seek to maximise the **independence** of the team undertaking evaluation, and any potential conflicts of interest should be transparent.

9. The DH should make a **commitment to publish** all evaluation reports and to make them freely publicly available, irrespective of the findings.

10. The DH needs to develop a **mechanism** for ensuring that evaluations follow their proposed plan (or changes are fully justified) and that the findings of evaluation are effectively disseminated to decision makers at local and national levels, given that the original policy leads who commissioned the evaluation may have moved on.

8. Recommended guides to evaluation

Research Methods for Policy Evaluation

Purdon S, Lessof C, Woodfield K, Bryson C. 2001. *Research Methods For Policy Evaluation*. Department for Work and Pensions Analytical Services Division Working Paper no 2: London.

Useful overview of main evaluation methods used within the Department for Work and Pensions. This working paper is intended as an introduction and thus intentionally lacks detail on design issues which are covered elsewhere in the same working paper series. This paper contextualises and defines key areas, and offers guidelines to when different approaches are useful or appropriate. Includes chapters on process evaluation and impact evaluation, and gives a clear summary of the differences between the two.

Available electronically from <http://research.dwp.gov.uk/asd/asd5/WP2.pdf>

The Magenta Book 2003

An electronic publication and set of resources available from the Civil Service website (www.civilservice.gov.uk), Policy Hub website (www.policyhub.gov.uk) and Government Social Research website (www.gsr.gov.uk) offering a set of guidance notes for policy evaluators and analysts and people who use and commission policy evaluation. It provides a user friendly guide for specialists and generalists on methods used by social researchers when they commission, undertake and manage policy research and evaluation. The resource includes guidance on how to refine a policy question to get a useful answer, the main evaluation methods used to answer policy questions, the strengths and weaknesses of different methods of evaluation, the difficulties which arise in using different methods of evaluation, the costs involved in different methods of evaluation and the benefits that are to be gained. Comprehensive links to further resources are also provided.

The Magenta Book was linked to a series of training and professional development modules in policy evaluation and analysis developed by the Government Chief of Social Researchers Office for government analysts and policy makers who use and commission policy evaluation.

United Kingdom Evaluation Society website

www.evaluation.org.uk

The UK Evaluation Society is a professional membership organisation that seeks to promote and improve the theory, practice, understanding and utilisation of evaluation. This useful website includes a 'Resources' section which includes good practice guidelines, a bibliography and links to a number of on-line resources.

W.K. Kellogg Foundation Evaluation Handbook

This handbook is a theoretical and practical guide to evaluations, divided into two parts. The first part gives a background to evaluation approaches, methodologically and historically contextualising the Foundation's approach to evaluation, concluding with an overview of the Foundation's three levels of evaluation, with a particular focus on project-level evaluation which is the primary subject of the handbook. The second part of the handbook gives a step by step guide to planning, implementing and utilizing evaluations, illustrated with case study examples. The aim of this is to provide a blueprint of the process of evaluation, stopping short of being a fully comprehensive how-to manual.

Available electronically from www.wkkf.org/Pubs/Tools/Evaluation/Pub770.pdf

9. References

- (1) Department of Health. Our health, our care, our say: a new direction for community services. The Stationery Office; 2006.
- (2) Department of Health. Making it happen: pilots, early implementers and demonstration sites. London: Department of Health; 2006.
- (3) Department of Health. Our health, our care, our say: making it happen. London: Department of Health; 2006.
- (4) Department of Health. The NHS Improvement Plan. Putting People at the Heart of Public Services. London: Department of Health; 2004.
- (5) Department of Health. Independence, Well-being and Choice: Our Vision for the future of Social Care for Adults in England: Social Care Green Paper (Cm 6499). London: Department of Health; 2005.
- (6) Black N. Evidence based policy: proceed with care. *Br Med J* 2001; 323:275-279.
- (7) Nutley S, Davies H. 'Making a reality of evidence-based practice: some lessons from the diffusion of innovations'. *Public Money & Management* 2000; 20(4):35-42.
- (8) Dobrow MJ, Goel V, Upshur RE. Evidence-based health policy: context and utilisation. *Soc Sci Med* 2004; 58(1):207-217.
- (9) Cabinet Office. Modernising Government. CM4310. London: The Stationery Office; 1999.
- (10) Strategic Policy Making Team. Professional Policy Making for the 21st Century. London: The Cabinet Office; 1999.
- (11) Coote A, Allen J, Woodhead D. Finding out what works: building knowledge about complex, community-based initiatives. London: King's Fund; 2004.
- (12) Department of Health. The New NHS. Cm 3807 ed. London: HMSO; 1997.
- (13) Performance and Innovation Unit. Adding it up. London: Cabinet Office; 2000.
- (14) Jowell R. Trying It Out. The Role of 'Pilots' in Policy-Making. London: Cabinet Office; 2003.
- (15) Allen P, Peckham S, Anderson A, Goodwin N. Commissioning research which is used: The experience of the NHS Service Delivery and Organisation Research and Development Programme. *Evidence and Policy* 2007; 3(1):119-134.
- (16) Weiss C. 'The many meanings of research utilization'. *Public Administration Review* 1979; 39:426-431.
- (17) Weiss C. What kind of evidence in Evidence-Based policy? Third International Inter-disciplinary Evidence-Based Policies and Indicator Systems Conference
- (18) Weiss C. The Interface between Evaluation and Public Policy. *Evaluation* 1999; 5:468.
- (19) Klein R. 'From evidence-based medicine to evidence-based policy'. *Journal of Health Services & Research Policy* 2000; 5(2):65-66.
- (20) Innvaer S, Vist G, Trommald M, Oxman A. Health policy-makers' perceptions of their use of evidence: a systematic review. *Journal of Health Services & Research Policy* 2002; 7(4):239-244.
- (21) Hanney SR, Gonzalez-Block MA, Buxton MJ, Kogan M. The utilisation of health research in policy-making: concepts, examples and methods of assessment. *Health Res Policy Syst* 2003; 1(1):2.
- (22) Graham I, Tetroe J, Robinson N, Grimshaw J, the International Funders Study Research Group. An International Study of Health Research Funding Agencies' Support and Promotion of Knowledge Translation. Boston: Academy Health Annual Research Meeting; 2005.
- (23) Clark G, Kelly L. New Directions for Knowledge Transfer and Knowledge Brokerage in Scotland. Scotland: Scottish Executive; 2005.
- (24) Pawson R, Tilley N. Realistic evaluation. London: Sage; 1997.
- (25) Elliott H, Popay J. How are policy makers using evidence? Models of research utilisation and local NHS policy making. London: *Journal of Epidemiology and Community Health*; 2000.
- (26) Russell J, Greenhalgh T, Byrne E, McDonnell J. Recognizing rhetoric in health care policy analysis. *Journal of Health Services & Research Policy* 2008; 13(1):40-46.
- (27) WK Kellogg Foundation. WK Kellogg Foundation Evaluation Handbook. Battle Creek MI: Kellogg Publications; 1998.

- (28) Freeman R. Epistemological bricolage: how practitioners make sense of learning. *Administration & Society* 2007; 39(4):476-496.
- (29) Sabatier P, Jenkin-Smith J. *Policy Change: an advocacy coalition approach*. Boulder: Westview Press; 1993.
- (30) Burns A. Recognizing the needs and roles of the key actors. *Journal of Health Services & Research Policy* 2005; 10 Suppl 1:53-54.
- (31) Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004; 82(4):581-629.
- (32) Lavis J, Davies H, Oxman A, Denis JL, Golden-Biddle K, Ferlie E. Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Policy* 2005; 10 Suppl 1:35-48.
- (33) Patton MQ. *Utilisation focussed evaluation*. 3rd ed. Thousand Oaks: Sage; 1997.
- (34) *Critical Appraisal Skills Programme Tools*.
<http://www.phru.nhs.uk/pages/PHD/resources.htm>
- (35) Yin RK. *Case Study Research: Design and Methods*. London: Sage; 2003.
- (36) Gray DE. *Doing Research in the Real World*. 2nd ed. London: Sage; 2009.
- (37) Rose H. Case Studies. In: Allan G, Skinner C, editors. *Handbook for Research Students in the Social Sciences*. London: Falmer Press; 1991. 190-202.
- (38) Silverman D. *Doing Qualitative Research*. London: Sage; 2000.
- (39) Britten N. Qualitative Interviews in Healthcare. In: Pope C, Mays N, editors. *Qualitative Research in Healthcare*. 2nd ed. London: BMJ Books; 1999. 11-19.
- (40) Barbour R. Education and debate - Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *BMJ* 2001; 322.
- (41) Miles MB, Huberman AM. *Qualitative Data Analysis*. 2nd ed. Thousand Oaks: Sage; 1994.
- (42) O'Cathain A. Mixed Methods Research in the Health Sciences. A quiet revolution. *Journal of Mixed Methods Research* 2009; 3(1):3-6.
- (43) Pope C, Mays N, Popay J. *Synthesizing Qualitative and Quantitative Health Evidence. A guide to methods*. Maidenhead: Open University Press; 2007.
- (44) Department of Health. *Choosing Health: Making healthier choices easier*. 6374 ed. London: The Stationery Office; 2004.
- (45) Department of Health. *Health reform in England: update and next steps*. London: Department of Health; 2006.
- (46) Department of Health. *Supporting People with Long Term Conditions. An NHS and Social Care Model to support local innovation and integration*. Leeds: COI; 2005.
- (47) Wanless D. *Securing Our Future Health: Taking a Long-Term View, Final Report*. HM Treasury; 2002.
- (48) Department of Health. *Reforming NHS financial flows. Introducing Payment by Results*. London: Department of Health; 2002.
- (49) Leese B, Bohan M, Gemmell I, Hinder S, Mead N, Pickard S et al. *Evaluation of 'Closer to Home' Demonstration Sites: Final Report*. National Primary Care Research and Development Centre, University of Manchester / Health Economics Facility, University of Birmingham; 2007.
- (50) Department of Health. *New guidance - Implementing care closer to home: convenient quality care for patients*.
http://www.dh.gov.uk/en/Policyandguidance/Organisationpolicy/Primarycare/Practitionerswithspecialinterests/DH_074419
- (51) Department of Health. *Payment by Results Tariff Information 2007/08*.
http://www.dh.gov.uk/en/Policyandguidance/Organisationpolicy/Financeandplanning/NHSFinancialReforms/DH_077279
- (52) Coast J, Noble S, Noble A, Horrocks S, Asim O, Peters TJ et al. Economic evaluation of a general practitioner with special interests led dermatology service in primary care. *BMJ* 2005; 331(7530):1444-1449.
- (53) Powell J. Systematic review of outreach clinics in primary care in the UK. *J Health Serv Res Policy* 2002; 7(3):177-183.
- (54) Salisbury C, Noble A, Horrocks S, Crosby Z, Harrison V, Coast J et al. Evaluation of a general practitioner with special interest service for dermatology: randomised controlled trial. *BMJ* 2005; 331(7530):1441-1446.
- (55) Layard R. *Mental health: Britain's biggest social problem? Paper presented at the No.10 Strategy Unit Seminar on Mental Health on 20th January 2005*. London: LSE; 2005.
- (56) Layard R. *The Depression Report. A new deal for depression and anxiety disorders*. London: The Centre for Economic Performance, London School of Economics; 2006.

- (57) National Institute for Clinical Excellence. Management of depression in primary and secondary care. London: National Institute for Clinical Excellence; 2004.
- (58) National Institute for Clinical Excellence. Depression (amended). Management of depression in primary and secondary care. CG23 ed. London: National Institute for Clinical Excellence; 2007.
- (59) National Institute for Clinical Excellence. Anxiety: management of anxiety (panic disorder, with or without agoraphobia, and generalised anxiety disorder) in adults in primary, secondary and community care. London: National Institute for Clinical Excellence; 2004.
- (60) Department of Health. A National service Framework for Mental Health: Modern Standards and Service Models for Mental Health. London: The Stationery Office; 1999.
- (61) Department of Health. The National Service Framework for Mental Health - Five Years On. London: The Stationery Office; 2004.
- (62) Blythe B. Incapacity Benefit reforms - Pathways to Work Pilots performance and analysis. 26 ed. Department for Work and Pensions; 2006.
- (63) CSIP Choice and Access Team. Commissioning a brighter future: improving access to psychological therapies - positive practice guide. London: Department of Health; 2007.
- (64) Improving Access to Psychological Therapies: About us. <http://www.iapt.nhs.uk/about/>
- (65) Clark DM, Layard R, Smithies R. Improving access to psychological therapies: evaluation of the two demonstration sites. Working Paper 1648 ed. London: LSE Centre for Economic Performance; 2007.
- (66) DH/Mental Health Division/Improving Access to Psychological Therapies programme. The IAPT Pathfinders: Achievements and Challenges. London: Department of Health; 2008.
- (67) Community Care (Direct Payments) Act 1996 (c.30). London: HMSO; 1996.
- (68) Spandler H. Friend or Foe? Towards a Critical Assessment of Direct Payments. Critical Social Policy 2004; 24(2):187-209.
- (69) Glendinning C, Haliwell S, Jacobs S, Rummery K, Tyrer J. Bridging the gap: using direct payments to purchase integrated care. Health and Social Care in the Community 2000; 8(3):192-200.
- (70) Glendinning C, Haliwell S, Jacobs S, Rummery K, Tyrer J. Buying Independence: using direct payments to purchase integrated care. Bristol: Policy Press; 2000.
- (71) Concannon L. Inclusion or control? Commissioning and contracting services for people with learning disabilities. British Journal of Learning Disabilities 2006; 34(4):200-205.
- (72) Leece D, Leece J. Direct payments; creating a two tiered system in social care? British Journal of Social Work 2008; 36(9):1379-1393.
- (73) Glendinning C, Arksey H, Jones K, Moran N, Netten A, Rabiee P. Individual Budgets Pilot Projects: Impact and outcomes for carers. York: University of York; 2009.
- (74) Ellis K. Direct payments and social work practice; the significance of street level bureaucracy. British Journal of Social Work 2007; 37(3):215-227.
- (75) Glendinning C, Challis D, Fernandez J, Jacobs S, Jones K, Knapp M et al. Evaluation of the Individual Budgets Pilot Programme: Summary / Full report. York: Social Policy Research Unit, University of York; 2008.
- (76) Cabinet Office. Improving the life chances of disabled people. London: Cabinet Office Strategy Group; 2005.
- (77) Department of Health. Independence, Well-being and Choice. London: Department of Health; 2005.
- (78) Lewis I. Speech at the Individual Budgets and Modernising Social Care National Conference. London: 2007.
- (79) Putting people first: a shared vision and commitment to the transformation of adult social care. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_081118
- (80) Department of Health. Transforming health care. London: Department of Health; 2008.
- (81) Darzi A. High quality care for all. London: Department of Health; 2008.
- (82) Brown G. Speech on the National Health Service. London, King's College London/Florence Nightingale School of Nursing; 2008.
- (83) Glasby J, Duffy S. Our health, our care, our say: What could the NHS learn from individual budgets and direct payments? Birmingham, Health Services Management Centre: Department of Health; 2007.
- (84) Glasby J. Individual Budgets and the interface with health: a discussion paper for the Care Services Improvement Partnership (CSIP). Birmingham, Health Services Management Centre: 2008.

- (85) Poll C, Duffy C, Hatton C, Sanderson H, Routledge M. A report on in Control's first phase, 2003-2005. London: In Control Publications; 2006.
- (86) Manthorpe J, et al. Safeguarding and Systems Change. *British Journal of Social Work* 2008;1-16.
- (87) Glendinning C, Challis D, Fernandez J, Jacobs S, Jones K, Knapp M et al. Evaluating the Individual Budgets Pilot Project. *Journal of Care Services Management* 2007; 1(2):123-128.
- (88) Calnan M, Ferlie E. *Analysing process in health care: the methodological and theoretical challenges*. 131 (2) ed. London: Policy and Politics; 2003.
- (89) Coventry City Council. Individual Budgets in Coventry; Our Stories. http://www.coventry.gov.uk/ccm/cms-service/download/asset/?asset_id=17779025
- (90) Daly Gea. Evaluating individual budgets in Coventry; Final report. Coventry University: Faculty of Health and Life Sciences; 2008.
- (91) Department of Health. *Building on the Best Choice, Responsiveness and Equity in the NHS*. London: Department of Health; 2003.
- (92) Department of Health. Better information, better choices, better health: Putting information at the centre of health. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4098576
- (93) Department of Health. Your health, your care, your say. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4127357
- (94) Department of Health. Our NHS, our future: NHS Next Stage Review - interim report. http://www.dh.gov.uk/dr_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_079087.pdf
- (95) The handbook to the NHS constitution to England. http://www.nhs.uk/NHSEngland/aboutnhs/Documents/NHS_Constitution_interactive_9_Mar09.pdf
- (96) King E, Rudat K, Andersen B, Grabarz A, Glendinning R, McHugh S et al. Evaluation of Information Prescriptions. Final Report to the Department of Health London: OPM. London: OPM; GfK; University of York; 2008.
- (97) Information prescriptions online resource. <http://www.informationprescription.info/resource/>
- (98) Read S. *Exploring New Roles in Practice*, final report. <http://www.shef.ac.uk/content/1/c6/01/33/98/enrip.pdf>
- (99) Department of Health. *The NHS Plan. A plan for investment. A plan for reform*. London: Department of Health; 2000.
- (100) Department of Health. *Changing workforce programme new ways of working in health care*. London: Department of Health; 2002.
- (101) Guest DPRRP. *An evaluation of the impact of nurse, midwife and health visitor consultants*. London: Kings College, London; 2004.
- (102) Read S, Graves K. *Reduction of Junior Doctors' Hours in Trent Region: the Nursing Contribution*. Sheffield: University of Sheffield; 1994.
- (103) Dowling S, Martin R, Skidmore P, Doyal L, Cameron A, Lloyd S. Nurses taking on junior doctors work: a confusion of accountability. *BMJ* 1996;(312):1211-1214.
- (104) Kessler IBS. *The Skills for Care New Types of Worker Programme: Stage 1 Evaluation Report*. London: Skills for Care; 2007.
- (105) Wild D, Nelson S, Szczepura A. *An Evaluation of an In Reach Model Of Care in LA Care Homes, Report 1*. 2007.
- (106) Flynn M, Pearson J, Paget B, Grant G. *Learning from Skills for Care's New Types of Worker Programme 2003-2007*. Leeds: Skills for Care; 2008.
- (107) Flynn M, Pearson J, Paget B. *New Types of Worker Fund: Evaluation of Phase 3*. Leeds: Skills for Care; 2009.
- (108) Imperial Cancer Research Fund OXCHECK Study Group. Effectiveness of health checks conducted by nurses in primary care: final results of the OXCHECK study. *BMJ* 1995; 310(6987):1099-1104.
- (109) Wood DA, Kinmonth AL, Davies GA, Yarwood J, Thompson SG, Pyke SDM et al. Randomised controlled trial evaluating cardiovascular screening and intervention in general practice: principal results of British family heart study. *BMJ* 1994; 308(6924):313-320.
- (110) Wonderling D, Langham S, Buxton M, Normand C, McDermott C. What can be concluded from the Oxcheck and British family heart studies: commentary on cost effectiveness analyses. *BMJ* 1996; 312(7041):1274-1278.

- (111) Mant D. Health checks - time to check out? *British Journal of General Practice* 1994; 44(379):51-52.
- (112) Warburton D. Evaluation of Your Health, Your Care, Your Say: Final Report. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4138622
- (113) Department of Health. Putting Prevention First, Vascular Checks: risk assessment and management. London: Department of Health; 2008.
- (114) DH/HIPD/SM&HRB/NHS LifeCheck. Developing the NHS LifeCheck: a summary of the evidence base. London: Department of Health; 2008.
- (115) O'Brien S. NHS Early Years LifeCheck Evaluation Reports: Summary and recommendations. http://www.dh.gov.uk/en/PublicHealth/Healthimprovement/NHSLifeCheck/DH_082697
- (116) HM Treasury. Stability, Security and Opportunity for All: Investing in Britain's Long Term Future: New Public Spending Plans 2005—2008, Spending Review 2004. London: HMSO; 2004.
- (117) Department of Health. National Service Framework for Older People. London: Department of Health; 2001.
- (118) Department for Work and Pensions. LinkAge Plus: Capacity building - enabling and empowering older people as independent and active citizens. London: HMSO; 2009.
- (119) PSSRU. National Evaluation of Partnerships for Older People: Interim Report of Progress. Discussion Paper 2612 ed. 2008.
- (120) Salisbury C, Munro J. Walk-in centres in primary care: a review of the international literature. *British Journal of General Practice* 2003; 53(486):53-59.
- (121) Rosen R, Florin D, Dixon J. Access to Health Care - Taking Forward the Findings from the Scoping Exercise. Report of a rapid appraisal of stakeholder views and review of existing literature. London: King's Fund; 2001.
- (122) O'Cathain A, Munro JF, Nicholl JP, Knowles E. How helpful is NHS direct? Postal survey of callers. *BMJ* 2000; 320(7241):1035.
- (123) Kinnersley P, Anderson E, Parry K, Clement J, Archard L, Turton P et al. Randomised controlled trial of nurse practitioner versus general practitioner care for patients requesting "same day" consultations in primary care. *BMJ* 2000; 320(7241):1043-1048.
- (124) Galley P. Physiotherapists as first-contact practitioners--New challenges and responsibilities in Australia. *Physiotherapy* 1977; 63(8):246-248.
- (125) Massey Jr BF. What's all the fuss about direct access? *Physical Therapy* 2002; 82(11):1120-1123.
- (126) Ferguson A, Griffin E, Mulcahy C. Patient self-referral to physiotherapy in general practice - A model for the new NHS? *Physiotherapy* 1999; 85(1):13-20.
- (127) Holdsworth LK, Webster VS. Direct access to physiotherapy in primary care: now?--and into the future? *Physiotherapy* 2004; 90(2):64-72.
- (128) Holdsworth L.K., Webster VS, McFadyen AK. Are patients who refer themselves to physiotherapy different from those referred by GPs? Results of a national trial. *Physiotherapy* 2006; 92(1):26-33.
- (129) Holdsworth LK, Webster VS, McFadyen AK. What are the costs to NHS Scotland of self-referral to physiotherapy? Results of a national trial. *Physiotherapy* 2006; 93:3-11.
- (130) Holdsworth LK, Webster VS, McFadyen AK. Self-referral to physiotherapy: deprivation and geographical setting: Is there a relationship? Results of a national trial. *Physiotherapy* 2006; 92(1):16-25.
- (131) National Statistics DoH. NHS Physiotherapy Services Summary Information for 2004-2005, England. <http://www.ic.nhs.uk/statistics-and-data-collections/hospital-care/outpatients/nhs-physiotherapy-services:-summary-information-for-2004--05-england>
- (132) Chartered Society of Physiotherapy. Briefing paper: self referral to physiotherapy services. London: CSP; 2004.
- (133) Department of Health. Self-referral pilots to musculoskeletal physiotherapy and the implications for improving access to other AHP services. 2008.
- (134) Dees J, Emerson J, Economy P. Enterprising Nonprofits: A Toolkit for Social Entrepreneurs. New York: Wiley; 2001.
- (135) Emerson J, Twersky F. New Social Entrepreneurs: The success, challenge, and lessons of non-profit enterprise creation. San Francisco: Robert Foundation; 1996.
- (136) Lupton C, Peckham S, Taylor P. The Management of Public Involvement in Health Care Purchasing. Buckingham: Opn University Press; 1998.
- (137) Office of the Deputy Prime Minister (ODPM). Factsheet 9, new deal for communities. www.neighbourhood.gov.uk

- (138) Department of Transport and Industry. Leading lights: Experiences from the phoenix development fund. <http://www.berr.gov.uk/files/file37786.pdf>
- (139) Thake S. Primed for growth: Adventure capital fund baseline report. www.neweconomics.org/gen/
- (140) Totner A, Lyne I, Ryan P. Reaching the promised land: Can social enterprise reduce social exclusion and empower communities? *Education, Knowledge and Economy* 2008; 2(1):1-14.
- (141) Department of Transport and Industry. Social enterprise: A strategy for success. http://www.cabinetoffice.gov.uk/third_sector/social_enterprise/?/media/assets/www.cabinetoffice.gov.uk/third_sector/se_strategy_2002%20pdf.ashx
- (142) Lewis R, Hunt P, Carson D. Social enterprise and community based care. Is there a future for mutually owned organisations in community and primary care. London: Kings Fund; 2006.
- (143) Managing your organisation. <http://www.dh.gov.uk/en/Managingyourorganisation/Commissioning/Socialenterprise/index.htm>.
- (144) Darkins A, Ryan P, Kobb R, Foster L, Edmonson E, Wakefield B et al. Care Coordination/Home Telehealth: the systematic implementation of health informatics, home telehealth, and disease management to support the care of veteran patients with chronic conditions. *Telemedicine and e-Health* 2008;1118-1126.
- (145) Department of Health CSIP. Building an evidence base for successful telecare implementation - updated report of the Evidence Working Group of the Telecare Policy Collaborative chaired by James Barlow. http://www.asap-uk.org/information/45779/46315/46331/46347/telecare_evidence/
- (146) Barlow J, Singh D, Bayer S, Curry R. A systematic review of the benefits of home telecare for frail elderly people and those with long-term conditions. *J Telemed Telecare* 2007; 13(4):172-179.
- (147) Deshpande A, Khoja S, McKibbin A, Jadad AR. Real-time (synchronous) telehealth in primary care: systematic review of systematic reviews. *Technology Report No 100*. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2008.
- (148) Pare G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. *Journal of the American Medical Informatics Association* 2007; 14(3):269-277.
- (149) Government Social Research Unit Policy Hub. http://www.nationalschool.gov.uk/policyhub/evaluating_policy/
- (150) Department of Health. Our health, our care, our say - one year on. Research Report. London: Department of Health; 2007.
- (151) Government Office for Science. GO-Science Review of the Department of Health. London: Government Office for Science; 2008.
- (152) Moher D, Schulz KF, Altman DG, CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet* 2001; 357:1191-1194.
- (153) Des Jarlais C, Lyles C, Crepaz N, the TREND Group. Improving the Reporting Quality of Nonrandomized Evaluations of Behavioral and Public Health Interventions: The TREND Statement. *American Journal of Public Health* 2004; 94(3):361-366.
- (154) Elm Ev, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007; 335(7624):806-808.
- (155) National Knowledge Service. EQUATOR - Enhancing the QUALity and Transparency Of health Research. <http://www.nks.nhs.uk/equator.asp>

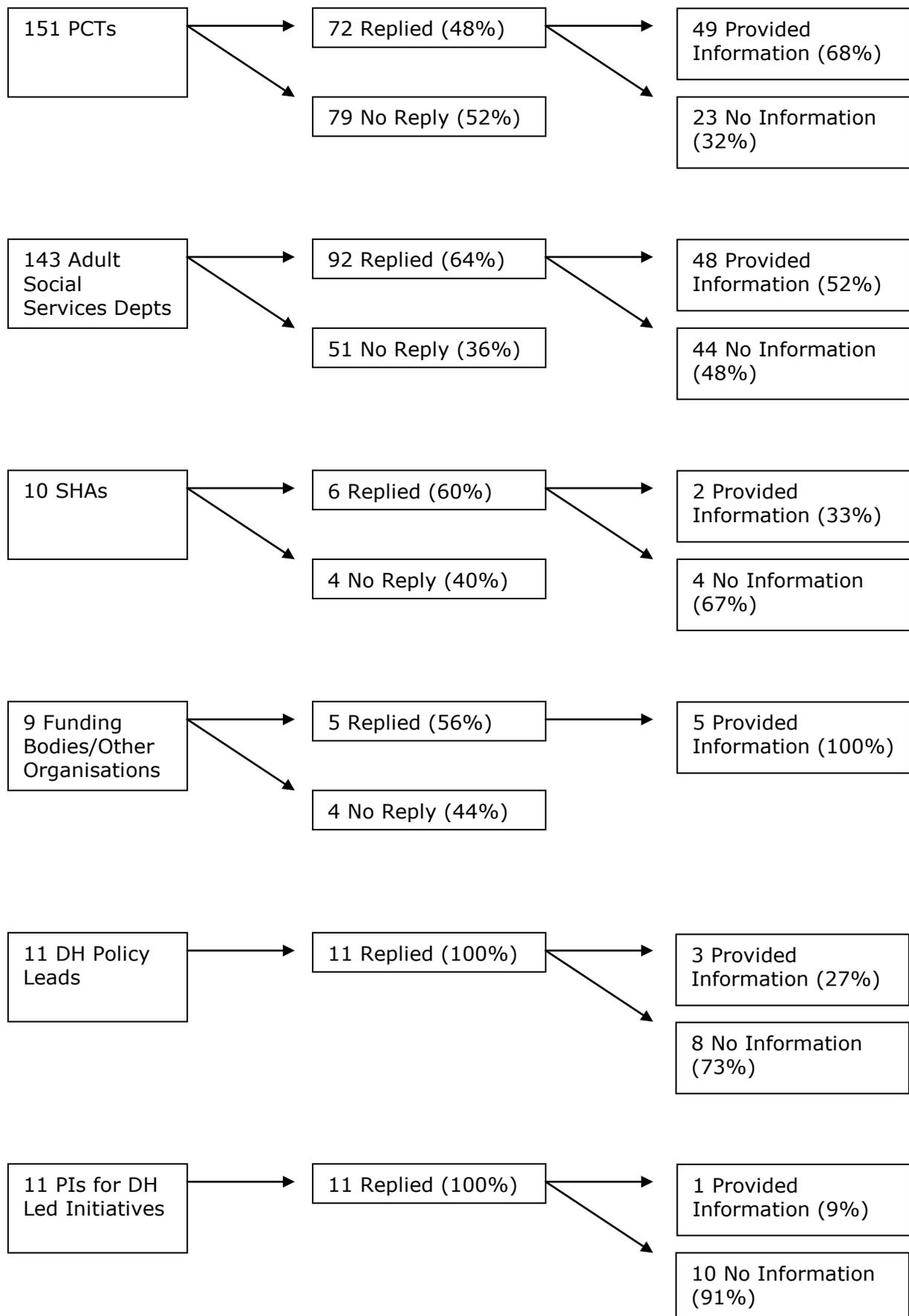
10. Appendices

Appendix A Mapping responses according to contact type

Appendix B Survey Questionnaire (MS Word format)

No table of figures entries found.

Appendix A Mapping responses according to contact type



Appendix B Survey Questionnaire (MS Word format)



Making the Most of Policy Evaluations Questionnaire

You have been invited to complete this questionnaire because we have been told that you are conducting an evaluation which is relevant to the initiatives highlighted in the White Paper 'Our Health, Our Care, Our Say'. If you require any clarification as to why you have been identified, please contact us by email at kate.stewart@bristol.ac.uk, or by phone on 0117 331 3852.

Before completing the questionnaire, please read the MMOE Questionnaire Participant Information Sheet, which you will either have been sent, or can be found at www.bris.ac.uk/primaryhealthcare/MMOE.htm

This questionnaire has been designed to be completed in Microsoft Word, or printed out and completed by hand.

If you are completing the questionnaire electronically, use the arrow keys or tab key on your keyboard to navigate between questions, and either click on the boxes provided to insert a cross , or type your answer in the space provided. Please return your completed questionnaire by email to kate.stewart@bristol.ac.uk

If you are completing the questionnaire by printing it out and filling it in by hand, please return it to:

Dr Kate Stewart
Academic Unit of Primary Health Care
University of Bristol
25 Belgrave Road
Bristol
BS8 2AA

If at any stage you require any clarification or help with the questionnaire, please contact us by email at kate.stewart@bristol.ac.uk, or by phone on 0117 331 3852.

Important, please read carefully

Answers to sections 1, 2 and 3 will be included in a database of projects and may be made available on request to those involved in providing, commissioning or evaluating health or social care. You will be asked to agree to this at the end of section 3.

The information provided in Section 4 will not be included in the database, and will not be identifiable in any reports which make use of this information. These questions are to help us understand the factors which help or hinder evaluation of this type.

Appendix B: Questionnaire

Section 1: Details of evaluation

1. Please Enter the ID number you were sent with your invitation to participate in this survey

2. Title of evaluation

3. Principal investigator

4. Key contact for this evaluation (if different from above)

5. Start date

6. End date

7. Please briefly outline your timetable below, or attach a GANTT chart

8. Please indicate the cost of the evaluation

No specific funding	<input type="checkbox"/>		
Less than £1,000	<input type="checkbox"/>	Between £100,001 and £200,000	<input type="checkbox"/>
Between £1,001 and 10,000	<input type="checkbox"/>	Between £200,001 and £500,000	<input type="checkbox"/>
Between £10,001 and £50,000	<input type="checkbox"/>	Over £500,000	<input type="checkbox"/>
Between £50,001 and £100,000	<input type="checkbox"/>	Don't know	<input type="checkbox"/>

9. Funding body for the evaluation

Section 2: Details of the initiative being evaluated

10. What type of initiative is being piloted?

- Care Closer To Home
- Common Assessment Framework
- Individual Budgets
- Information Prescription
- New Types of Social Care Workers
- NHS Life Check
- Partnership for Older People Programme (POPPs)
- Physiotherapy Self Referral
- Psychological Therapies
- Social Enterprise Pathfinders
- Whole System Long Term Condition

Please add comments if the initiative does not clearly fall within one of these categories

11. Please briefly describe the nature of the service that you are evaluating:

12. What are the key objectives for this initiative. (These are the objectives of the initiative, not your research objectives)

13. Are these objectives written down and explicit? Yes
(please indicate where below)

No
Don't Know

14. Which of the white paper goals is this initiative intended to address?
(tick as many as apply)

- Better prevention and early intervention for improved health, independence and well-being
- More choice and a stronger voice for individuals and communities
- Tackling inequalities and improving access to services; more local and convenient care
- More support for people with long-term needs
- Other/none of the above (please give comments below)

Comments

Appendix B: Questionnaire

15. **What is the geographical area covered by the initiative you are evaluating?**
16. **If you are implementing a national initiative within a specific area(s) (e.g. specific PCTs), please describe the area(s) you are evaluating.**

17. **Please briefly characterise the patients or clients towards whom the initiative is targeted**

(e.g adults with musculoskeletal disorders requiring physiotherapy; adults with mild/moderate depression and anxiety in receipt of incapacity benefit etc)

18. **Please describe any particular features about the local area or other contextual factors which may impact on the generalisability of your evaluation for other areas:**

These might include contextual factors about the local population, other local services available, geographical location, or factors about the local history of service delivery, politics, opinion leaders etc.(e.g. strong links between health and social services; recent closure of an A&E department or health service re-organisation; national opinion leader has links to the area etc.)

Section 3: Approach to evaluation

19. Does your evaluation have a written protocol? Yes
No

20. What are the aims/objectives of the evaluation?

21. Please give us an overview of your evaluation

(You may wish to copy and paste the summary section from your protocol)

22. What are the research components, or stages (if applicable)?

23. Which of the following terms describes your research approach?

(please tick all that apply)

- | | |
|--|---|
| Quantitative <input type="checkbox"/> | Case study approach <input type="checkbox"/> |
| Qualitative <input type="checkbox"/> | Documentary analysis <input type="checkbox"/> |
| Mixed Methods <input type="checkbox"/> | Rapid appraisal <input type="checkbox"/> |
| Action research <input type="checkbox"/> | None of the above <input type="checkbox"/> |
| | Other (please specify) <input type="checkbox"/> |

24. Does your evaluation involve a comparison between people, groups or areas which did or did not receive the intervention? Yes

Please describe your comparison

No, the evaluation only provides information about people or places receiving the initiative

(Please go to question 27)

Appendix B: Questionnaire

25. If you are comparing intervention and control groups, please describe your unit of allocation

(Here, 'intervention group' means those people having access to the initiative you are evaluating, and 'control group' means those not having access to it)

- Patient or client Community
Health or social care provider Clinic day
(e.g. general practice, hospital, social services department)
Institution Geographical area
Other – please describe

26. How were participants allocated to 'intervention' and 'control' groups?

Random allocation of individuals

Random allocation of clusters

Please describe:

People or sites allocated prospectively by the research team, but not randomly

Please describe:

Observational study – no allocation by the evaluation team

27. Observational studies

If your study is an observational study please tick the box(es) below that describe your study:

Controlled before and after study

(Data is collected on both the intervention and control groups before the intervention is introduced and then further data is collected after the intervention is introduced)

Before and after study, without other controls

(Comparison of a group of people before and after they had access to or received the intervention, but without any other comparison group)

After-only study with non-randomised control group

(Data is collected on both the intervention and control groups after the intervention)

Interrupted time series.

(Collection of data on at several different time points both before and after the intervention is introduced, to examine the impact of the intervention on a trend over time.)

Other – please specify

Appendix B: Questionnaire

28. How were participants selected for your evaluation?

(please tick all that apply)

Inclusion criteria
(please describe)

Exclusion criteria
(please describe)

Don't Know

29. How was the sample of participants chosen?

(Please tick all that apply)

Convenience sample

Purposive sample

Quota sample

Random sample

Snowballing

Theoretical sampling

No sampling: all involved were included

30. What types of data collection method are being used?

(Please tick all that apply)

In-depth qualitative interviews

Focus groups

Direct observation/ ethnographic studies

Documentary analysis

Systematic literature review

Questionnaire survey

Collection of quantitative data specifically for this evaluation (other than questionnaires)

Extraction of routinely collected data

Other
please describe

31. How many sites are implementing the initiative, nationally?

32. How many sites did you approach to participate?

33. How many sites are participating in the evaluation?

Appendix B: Questionnaire

34. Does your evaluation have a written analysis plan? Yes
No
Don't Know

34a Was the analysis plan written before the evaluation started? Yes
No
Don't Know

35. What unit(s) of analysis will be used for analysis and comparisons?
(Please tick all that apply)

Patient or client
Health or social care provider (e.g. general practice, hospital, social services department)
Institution
Community
Clinic day
Geographical area
Other
please describe:

36. Does your study involve a pre-planned sample size calculation? Yes
No
Don't know

37. How many individuals are/were in your planned sample size?

a. Intervention
b. Controls (if applicable)
c. Total

38. If recruitment has been completed, please give details of the numbers of patients/clients/staff/sites you have recruited

Appendix B: Questionnaire

39. What are the key outcome measures or measures of success that you are evaluating?

Please tell us the outcome you are assessing and the instrument you are using to assess this outcome.

	Outcome	Measure	Tick if baseline data was collected before clients received the intervention	Tick if this is the primary outcome
e.g.	<i>Client satisfaction</i>	<i>Client satisfaction Questionnaire 8</i>	✓	
			<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>
			<input type="checkbox"/>	<input type="checkbox"/>

40. Did the people collecting data about outcomes know whether or not the respondents had received the intervention?

Yes: they knew whether or not people had received the intervention

No: they did not know – they were 'blind' to the intervention

Neither: the outcomes were based on questionnaires completed by patients/clients

Neither: the outcomes were based on routinely collected data

Other (please specify)

41. For how long will outcomes be measured/were outcomes measured after patients/clients receive the intervention?

Appendix B: Questionnaire

47. Plans for dissemination (please tick all that apply)

Peer-reviewed academic papers

Articles for professionals in magazines
e.g. Health Service Journal

Articles for the public

Presentations at academic conferences

Presentations to policy makers or commissioners

Report for local dissemination only
(*e.g. a PCT or social services department*)

Report for national dissemination
(*e.g. for a national funding body*)

No plans

Other (please specify)

48. Please provide details of any arrangements for interim reports: both those completed so far, and when any future reports are anticipated.

49. Will there be a final project report about your evaluation? Yes

When will this be completed?

No

49a Will your main project report be publicly available ?

Yes

No

No final report

50. Has anyone in your research team or organisation had any involvement in designing or promoting the service that you are evaluating?

Yes

No

Don't Know

51. Does anyone in your research team or organisation have any financial interest or professional interest in the success of the service that you are evaluating?

Yes

No

Don't Know

52. Have any of the bodies funding your evaluation had any involvement in designing or promoting the service that you are evaluating?

Yes

No

Don't Know

Appendix B: Questionnaire

53. Does the body funding the evaluation have the right to prevent or influence the publication of the findings from your evaluation?

Yes

No

Don't Know

Please expand on any 'yes' answers to questions 50-53

Answers to sections 1,2 and 3 will be included in a database of projects and may be made available on request to those involved in providing, commissioning or evaluating health or social care. If you do not agree to this, the database will only include the information in section 1, which is publicly available information.

Please sign below

54. I AGREE / DO NOT AGREE (select as applicable) that the information above may be included in this database.

Name

date

Appendix B: Questionnaire

Section 4: Additional information

The information provided in the remaining questions will not be included in the database, and will not be identifiable in any reports which make use of this information. These questions are to help us understand the factors which help or hinder evaluation of this type.

55. In what ways, if any, did your project design change from the original research plan or protocol?

56. Did your project start on time?

Yes

No

If not, why not?

57. Is your project progressing according to the timetable, or did it finish on time (where applicable)?

Yes

No

If not, why not?

58. Have you had to, or do you anticipate, requesting an extension to the original timetable?

Yes

No

If yes please give details

59. Please add any comments about any factors which are making or made your evaluation progress run particularly smoothly:

Appendix B: Questionnaire

60. Please add any comments about any factors which have made your evaluation particularly difficult to conduct:

61. Please tick here if you would like a copy of the final report of this project, and give details of where to send it below:

Thank you for completing the questionnaire. Please return it **either** by email to kate.stewart@bristol.ac.uk

or by post to

Dr Kate Stewart

Academic Unit of Primary Health Care

University of Bristol

25 Belgrave Road

Bristol

BS8 2AA