King's Fund

National Evaluation of Total Purchasing Pilot Projects
Working Paper

Determining Success Criteria for Total Purchasing Pilot Projects

Ann Mahon
Brenda Leese
Kate Baxter
Nick Goodwin
Judith Scott

National Evaluation of Total Purchasing Pilot Projects
Working Paper

Determining Success Criteria for Total Purchasing Pilot Projects

Ann Mahon
National Primary Care R&D Centre
University of Manchester

Brenda Leese
National Primary Care R&D Centre
University of Manchester

Kate Baxter
Division of Primary Health Care
University of Bristol

Nick Goodwin
King's Fund
London

Judith Scott
Department of General Practice
University of Edinburgh

For further information on this part of the national evaluation contact Brenda Leese (tel 0161 275 7619/fax 0161 275 7600/email bleese@.man.ac.uk).

This working paper forms part of the output of the National Evaluation of Total Purchasing Pilot Projects which is led by the King's Fund

Published by
King's Fund
11-13 Cavendish Square
London W1M 0AN

ISBN 1 85717 191 8

A CIP catalogue record for this book is available from the British Library.

Further copies of this report can be obtained from Grantham Book Services Limited, Isaac Newton Way, Alma Park Industrial Estate, Grantham, Lincolnshire NG31 9SD. Tel: 01476 541080, fax 01476 541061; or from the King's Fund Bookshop. Tel: 0171 307 2591.

This report has been produced to disseminate research findings and promote good practice in health and social care. It has not been professionally copy-edited or proof-read.

The Total Purchasing National Evaluation Team (TP-NET)

The national evaluation of total purchasing pilots in England and Scotland is a collective effort by a large consortium of health services researchers. The study is led by the King's Fund, but also involves the National Primary Care R&D Centre; Universities of Edinburgh, Bristol, Southampton, York and Birmingham; the London School of Hygiene and Tropical Medicine; and the London School of Economics and Political Science. More information about the evaluation as a whole is available from: Nick Goodwin, King's Fund, 11-13 Cavendish Square, London W1M 0AN.

Acknowledgements

The national evaluation was commissioned and funded by the Department of Health in England (1995-98) and the Scottish Office Health Department (1995-97). However, the views expressed in this paper do not necessarily represent the policy of the two Departments.

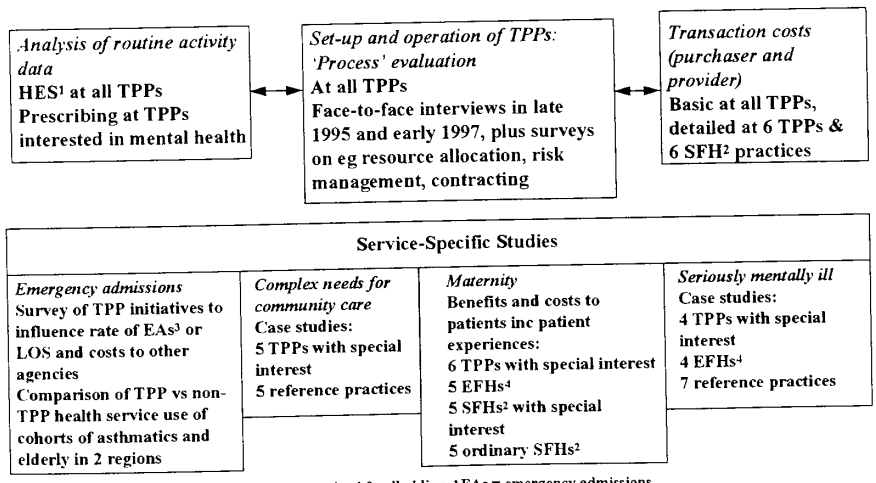
CONTENTS

Preface: The National Evaluation of Total Purchasing Pilot Projects

Total Purchasing Pilot Projects allow for the purchasing of potentially all hospital and community health services by fundholding general practices which began their preparations for contracting in April 1995. Since 'total purchasing' (TP) represented an important extension of the already controversial fundholding scheme, the Department of Health decided to commission an assessment of the costs and benefits of this NHS Executive initiative. This working paper represents part of the interim reporting of the evaluation which began data collection in October 1995 (mid-way through the total purchasing pilots' (TPPs') preparatory year) and which is due to produce final reports in Autumn 1998, by which time the TPPs will have completed two full purchasing years. Other titles in this series of working papers are listed on page iii.

The evaluation amounts to a programme of inter-linked studies and is being undertaken by a large consortium of researchers from different universities led from the King's Fund. Full details of the participants are given on the back cover of this report. All 53 of the 'first wave' TPPs and the 35 'second wave' pilots which began a year later are being studied. The diagram below summarises the main elements of the research which has at its core an analysis of how TP was implemented at all projects and with what consequences, for example, in terms of hospital activity changes. These elements are linked to a series of studies at sub-samples of TPPs which attempt to compare the costs and benefits of TP with conventional health authority purchasing for specific services (emergency admissions, community care, maternity and mental health). In these parts of the evaluation, comparisons are also made between extended fundholding (EFH), where practices take on a new responsibility for purchasing in a single service area (e.g. maternity or mental health) and TP, where practices purchase more widely.

Main components of National Evaluation of First Wave Total Purchasing Pilot Projects



[1] HES = hospital episode statistics, [2] SFH = standard fundholding, [3] EAs = emergency admissions, [4] EFH = extended fundholding pilot

Further details about the evaluation design and methods are available in a leaflet available from the King's Fund and in the preliminary report of the evaluation which was published by the King's Fund early in 1997 and entitled *Total purchasing: a profile of national pilot projects*.

The evaluation would not have been possible without the co-operation and interest shown by all the staff involved in the TPPs. We are very grateful, principally for the time people have given up to be interviewed, whether in practices, health authorities, Trusts, social services departments or elsewhere in the health and social care system.

Nicholas Mays
Co-ordinator, Total Purchasing National Evaluation Team (TP-NET)
King's Fund, London
January 1998

National Evaluation of Total Purchasing Pilot Projects
Main Reports and Working Papers

Title and Authors — *ISBN*

Main Reports

Nicholas Mays, Nick Goodwin, Gwyn Bevan, Sally Wyke on behalf of the Total Purchasing National Evaluation Team (1997). *Total purchasing: a profile of the national pilot projects* — 1 85717 138 1

Nicholas Mays, Nick Goodwin, Amanda Killoran, Gill Malbon on behalf of the Total Purchasing National Evaluation Team (1998). *Total purchasing: a step towards primary care groups* — 1 85717 187 X

Working Papers

The interim report of the evaluation, *Total purchasing: a step towards primary care groups*, is supported by a series of more detailed Working Papers available during the first half of 1998, as follows:

Nicholas Mays, Nick Goodwin, Gill Malbon, Brenda Leese, Ann Mahon, Sally Wyke
What were the achievements of total purchasing pilots in their first year and how can they be explained? — 1 85717 188 8

Gwyn Bevan
Resource Allocation within health authorities: lessons from total purchasing pilots — 1 85717 176 4

Ann Mahon, Brenda Leese, Kate Baxter, Nick Goodwin, Judith Scott
Developing success criteria for total purchasing pilot projects — 1 85717 191 8

Ray Robinson, Judy Robison, James Raftery
Contracting by total purchasing pilot projects, 1996-97 — 1 85717 189 6

Kate Baxter, Max Bachmann, Gwyn Bevan
Survey of budgetary and risk management of total purchasing pilot projects, 1996-97 — 1 85717 190 X

Ann Mahon, Helen Stoddart, Brenda Leese, Kate Baxter
How do total purchasing projects inform themselves for purchasing? — 1 85717 197 7

John Posnett, Nick Goodwin, Amanda Killoran, Gill Malbon, Nicholas Mays, Michael Place, Andrew Street
The transaction costs of total purchasing — 1 85717 193 4

Jennifer Dixon, Nicholas Mays, Nick Goodwin
Accountability of total purchasing pilot projects — 1 85717 194 2

James Raftery, Hugh Macleod
Hospital activity changes and total purchasing
1 85717 196 9

Sally Wyke, Jenny Hewison, James Piercy, John Posnett, Linda Macleod, Lesley Page, Gavin Young
National evaluation of general practice-based purchasing of maternity care: preliminary findings.
1 85717 198 5

Linda Gask, John Lee, Stuart Donnan, Martin Roland
Total purchasing and extended fundholding of mental health services
1 85717 199 3

Susan Myles, Sally Wyke, Jennie Popay, Judith Scott, Andrea Campbell, Jeff Girling
Total purchasing and community and continuing care: lessons for future policy developments in the NHS
1 85717 200 0

Gill Malbon, Amanda Killoran, Nicholas Mays, Nick Goodwin
A profile of second wave total purchasing pilots: lessons learned from the first wave
1 85717 195 0

Summary

Total purchasing was officially introduced following an Executive Letter in October 1994 entitled *Developing NHS Purchasing and GP Fundholding* (NHSE, 1994).

Total purchasing pilots (TPPs) developed through local negotiation with their host health authorities (HAs) and consequently many different styles of pilots have resulted (Mays and Dixon, 1997; Mays *et al*, 1998). Unlike standard fundholding, the TPPs were to be evaluated and, following a tendering exercise, this evaluation was to be undertaken by a consortium of research institutions led by the King's Fund Policy Institute.

There were no set criteria against which to assess the merits of the various ways in which total purchasing developed nationally. An important part of the first year's evaluation therefore was an attempt to derive success criteria which would enable comparison between all the TPPs. This paper is the result of these deliberations and examines the different criteria that could be used to judge the success of the TPPs. The paper describes the policy objectives associated with total purchasing and it examines the small amount of literature that has considered what the outcome measures of policy innovations, such as total purchasing, should be. It then describes and compares the views of stakeholders within the TPPs before looking at the different success criteria formulated inductively by the national evaluation. Three sets of success criteria developed by the evaluation team allow comparison of TPPs' progress. Finally, the main findings are discussed with special reference to their implications for the evaluation of the various policy initiatives currently being introduced into the NHS such as the new Primary Care Groups (Secretary of State for England,1997).

1 Introduction

Total purchasing (TP) is the extension of standard fundholding whereby a general practice, or consortium of practices, are delegated a budget by their local health authority (HA) to purchase potentially all hospital and community health services not included in standard fundholding. Thus, TP allows general practices to extend the range of services they can buy to include non-elective services such as emergency inpatient care, accident and emergency services, inpatient services for people with severe mental illness, maternity services, community care, palliative care, regional specialisms (such as dialysis) and health promotion.

Total purchasing was officially introduced following an Executive Letter in October 1994 entitled *Developing NHS Purchasing and GP Fundholding* (NHSE, 1994). In a section on expanding general practice fundholding, the NHSE explained that 'extensive consultation' over the summer of 1994 had 'underlined the benefits of increasing both the numbers of general practitioners involved in fundholding and the range of goods and services they can buy'. Total purchasing was promoted on these grounds and introduced on a pilot basis.

It is important to emphasise that total purchasing differs from standard fundholding in many respects. Most importantly, the purchasing budget devolved to total purchasers remains the legal responsibility of its HA. It is thus a delegated budget rather than a devolved one. As a consequence, total purchasers have a direct link with HA and, officially, the total purchasing pilots (TPPs) are subcommittees of the HA, and therefore a component part of the larger body. Total purchasing is thus a hybrid of, on the one hand, GP-led purchasing in the fundholding tradition and, on the other, a form of locality commissioning. Moreover, since TPPs developed through local negotiation with their host HAs, many different styles of pilots have resulted (Mays and Dixon, 1997; Mays *et al*, 1998).

Much of the impetus for TP had come from the experience of four pioneering fundholding projects (Bromsgrove, Berkshire, Runcorn and Worth Valley) that had expanded the scope of their purchasing through local negotiation with their respective HAs. In April 1995, these pioneers were joined by 53 new projects from England and Scotland that had volunteered to take part in the scheme and, subsequently, by a 'second-wave' of volunteer pilots in April 1996.

Unlike standard fundholding, the TPPs were to be evaluated 'to identify the most appropriate models' (NHSE, 1994) and, following a tendering exercise, this evaluation was to be undertaken by a consortium of research institutions led by the King's Fund Policy Institute.

2 Aims of the Paper

There were no set criteria against which to assess the merits of the various ways in which total purchasing developed nationally. An important part of the first year's evaluation therefore was an attempt to derive success criteria which would enable comparison between all the TPPs. This paper is the result of these deliberations and examines the different criteria that could be used to judge the success of the TPPs.

First, the paper describes the policy objectives associated with total purchasing and the nature of the national evaluation of the pilot projects. It then examines the small amount of literature that has considered what the outcome measures of policy innovations, such as total purchasing, should be. Following this, the paper describes and compares the views of stakeholders within the TPPs of how success should be defined and measured. Finally attention is turned to the different success criteria formulated inductively by the national evaluation.

The policy objectives of total purchasing

The policy documentation from the NHSE shows a distinct absence of any detailed description of the aims and objectives of the TPPs. Indeed, it was the policy of the NHSE to have no formal blueprint for total purchasing, instead, making a virtue of the fact that a 'hands-off' approach would promote and encourage innovation. Such an approach, however, clearly made evaluation difficult since no objective success criteria were stipulated.

To determine the success criteria to be used to assess total purchasing, the evaluation team was reliant on a research brief, prepared by the Department of Health Research and Development Division, for an indication of the effects which the architects of TP believed that the TPPs might produce (Department of Health, 1995). Since the brief described TP as an extension of fundholding, the implication was that the Department and the NHSE expected similar consequences to standard fundholding and were looking with particular interest at those types of pilots which appeared most cost-effective and which delivered benefits to patients.

The focus on cost-effectiveness and patient benefits would suggest that a successful pilot would be able to demonstrate one or both of the following:

- that it had been able to purchase better services at the same costs as its HA;
- that it had delivered services of equal quality but at lower cost to that purchased by its HA.

The implicit link between total purchasing and fundholding would suggest that TP was introduced with similar objectives. Fundholding had been introduced as part of a wide ranging set of NHS reforms with the central aim of improving efficiency and creating value for money through a process of devolved responsibility, greater competition and enhanced consumerism, to be nurtured by the NHS internal market. By giving GP practices the option of holding a budget, it was hoped that efficiency, quality and choice would be improved. Specifically, the architects of fundholding expected a number of tangible benefits to emerge: reduced inefficiencies in provider organisations; better quality in secondary care provision; downward pressure on drug costs and unnecessary referrals; enhanced practice-based care and facilities; and promotion of greater choice and responsiveness to local health needs (Goodwin, 1996). Despite the fact that the evidence available to assess the true impact of fundholding shows a somewhat eclectic mix of success and failure (Audit Commission, 1995; Coulter, 1995; Glennester *et al*, 1994; Goodwin, 1996), total purchasing was introduced to extend the perceived benefits of a GP budget-holding system to a wider number of services. However, unlike fundholding and the NHS reforms more generally, the objectives of total purchasing are far more vague and in many respects different. Total purchasing is not a simple extension of fundholding and, in particular, the original version of total purchasing as a budget-holding extension of fundholding has been diluted over time through the increasing prevalence and favouritism from the centre towards non-budget forms of commissioning to which total purchasing could be easily transformed.

Since the TPPs were fundamentally new organisations with no formal blueprint for development there has been great scope for innovation and a heterogeneity of approaches. This was confirmed in the initial analysis of TPPs which showed that many different types of TPPs were developing along a spectrum from single practice 'fundholding' models to multi-practice locality commissioning-type approaches (Mays and Dixon, 1997; Mays *et al*, 1998). Moreover, it was clear that the aims and objectives of the projects within total purchasing varied in scope (TP-NET, 1997). The challenge for the national evaluation team has been to derive success criteria that are relevant to all the TPPs.

3 The Evaluation of Total Purchasing

The evaluation (TP-NET, 1997) addressed three main questions:

- what factors are associated with the successful set up and operation of total purchasing?
- what are the costs and effectiveness of total purchasing?
- what are the benefits to patients?

A range of interrelated components, some of which are being carried out at all projects, others at sub-samples of TPPs, were used to address these questions. The components of the evaluation include:

- *The set-up and operation of total purchasing (process evaluation)* - undertaken at all projects, this part of the evaluation describes how TP has been implemented and is being undertaken through a combination of face-to-face interviews, diary cards, postal questionnaires, telephone interviews, analysis of routine data and the analysis of documents.
- *Transaction costs* - at a sample of TPPs this component is describing and quantifying incremental transaction costs associated with TPP.
- *Activity changes* - this component is examining the changes in activity (e.g. patient episodes, lengths of stay, levels of prescribing costs etc) before and after the advent of TP and comparing TP with SFH and non-SFH populations at all TPPs through the use of routine NHS data.
- *Purchaser efficiency and costs of services* - the TPPs' ability to negotiate lower cost, higher volume or better quality services are being compared with that of the local HA, using routine activity and cost data at all TPPs.
- *Benefits and costs to patients and their experiences of specific services* - four separate sub-studies are examining the patterns of care, service costs and patient's reactions to specific services purchased by the TPPs in comparison with the same services purchased by HAs. The particular emphasis here is an attempt to assess directly the benefits to patients of TP. The four services are community and continuing care for people with complex needs, accident and emergency services together with emergency admissions, services for people with serious mental health problems and maternity care.

The development of total purchasing by the pilot projects was diverse and the aims and objectives of TPPs also differ. Consequently much of the evaluation's first year therefore set out to answer the questions 'What is total purchasing?' (Mays *et al*, 1997) and 'How do we assess the success of TP?'.

4 Outcome Measures of Policy Innovations

The evaluation of health policy innovations is complex and it is often argued that it is not feasible to evaluate the impact of such initiatives because of the length of time that needs to elapse before measurable change occurs (Bartlett and Le Grand, 1993; Robinson and Le Grand, 1994). The official endorsement of policy evaluation has been seen as an unnecessary obstacle to policy implementation, as well as a sign of weakness and uncertainty. This was certainly the case when a range of policy initiatives was introduced following the publication *Working for Patients* in 1989 (Department of Health, 1989). Kenneth Clarke, the former Secretary of State for Health, stated this view quite explicitly in his responses to the social services select committee; 'I'm not prepared to say the whole idea of changing the way the NHS is managed and money distributed is some sort of experiment and we're going to spend several years looking at it' (Davies, 1989). Against this background of official opposition to evaluation the King's Fund launched a programme of research evaluating various aspects of the reforms (Robinson and Le Grand,1994).

The identification of appropriate criteria for measuring success is difficult and three particular sets of issues can be identified (Bartlett and Le Grand, 1993). The first set of issues relate to the scope of the evaluation. Policy initiatives are often characterised by broad aims and objectives; what outcomes should or should not be included in an evaluation therefore needs to be considered. The second set of issues relate to definition. Having identified a set of outcomes relevant to the policy initiative, how should the evaluation criteria be specified or operationalised? This is a crucial question which shall be considered in more detail later in this paper. The third set of issues relate to ownership. Whose criteria are being used in the evaluation? Again this is a crucial question since it has been demonstrated that different groups of stakeholders tend to emphasise different outcome criteria (Palfrey and Thomas, 1996). Unfortunately Bartlett and Le Grand (1993) did not attempt to resolve these issues but took a more pragmatic approach and proposed their own set of criteria against which to assess the empirical consequences of quasi-markets.

Further problems facing policy evaluation relate to the policy context itself. Often, innovative policies are introduced alongside a series of other national and local policy initiatives. The complexity of some innovations also makes the evaluation process more problematic and this is further complicated by the trend towards central policy decision-making being subject to local interpretation and implementation ie simultaneous devolution and centralisation. Both of these problems - context and complexity - are features of total purchasing.

The problems identified do not mean that one should not attempt to evaluate policy initiatives. They mean that those involved in the evaluation process should be explicit about what criteria they have included (and excluded) and why, and how they have attempted to deal with the problems of timing, complexity and of the policy context.

A number of possible sets of success criteria that could be adapted for total purchasing have been identified. Bartlett and Le Grand (1993), for example, identified a set of broad criteria that, they argue, capture the principal interests of all those concerned with the progress of the NHS reforms:

- efficiency
- user choice
- responsiveness to clients
- equity

Having specified their success criteria they go on to identify *conditions for success*, such as market structure, information, transaction costs, motivation and cream skimming. Thus it is possible, and indeed necessary, to distinguish between the criteria for success and criteria of success. Drawing on this and other studies (Robinson and Le Grand, 1994; Maxwell 1984), Thomas and Palfrey (1996) produced a list of ten possible success criteria reproduced below:

- effectiveness
- efficiency
- equity
- acceptability to patients
- accessibility
- appropriateness
- accountability
- ethical considerations about the design of the study
- responsiveness
- choice

Thomas and Palfrey believe that these ten criteria will attract different weightings, or degrees of relevance, depending on particular policy settings.

Other sets of criteria specifically relating to total purchasing have been produced by 'opinion leaders'. Explicit statements concerning specific success indicators of the total purchasing pilots have been made by proponents of fundholding, characterised by their almost unreserved enthusiasm for fundholding and the expansion of the model to total purchasing. Carruthers, for example, lists twelve indicators of success for total purchasing (Carruthers, 1994). These indicators range from very broad and vague statements such as 'more and better quality care' and 'improved choice for patients' to more specific statements that could be subject to empirical investigation, such as 'changes in referral and treatment patterns planned and based on audit'. However, in contrast to Bartlett and Le Grand (1993), no attempt is made by Carruthers, or for that matter the NHS Executive, to specify the mechanisms for achieving success or the conditions that must be in place to realise specific aims.

There are crucial conceptual and empirical issues that need to be addressed in devising meaningful success criteria for total purchasing. These issues are not always addressed explicitly by policy analysts and researchers and are often ignored altogether by 'opinion leaders'. Conceptual issues relate to defining the nature of policy initiatives and identifying what they are expected to achieve. In other words, what is the specific intervention and what is the anticipated outcome? The anticipated outcomes of policy initiatives are often stated in broad and vague terms; thus the various criteria identified above referred to effectiveness, efficiency, quality and so on. These are all global concepts that require further definition and clarification to become meaningful and, indeed, to become measurable. The problem of measurement is not insurmountable but is considerable. The broad concepts, once defined in relation to the policy initiative, then need to be 'operationalised' or translated into measurements or empirical indicators. The policy context in relation to total purchasing, described earlier, highlights the limitations and at times unrealistic nature of the outcome criteria identified in this section. Complexity of the policy context, timescale and the problem of attribution are such that serious questions are raised about appropriateness of outcome based evaluation of complex policy situations. Pre-determined success criteria have not been and cannot be set in a dynamic and evolving policy context. The strength of the ongoing process evaluation of TP lies in its sensitivity to local implementation and the views of the various stakeholders involved in the process.

5 The Views of Stakeholders Within the TPPs

As part of the process of evaluating TP, semi-structured interviews with stakeholders from each TPP site were carried out. A range of participants were interviewed but, for the purpose of this paper, analysis of the responses of three main stakeholders is included; the lead general practitioner (GP), the site manager and the HA lead. These were the key personnel involved in the development of TP at each pilot site and may be considered to be those with the main responsibility for the success of TP. Interviews were undertaken during the preparatory year and during the first 'live' year of TP. Sections of the interviews relating to the stakeholders' experiences were analysed to elicit concepts of success in TP.

Many of the lead GPs were able to point to specific ways in which their TPP had been successful, but others gave a more circumscribed response, indicating their awareness that measuring success is not easy, and that much may not be quantifiable. A comment, typical of the response from several GPs was; 'It could take many years to assess whether organisational change could improve the health of the population'.

This indicates that GPs felt that the more concrete manifestations of success, such as specific service changes leading to better patient services, could take some time to become apparent. The success criteria most frequently identified by the lead GPs centred around relationships between the main players and stakeholders and referred to cohesion and working together amongst GPs, in particular getting on with those in other practices; working with HAs and providers; developing a shared philosophy and understanding of TP. It was apparent that these less tangible issues were felt to be very important and necessary for TP to move ahead and for its success. Less frequently, GPs cited specific successes that they had been able to achieve, such as having achieved their stated aims, completed budget setting, made major service changes, improved clinical care, developed good documentation and having set up good organisational arrangements between practices.

In contrast to GPs, the most frequent success criteria mentioned by site managers related to the extent to which service provision had changed as a result of TP. Site managers also identified specific success criteria, most frequently those relating to achieving the TPP's own stated aims and to issues relating to purchasing, contracting and commissioning. In the case of the latter, success was measured by the degree to which the TPP had acted independently or by the extent to which GPs were strategically involved in purchasing. Specific success criteria that were mentioned less frequently by site managers related to costs and savings achieved by the TPP and to the degree to which information was being gathered and information systems

were in place. Like the lead GPs, many of the site managers identified success criteria that related to the development of relationships both within and outwith the TPP. However, unlike the GPs, the success criteria identified by the site managers did not have such an overwhelming focus on the development of relationships. In contrast the success criteria identified by site managers were more balanced in their focus between the practical aspects of total purchasing, the achievement of changed service provision and the less tangible aspects of organisational development.

The HA stakeholders, like the site managers, acknowledged that it would take time to determine the overall success of TPP, but, in contrast to the lead GPs and the site managers, they put a greater emphasis on the timespan of the pilot itself and felt that success should involve the TPPs having something positive to show at the end of the pilot period. This could relate to physical changes in service provision or to developments that facilitate future change. Like the site managers and the lead GPs, the success criteria relating to developments facilitating future change identified by the HAs related to the setting up of appropriate IT systems, the holding of a 'live' budget, and setting up appropriate management systems. Similarly, the HAs saw the development of improved relationships and a greater understanding of the NHS as being a measure of success. A major theme, that was identified to the greatest extent by HA stakeholders, was that of success being measured in terms of the achievement of an appropriate balance between operational and strategic level approaches within a TPP site. However, where the appropriate balance was thought to lie differed between HAs, with some HAs viewing success as being the GPs taking on a wider strategic viewpoint, leaving operational matters for project and fund managers; the opposing view being that TPPs as a whole should concentrate on the operational aspects and on making clinical and organisational improvements within the system, on a level at which the HA find it difficult to engage. The HA stakeholders also identified a wider dimension of success in which a successful TPP would facilitate the development of local purchasing in general, by sharing knowledge and experience, and by acting as a stimulus for other groups or purchasing models, such as locality commissioning.

A common theme running through all three stakeholder perspectives was that it was too early, after one year, to determine the overall success of TP. They all, implicitly, held the view that the more concrete manifestations of success (such as specific service changes leading to improved patient services) would take some time to become apparent. However, they also felt that success could be described in terms of factors relating to the climate as a necessary pre-condition, before the more measurable aspects can be achieved. The inherent problem within

this is that TPP sites may set up the pre-conditions of success, but these in themselves may not lead to service changes or patient benefit.

Although all stakeholders referred to specific achievements, such as physical changes in service provision, as being criteria defining success, there was a greater emphasis, among them all, on success being measured in terms of factors relating to facilitation of future change. The development of good internal and external relationships was a feature within all the stakeholders' approaches to creating successful pre-conditions for future change. This included the development of new relationships, particularly those where there had been little contact historically, such as relations between the HA and GPs, and relations between individual practices. However stakeholders placed a different emphasis on which relationships they felt were important to the success of TP. The HA stakeholders identified relations among GPs themselves and relations between GPs and clinicians as being significant, and viewed themselves as having a more distant role in the TP process. Cohesion and working together amongst GPs and between practices were important to the lead GPs, but GPs also felt that developing relations with HAs was an important criterion of success. Site managers mentioned the greatest variety of relationships outwith TP, and included developing relationships with social services and patient groups and representatives that were not specifically mentioned by other stakeholders.

The different approaches described above highlight the need for clarity in any development of success criteria and emphasise the need to recognise that success can come in different forms; what is seen as success for an individual TPP site may not be seen as a success for TP as a policy. This is seen most clearly from contradictions raised within some HA stakeholder interviews, where the data gave specific detail about the indicators of failure, rather than success, of TP, and some definitions of failure contradicted the criteria given for measuring success. A definition of failure was that of a TPP trying to work in isolation from the HA, which contradicted the views of HAs who saw success as their TPPs working as independently as possible in order to stimulate innovation and develop new purchasing methods. Increasing inequity was also considered to be a failure by some HAs, the danger being that TPPs may provide better services at the expense patients elsewhere. The contradiction being that changing services, and therefore implicitly providing better services for patients, was the raison d'être of TP. This raises the conundrum of how to pilot innovative schemes with the aim of providing better services for patients, without some type of inequity developing at the same time.

6 Success Criteria Formulated by the National Evaluation

The task of identifying a set of success criteria applicable to ALL total purchasing projects poses a particular challenge given their characteristic heterogeneity. How should success be defined when the participating projects are multi-dimensional with no single definitive approach to total purchasing and the participants within the projects have such different ideas as to what constitutes a successful project? Such a question is difficult to answer, not least because the concept of total purchasing itself is interpreted in various ways and eludes a precise definition.

Returning to the objectives of the evaluation, three main questions were asked:

- what factors are associated with the successful set-up and operation of total purchasing?;
- what are the costs and effectiveness of total purchasing?;
- what are the benefits to patients through total purchasing?

The last two questions concerning the outcomes of total purchasing - the costs and effectiveness of total purchasing compared to alternatives and benefits to patients - cannot be answered in the early stages of the evaluation since the evidence for changes in outcomes will only be apparent in the longer term. As a result, the success criteria developed by the national evaluation to examine and compare the TPPs' performance in the first 'live' year of total purchasing are necessarily limited to measures of progress in setting-up and operating as a TPP and the information used to make these judgements has been gained from the 'process evaluation' component of the study. The other components in the study will be the main methods by which outcomes will be measured and some of their interim findings are contained in sister working papers (Bevan, 1997; Gask *et al* 1997; Myles *et al*, 1998; Raftery and McLeod, 1997; Street *et al*, 1997; Wyke *et al*, 1998).

Having determined that the success of the TPPs should be examined using operational criteria, the next step in the process was to devise criteria which could make useful comparisons between the diverse set of projects. Three sets of success criteria measurements were developed: first, the extent to which total purchasers met basic development criteria; second, the ability of total purchasers to achieve their own objectives; and third, the extent to which total purchasers had influenced TP-related services.

The extent to which total purchasers met basic development criteria

A most objective, and basic, measure of achievement can be made in terms of how the projects have progressed in becoming fully-fledged TPPs. In other words, every successful TPP should have developed, during the set-up and operation stage of their development, a number of generic characteristics that classify them as operational total purchasers.

The evaluation has used six process indicators to show how successful a pilot project has been in setting up and operating like a functioning total purchaser. The examination of these process indicators is considered in more detail in a separate working paper on achievements (Mays *et al*, 1998) and will not be discussed here, but will be described briefly:

(1) has survived as a project in its original form;
(2) has been able to purchase services through contracting with providers either directly or through a co-purchasing arrangement with the HA;
(3) has been able to change service provision;
(4) has enhanced the primary care function, including some shift of activity or resources from the secondary to the primary sector;
(5) has stayed within its budget and/or made savings and;
(6) has developed external links with providers and social services.

Examples of the above indicators can be found in other reports in this series, for example the study by Mays and colleagues (Mays *et al*, 1998) showed that some TPPs have not survived in their original form (indicator 1). This study also showed that many TPPs had been successful in changing service provision (indicator 3) and had also shifted the location of care (indicator 4). Robinson and colleagues (Robinson *et al*, 1998) were able to show that 28 (62%) of TPPs had contracted independently (indicator 2). Baxter and colleagues (Baxter *et al*, 1998) found that whilst many projects had successfully managed a budget, multi-practice projects found this more difficult than single practice ones (indicator 5). External links with providers and social services, as exemplified in the maternity and community care studies respectively, (Wykes S *et al*, 1998; Myles S *et al*, 1998), varied considerably but are pre-requisites for successful purchasing (indicator 6).

The ability of total purchasers to achieve their own objectives

In an attempt to move forward from basic development indicators, but still with the problem of comparability in mind, the ability of projects to achieve their own objectives was a second

approach taken to indicate the success of the projects after their first 'live' year. During the process evaluation in the TPPs' set-up stage all projects were asked to set out their main objectives for the first 'live' year of total purchasing. Thus, this criterion measures the TPPs progress towards achieving their own objectives and reflects on general abilities and rate of progress, without taking into consideration the level of ambition or the potential scope of the project. This approach, rather than one based on the overall impact of a TPP, has the advantage of comparing projects which are so radically different in terms of their size, scope and location.

The problem with this approach is that scope and ambition not included. It could be argued that a particularly ambitious project that had achieved only one or two of its objectives is more successful than a project which has achieved all of its main objectives, but which has had less overall impact. Moreover, the ability of the TPPs to achieve their main objectives is further influenced by the local context within which the TPP operates.

The extent to which total purchasers had influenced TP-related services

A final criterion for success used in the evaluation has been the extent to which the TPPs have influenced TP-related services. This measure takes more account of the scope of the projects since TP-related services refer to those services, such as maternity and mental health care, which could not have been purchased through standard fundholding. This criterion, therefore, plays down the success of those TPPs who had managed to fulfil their own objectives, but not in TP-related areas. As the working paper on achievements shows (Mays *et al*, 1998), not all projects used their total purchasing status to purchase or commission services in TP-related service areas.

The three sets of success criteria developed by the evaluation team allow comparison of TPPs' progress and, to some extent, a method for including the scope of their achievements. What the success criteria have not been able to do is include important contextual information (such as the nature of the provider market or local political cultures) which may be important factors influencing success. Moreover, the success criteria are measures of operational progress and do not assess outcomes in terms of the effectiveness or usefulness of the changes made. Nevertheless, the criteria developed are objective and provide a useful approach to compare the development of these heterogeneous organisations in their first 'live' year of total purchasing.

7 Discussion and Implications

One of the most important findings from this work is that, in the short time period of the pilot - three years - achieving and measuring 'success' are problematic and are perceived differently by the various stakeholders. The overall aim of organisational change should be to improve patient care, but it was recognised by many of the interviewees that such an aim may take many years to be realised and may, in any case, be difficult to measure. Success, at this stage in the evaluation, must, therefore, be measured in terms of intermediate outcomes, rather than as specific service changes of direct benefit to patients.

There are two main issues associated with defining success - firstly, how is success judged, and secondly, what factors are necessary for an innovation to be successful? In manufacturing industry success can be judged by profit for the company if a product sells, but for service organisations, even defining the 'product' may be problematic. The second issue follows from the first. Having defined the nature of success, it is possible to look at the factors which are necessary for an innovation to be successful. The latter is what has been happening in the TPP projects.

Many stakeholders cited factors which they regarded as markers of success but which we have called 'preconditions' for success to be achieved. These generally include the more intangible factors such as creating the climate in which total purchasing can become acceptable and in which it can be seen to be achieving change or at least to have the potential for achieving change. Many sites had made a start on the preconditions but lack of time had meant that more concrete manifestations of success, such as changed and improved patient services, were not yet obvious.

Stakeholder views differed and were heavily dependent on their own specific standpoint. The GPs pointed to the advances they had been able to make in terms of improved relations with the HAs and trusts, organisations with which they had previously had little contact. Similarly, getting on with GPs from other practices was also important to GPs since, again, this is something of which they had, in the past, had little need for. There was the recognition that these were important steps in achieving change and indeed in setting the stage for a move towards other organisational changes such as Primary Care Groups and taking part in PCAPS, all of which will succeed only where the local climate is appropriate. Site managers shared these concerns but tended to place greater emphasis on the concrete achievements, as would be expected from those whose role, in part, is to improve relations between the main players and develop an environment in which total purchasing can be seen as having succeeded. After

all, their job could be coming to an end and they are likely to be looking elsewhere and wishing to show that their management had resulted in some tangible benefits. HAs, on the other hand, had a different agenda. They were keen to show that specific achievements had taken place and to have something positive to show for TPP at the end of the pilot period since, technically, HAs are the budget holders. Despite these different standpoints, common themes have emerged relating to whether sites are perceived, even at this early stage, as having been successful.

Since there were no pre-stated aims and no agenda was set centrally for sites to judge their success, they had to set these up for themselves. There is, therefore, considerable disparity in the goals which the individual sites had set for themselves. Some had set out to achieve tasks which were seen as difficult at the outset, but with the feeling that TPP represented an opportunity for experimentation and innovation which should be grasped with both hands. Others had taken a more cautious approach, wanting to create a climate in which TPP might eventually flourish, but placing the emphasis on getting things right first, such as information and internal relationships. There is, therefore, a spectrum from those sites which set themselves modest aims and achieved relatively little to those which more ambitious aims and achieved most of these. Generally, concrete changes, in terms of service changes leading directly to improved services for patients, were few and far between. Even where sites had set out with such ends in mind, many had succeeded 'only' in achieving some of the pre-conditions along the way. Others had set out only to accomplish some of the precursors and had succeeded in this. It is, therefore, only possible to state the TPP had achieved some small success and had acted as a vehicle for change, the eventual outcomes of which may be far into the future and not appreciated for many years. TPP has allowed the various stakeholders to think the unthinkable and has the potential to achieve far reaching changes, but not yet. Even those sites which had achieved little in their first live year may be on the brink of great things in their second year of purchasing, having set the scene in year one.

There are lessons to be learned from this study for the additional organisational changes which are shortly to be introduced within the NHS, particularly for the various types of primary care act pilot sites (PCAPS), Primary Care Groups and Health Action Zones. If goals by which these organisations can judge their success are not made explicit from the outset, it would be unacceptable for them not to draw upon the findings of the TPP evaluation. These new organisations will be at least as diverse, perhaps more so, than the TPPs. They should therefore take note of the findings of this study - that there are a number of ways by which success can be defined, and that in many instances much effort has to be put into developing the appropriate pre-conditions which allow organisations to flourish and hence achieve

change. Some of these have been identified (Mays *et al*, 1998). It was apparent that it took time to plan and implement changes through devolved purchasing and that projects were able to move at different speeds because of a mixture of internal and external circumstances. The participants stressed that the first eighteen months had been spent establishing the pre-conditions for successful purchasing, such as improving information, improving understanding between the practices, the HA and local providers, and improving relationships between all the local participants. The new organisations can use these findings to start at a higher level than has been possible for many of the TPPs, so speeding up the process of achieving change. Indeed, the NHSE has produced a set of guidelines for local evaluation of PCAPS, but this falls short in setting out how to measure success (NHSE, 1997).

As noted earlier, specific development indicators have been drawn up which can be used by TPPs to give an indication of success (Mays *et al*, 1998). Clearly many of these are specific to TPP, but can be drawn upon by the other new organisations as an aid to devising their own specific criteria. The pre-conditions for success are more likely to be applicable to most types of organisations i.e. the importance of climate, relationships, and working together.

Assessing the implementation and development of TP has relied heavily on the opinions of comparatively few stakeholders in each site, but it has enabled an overall impression to be obtained of the ways in which the various stakeholders view the changes. As well as collecting much qualitative data, this has then been summarised and interpreted by the interviewers and condensed to produce a final commentary. Clearly such impressions can only give a broad summary of how TPP has evolved and what it has managed to achieve in its short existence. However, the methods and findings from the early stages of our evaluation allow us to measure and determine pre-conditions associated with the successful implementation and development of total purchasing. As time, and the evaluation, progress we will be able to examine the outcomes of this new model of primary care purchasing.

References

Audit Commission. *Briefing on GP fundholding.* London: HMSO, 1995.

Bartlett W, Le Grand J. The Theory of Quasi-Markets. In Le Grand J, Bartlett W, eds. *Quasi-Markets and Social Policy.* London: MacMillan, 1993.

Baxter K, Bachmann M, Bevan G. *Survey of Budgetary and Risk Management.* London: King's Fund Publishing, 1998.

Bevan G. *Resource Allocation Within Health Authorities: Lessons from the Total Purchasing Pilots.* London: King's Fund, 1997.

Carruthers I. Total fundholding in the mainstream of the NHS. *Primary Care Management* 1994; **11**: 7-9.

Coulter A. Evaluating general practice fundholding in the United Kingdom. *European Journal of Public Health* 1995; **5**: 233-239.

Davies P. Ready for take off without the pilots. *The Health Services Journal* 1989: 1 June.

Department of Health. *Working for Patients.* Cm. 555. London: HMSO, 1989.

Department of Health Research and Development Division. *National Evaluation of Total Purchasing Pilot Schemes: Research Brief.* London: Department of Health, 1995.

Gask L, Lee J, Donnan S, Roland M. *Total purchasing and extended fundholding of mental health services.* London: King's Fund Publishing, 1998.

Glennerster H, Cohen A, Bovell V. *Implementing General Practitioner Fundholding: Wild card or Winning Hand?.* Buckingham: Open University Press, 1994.

Goodwin N. GP Fundholding: A Review of the Evidence. In Harrison A, ed. *Health Care UK 1995 6*, London: King's Fund Publishing, 1996, p 116-130.

Maxwell RJ. Quality Assessments in Health. *BMJ* 1984; **288**:166-203.

Mays N, Dixon J. *Purchaser Plurality in UK Health Care: Is a Consensus Emerging and Is It The Right One?* London: King's Fund Publishing, 1996.

Mays N, Goodwin N, Bevan G, Wyke S on behalf of the Total Purchasing National Evaluation Team. What is total purchasing? *BMJ* 1997; **315**: 652-5.

Mays N, Goodwin N, Malbon G, Leese B, Mahon A, Wyke S. *What were the achievements of total purchasing pilots in the first year and how can they be explained?* London: King's Fund Publishing, 1998.

Mays N, Goodwin N, Bevan G, Wyke S on behalf of the Total Purchasing National Evaluation Team. *A profile of total purchasing pilot projects in England and Scotland.* London: King's Fund Publishing, 1997.

Myles S, Wyke S, Popay J, Scott J, Campbell A, Girling J. *Total purchasing and community and continuing care: lessons for future policy developments in the NHS.* London: King's Fund Publishing, 1998.

NHS Executive. *Developing NHS Purchasing and General Practitioner Fundholding.* (EL(94)79). Leeds: NHS Executive, 1994.

NHS Executive. *Personal Medical Services Pilots Under the NHS (Primary Care) Act 1997: A Guide to Local Evaluation.* Leeds: NHS Executive, 1997.

Raftery J, McLeod H. *TPPs and Routine Data: Progress Report and Preliminary Analysis.* Unpublished paper. Birmingham: Health Economics Faculty, HSMC, University of Birmingham, 1997.

Robinson R, Le Grand J, eds. *Evaluating the NHS Reforms.* London: King's Fund Institute, 1994.

Robinson R, Robison J, Raftery J. *Contracting by total purchasing pilot sites.* London, King's Fund Publishing, 1998.

Secretary of State for England. *The New NHS.* London: The Stationary Office, 1997.

Posnett J, Goodwin N, Killoran A, Malbon G, Mays N, Place M, Street A. *The transaction costs of total purchasing.* London: King's Fund Publishing, 1998.

Thomas P, Palfrey C. Evaluation: Stakeholder-Criteria. *Social Policy and Administration* 1996; **30:2**: 125-142.

Wyke S, Hewison J, Piercy J, Posnett J, Macleod L, Page L, Young G. *National evaluation of general practice-based purchasing of maternity care: preliminary findings.* London: King's Fund Publishing, 1998.

Total Purchasing National Evaluation Team (TP-NET)

The evaluation is led by Nicholas Mays, Director of Health Services Research at the King's Fund, London. The different consortium members are listed below, together with their research responsibilities.

KING'S FUND
11–13 Cavendish Square, London, W1M 0AN
T: 0171 307 2400 F: 0171 307 2807

Lead: Nicholas Mays
Other members: Nick Goodwin, Gill Malbon, Julian Le Grand, Jennifer Dixon, Amanda Killoran, Jo-Ann Mulligan

Project Responsibilities: Hertford, Hemel Hempstead, Hillingdon, New River, St Albans, Stevenage, Attleborough, South Bucks, Belper, Keyworth, Long Eaton, Melton Mowbray, Wakefield.

Other Main Responsibilities: Process evaluation co-ordination (Mays, Goodwin); A&E services and emergency admissions (Dixon, Mays, Mulligan); monitoring at all TPPs (Mays and Malbon); case studies (Mays, Goodwin, Killoran, Malbon).

NATIONAL PRIMARY CARE R&D CENTRE
Manchester: University of Manchester, 5th Floor, Williamson Building, Oxford Road, Manchester, M13 9PL
T: 0161 275 7600 F: 0161 275 7601
Salford: PHRRC, University of Salford, Davenport House, 4th Floor, Hulme Place, The Crescent, Salford, M5 4QA
T: 0161 743 0023 F: 0161 743 1173
York: YHEC, University of York, YO15 4DD
T: 01904 433620 F: 01904 433628
CHE, University of York, York, YO1 5DD
T: 01904 433669 F: 01904 433644

Leads: Brenda Leese (Manchester and CHE), Linda Gask (Manchester), Jennie Popay (Salford), John Posnett (YHEC)
Other members: Ann Mahon, Martin Roland, Stuart Donnan, John Lee, Andrew Street

Project Responsibilities: High Peak, North Lincolnshire, Rotherham, Sheffield South, Ellesmere Port, Knutsford, Liverpool Neighbourhood, Newton le Willows, Wilmslow, Ribblesdale, Southbank, North Bradford, York.

Other Main Responsibilities: Transaction costs (Posnett and Street); service provision for the seriously mentally ill (Gask, Roland, Donnan and Lee); service provision for people with complex needs for community care services (Popay); relations with health authorities (Leese and Mahon); maternity (Posnett).

DEPARTMENT OF SOCIAL MEDICINE, UNIVERSITY OF BRISTOL
Canynge Hall, Whiteladies Road, Bristol, BS8 2PR
T: 0117 928 7348 F: 0117 928 7339

Lead: Kate Baxter
Other members: Max Bachmann, Helen Stoddart

Project Responsibilities: Bewdley, Birmingham, Bridgnorth, Coventry, Solihull, Worcester, Saltash, South West Devon, Thatcham.

Other Main Responsibilities: Budgetary management (Baxter); risk management (Bachmann); use of evidence in purchasing (Stoddart); case studies (Baxter).

DEPARTMENT OF GENERAL PRACTICE, UNIVERSITY OF EDINBURGH
20 West Richmond Street, Edinburgh, EH8 9DX
T: 0131 650 2680 F: 0131 650 2681

Lead: Sally Wyke
Other members: Judith Scott, John Howie, Susan Myles

Project Responsibilities: Durham, Newcastle, Tynedale, Aberdeen West, Ardersier & Nairn, Grampian Counties, Lothian, Strathkelvin

Other Main Responsibilities: Maternity (Wyke); monitoring of participants' views (Wyke); prescribing (Howie); community care (Wyke and Scott).

INSTITUTE FOR HEALTH POLICY STUDIES, UNIVERSITY OF SOUTHAMPTON
129 University Road, Highfield, Southampton, SO17 1BJ
T: 01703 593176 F: 01703 593177

Lead: Ray Robinson
Other members: Philippa Hayter, Judy Robison, David Evans

Project Responsibilities: Dorset, Romsey, Trowbridge Bath & Frome, Winchester, Bexhill, East Grinstead, Epsom, Kingston & Richmond, Merton Sutton & Wandsworth, West Byfleet.

Other Main Responsibilities: Contracting methods (Robinson, Raftery, HSMC and Robison; case studies (Evans).

HEALTH ECONOMICS FACILITY, HSMC, UNIVERSITY OF BIRMINGHAM
40 Edgbaston Park Road, Birmingham, B15 2RT
T: 0121 414 6215 F: 0121 414 7051

Lead: James Raftery
Other member: Hugh McLeod

Main Responsibilities: Activity changes in inpatient services; contracting methods (with Robinson and Robison, IHPS); service costs and purchaser efficiency (with Le Grand).

HEALTH SERVICES RESEARCH UNIT, LONDON SCHOOL OF HYGIENE AND TROPICAL MEDICINE
Keppel Street, London, WC1E 7HT
T: 0171 927 2231 F: 0171 580 8183

Lead: Colin Sanderson with Jennifer Dixon, Nicholas Mays and Jo-Ann Mulligan (King's Fund), James Raftery (HSMC)
Other member: Peter Walls

Main Responsibility: A&E services and emergency admissions.

LSE HEALTH, LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE
Houghton Street, London, WC2A 2AE
T: 0171 955 7540 F: 0171 955 6803

Lead: Gwyn Bevan

Main Responsibilities: Resource allocation methods.

ISBN 1-85717-191-8