King's Fund

Intermediate Care

Approaches to Evaluation

Andrea Steiner, Barbara Vaughan and
Linda Hanford

Intermediate Care Series - Three

INTERMEDIATE CARE

Approaches to Evaluation

prepared by

Andrea Steiner PhD
Senior Lecturer in Gerontology and Health Policy
University of Southampton

Barbara Vaughan MSc, RGN
Programme Director
King's Fund

Linda Hanford MSc (A), BN
Project Manager, Intermediate Care
King's Fund

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

ISBN 1 85717 222 1

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

Further copies of this report can be obtained from the King's Fund bookshop.
Telephone 0171 307 2591

Table of Contents

List of Tables, Figures and Boxes

Acknowledgements

There are numerous people to whom we are grateful for their help with producing this report. First and foremost are the seminar participants, whose names and affiliations are listed in Appendix 4. It is their perceptions which have defined or framed the issues to be raised in the coming pages. The seminar speakers also deserve special mention. They are: Sue Dopson, Sarah Furlong, Alastair Gray, Peter Lees, and Bronagh Walsh. Andrew Lynch assisted with the preparation of this document, and many other people provided useful comments. We are particularly grateful to Janice Robinson, who also acted as workshop facilitator. Many other people, too numerous to name, have helped to inform our thinking and to them we offer our thanks.

Section One: Introduction

This report, which focuses on the evaluation of intermediate care services, is the third in a series of papers. Previous work has provided a conceptual framework to clarify understanding of intermediate care and examined issues related to implementation.[1,2]

Intermediate care is defined as a function concerned with transition from medical dependence to personal independence, focusing on restoration of self-care abilities. It is aimed at meeting the needs of those people who are physiologically stable or predictable following an acute episode of illness but who could improve the quality of their lives, increase their ability to live independently, and minimise their longer term dependence on health services through timely therapeutic input. Alternatively, it can provide care to prevent inappropriate admission to acute hospital for people who have a short term episode of ill health or a short-term social care need which does not require specialist medical intervention.

Although the concept is not a new one, interest in intermediate care services has grown rapidly over the last few years. In the last year alone, that interest has seemed to make another exponential leap. In the recent King's Fund report on health services in London, intermediate care was seen as an essential part of service provision.[3] Throughout the UK, there is now a range of different services specifically offered to meet intermediate care needs - encompassing nurse-led post-acute units in hospital, rapid response teams in the community, primary care-led initiatives using home, community hospital and nursing home beds to provide intermediate care services - and much of this is newly implemented. More than 40 new schemes, many of which could be defined as intermediate care, were supported through funds supplied to meet the anticipated winter beds crisis. Similarly, Anglia and Oxford Health Authority now hold a regional database of interested parties and/or initiatives in intermediate care numbering more than 800.

In many cases, however, implementation has occurred without pre-testing or with an absence of hard evidence that the new service will work. This is troubling, since resources are limited and patients (or clients) deserve the most effective, cost-efficient and acceptable services possible. Nonetheless, services often go forward without any

[1] Steiner, A. (1997) *Intermediate Care: A conceptual framework and review of the literature*, London: King's Fund.
[2] Steiner, A., Vaughan, B. (1997) *Intermediate Care: A discussion paper arising from the King's Fund seminar held on 30th October 1996,* London: King's Fund.
[3] King's Fund London Commission (1997) *Transforming Health in London*, London: King's Fund.

plan to evaluate. In other cases, there is a strong burden of proof placed upon the clinical staff who have persuaded managers to agree to an intermediate care scheme. Yet the form of evaluation demanded does not necessarily bring out what is best about the service being delivered. It does not always meet the information needs of those developing the service; nor, on the whole, are the evaluations themselves decently resourced.

This report is the product of an invited seminar convened by the King's Fund to discuss strategies for evaluation of intermediate care and to compare participants' experience with different approaches, evaluation designs, and measures. Participants were asked to identify key questions about intermediate care that they thought each of four types of evaluation - process, implementation, outcome and economic - was particularly well-equipped to answer. They were also asked what tools they used to assess each type; what drivers existed to conduct this sort of evaluation of intermediate care and what barriers prevented it; what data were readily available; and how they thought methods should be matched to resources.

Inevitably, and happily, the discussions were wide-ranging and did not always maintain a rigid focus on the questions at hand. At times, instead of talking about evaluation, participants spoke of the intermediate care services themselves, and the challenges involved in developing them. This is a confusion which bears mentioning, because it is easy to stray from developing a service to developing a plan to evaluate it, and back again. But equally, they raised key questions, not about intermediate care *per se*, but about how to evaluate it. Although those invited to participate were a select list of people with experience in intermediate care evaluation, many expressed genuine confusion about how to answer some of the questions we posed. They had come to the seminar as much to gain clarity as to offer their views and experience.

In response to this, we have chosen not simply to summarise the proceedings of the day, but to amalgamate the body of information produced, and to report and assess the concerns expressed about evaluating intermediate care appropriately. We present lists of the questions that different types of evaluation are best equipped to answer; we also report the questions raised about each type of evaluation, and offer guidance where possible. Thorny issues will not necessarily be resolved but they will be clarified and some alternative ways forward will be offered. Participants' opinions on the other questions - about drivers and barriers, and matching resources to methods - are embedded in the text at various stages. Our intent is to offer both a discussion paper on evaluation of intermediate care and a toolkit for those who wish to judge the merit

of their intermediate care services systematically from the outset. It should also be of value to those who are further along in programme development but, like the people attending the seminar, continue to seek support for conducting rigorous, sensible evaluation in a challenging area.

The report is organised as follows. In Section Two, general comments about evaluation and particular issues about evaluation of intermediate care *per se* are raised and discussed. Sections 3-6 tackle the specifics of process, implementation, outcome and economic evaluation. Section 7 discusses how and when to bring these forms of evaluation together, to create a coherent picture and find a way forward. In all sections, the focus will be on practical considerations as well as on good science. Finally, we present a series of appendices, including an annotated inventory of pre-validated assessment tools; a glossary of terms; recommended reading; and a list of those who attended the seminar.

Section Two: Central Issues in Evaluating Intermediate Care

Box 2.1

Highlights of this Section

- Intermediate care evaluation is complicated by the many care sectors and stakeholders involved, the diffuse and variable quality of the data required, multiple user groups ranging from children to the elderly, and a holistic model of care that is often at odds with conventional medicine and conventional ways of assessing success in a medical environment.
- Evaluation should meet the right need at the right time. This will mean varying the approach according to what is most important to learn.
- Success criteria should be set with reference to the objectives of the intermediate care intervention under evaluation. Different stakeholders will have different views of what matters, and what is most desirable to achieve.
- The perspective to be taken (that of the commissioner, patient, carer, staff, manager, consultant, etc.) should be agreed as clearly and publicly as possible.
- Although intermediate care evaluation is likely to be a small-scale, local enterprise, generalisability is still possible - either through conceptual generalisability achieved through qualitative methods or statistical generalisability achieved through comparison-group designs and based on power calculations using appropriate outcomes.
- Reliance on validated measures is recommended.
- Planning a series of investigations in pre-identified phases will lead to evaluation results most likely to be useful to practitioners and decision makers alike.

2.1 Why evaluate?

Evaluation is an essential tool for developing innovative and evidence-based services.[4] By taking a systematic look at the model of care, the way in which it is implemented, its effectiveness and the costs and benefits associated with it, and by comparing these with other models or services, it is possible to gain information that can be crucial to determining whether a new programme should be continued and extended, or not. People working in an area develop their own impressions, and often

[4] Evans, D., Steiner, A. (1997) A guide for local evaluation of PMS pilots. In McKeon, A., *Personal Medical Services Pilots under the NHS (Primary Care) Act 1997: A guide to local evaluation.* NHS Executive, HMSO, catalogue no. 97PP0130, 11-47.

their impressions are accurate. But sometimes they are not. More importantly, people's impressions will often depend on their personal situation and their position within the organisation. Evaluation can help sort through the complexity and diversity of impressions to produce a clearer view of what is being invested, what produced, and to what effect.

Because they are new, intermediate care services often have to prove themselves. Evaluation is the tool by which such proof is produced. Equally, if a new development is having difficulties, evaluation can identify the strengths and weaknesses of a programme and can highlight the stress points in a system, or the tensions between levels of organisation, so that the service can be improved. A different kind of evaluation can then be used to determine whether, after improvements have been made, the service is able to meet desired objectives, or can produce desired outcomes.

2.2 Some rules of thumb

It is important, then, that the evaluation is designed in such a way that it ***meets the appropriate need at the appropriate time***. An outcome-based summative evaluation is probably premature when a service has barely been introduced and there is still much to learn about skill mix, targeting and real-world implementation. A case-study approach will not be adequate when locality commissioners, for example, request the evaluation, because their intent may be to choose between the intermediate care service under offer and some other model or some other service - possibly for a different set of users, even in a different care setting. In this case, comparative assessment is essential.

It is also important that the evaluation be ***designed with the objectives of the intermediate care programme clearly in mind***. For example, if the objective of a post-acute intermediate care unit in hospital is to shorten overall length of stay, only one outcome matters. If, instead, it is to improve the transition from hospital to home, length of stay is less important than coordination of care pre- and post-discharge, or patients' and families' assessments of the discharge process, or readmissions. If, as a final example, it is to restore patients' functional independence, then a functional status measure is the most appropriate yardstick of success. The reason to emphasise this point is that if the wrong thing is measured, an intermediate care service perceived internally as a success can be judged externally as a failure, and vice-versa. It will also have implications for the running of the service, even before evaluation begins, because - to continue with the examples just provided - a unit with the

objective of reducing length of stay would select patients who had a positive prognosis and were capable of intensive rehabilitation; in contrast, a unit with the second objective might recruit patients expected to have continuing care needs in order that it could demonstrate effective coordination of care; a unit with the third objective would do well to reject patients who could not benefit from occupational therapy (for example, cognitively impaired elderly people). See Table 2.1.

Thus, it is crucial to ***agree whose perspective will dominate the evaluation design***. There is a perception among providers that those purchasing services demand quantitative measures of economic efficiency, ideally using a randomised controlled trial (RCT) design. Under such circumstances, the only outcomes of interest would be those whose currency can be easily translated into pounds and pence; namely, length of stay, number of readmissions, number of admissions avoided, or number and type of supportive services delivered. Although few would dispute the importance of such measures, there may be others which those who organise the service value more; for example, ability of staff to carry out responsibilities, staff morale, patient/user satisfaction, or safety (e.g., number of complications or requests for consultant support). The kind of evaluation questions asked may depend on who has commissioned the work, and to what end. The basic rule of thumb here is to ***consult as widely as possible, as soon as possible***, and to agree explicitly what the questions are, what the specific objectives of the evaluation are, who will fund the evaluation or review its findings, and what design will be used to answer the key questions specified.

We believe that the most important factor in devising your evaluation plan is to decide, *'What do <u>you</u> want to know?'* ***The best evaluations will be those that serve something more than an external authority's requirements***. To the extent that you can ask your own questions, and develop an approach that will meet both your needs and those of other judges, your evaluation will be more carefully executed, with clearer results. At the same time, your evaluation will be strengthened by choosing methods and measurement tools that enable the local work to feed into, and help create, a body of national evidence on the effects of intermediate care.

2.3 What makes intermediate care evaluation special?

Intermediate care shares with other bridging services a complexity that sometimes seems to defy evaluation. In this section, some of the specifics of this complexity are aired.

Table 2.1
Matching outcomes to objectives

Possible intermediate care objectives	Relevant success criterion	Implications for implementation
To unblock beds	Throughput in acute ward	Case management or long-term care placement likely to dominate; need to bring in social services; will accept wide range of patients with complex needs
To teach self-care	Functional status score, using a scale that measures high end performance; or measure of adjustment to severe trauma, such as Wood-Dauphinee's Reintegration to Normal Living scale (also used as a proxy for quality of life in the elderly)[5]	Unit should not accept cognitively impaired patients
To prevent admission	Number of people approaching crisis successfully treated in alternative way; description of interventions made; if RCT, number of admissions	Need for patient tracking systems between primary, secondary and social care sectors, in order to identify potential users and monitor treatments

2.3.1 Many interfaces

Almost by definition, intermediate care services have direct implications for multiple sectors along the health and social care continuum. This, in turn, has direct implications for evaluation in that a good evaluation will examine the knock-on effects between care sectors. For example, an admission avoidance scheme may be good news for acute trusts, but only at the expense of increasing GPs' and district nurses' workload. For the acute trust to produce a complete picture, its evaluation would have to set that increase against its own reductions in A&E use or its estimates of change in inappropriate admissions. Although this is very important, it increases the cost and complexity of evaluation. Moreover, if the acute trust has no financial

[5] Wood-Dauphinee S., Williams, J.I. (1987) Reintegration to normal living as a proxy to quality of life. *Journal of Chronic Disability*, 40:491-499.

stake in the community care situation, it may prefer to keep its evaluation resources focused in-house.

The involvement of multiple sectors also means that there will be a large number of players potentially interested in the intermediate care service, many of whom would wish to be involved in planning an evaluation. These players may include general practitioners, geriatricians, other consultants, managers, nurses, therapists, social services, community pharmacists, family, friends, the voluntary sector, the private sector, the commissioners or purchasers of health services, and patients themselves. Although many would want to play a part in assessing the costs and benefits of a new development in transitional care, since they would all be affected by it, not all will be appropriate members of the evaluation team. A good evaluation plan will identify a research team that has appropriate skills and clearly designated responsibilities. The larger the team, the more expensive and cumbersome the evaluation. So a balance must be struck between inclusiveness and practicality (see Box 2.2).

Box 2.2

Checklist for choosing the evaluation team

- Who will conduct the evaluation? Is this person from the same source that has commissioned it? Are members of both parties included on the team?
- Who will be included in the evaluation team? Practitioners? Health authority staff? Service users? Voluntary sector staff?
- Is it clear who is leading or coordinating the team and what the respective responsibilities of team members are?
- Will an independent evaluator be used? If not, how has the risk of bias been addressed?
- Have the costs of the evaluation team, e.g. fees or expenses, been included in cost estimates for conducting the evaluation?
- Has there been explicit agreement between commissioner and evaluator as to the purpose and objectives of the evaluation?
- Does the evaluation team possess all the required skills? For example,
  - ◊ questionnaire design?
  - ◊ statistical expertise?
  - ◊ qualitative methods?
  - ◊ health economics?
- If not, does the team have access to advice as needed?

(Adapted from Evans and Steiner, *op. cit.*, footnote 1)

A third aspect of interface evaluation is that data collection is much more complicated than for evaluation of a simple, discrete service in a well-defined setting. The data will vary in its quality and reliability, and also in the ease with which evaluators can obtain the information they want. There may be differential costs associated with acquiring information, which raises questions about the ethics of treating one sector differently from another in the process of gathering information. A pragmatic approach may dictate that the most accessible data is what gets collected, but this will not lead to the comprehensive evaluation desired. At times, it will be more important to invest resources in devising high quality information systems, so that data that should feed into an evaluation process, can do so - and can be trusted as reliable.

2.3.2 Many potential user groups

Intermediate care can be appropriate for a wide range of users, including children, adults recovering from surgery or adjusting to a new diagnosis of a serious chronic illness such as diabetes or heart disease, physically disabled people of any age, and the elderly. Each group - particularly children and older people - pose particular challenges for evaluation.

For children, there are questions regarding whose impressions of care should be sought - the patient's (i.e., the child's) or the parents'? Are parents good proxies for their children? There is evidence to suggest they are not, and even stronger evidence to suggest that health practitioners, such as doctors, are particularly poor judges of what young patients think.[6] The power of children to describe their satisfaction with care or their quality of life has not been explored in the development of most social science quality of life scales, for example, and there are few measurement tools validated for young people.[7]

Older people are expected to form a large proportion of the intermediate care population; however, only a subset of the elderly are ideal candidates for it. These are people who are non-institutionalised but with some functional problems; cognitively healthy; and at high risk of readmission. In other words, they need care, have a capacity to benefit from it, and are at risk if they fail to get it. Experience suggests,

---

[6] Payne S., Betts P. (1995) Explanatory models of diabetes. Project funded by the Nuffield Foundation, as described in Steiner A. (Ed.) *Health in Social Sciences*, Special issue of *Research News*, Southampton: University of Southampton.

[7] One exception is a children's version of the 9-item Dartmouth COOP chart, used to assess health status and including a quality of life item; see Nelson E.C., Wasson J.H., Johnson D.J., Hays R.D. (1996) Dartmouth COOP functional assessment charts: brief measures for clinical practice, in Spilker B. (Ed.) *Quality of Life and Pharmacoeconomics in Clinical Trials (2nd edition)*, Philadelphia: Lippincott-Raven.

however, that many practitioners equate intermediate care with non-acute health services for patients of a certain age.

One of the most difficult problems in evaluating older patients' intermediate care stems from the fact that this is the most heterogeneous age group of all. As a result larger samples are required in order to detect changes in the outcome of most interest, as well as to enable assessment of differential effects which will be crucial for targeting intermediate care appropriately. A second issue is that, on average, age does bring decline, so evaluation of the effects of treatment on older people's health must use methods, such as repeated measures analysis, that will be able to tease apart effects due to time and effects due to treatment. A related evaluation question to be resolved early on is whether the objective of care is to improve health status, or to prevent or slow the onset of decline. Again, the sample sizes required will vary depending on the measure of success.

2.3.3 A non-biomedical model

Intermediate care is not alone in this - health promotion and complementary medicine are other examples - but, despite operating in the context of the NHS, its philosophical and therapeutic underpinnings are not biomedical. It is not based on isolating a pathogen and eradicating it; rather intermediate care is based on a holistic view of patients and a belief that a comprehensive, empowering approach is therapeutic in itself. Lydia Hall, a pioneer in the development of primary nursing in the US, conceptualised this area of need as three overlapping circles of "care (focused on the patient's bodily needs), core (the engagement of the nurse her- or himself) and cure (emphasising multidisciplinary activities, with the nurse as patient advocate)".[8]

Some believe that the holistic approach itself precludes evaluation, because the process is too diffuse and contextual to be captured in a systematic manner. Others believe that it is eminently possible to demonstrate an association between an intermediate care programme and improved outcomes, where such exist. Clearly there is a challenge for proponents of intermediate care to devise evaluation plans that are capable of isolating the effect of their intervention, when that intervention is likely to vary - at least to some degree - from patient to patient and possibly from practitioner to practitioner. If such an effect can be isolated, moreover, there is the further problem of how to open the 'black box'. An evaluation which treats 'intermediate care' as the independent variable (the supposedly causal factor in

[8] As described in Pearson A., Punton S., Durant I. (1992) *Nursing Beds: An Evaluation of the Effects of Therapeutic Nursing*, Royal College of Nursing Research Series, Harrow: Scutari Press.

producing an effect) may be "sufficient for prediction, but it does not advance understanding or control" (of what has occurred in practice).[9] One way to manage this is to plan a staged, or phased, evaluation which carefully describes the model, its implementation and only then its effect on the outcomes of interest. This is discussed below.

2.4 Is generalisability possible?

It would be utopian to suggest that funds for evaluation of intermediate care are thick on the ground, and that is not the purpose of this report. Indeed, funds for intermediate care itself are often difficult to obtain. In such an environment, the only evaluation feasible is often a local effort. Multi-site studies are not the order of the day; instead, single site studies may be all that is possible. It should also be said that, even if large-scale evaluations were more easily undertaken, many organisations' interests will be limited to their own activities and strategic planning needs. Practitioners themselves may be more interested in fully understanding their own local dynamics, in order to provide the most effective and acceptable service possible, than they are in knowing where they sit in relation to some national norm or Royal College-based gold standard (not that there is one as yet).

2.4.1 Conceptual and statistical generalisibility

Is it possible, then, to conduct an evaluation which is at once locally relevant and generalisable to other settings? The answer can be formed in at least two ways. First, there is a distinction between conceptual generalisability - the sort attempted in qualitative research - and statistical generalisability. To achieve conceptual generalisability, the use of small and varied samples to explain people's perceptions and beliefs, and the way that these may vary under different circumstances, is considered best practice. By identifying the range of attitudes or impressions and establishing commonalities as well as distinctions among them, evaluators can produce case study-type findings which clinicians often find extremely useful. If they recognise their own patients in the sample described, and if they are delivering services in a comparable manner, they can trust that the conclusions drawn from the evaluation are likely to generalise to their own situation. Thus the clinical applicability of any evaluation can be taken as a starting point from which to consider and, in some instances, initiate changes in practice.

---

[9] Hornbrook M.C., Goodman, M.J. (1991) Managed care: Penalties, autonomy, risk and integration. In Hibberd H., Nutting P.A., Grady M.L. (Eds.) *Primary Care Research: Theory and Methods* Washington, DC: US Dept. of Health & Human Services, Public Health Service, AHCPR,107-126.

For decision-makers, such as health authorities, who must identify the optimal allocation of resources across a wide spectrum of needs, conceptually generalisable results may suffice; however, they may require population-based findings that are statistically significant as well as practically important. To this end, a comparison group is often needed and power calculations must be made to identify the sample size required to have a chance (how good a chance is pre-specified) to detect change (how much change is also pre-specified) if that change does occur. Even in a single-site study, this is often feasible. For example, Walsh is conducting an evaluation of a nurse-led unit in Southampton using an RCT design. Taking as the main outcome the proportion of patients returning to their pre-admission residence or one requiring less support, 105 people in each group were required to have an 80 percent chance of detecting a 20 percentage point difference between groups. Taking length of stay as the outcome of interest, 85 patients in each group would be needed.

Although Walsh's evaluation will not be able to say what it was about the intervention that did or did not work, it will be able to report that intermediate care is, or is not, associated with more independent discharge destinations and/or shorter length of stay than usual care. That sort of bottom line finding is extremely valuable to managers and planners. As to generalisability, it will form a sort of bounding argument for those considering the introduction of intermediate care in their locality - at least some form of nurse-led intermediate care unit does or does not work. Further, the type of hospital (teaching) and patient population mix will give external consumers of the evaluation other information with which to judge its relevance to them.

A final way to enhance generalisability is to select measures and approaches that have been validated in the literature. In that way, it is most likely that the evaluation will measure what it intends to. It will not, for example, attempt to assess staff morale but mistakenly investigate coping skills. It is also possible to extend the assessment of validity by evaluating it in your own study population since validity is never completely established but always relies on accumulated evidence. Importantly, use of a standard set of accepted measures will help to develop a body of research results which could be amenable to systematic review or meta-analysis. This will be a crucial step in the process of establishing the usefulness and effectiveness of intermediate care interventions. One caveat will temper this advice however; the wrong measure, no matter how carefully validated, will remain the wrong measure. Many seminar participants, for example, argued that the Barthel Index was an inadequate measure of function for the quintessential intermediate care candidate; although it dominates functional status measurement in the UK, what is needed here are measures of

intermediate, or even advanced, levels of function instead of the basic assessments captured by the Barthel.

2.5 An argument for phased evaluation

As the sections on process and implementation evaluation will demonstrate, developing a service has numerous phases. The snapshot approach cannot possibly present an accurate summation of the effectiveness of the new development. For this reason, the best evaluations will have planned in advance ***a series of investigations*** designed to capture the dynamics of introducing a new service. Only when the delivery of care has stabilised is it appropriate to conduct summative evaluations using an RCT design or, more pragmatically, an observational study with comparison group. At a minimum, process evaluation should accompany outcome evaluation wherever possible. To return to the example of the Walsh RCT, a careful analysis of the model of intermediate care as implemented will greatly strengthen the basic evaluation design. Without it, a successful model (as defined by Walsh's specified outcomes) will be difficult to replicate. An unsuccessful one risks throwing the baby - here, a potentially valuable intermediate care function in the continuum of hospital-based treatment - out with the bathwater of a flawed intervention.

Phased evaluation can be conducted with a formative or summative goal in mind. In the first case, the objective of evaluation is to aid in identifying strengths, weaknesses and stress points so as to improve the service being developed. There is likely to be feedback along the way, which is expected to alter the model of care or the people deemed appropriate for it. (Formative evaluation carries a range of challenges regarding the politics of dissemination; these will be addressed in Section Four.) In the second case - summative evaluation - the information gathered along the way will not be fed back to the people involved in the service until all stages of evaluation are complete. The purpose of phased evaluation in this instance is to document the dynamics of introducing a particular model of intermediate care into a particular setting, to describe the extent to which the model changes over time, and to make a 'final' assessment of its effectiveness, once delivery patterns have stabilised.

Although this is a time-consuming approach, and potentially a resource-intensive one, it is a standard to strive for as it has the best chance of producing results that will be useful to practitioners and planners alike. Moreover it is not known how long it really will take; it is plausible that six months to one year would be adequate to collect data for the process and implementation analyses. The amount of time needed for outcome analyses would depend on which outcome is considered most appropriate to measure,

the power calculations to determine sample size, and the turnover or throughput rates of intermediate care recipients. The first set of analyses would not require a comparison group, but the second set would. It would also be possible, and desirable, to continue the process/implementation evaluation in parallel with the outcome evaluation phase. Note that, in thinking about evaluation phases, it is crucial to allow time for creating and agreeing the evaluation plan and piloting such aspects as requesting consent (if appropriate) and data collection, to confirm that your plans are feasible.

For those who must work to annual budgets and need evaluative results 'now', the best approach may be to agree in advance which part of the picture is most important to clarify, and to acknowledge that only a portion of what needs to be known can be produced. Again, such decisions relate to who is commissioning the work, and for what purpose. A programme starting from scratch may be forced to take a *few hard measures* like before-and-after admission counts or length of stay, but could also focus on process issues such as referral procedures, staff training and acceptance, and patients' views. On the other hand, a practitioner group whose working philosophy conforms closely to the intermediate care model already, and who have high expectations of working smoothly to it, may be able to target their evaluation efforts at a point farther along the continuum.

Section Three: Process Evaluation

Box 3.1

Highlights of this Section

- Process evaluation is concerned with exploring exactly 'what is happening' in terms of clinical care and organisational developments. The main questions it can answer have to do with careful description, especially of activities and relationships.
- Usually qualitative methods will be required, such as observation, open-ended interviews, focus groups or use of narratives and story-telling; but more quantitative tools, e.g. structured questionnaires or scales to measure workload, may also be used.
- It is critical that the manner in which the data will be analysed is decided before data collection is commenced.
- Study designs in process evaluation should allow for collection of baseline data.
- Longitudinal studies have the advantage of identifying changes that occur over time; for example, confidence in the service may increase or referral patterns may alter.
- Pluralistic approaches - accounting for multiple perspectives - are also recommended in order that a more holistic understanding can be gained.
- When undertaking process evaluation, it is important to differentiate between 'care needs' and 'medical needs'.
- Remember - if processes of care are not understood, then it will not be possible to replicate services, either locally or in a wider arena.
- Ideally what will be required over time is the link between process and outcomes.

3.1 What is meant by 'process evaluation'?

Although it may seem a simple enough concept, professional evaluators find process evaluation the most difficult to define. The words can describe an approach to evaluation - a sort of action research approach - which uses the process of evaluation to help develop and improve the service under review. This is also referred to as 'formative evaluation'. Alternatively, process evaluation can refer to any type of evaluation whose purpose is to describe the processes of, in this case, delivering

intermediate care: what exactly is done, to whom, by whom, in what order and under what circumstances.

Both characterisations are useful, and either approach can lead to improved understanding of an intermediate care service. This section will emphasise the latter definition - process evaluation as evaluation designed to capture as much information as possible about the model of care being organised, delivered and received. (Implementation evaluation may be seen as a subset of process evaluation.) It should be recognised, however, that comprehensive assessment of the processes of care may occur in a formative, developmental context if desired.

In intermediate care, there are certain drivers encouraging process evaluation. These derive mostly from intermediate care's status as a new - hence poorly understood - development in health services. First, health authorities are responsible for informing users about service changes. They need the sort of descriptive information that process evaluations produce to enable them to understand the intermediate care approach. Second, at the national level, the private sector has expressed interest in evaluating models of intermediate care from a process perspective in order to improve their strategic thinking about the bundle of services they provide.

3.2 Approaches to process evaluation

Process evaluation will usually require ***qualitative methods*** of collecting data, because the kind of information that is of most interest will not fit neatly into pre-set categories. Qualitative analysts assume that knowledge is based on context; that is, that people will form opinions and judgments based on their social, economic, demographic, cultural and psychological positions. By identifying both context and perspective, qualitative evaluators allow the full range of views to emerge, and for common views to be recognised. This can be critical to successful implementation - or to concluding that implementation is not advised. The qualitative approach is relevant to obtaining both patients' views and those of professionals; in the first case, because satisfaction with care is related to health outcome and in the second case, because it is professionals who will determine the extent to which a development is adopted and how it is run. The goal, then, is to go beyond surface appearances, to explore the meaning of 'lived experience' and to discover ***not only what was done, but also how and why***.

Qualitative techniques of data collection include in-depth interviews, focus groups, use of diaries, photographs, film or videotape, and participant observation. For interviews,

it is important to think about getting a wide range - termed 'maximum variety' - of people in your sample, to represent different perspectives. Large numbers are neither feasible nor desired here; the goal intended is conceptual generalisability. For focus groups, where an atmosphere of trust is important, you may want to hold several meetings with different sets of (more or less) similar people; you will also want to think carefully about whether the group you are interested in is likely to accept the facilitator you plan to use. Sometimes researchers will engage somebody who is of a similar background to the members of the focus group. Another approach suggested by participants is the use of post-discharge narratives or story-telling, allowing those concerned to 'tell it as it was', then subjecting the data to content analysis.[10] [11] In this way, insights into the issues that are important to users as well as professionals can be brought to light (see Box 3.2).

Some ***quantitative tools*** can also be applied to collecting data about process. ***Bar-coding*** is one such technique.[12] Barcoding in essence quantifies the content of clinical activity. The barcode devices are identical to those used in shops for stock checking and check-out pricing, but for intermediate care the codes would be developed by nurses or therapists to describe what they do. The methodology has been validated for use in measuring nursing workload in a Southampton neurosurgical unit. Although bar-coding has been questioned in terms of its reliability and, to some degree, its validity (for example, how can non-verbal forms of support be categorised, acknowledged *and* verified?), it remains a highly promising tool for describing the model and for answering some process-related questions about skill mix.

Another method which has been suggested is the Quality Patient Care Scale (Qualpacs).[13] In ***Qualpacs***, patients are observed for specified periods during which time all their care is rated against explicit quality criteria, using a range of care dimensions. Developed in the US in the 1970s, the instrument has since been adapted for British use. Another approach to measuring process is to use patient satisfaction questionnaires, but again, their validity has been questioned, particularly among older patients who may be concerned that complaining in any way might adversely affect their treatment. Patient evaluation questionnaires, where service users judge others

---

[10] Williams G. (1984) The Genesis of Chronic Illness: Narrative Reconstruction, *Sociology of Health and Illness* 6 (2) 175-200.

[11] Bruner J. (1987) Life as Narrative *Social Research* 54 (1) 11-32.

[12] Connell N.A.D., Lees, D., Lees, P., Powell, P., Stafford, P., Sutcliffe, C. (1996) Costing and Contracting in the NHS: A Decision Support Approach. In Bourn, M., Sutcliffe, C. (Eds.), *Management Accounting in Healthcare*. London: CIMA, pp. 31-40.

[13] Wandelt M., Ager J. (1974) *Quality of Patient Care Scale,* New York, Appleton Century Croft.

rather than reveal something about themselves, perform better; however, no such instrument exists yet for evaluating intermediate care.

Box 3.2 Exploring experiences of care - the use of narratives

Story telling as a powerful means of understanding history, culture and tradition, has long been acknowledged. However, it is only in recent times that it has been recognized as a technique that can be used as a means of understanding experiences of care from the perspective of patients or service users. This shift may partly be attributed to the growing criticism of satisfaction surveys, which commonly record very high performances[14] but whose content is limited to a preset agenda of what is *thought* to be of interest to the participant.

The origin of story telling as a means of data collection can be tracked back to the work of Flannagan who developed the 'Critical Incident Technique' as a means of identifying training needs for pilots[15 16]. Since that time, it has been used widely in health services research and has, over the intervening years, broadened out to encompass analysis of whole stories which are acknowledged as a legitimate means of accessing individuals' perceptions of life events.

The use of story telling falls into the school of naturalistic enquiry. It allows participants to 'tell it as it was', highlighting what was important and meaningful to them. Stories are most commonly told verbally and taped so that the data can be subjected to rigorous content analysis. Skilled interviewing technique is an essential pre-requisite to ensure that the agenda is set by the interviewee rather than the interviewer.

Moving to more generic methods, one approach is the use of ***care pathways***, where usual patterns of progress are predicted, allowing the levels and range of services to be planned. Some seminar participants recommended care pathways as a method to describe process systematically. However, the wide variation in individual needs and the underlying emphasis on holistic treatment may make the application of a care pathway problematic for intermediate care services. Also, many people who are

[14] Carr-Hill R., Dixon P., Thompson. A (1989) Too Simple for Words *Health Services Journal* 99 (1) 728-729.
[15] Flanagan J. (1954) The Critical Incident Technique, *Psychological Bulletin* 51 327-358.
[16] Norman I., Redfern S., Tomlin D., Oliver S. (1992) Developing Flanagan's Critical Incident Technique to Elicit Indicators of High and Low Quality Nursing Care from Patients and their Nurses, *Journal of Advanced Nursing* 17 590-600.

candidates for intermediate care will not have a linear pathway to recovery; instead, there will be fluctuations in their progress. Thus, if care pathways are employed, a distinction must be made between fluctuations in progress and variations in progress; fluctuations should not necessarily trigger a new level of input. Those participants in favour of using care pathways in evaluation emphasised that they needed to be related to decision pathways, with shared assessment across the acute and primary care interface; they commented that this was not always the case.

Usually a process evaluation does not require a comparison group and, sometimes, a comparator will actually be inappropriate. For example, if you want to know how patients view their care, or nurses view their service, it would be more than difficult - it would be silly - to try to persuade them to conceptualise what they do in terms of an acute and post-acute phase.

3.3 Which questions are best answered by process evaluation?

Table 3.1 lists key questions identified by seminar participants which are best answered by process evaluation. As expected, they are questions relating to description and analysis of dynamics; often they are about relationships between people (i.e. patients or professionals). In general, there was consensus that careful description must precede any summative evaluation.

There was also the sense that process evaluation could protect against drawing misleading conclusions. The interest in making the assessment, referral and admission process transparent is an example of using process evaluation both formatively and in this protective manner. Many participants noted that the verdict on intermediate care might well be contingent on who was referred for care, and whether they were appropriate candidates. The possibility of a mismatch between the patients referred by acute-ward consultants and those desired by the professionals providing post-acute care was raised. If, for example, the patients were especially frail, the intermediate care services might be comforting to patients but would not succeed in restoring function. Thus an outcome-based evaluation would have to know about the referral pattern in advance, so that it could identify as an appropriate outcome *not* functional status but something like perceived well-being or satisfaction with care. A process evaluation, by contrast, would be able to describe the type of patients transferred, the sort of care they received and any feelings of misunderstanding between acute and post-acute care providers.

A further set of issues was identified regarding the people who provide intermediate care. Here, qualitative methods can work in concert with quantitative methods such as bar coding to identify who does what, how much, and how well. In this case, it could be relevant to examine process carefully in a comparison setting such as the acute ward in a hospital as well.

Table 3.1
Key questions about intermediate care that process evaluation can answer

Area of Interest	Key Questions
Organisational model	• What precisely do different models look like? • What do those delivering an intermediate care service think they are doing? • Where does intermediate care fit into an organisation's overall strategy or service mix?
Assessment and referral	• What is the process of assessment and referral to intermediate care?
Staffing	• What do different professional groups and/or grades (e.g. occupational therapists vs. nurses, qualified vs. unqualified) do in the context of offering intermediate care, and how well do they do it? • How much overlap of function is there between the work that different occupational groups do? Is there a way to rationalise cross-discipline work to make better use of resources? • What are the skill mix requirements of intermediate care, and what are the differences between professional groups in terms of their acquisition of appropriate skills and the strength of their professional boundaries?
User views	• What is the patient's (and carer's) perception of intermediate care? Do they see it as the 'cheap option' and therefore second-rate? Do patients want intermediate care? • Where do patients really want to be cared for? Is the ideology 'home is best' a true interpretation or a professional dream?

Finally, seminar participants raised a series of questions regarding the ways that patients viewed intermediate care. There was some concern that rather than seeing intermediate care as a progressive service which represented the best in therapeutic approach, patients might worry that they were being hurried out of specialists' hands into a low-tech setting and being isolated from the best that medicine has to offer. It

was suggested that not everyone preferred home care to an institution, and that evaluation should explore this issue before making any decisions as to the acceptability of an intermediate care service.

3.3.1 Using process evaluation to define intermediate care

Implicit in the key set of questions identified is uncertainty about what, exactly, the term intermediate care is meant to denote. Many participants reported a lack of clarity about such basic issues as the purpose, target group, and appropriate professionals to deliver intermediate care. Others felt that they were already fulfilling the intermediate care function, although their service was not termed 'intermediate care.' They wanted to use evaluation to prove that they were, indeed, offering intermediate care. The lack of a common language, both within and between sectors, was seen as a major difficulty in coming to an understanding about the usefulness of such a model.

Process evaluations can make a positive contribution to resolving this problem by making explicit the theoretical or conceptual underpinnings of the care intended and by studying the delivery of services that are thought to be intermediate care (including the identification of which types of patients are offered them). This could be a new service, actually called intermediate care, or an existing service believed to *be* intermediate care. What matters most is that the plan and the process be thoroughly described; this is the groundwork which can then be compared across sectors and sites in order to develop a common language and a common framework for examining a new development in the organisation of care. In this context, one participant noted that in beginning to plan an evaluation by holding stakeholder meetings to agree objectives and criteria for success, all those involved rapidly discovered that hospital-based consultants differed strongly from nurses, GPs and others on these points. In a sense, he had his first finding in a process evaluation of the service - and so did the stakeholders.

3.4 Which questions can process evaluation help to answer?

In addition to the key questions in Table 3.1, Table 3.2 lists other questions that would benefit from process-focused investigation, but which could not rely solely upon that approach. Instead, resolution of these questions would probably require a structured research design that used appropriate comparison groups. In these cases, multiple methods can work in concert - either sequentially or in parallel - to explore key issues. What should be outstanding in Table 3.2, is that the questions listed are more likely to imply a reference point that is external to the intermediate care service itself.

Table 3.2
Key questions about intermediate care that process evaluation can help to answer

Area of Interest	Key Questions
Relative value of intermediate care	• Is intermediate care necessary? Is it feasible?
Comparison of models	• Which models are best? Which mixture of models (hospital/community-based) should be considered if a purchaser wants to contract for a range of services? • Does intermediate care intervention differ in type, quantity and quality of care in a nurse-led compared to an acute ward? Can IC be delivered on an acute ward? • Does intermediate care intervention differ depending on location and who is providing the service?
Targeting	• Who should be eligible for services?
Substitute or complement?	• Does intermediate care substitute for other forms of care, such as days on an acute care ward, or is it an add-on to existing services? That is, does intermediate care meet an already identified need in a different way, or does it provide a service responding to previously unmet needs?

3.5 Ways to improve process evaluation in intermediate care

In this section, we offer advice culled from the seminar participants (including speakers) on how to strengthen your approach to process evaluation of intermediate care.

3.5.1 Begin with an information audit

All evaluation centres on collecting information in a systematic manner in order to aid in judgments of merit. As noted in Section Two, intermediate care is characterised by a diffuse and uneven network of information. Thus the best way to begin - in parallel with other tasks involved in drawing up an evaluation plan - may be to conduct an information audit (see Table 3.3). Seminar participants reported difficulty not only with obtaining information, but even with knowing how to find out what data are

available. One person commented that the information she found in the literature, or got from her health authority, never seemed to be in quite the right currency; for example, if she needed data on people aged 75 and over, she could only find data for those aged 65-plus or 85-plus. This will be a recognisable problem, and one which needs resolution.

The information audit will signal at an early stage which questions it is feasible to answer simply, and which will require more thought, effort or expense. It can have a formative function too, in that it may start stakeholders thinking about what is important to know if intermediate care is to be properly implemented (and assessed). It may also start evaluators thinking about how, exactly, they will be able to collect data. For example, in one locality, Psions® (hand-held PCs) were used to collect bed-side patient data. Although it seemed a neat technology, Psions® were vulnerable to the same shortcomings as other methods. People lost them, forgot to use them, were inconsistent in their recording habits, and so on. Bar coding pens have been suggested as another real-time data collection tool. But seminar participants expressed concerns about them too, especially in relation to inter-rater reliability and the difficulty of establishing valid coding categories.

3.5.2 Know your baseline

Whatever form of evaluation you choose, it is essential to take some baseline measurements. Who is the current user group or groups? What are their needs as currently identified? (These will usually be expressed in medical terms; sometimes they will be described in the language of social services instead.) What is the turnover or throughput rate? How well do patients fare under the conventional system? In the case of staff, what is the experience level and skill mix of staff treating patients under usual care arrangements? And so on. In this way, it will be possible to chart changes associated with the intermediate care programme. It will also be possible - in the spirit of using process evaluation to protect against misleading conclusions being drawn - to observe whether the patient casemix becomes more or less severe or the numbers one is expected to treat differ significantly from the baseline situation.

Baseline measures allow estimation of counterfactuals, such as '*What would happen to patients if the acute beds and/or community hospital beds were not there?* And *what might patients get if they didn't get (a) what they have now or (b) intermediate care?*' This is a critical point. Numerous seminar participants considered these questions to be at the heart of intermediate care evaluation. But it is also important to be able to

Table 3.3
What to ask in an information audit

Question to ask	Considerations and ideas
What information might you want?	Patient-level data such as demographics, length of stay, diagnosed conditions, procedures performed, number and type of services given; staff-level data such as number and type of medical, nursing and therapy staff, training levels; organisational data such as costs
Where do you think you can get it?	Trust or hospital computer, medical charts, health authority, library, local authority, advocacy organisation
Is it there?	Check first by asking, then by piloting a few real or imaginary cases
Are those who hold the data willing to share it with you?	It is useful to confirm this in writing, because of potential personnel changes over the course of evaluation
Is there a fee?	Find out in advance and confirm in writing
What routine data does your health authority hold?	Include the public health or primary care divisions when requesting routine data
What information do district nurses in your area record as a matter of routine? Is it available? At what time/money cost?	May be objective (number of contacts, procedures performed, medical diagnosis) and subjective (patient's mood, social support, cognitive status). Would have to negotiate a convenient way to collect data.
How accurate is the information (wherever you get it)?	Pilot this in advance, by looking for something where you already know the correct information, or observing care to check practice against what is recorded, or test for inter-rater reliability.
Who will be responsible for abstracting data from existing sources?	Decide in advance whether it will be somebody on the evaluation team, or the person providing the information (e.g. therapist, doctor, nurse, patient). If training is required, provide it. Also consider double-entry for a subset of records, so that you can confirm reliability.

Table 3.3, continued

Question to ask	Considerations and ideas
Are there systematic differences in recording/data entry quality that you can identify?	For example, if nurses are better recorders than doctors, you may choose to use data from the former but not the latter
What are the costs and benefits of different methods of data collection?	If, for example, it is easy to obtain data with high levels of error, it might be preferable to invest in collecting your own information more carefully; if, on the other hand, that proves too costly, you may choose to omit that area entirely from your evaluation
When data from multiple sources are involved, can it be merged to your satisfaction?	Find out whether there are common codes (for example, to do with patient's diagnosis) and, if not, whether they can be created
Will it be necessary to obtain ethical committee approval before data can be collected?	Local research ethics committees vary widely in their requirements. Be sure to find out what is needed well in advance of when you want the evaluation to begin.

identify the development over time in the way that the intermediate care programme itself is able to meet its objectives (see Section Four, Implementation Evaluation). Establishing the baseline will allow for this.

3.5.3 Distinguish between care needs and medical needs

At the point that intermediate care is introduced, the available baseline measures may be all - or at least mostly - expressed in medical language; namely, diagnoses, clinical signs and symptoms, and procedures carried out. But seminar participants felt that what made intermediate care distinctive was its emphasis on the patient as a whole person, and its therapeutic objective of helping people move from the identity of 'patient' back to the identity of 'person'.

Some seminar participants reported experiences of having had what they saw as inappropriate evaluation frameworks foisted upon them - for example, community-based programmes were forced to run a hospital-style evaluation (with performance-based rather than qualitative outcomes). As a result, they felt that patients were inappropriately medicalised in the course of their evaluation.

It is worth acknowledging early on the pressure to conform to the medical measures of output and outcome, in order to negotiate acceptance of a set of measures that includes subjective assessments felt to be more appropriate to the intermediate care model. To the extent that intermediate care delivered in different settings will emphasise different therapeutic goals, process evaluations should also differ in their definitions and measurement of success. The need for nursing care or rehabilitation care is not identical to the need for medical care, and this should be made explicit in all intermediate care evaluations.

At the same time, however, evaluators should recognise that many intermediate care projects are instituted because consultants in hospitals hope for improved outcomes *as they define them*. If the intermediate care service under review is not organised to shorten length of stay, reduce readmission rates or increase the rate at which patients are able to engage in self-care, then it should be a finding of the evaluation that the consultants' goals cannot be met with the model as implemented.

3.5.4 Take a pluralistic approach

The discussion in the previous section reinforces the importance of identifying whose perspective will dominate the evaluation. In participants' experience, the focus tends to be defined by the commissioners of the evaluation. However, in every working group at the seminar, participants expressed strong preference for a pluralistic approach. They believed that a 'good' process evaluation should encompass multiple perspectives (including those of patients, carers, staff and managers) rather than focusing on any one perspective in particular. They reasoned that this was crucial because of the many important inter-relationships in the care process. The goal would be to identify the range of views, including common ground as well as differences. Some went further, and suggested that process evaluation should study the relationships themselves.

In addition to preferring a multi-faceted evaluation focus to a more narrow one, participants wanted to discuss the relative merits of different analytic methods. In particular, they sought guidance as to whether a single method (any single method) was better or worse than using multiple methods in a single evaluation. Again, there was a clear preference for multi-method approaches. Regarding which aspects of intermediate care were most important to evaluate, participants felt the pluralistic approach was best suited to bringing out the more complex - and more reality-based - lessons. For example, intermediate care may be preferred by patients but found to be less effective than standard care from a medical perspective. What should be done?

The answer must be driven by values, but at least a comprehensive evaluation will highlight the conflict, where a simplistic one would not.

It must be noted that the more multi-faceted the evaluation, the more complicated and expensive it becomes. It is difficult to obtain all the appropriate expertise without mounting a very ambitious effort. Several options are possible. It may be possible for the multiple sectors potentially affected by the intermediate care service to jointly contribute to an evaluation that will attend to their various interests. Alternatively, it may be possible to seek external funding. There may be expertise in-house in the absence of adequate resourcing. However, it is better to keep it simple and clean - acknowledging the limitations of such an approach - than to muddle through an inappropriately grandiose plan.

3.5.5 Use a longitudinal design

Participants agreed that it would be a very good use of evaluation funds to follow a longitudinal design in any process evaluation, since processes change over time as a function of increased understanding, increased skill, changes in staffing or leadership, or external shifts in the environment of health care delivery (to mention but some of the factors). In light of this reality, evaluators are advised to plan for several rounds of data collection during the first one to two years of implementation.

For example, some participants commented that it is often difficult to predict what patients' preferences will be, how they will progress, or what sort of benefit they will experience from intermediate care. Because of this, they resisted evaluations that used pre-set criteria of success. It is for this reason that we recommend phased evaluations (introduced in Section Two and taken up again in Section Seven). However, a longitudinal process evaluation will contribute importantly to understanding what influences patients in their preferences, prognoses and capacities to respond to care.

In a slightly different context, one participant has found focus groups to be a source of longitudinal information since many users of intermediate care services are people who have been involved with the health service for many years. They can offer a perspective on the organisation of care which takes into account a whole range of experiments and modifications, and can provide their own interpretations of different programmes' successes or failures.

3.5.6 Focus on analysis

Finally, it is essential to note that the most important aspect of process evaluation is analysis. There is often an emphasis on data collection, because process data can be time-consuming and expensive to gather. But they are equally time-consuming, difficult and expensive to analyse. It is all too easy to find oneself with masses of information and little feel for how to use it. For this reason, and possibly to a greater extent with process evaluation than any other evaluation approach, it is critical that consideration is given to how the interview transcripts, focus groups, audio or videotape proceedings, and printed documentation will be used to answer your evaluation questions. Because we live in a verbal society, we tend to think that expert support is required for quantitative evaluation but that we can manage the qualitative analyses ourselves. Be forewarned: this is highly skilled work. If you don't know in advance how you will conduct your analysis, seek expert support.

3.6 Limitations of process evaluation

Not only is it important to understand what process evaluation is uniquely suited to assessing, and what it can contribute to understanding; it is also important to acknowledge its limitations. Indeed, although some participants preferred process evaluation to the exclusion of all else, others felt that - however interesting it might be - process evaluation was an inadequate approach because it could not answer the question 'does it work?'. In reality of course, both questions must be addressed since we must know what it is that is working in order to be able to replicate.

Although good qualitative research seeks explanatory models to answer research questions, process evaluation is fundamentally a descriptive methodology. The explanations it achieves will be in the form of testable hypotheses; it will not be able to confirm or deny that the hypotheses are borne out by empirical evidence. The richness of the process evaluation approach must be traded against the scientific power of the outcome evaluation approach (see Section Five - it has its own limitations!). Participants expressed concern that because process evaluations tended not to use comparison group designs and were usually single-site, the import of what they observed was difficult to assess.

They also emphasised the importance of identifying and confirming process-outcome links. This was seen as more important than examining either one or the other in depth. Thus the format of the seminar - and, to some extent, of this report - was seen as somewhat arbitrary or even artificial. The combining of process and outcome-orientated evaluation will be discussed in Section 7; however, although the most

powerful results will indeed come from associating clearly identified practice with desired and well-defined outcomes, each approach to evaluation stands alone at times. For process evaluation, its moments to stand alone are when the model of service and the perceptions of those involved with it are still unknown.

Section Four: Implementation Evaluation

Box 4.1

Highlights of this Section

- Implementation evaluation is a special type of process evaluation primarily concerned with plotting or mapping the course of change.
- The most commonly used method is that of the case study, drawing data from multiple sources including documentations, interviews and observation.
- Even more than offering insights into intermediate care *per se*, implementation evaluation can illuminate the strategic and other interests of those involved. Thus a multidimensional perspective is essential.
- Analysis of such data is complex and time-consuming.
- Implementation of a new service takes time. Evaluators should resist imposing an image of structure prematurely, but might consider when it starts to take form.
- It will be useful to focus on the unanticipated consequences of implementation.
- The advantages of an internal evaluation (e.g. prior knowledge of the organisation and the context in which it operates) may be outweighed by the external evaluators' impartiality and ability to convey what may be unpopular messages.
- Implementation evaluation is likely to produce information leading to conceptual generalisability rather than statistical generalisability.
- Feedback sessions at several points along the way can be used to increase internal validity. However, information gained through implementation evaluation can be very sensitive and feedback must be handled carefully.
- Remember - it is critical that all stakeholders who should be involved in the evaluation gain a sense of ownership from the outset.

4.1 Analysing change

At the heart of all implementation evaluation lies the concept of change, and intermediate care represents a change - in name and philosophy at a minimum, and often in the basic structures of service delivery. As a result, its introduction is likely to be met with resistance by some and possibly excessive enthusiasm by others. It can upset, or threaten to upset, conventional power arrangements, and it will have implications - some anticipated, others unforeseen - for the whole system of care

where it is introduced. Such impacts need to be identified and evaluated, so that the people involved can understand what factors make a new programme succeed, and what factors block its development.

There are numerous theories of how organisations change. Pettigrew, for example, considers change a natural and rational process with four clearly defined stages, namely; diagnosis of the problem, planning for change, launching the change, and following up.[17] Lewin uses a schematic of "unfreeze, change, refreeze" to describe organisational change.[18] Lindblom usefully depicts organisations as simply "muddling through".[19] A popular notion in NHS research and development is that of the 'virtuous circle' by which R&D supports work that produces evidence, which leads to changes in practice, which requires evaluation, which produces evidence, and so on.

4.2 What is implementation evaluation?

Implementation evaluation fits nicely into the virtuous circle concept. The path from evidence to practice is rarely straight and smooth, and evaluation can help to account for curves and bumps along the way. Although the approach can be used to assess the *effectiveness* of implementation (was the change introduced well?), most people use implementation evaluation as a way to *identify and analyse organisational dynamics* when a new service is created. In one participant's words, implementation evaluation is more about lessons than judgments. It is expected that difficulties as well as successes will be identified; sometimes, ideas for the resolution of problems will also become evident.

Implementation evaluation can be seen as a special case of process evaluation. It is best used for formative purposes, so that the lessons learned can be applied. The particular strength of implementation evaluation is its capacity to examine systems and relationships, to focus on the dynamic interaction between elements and not simply on the elements themselves, for example, nurses' and doctors' attitudes. Implementation evaluation will not involve comparison group designs. Rather, the focus is on the organisation and how it functions to meet its goal - in this case, to effectively deliver intermediate care.

---

[17] Pettigrew, A. (1985) *The Awakening Giant, continuity and change in ICI* Oxford, Blackwell.
[18] Lewin K. (1958) The Group Decision and Social Change in: Macoby E (Ed) *Readings in Social Psychology* London, Holt, Rinehart and Winston.
[19] Lindblom C.E. (1959) The science of muddling through. *Public Administration Review*, 19:79-88.

4.3 Which questions are best answered by implementation evaluation?

As Table 4.1 demonstrates, implementation is widely perceived to be influenced by organisational politics. The sorts of questions that seminar participants considered key elements in an implementation evaluation are concerned not with the nature of the intermediate care intervention itself (as in Section Three) but with the strategic interests involved in its introduction. They also reveal the importance of examining innovation in a whole systems context.

In thinking about evaluations of intermediate care, participants said they would want to know what the impetus for change had been. For example if it were a top-down initiative the new service would have the advantage of strategic and managerial support. However clinicians, nurses or therapists might lack awareness that a new philosophy of treatment was being tested. They might also lack the skills, which could jeopardise the service's chances of success. If, instead, the intermediate care programme were a grass-roots initiative, it might be met with considerable resistance from the existing power structure - for example, by consultants used to a strictly medical model of care. Certainly, a budget-driven impetus for change would lead to a different sort of intermediate care service than if the change were driven by quality concerns, and a strategy that sought to meet both objectives would be even more complex. An appropriate evaluation would begin by clarifying this question.

To be useful, evaluations focused on the implementation of a service should attempt to *uncover the organisational politics and power arrangements* that shape an organisation's culture. In intermediate care, this would extend to assessing the links between different service sectors, because it is often the quality of such connections that make or break a boundary-crossing intervention. This will be a challenging but worthwhile enterprise. First, the data collection may be complicated by having multiple sites to visit and numerous interviews to arrange. Second, the analysis will require considerable attention so that the relationships within and between systems can be mapped. Third, and possibly most important, the dissemination of findings may be particularly sensitive.

The final area of investigation that seminar participants considered the remit of implementation evaluation focused on the consequences of making a change. They wanted their evaluations to tell them what people would do differently, were they to start again; what happened that they had not anticipated - both positive and negative events; and they wanted to get a feel for the factors which either helped or hindered

the service from getting under way as initially planned. In other words, if they were to undertake this type of evaluation, they wanted to gain the benefit of hindsight, and come away with lessons for the future - or, at a minimum, a well-documented understanding of what went right or wrong.

Table 4.1
Key questions about intermediate care that implementation evaluation can answer

Area of Interest	Key Questions
Impetus for change	• Where (from what sector or level) does the desire for change come from? What is driving the change? • Did the intermediate care service come about in response to a problem, such as bed shortages in hospital, or was it somebody's 'good idea' looking for a niche? • What is the level of unmet need, and who is defining it? • What are the vested interests and what do they expect from this change?
Power	• What are the politics and power arrangements informing implementation of the intermediate care service? (This reflects seminar participants' perception that intermediate care must be negotiated and implemented in the 'real world') • In developing an intermediate care service, what is the service's relationship with GPs, nurses and others in the locality?
Stakeholders	• What is the involvement of consultants? (e.g., supportive, absent, obstructive) • What is the involvement of consumer groups? (as above) • What is the involvement of the voluntary sector, and of social services? • What is the involvement of senior management and the Trust board?
Effects of change	• What are the barriers to creating and supporting intermediate care? • What would be done differently if it could done over again? • What were the unanticipated consequences of introducing intermediate care?

The final area of investigation that seminar participants considered the remit of implementation evaluation focused on the consequences of making a change. They wanted their evaluations to tell them what people would do differently, were they to start again; what happened that they had not anticipated - both positive and negative

events; and they wanted to get a feel for the factors which either helped or hindered the service from getting under way as initially planned. In other words, if they were to undertake this type of evaluation, they wanted to gain the benefit of hindsight, and come away with lessons for the future - or, at a minimum, a well-documented understanding of what went wrong.

Box 4.2 Is it a struggle to implement intermediate care?

Many seminar participants who were 'believers' in the value of intermediate care reported difficulties in implementing a service that had not been in place before. They felt hampered by top-down pressures not to introduce any new costs, which sometimes translated to a reluctance to hire specially trained staff or train existing staff; and not to alienate the more powerful players in the organisation, who might view any change they did not control as a potential threat.

Those who had intermediate care foisted upon them noted that a lack of openness about strategic goals could seriously limit their capacity to succeed, but that despite this, some stakeholders would not be open about their goals. Participants also felt pressured, on the one hand, to couch all their efforts in the language of the medical model - which did not match their philosophy of care; or, on the other hand, to treat intermediate care as a kind of social care management approach - which also did not match their ideas about rehabilitation, restoration and self-empowerment. When it was time to evaluate their service's value, these pressures were translated into what they felt to be inappropriately specified outcomes against which the service was to be judged. (See Section 4.6.1 'The importance of ownership'.)

4.4 Methods of implementation evaluation

The central research design for implementation evaluation is the *case study*, which painstakingly examines a single example of (in this case) an intermediate care service. Box 4.3 lists some ways to collect information for the case study. Somewhat unusually, implementation evaluation relies on multiple methods, not so much for triangulation[20] as to map the range of perceptions and attitudes existing, and to learn the extent to which they work in concert, or in conflict, with each other. Thus *a multi-dimensional perspective is essential.*

[20] Fielding N. , Fielding J. (1986) *Linking data, Qualitative Research Methods Series No.4. London: Sage Publications.*

Box 4.3 Methods used to evaluate implementation

- Contextual analysis (mapping of stakeholders' domains and interests)
- Semi-structured interviews (telephone and in-person)
- Questionnaires (postal, self-administered on site, interviewer-administered)
- Observation of key meetings
- Analysis of documentary material (mission statements, contracts, etc.)
- Focus groups
- Feedback to study sample (an action research component, in that the process itself can be evaluated; how was the information received?)
- Single case studies

(Adapted from Dopson, S. 'Implementation evaluation', paper given at the King's Fund Intermediate Care Evaluation Seminar, 29.10.97)

Most of the methods listed above would be used in a number of contexts. For example, to get a feel for the macro- or meta-issues involved in creating an intermediate care service, seminar participants felt that it would be extremely useful to interview people 'at the top'; was the service being introduced to relieve pressure on overall costs, or a particular acute-care ward, or was it simply someone's brainchild being tried out? Equally, however, the evaluation should include interviews with people 'at the sharp end', who might be unaware of the strategic issues but would be highly sensitive to the practical issues affecting delivery of care. It is not uncommon to get a variation of views of these two perspectives. A good implementation evaluation will include interviews with as many of those involved as possible. It is essential, however, to have a clear understanding in advance of how the questions you ask will be incorporated into the analysis; otherwise, it is all too tempting to ask questions of only peripheral interest which will confuse the analysis later on.

4.4.1 A time-consuming analysis

Because the objective of implementation evaluation is to use an array of information sources to derive a picture of what the organisation, as a whole, is doing, it can be a demanding and time-consuming approach. In addition to the bulk of information to be gathered, the analysis requires care and thoughtfulness. First, it is a given of implementation evaluation that the key to analysing change is understanding attitudes. But it can be complicated to sort through documents, words and actions to arrive at a judgment regarding one individual's attitudes. In implementation evaluation, many people's attitudes must be understood, and it is expected that attitudes will differ.

Second, Pettigrew notwithstanding, it is a reasonable assumption that rational planning does not equate with real-world implementation. The decision to deliver a

new service, no matter how detailed the planning, is merely a starting point. After that comes bargaining, negotiation and compromise. Even the objectives of a programme can - and often do - change over time, as those involved come to understand more clearly what the programme's strengths and weaknesses are, what the politics of the situation are, and therefore what is feasible. Implementation evaluation is strongest when it maps these progressions.

Finally, both implementation and its evaluation are iterative processes. We have already suggested that the best approach to evaluation is longitudinal, so that the changes expected over time can be identified and assessed. However, longitudinal evaluation is complex to analyse. Both the observer and the observed change; the analyst may become less objective - albeit more knowledgeable - over time. These cautions are not offered to confuse or frustrate; rather to affirm that it is good to be aware of the pitfalls before starting out, and to have a plan for coping with those that arise. For example, many evaluators build feedback meetings into their evaluation plan, so they can test their observations against other people's perceptions of the same experience and thereby increase internal validity.

4.5 Questions about implementation evaluation

Most seminar participants felt that a comprehensive evaluation would be incomplete without some focus on implementation; that, they believed, would be the best way to understand intermediate care in a real-world context. By 'real world' it was clear that they referred to the economic and political pressures in their organisations that shaped, and at times constrained, their decisions about the service they would offer. However, they raised questions about how best to undertake implementation evaluation and whether it was a good use of their limited resources. In this section, we deal with their questions.

4.5.1 Timing of the evaluation

Some workshop participants questioned whether a new intermediate care service would have enough structure to be evaluated. This question is fundamentally about the timing of an implementation evaluation. Often with something as multidisciplinary and potentially diffuse as an intermediate care service, the patterns of assessment, referral and treatment take time to become established. We would argue that, provided a longitudinal evaluation design is used, *the early days of implementation are very important to describe* because that is where the first negotiations occur and the first identification of who resists and who supports the service becomes possible to observe. Such early relationships set the context for how

the service progresses and, indeed, an evaluation may well bring out the ways in which players change their attitudes over time. For example, it may be possible to demonstrate that nurses who began their assignment to an intermediate care service with a diagnosis-orientated, medical-model approach learned, over time, to take a more holistic and empowering view of their patients. This would be visible from the notes they took or the types of observations made during team meetings.

Even a service with a weak structure can - and probably should - be evaluated, because the benefits of early observation outweigh the difficulties of discerning structure where there is none. However, those conducting the evaluation should resist imposing an image of structure prematurely. Instead, they might want to set as one evaluation question, 'At what point did the structure of the intermediate care service come into place?'.

4.5.2 Internal versus external evaluators

A second question was whether evaluators could be internal to the organisation, or had to be external. This question applies to all types of evaluation but is particularly acute for implementation evaluation, where external evaluators are vulnerable to the accusation that 'they just don't understand' and internal evaluators are all too aware of the political pressures they face as they shape their analysis and feedback.

Seminar participants debated this question vigorously. Some said that if your organisation has people with evaluation skills internally (for example, an audit department) then evaluate internally. It will be cost-saving and possibly time-saving because the internal people will have a better baseline understanding of organisational structure, politics and information management. Others disagreed, maintaining that internal evaluation will be biased by definition and that the additional costs of having an external evaluator were well worth it. External evaluators were viewed as better able to convey unpleasant or unpopular information than internal evaluators.

On the other hand, external evaluators were seen to be more vulnerable to response bias, and more at risk of being manipulated by internal politics. Indeed, senior managers aware of problems in their organisations will at times hire external evaluators to validate what they already know, take the flak when reporting the problems, and give them the credibility to remove or change staff or to reorganise activity and outputs.

Both sides of the debate are reasonable enough, so the decision should come after assessment of the particular situation. If the necessary skills are available in house, and there is good communication and agreement among stakeholders regarding the objectives of intermediate care, internal evaluation may be a feasible alternative. However, if the evaluation skills are lacking (see Box 2.2) external support will be required. If budgets are limited, that support may have to come in the form of short-term expert consultancy; however, ideally, an externally contracted team would be used to conduct the entire evaluation (see Box 4.4). Certainly, in a highly charged political environment, an external evaluator who will be seen as an 'honest broker' may be the only source of objective judgment. Even then, those involved in organising or delivering intermediate care should be aware of who has commissioned the evaluation and how the goal of the programme has been described to the evaluation team. In some situations, a team composed of internal and external assessors will confer the most credibility upon the evaluation.

Box 4.4 Possible evaluation team collaborators

- Practitioners in other Trusts or practices
- Research network colleagues, if appropriate
- Health authority public health, policy, or secondary, primary or community care professionals
- Academic units
- Community health council
- Voluntary sector organisations and consumer health groups
- Management consultants
- National agencies such as the British Medical Association, Royal College of Nursing, Royal College of Physicians.

(Adapted from Evans and Steiner, *op. cit.*, footnote 1)

4.5.3 A changing model of delivery

Participants also wanted to know what methods could be used to take account of the reality that, as it develops, the delivery model and its application may change radically. The best way to manage this is to plan for it in the evaluation. Indeed, whereas outcome evaluation does not allow for mid-course modification of the intermediate care model under assessment, implementation evaluation does. It is for this reason that implementation evaluation is well suited to assessments of progress, as compared to effectiveness *per se*.

Evaluations are made with reference to specific, measurable, appropriate objectives. It is a feature of implementation evaluation that it can be designed to identify objectives, revisit their specification, and note whether they have been modified - without reference, necessarily, to whether the objectives have been met. By taking a qualitative approach, evaluators may be as interested in describing the ways in which - and the reasons that - objectives shift over time, as they are in discovering the extent to which those objectives have been met.

4.5.4 Local relevance versus external validity

Finally, a question was raised as to whether implementation evaluations would always be 'local' in orientation, or whether they could be designed to have external validity as well. As discussed in Section Two, evaluators vary in the sort of generalisability they attempt. Quantitative researchers seek statistical generalisability, based on the random allocation of service users to intermediate or conventional care, along with sufficient numbers to detect benefits, if benefits do, in fact, occur. Qualitative researchers seek conceptual generalisability, where the ideas that are generated are based upon in-depth analysis of a varied enough (though small) group of players - or patients - that theories are generated which seem plausible beyond the immediate situation.

Implementation evaluation is most likely to use qualitative methods - for example, semi-structured interviews or focus groups - or to use quantitative tools such as questionnaires within a qualitative design. That is, even when data are collected with survey instruments, sample sizes will not have been calculated with reference to the numbers needed to detect differences between groups, or changes over time; nor will there be a control or comparison group used to estimate counterfactuals. Therefore, it is conceptual generalisability that should be attempted, if generalisable results are desired.

We would argue that they *are* desired, because it is only in this way that evidence can be developed by which the merit of intermediate care services can be judged. Providers and evaluators are urged to take this into consideration, even when the first business at hand is finding out about one's own organisational dynamics. There are always generalisable lessons, and some attention should be given to establishing which ones they are.

Box 4.5 Common issues with implementation of intermediate care services

It may be useful to investigate these issues as part of an implementation evaluation.

- Gaining and retaining ownership by the full complement of stakeholders is a basic requirement of implementing a new service.
- Commitment from health and social services is required, yet that may be overlooked.
- Communication is critical, but can be difficult because people coming from different backgrounds or occupational groups may have different understandings of the purpose of intermediate care. Even vocabulary is often different, so it is essential to confirm understandings and ensure that the channels and mechanisms of communicating about the new service are effective.
- Many people have a fear of the unknown, so those implementing intermediate care often need to be ready to persuade others of its value.
- Risk management is another important issue. By shifting authority for patients to nurses or PAMs, or from hospital to community, new lines of accountability must be created and managed. Also, some say that the intermediate care model tolerates more risk-taking by patients than more paternalistic models of treatment; this question should be tackled explicitly and assessed in each local context.
- In the search for win-win situations, those implementing intermediate care will want to evaluate substitutions and trade-offs in treatment patterns. For example, consultants may be freed to focus attention on acute patients, and PAMs will take more responsibility but may have better career paths.
- It may be easier to demonstrate long-term benefits of intermediate care than short-term gains, yet new services often have to justify themselves in the short-term. Sometimes, a certain volume of patients is needed to achieve cost savings by, say, closing an acute ward; but starting up a new service may require accepting patients from many wards, which diffuses the effects on acute care in the short term. These sorts of dynamics need to be understood and acknowledged.
- Shifting sands, or the turbulent environment of the NHS, require flexibility. Purchasers might change their purchasing package, the new White Paper will produce changes in the organisation of care, and seasonal effects can influence what sorts of care are supported (for example, many places are funding their intermediate care with winter bed crisis funds). Those who would implement intermediate care services may need to juggle long-term goals with opportunistic funding mechanisms.

To enhance the external validity of implementation evaluation in intermediate care, consider discussing your organisation's plans with colleagues from another Trust or health authority; or, if this is not feasible, with colleagues in a another region. Alternatively, you can review the literature to see which issues continually arise and plan to investigate these. Box 4.5 lists some of the implementation issues raised by the intermediate care seminar participants.

In addition, should you choose to use structured questionnaires, it is wise to review the literature for measures with established reliability and validity and choose these if at all possible. Appendix 1 offers some scales that may be appropriate. If you use semi-structured interviews, record the questions and include them in a written report so that others can consider asking the same items. Either in the body of the report or in an appendix, note the range of answers to those pre-specified questions, along with any new items that arose so often that future evaluations should include them specifically. If any aspects are considered confidential, these can be held back; but the report should note that certain sensitive items have not been included.

The mapping exercises referred to earlier (page 32) are another way to enhance generalisability since they enable outside consumers of your evaluation to judge the extent to which their organisational structure resembles yours - and to begin to think about how the differences between organisations might affect their implementation of intermediate care. Evaluations focused on implementation *are* local events, but their usefulness can extend beyond local borders if conceptual generalisability is considered, measures are selected thoughtfully, and findings are shared.

4.6 Ways to strengthen implementation evaluation in intermediate care

In general, the advice we gave for process evaluation holds for implementation-focused assessments: take a pluralistic approach, establish the baseline situation, use a longitudinal design, and keep focused on the analysis. In this instance there are other important considerations as well, summarised in Table 4.2. Following that, we focus on two further crucial issues in implementation evaluation; namely, ownership and dissemination.

4.6.1 The importance of ownership

Seminar participants could not stress strongly enough the importance of ensuring that stakeholders help set the agenda for evaluation and, in particular, that they determine the central evaluation question. Implementation evaluation is a pluralistic enterprise. Different stakeholders will take different views of intermediate care and will have

different ideas about what is important. In addition, it is very difficult not to feel judged, policed and spied upon by evaluators - whether they are internal or external to the organisation. Such sentiments do not make for open communication, which is - of course - what evaluators desire.

Table 4.2
Things to consider in implementation evaluation

Consideration	Issue
Timing	Evaluation tends to be tacked on rather than built in, so often comes too late in the process and misses valuable start-up lessons. Avoid this!
Politics	What is the purpose of the evaluation? Is it to rubber-stamp a programme under way? Or to undermine it? Is it to send a message about the value the organisation places on evidence-based care, or intermediate care? Make every effort to find out.
Expectations	Evaluators need to manage the expectations held by those who commission the work about what is possible to deliver. Consider the time frame and budget and be specific about what you will do. Better to under-estimate and deliver more than the other way around.
Data collection	Decide whether you will design your own data collection tools and budget in the appropriate amount of time.
Data analysis	Similarly, allow plenty of time for analysis. A rule of thumb would be at least 1/3 of allocated time.
Reporting back	Be sure to report back, and negotiate in advance a structure for dissemination. How many times will you provide feedback? Will it be oral or written; to all involved with the service, or only to some?

(Adapted from Dopson, S. 'Implementation evaluation', paper given at the King's Fund Intermediate Care Evaluation Seminar, 29.10.97)

One way to minimise this problem is to engage with key people at an early stage, and to use them to help open doors. It may take some preliminary discussion to identify these key people but eventually it will become obvious whose name comes up most frequently, and in what context. By gaining the support of key stakeholders, the process of involving others in the evaluation - and of ensuring that they have a positive view of the evaluation - will be made infinitely easier. Key supporters can tell you how the political system operates and who else to meet; they can arrange such meetings and introduce you and your activities; they can smooth the way during the initial period of uncertainty about hidden agendas.

Another way to ensure ownership is to hold discussions early on in which stakeholders' objectives and their understanding of intermediate care and its evaluation can be aired in a group, ideally so that consensus can be reached, but if not, so that the range of views is made transparent. Spoken as well as unspoken communication will be valuable here. In general, by making your approach and your purpose as transparent as possible, it will be more realistic to hope for support, because people will know what they are supporting. Early discussions will also help to focus the evaluation task by identifying which issues are key to the people involved and agreeing on priorities. Again, this provides a way for those who form the focus of the evaluation to feel an ownership. It is helpful to encourage people to use the evaluation for their own purposes, and then work hard to make sure those purposes are clearly understood by all the relevant stakeholders.

4.6.2 Reporting back

There are three motives for reporting back in an implementation evaluation:

- to help improve the service;
- to validate findings; and
- to report on the strengths and weaknesses of the service as implemented.

A good feedback strategy will take into account the timing of, and reason for, feedback. That is, formative evaluation will take a different approach to summative evaluations. In all instances, implementation evaluation will be strengthened greatly by sensitive handling of dissemination.

Taking programme improvement first, implementation evaluation is the type best suited to meeting a formative purpose. Its capacity to uncover organisational dynamics may also enable evaluators to suggest interventions to improve them if necessary. To do this, of course, they need to report their findings at several stages during the evaluation. But this is a stage when the politics are most delicate. Although those providing the service will want to know how to improve what they are doing, they may not want others - either in or outside their organisation - to know that there are any glitches at all. They may be resistant to receiving interim reports in writing, or even disseminated to anyone but a small elite group.

Interim feedback mechanisms should be negotiated in advance. The goal is to create an environment safe enough to allow potentially difficult problems to be aired. Evaluators should not hedge their findings. To this end, it may make sense to agree

terms of confidentiality at different stages through the evaluation; for example to brief only a small group of stakeholders or to present results verbally but not in writing at interim points. Providers who feel threatened by senior management or the purchaser's perspective should negotiate strongly for keeping their findings confidential for, say, the first year of operation. However, for their own records - and the final report - it is wise for evaluators to draft internal interim reports that will form a series of analyses in a longitudinal evaluation.

Reporting back for validation purposes is rather different. Here, it is important to disseminate findings not only to the elites among the stakeholders (who may then decide to provide their staff with further information, but information that they control), but also to people 'on the ground.' Again, however, there are choices about the best way to do it. One approach is to use one or more focus groups. These are small groups (perhaps 5-10 people) where discussion is intended to be open and even free-wheeling. Groups can be selected to represent, for example, therapists, doctors and nurses; or patients and carers; or managers (if enough are involved). Although your findings will be reported back to the rest of staff - as gossip at least - the format will allow you to gather the information you need; namely, whether your findings seem credible to the group (and, if not, who/what sort of professional objects and why). A second approach is to conduct several one-to-one interviews with key players, to check the veracity and plausibility of your findings. This is more time-consuming but may be the only feasible approach in an organisation that is concerned about confidentiality. Negotiate this in advance as well.

Finally, at the end of the evaluation, dissemination should take several forms. There should be a written report, designed to be comprehensive in its presentation of the evaluation planning, the resultant plan, data collection measures, findings and analysis. Ideally, this would be distributed openly, both internally and externally. In addition, an open session within the organisation - for example, a seminar held by the relevant Trust department or directorate - could be held. This provides closure for the many participants in the evaluation; it also gives them a clear understanding of the lessons to be derived. Another approach is for dissemination in the form of papers presented to professional meetings and articles drafted for professional journals. These would need to be appropriately anonymised, if the commissioners of the evaluation desire.

It is very difficult for participants in an evaluation not to feel judged. Indeed, they *are* being judged, and the stakes are often high. If you end up singling out a particular

professional group for praise or, especially, for criticism, it must be done with great care. Be certain to link criticisms with constructive suggestions for improvement. If the problems you uncover are extremely grave, they may best be saved for private sessions rather than an open meeting. In all dissemination, be sure to reinforce the intrinsic value of evaluation; **to learn what works and to help with future service development.**

Section Five: Outcome Evaluation

Box 5.1

Highlights of this Section

- Outcome evaluation is concerned with establishing a 'cause and effect' relation between two variables. It should be differentiated from 'outputs' of a service such as number of patients seen, or number of visits made.

- Outcome evaluation is sometimes known as summative evaluation in that it 'sums up' the effects of change. Hence there is no 'mid point' at which results can influence or redirect developments.

- There are mixed views about the appropriateness of outcome evaluation for intermediate care because of the holistic nature of the service although there is strong support for seeking ways in which measurable outcome can be identified and evaluated.

- Methods of choice in outcome evaluation are ;
  i) the randomised controlled trial - where patients are randomly assigned to an experimental or control group and comparisons are made against agreed, pre-determined outcome
  ii) quasi experiments - where comparisons are made between an intervention group (in this instance those receiving intermediate care) and a 'comparison group' who are matched as closely as possible but not offered the service
  iii) before and after studies - where data is gathered prior to the introduction of a new factor (intermediate care) and after implementation, holding the agreed outcome measure constant.

- The RCT is seen as ideal when possible since there is a greater capacity to compare like with like and to estimate the counterfactual situations. It is less easy to control for extraneous variables in the quasi experiment. Before and after studies have the added problem of time lag, introducing other factors which may influence outcome but whose effects cannot be confirmed.

- Which outcome measures are used may depend on the specific interests of the commissioners (or stakeholders). They may focus on organisational, professional or clinical issues but care must be taken not to overload a study by attempting to respond to too many questions within a single investigation.

- There is a need to differentiate between 'equivalence' and 'gain'. For example the new service may have the same outcome as traditional care but be cheaper (equivalence) or there may bean actual gain such as fewer admissions, more satisfaction or a greater degree of independence.

Box 5.1 (cont.)

- It is important to remember that there may be a 'trade off' if results from outcome studies are positive for one party but problematic for another. For example, the knock-on effect for community services of a shorter length of hospital stay must be considered.

- Some outcome data may be readily available through records but care must be taken to ensure that the information is reliable (accurate) and valid in relation to the question posed.

- It is possible, with the help of a statistician, to calculate in advance the number of people who need to be included in an outcome study in order to obtain meaningful results. This may create difficulties if the population under study is small or the length of time it would take to gather sufficient numbers is greater than the length of time available for the study. In such situations, it may be better to leave outcome evaluation to others.

- It is also necessary to ensure that the nature of the service under investigation is stable for a sufficient length of time before evaluation to make the assessment a meaningful test of effectiveness.

- It is always worth considering whether there are any pre-validated tools which can be used in outcome studies, and to make use of these. However it is important to check that the tool has been developed to measure the desired outcome for the study in question.

- Bear in mind the need to pilot any outcome study, determine the sample size scientifically and ensure that the right comparison group is used.

- And finally - remember that any outcome evaluation will only paint part of the picture. Without process data it is not possible to replicate and, with services as complex as intermediate care, it is always necessary to take contextual issues into account.

5.1 The central role of outcome evaluation

At the seminar on which this document is based, one group was asked to vote for which sort of evaluation they would undertake, if they could afford to do only one. All but one person chose outcome evaluation. The dissenter said she would do first process, and then outcome evaluation - regardless of what could be afforded!

This anecdote is presented to convey the centrality - even the inevitability - of a summative or outcome-based approach to evaluation. This was the type of evaluation that raised the strongest feelings and the most criticism. Many people engaged in

developing or providing intermediate care services felt inappropriately forced to undertake outcome evaluation, and chafed under the pressure to do so. Despite difficulties however, there seemed to be consensus that sooner or later a thorough evaluation strategy would have to include outcome measurement, and a recognition that the optimal form of outcome measurement would be based on a quantitative comparative model.

5.2 What is an outcome, and what is meant by 'outcome evaluation'?

An outcome is an endpoint. In the intermediate care context, it is a measure of the effects of an intermediate care programme on, for example, users' health, function, quality of life, or use of services; carers' well-being; or staff morale. Outcomes should be distinguished from outputs, which refer to activities (for example, number of visits by a district nurse, number of sessions with a physiotherapist). Instead, they refer to what actually happens to people, or even to an organisation - *not what is done, but how it turns out.*

Outcome evaluation, then, is an approach designed to judge the effectiveness of an intervention (here, an intermediate care programme) in achieving desired, pre-specified outcomes, and to determine whether that intervention produces those outcomes more often than would be expected to occur just by chance. The jargon is that one is trying to discover whether the intervention is 'significantly associated' with the outcome of interest. *The purpose is to establish causality* - to eliminate or control for[21] possible other reasons for observed differences in patients' health (to take one outcome measure), until it is reasonable to conclude that it was the intermediate care programme that made those differences. Thus it becomes possible to estimate what would have happened to a different set of patients, had *they* received intermediate care (this is known as estimating the counterfactual). Because of its unique capacity to do this, and to develop analyses that are based on patient samples that are statistically representative of a larger population of interest, outcome evaluation can provide crucial information to clinicians, managers and policy-makers

[21] To 'control for', or 'adjust for', something means to use a statistical method, often one called *regression modelling,* to hold constant other factors that might influence the outcome you care about in order to isolate the influence of the supposedly causal factor (whether a patient received intermediate care or not). For example, if you would expect older people to have worse function than younger people (on average, and all else being equal) and if you would expect older people to be more likely than younger people to receive intermediate care, you would want to adjust for age in your evaluation of intermediate care's effect on patients' function. Otherwise, if you observed that intermediate care was associated with worse function, you would not know whether that result was owing to an ineffective service or to the fact that the intermediate care patients were mostly older than those who had usual care.

alike - but perhaps especially to those who must think in terms of populations rather than individuals.

5.2.1 What outcome evaluation is *not*

Judging from the comments arising at the seminar, there appears to be widespread confusion between an evaluation designed to determine effectiveness through an attempt to isolate the impact of intermediate care upon outcomes of interest, and an evaluation that reports as part of its findings numerical information - for example, mortality rates, readmission rates, length of stay or pounds spent per patient - of the sort typically referred to as 'hard outcomes.' Although a 'hard outcome' means one that is readily measurable, any form of evaluation can use such measures. What distinguishes outcome evaluation from other approaches is the use of an experimental design which makes statements of causality possible, or a quasi-experimental design which identifies associations between process (defined here, as 'intermediate care or not') and outcome.

5.3 Methods of outcome evaluation

Outcome evaluation has also been termed 'summative evaluation' because its purpose is to *sum up* the effects of a change. Therefore, in contrast to implementation evaluation, the data are not analysed until they have all been collected, and mid-course dissemination is to be avoided. For this test of effectiveness, the evaluators take care to remain outside the process of delivering care, and do not try to alter it along the way. Instead, they allow the process to unfold as it will, collect information about the patients, the service, and the effects (or outcomes) and, at a certain pre-determined time or times, compare what has happened to patients treated by the intermediate care service to what has happened to patients who were not.

5.3.1 The randomised controlled trial (RCT)

The randomised controlled trial (RCT) is "generally regarded as the most scientifically rigorous method of testing hypotheses available in... health services research."[22] It is a longitudinal experimental design that assigns people at random to 'treatment' or 'control' conditions. Here, the treatment would be intermediate care; the control condition would be usual care - although more will be said about that below (see section 5.8.3). The reason to use randomisation to decide who goes into which group is that it will tend to balance the various patient characteristics that might affect outcomes - age, gender, type of health problem, level of social support, other

[22] Shepperd S., Doll H., Jenkinson C. (1997) Randomised controlled trials. In Jenkinson C. (Ed.), *Assessment and Evaluation of Health and Medical Care: A methods text.* Buckingham: Open University Press, p.6.

things you haven't thought of or cannot readily measure - roughly equally in the treatment and control groups. This, in turn, will make the comparisons between groups more fair than if one group had all the women, or most of the severely ill, or none of those living alone. The jargon is that randomisation will 'minimise bias'; that is, it will give a fair view. Randomisation is also considered the best way to obtain a sample of people whose mix is most representative of the larger population of interest.

5.3.2 Observational studies

In some cases (for instance, for ethical reasons) it will not be considered feasible to randomise patients to one condition or another. An alternative approach to the RCT is an observational or quasi-experimental study design, in which two populations are distinct but not randomly allocated. For example, there may be a neighbouring trust with similar demography, which has not as yet implemented any intermediate care services. If you can follow patients there, who are broadly comparable to those receiving intermediate care in your trust, then you have a natural comparison group for your evaluation. Strictly speaking, these two groups are called 'intervention' and 'comparison' rather than 'treatment' and 'control.' With observational studies, evaluation can proceed much as with an RCT; however, it is even more important to consider in advance the potential confounding factors - the factors likely to be related both to the choice to provide intermediate care and to the outcome of interest - so that these may be measured and accounted for during analysis.

5.3.3 Before and after studies

A third study design is known as 'before-and-after', or 'pre-post'. It does not use a contemporaneous comparison group, but makes its comparison between two time points instead.[23] In these studies, patient and/or organisational characteristics, and the outcomes you expect to change, are measured for a period prior to the new service, that is the intermediate care programme, operating. These data can either be collected prospectively or retrospectively, from patient records. Otherwise, the data can be collected retrospectively from patient records. Subsequently, after the service has been operating for a while, the same set of measurements are taken for patients who received treatment under the new system. The 'after' group is compared to the 'before' group. This is a not uncommon approach. However, compared to RCTs and observational studies, before-and-after evaluations are much less satisfying.

[23] The word 'instead' may be a misnomer, because many RCTs include both the 'treatment vs. control' and 'pre vs. post' comparisons.

5.3.4 The relative advantage of comparison group designs

The problem with evaluations that do not use comparison groups is that when a change is observed, there is no way of knowing whether the intervention - the new service - caused that change or whether there were trends operating in the direction of change anyway. If, for example, other changes in health care delivery improved the coordination of services across secondary, primary and social care boundaries, your finding that patients who received intermediate care experienced better coordination than patients treated two years earlier would be appropriately open to criticism. There would be no way of knowing whether the intermediate care was beneficial or the improvement was due to better services across the board.

Similarly, during a two-year study period in which you observed no differences in, for example, patient satisfaction, there may have been a downward trend outside your unit which the intermediate care patients did not experience. In that case, a finding of no effect would be misleading. The intermediate care service would have been a success, relative to other care, but because nobody had measured the background phenomena, this success was missed.

Whatever the (many) benefits of process or implementation evaluation, only a comparison group design can establish effectiveness. Whether through randomisation or by means of strong statistical adjustment for potential confounding factors, comparison group designs form the basis for good policy development in a way that no other approach can manage. Only a comparison group design can be informative with regard to the crucial questions, 'What would have happened to *these* intermediate care patients if they had received usual care instead?' and 'What would have happened to the patients who did *not* receive intermediate care, if they had done?' By answering these questions, decision makers can draw valid conclusions about the likely benefit of providing or extending a service to others.

5.4 Which questions are best answered by outcome evaluation ?

Table 5.1 presents some of the key questions about intermediate care that outcome evaluation is well suited to answer. (Readers may recall that Table 3.2 also listed questions that would require outcome evaluation.) Implicit in these questions is the concept of a bottom line - of concrete, measurable markers that can be compared across populations, services, or settings. In addition, most of the questions can be rephrased in a comparison-based format; for example, 'Do patients who receive intermediate care have *better* results (such as function, mobility, or quality of life) than similar patients who do not receive intermediate care?'

Table 5.1
Key questions about intermediate care that outcome evaluation can answer

Area of Interest	Key Questions
Health outcomes	• Does intermediate care improve patients' function? mobility? general health? quality of life? coping skills? • Does intermediate care improve patients' discharge destination? For example, does intermediate care reduce admissions to residential care? • Does intermediate care reduce hospital length of stay? • Does intermediate care reduce hospital readmissions? • How sustainable are observed outcome improvements? • What are the appropriate admission criteria for intermediate care, i.e. which patient types benefit most from intermediate care?
Organisational effects	• For post-acute models, which mix of staff produces the best results? • Does intermediate care increase dependency in the community? • What is the effect of intermediate care on activities in other parts of the service (e.g. in acute wards)?
Best models	• How transferable are intermediate care services, i.e. will the outcomes observed in one setting generalise to another setting?

The first area of interest refers to patients/service users. Can intermediate care affect how people feel, or what they are able to do? Seminar participants emphasised function and quality of life as more important outcomes than mortality or clinical markers such as blood pressure or insulin levels. Implicitly, they also drew a connection between intermediate care and living arrangements, reflecting a consensus that intermediate care might be able to help people maintain less dependent living arrangements than would otherwise be possible. Certainly, process evaluation can also investigate users' health outcomes - it does so by exploring perceptions of health- but only outcome evaluation can standardise their responses and use them to compare the effectiveness of one service with that of another. Hospital-relevant outcomes such as length of stay and readmissions (number, rate, or days) were also seen as reasonable markers of effectiveness. However, some participants asserted that length

of stay in particular was of greater interest to consultants and management than to patients and their families.

The last questions in the health outcomes section of Table 5.1 target more specialised issues. There is, first, the important issue of how long it takes before benefits begin to appear, and how long they last. This has implications for the study design, especially in terms of how many measurement time points to include. Generally, the longer-term the outcomes, the more difficult it is to argue a causal relationship between process and outcome, as many things relevant to the outcome other than the original process intervention can occur during a long follow-up period. Second, the crucial question of which patients benefit was seen to strongly influence decisions about evaluation design. Participants observed that unless they knew how to set their admission criteria, it would be difficult to define appropriate measures of effectiveness. For example, if physically disabled patients were accepted, mobility at discharge would be an irrelevant outcome. If cognitively impaired patients were accepted, it might not make sense to include subjective measures of success (which rely on individuals reporting their own impressions) unless a dementia expert were available to help interpret responses.

A second category of key questions refers to organisational effects. The first of these, about optimal staffing, would require an evaluation plan that compared different mixes of professions, grades and/or years of experience. Techniques from operational research such as linear programming, in which a 'best' solution given pre-defined constraints is derived, may be useful.[24] If the question is somewhat simpler - for instance, which is better for patients' function at discharge, nurse-led care or multidisciplinary team-led care? - then a more straightforward comparison-group trial may be performed.

The other questions about the effect of intermediate care on the organisation targeted the important issue of trade-offs. Perhaps intermediate care shortens hospital length of stay, but are the discharged patients a more dependent group who require community health or social services to a greater extent than patients treated more conventionally in hospital? Inside the hospital, does the presence of an intermediate care unit increase or decrease workload, shorten average length of stay, or affect nurses' or consultants' morale on the referring wards? To answer these questions, data collection would have to extend beyond the intermediate care unit. Because this

[24] Stokey E., Zeckhauser R. (1978) *A Primer for Policy Analysis*. New York & London: W.W. Norton & Company.

will increase the costs and complexity of the evaluation, it is worth clarifying in advance that these really are the questions that are most important to resolve. There may be some capacity to explore questions of trade-off qualitatively. However, if this is the evaluation issue with top priority, it is best to budget for the additional requirements of cross-sector data collection (including adequate sample sizes).

Finally, seminar participants wanted to know, and thought that outcome evaluation could prove informative, about the portability of the intermediate care model. The question noted in Table 5.1 can be operationalised in numerous ways. One - 'what exactly *is* the intermediate care model?' - is better dealt with through process evaluation. However, outcome evaluation can certainly tackle questions such as 'Is a post-acute unit on the main hospital site as effective in producing desired outcomes as a post-acute unit on a satellite site?' or 'Does a community-based admission avoidance scheme in Region A reduce hospital admissions by as great a proportion as an admission avoidance scheme in Region B ?' Again, successful evaluation depends upon a relatively complicated data collection effort that must be carefully planned. For example, comparisons between northern and southern England may be of less interest, at least at first, than comparisons between two areas with quite similar population and disease profiles. A single trust may be unwilling to create two nurse-led units in order to test whether they are equally effective. Thus it may be that two trusts will want to negotiate a collaborative evaluation that will allow the desired comparison to be assessed. These arrangements will take time, but can be extremely fruitful. In fact, some seminar participants commented that the 'state of the art' seemed to have reached the stage where single-site studies could no longer satisfy, and that well-coordinated research of a larger scale was now required.

5.5 Discomfort with outcome evaluation in intermediate care

Although there was widespread recognition that outcome evaluation *is* the policy-relevant gold standard, there was considerable discomfort with taking this approach. Seminar participants were acutely sensitive to the limitations and challenges of undertaking outcome evaluation in an area like intermediate care. Some simply bemoaned the fact that their service was under pressure to prove itself, just because it was new (see table 5.2). Others questioned whether it was even possible to measure outcomes, especially the outcome called 'health gain.' This central issue is discussed further in section 5.7.

Table 5.2
Questions about Outcome Evaluation

Area of Interest and Question Asked	Answers and Comments
Justification for outcome evaluation	
• Why must we prove ourselves in this way? Is it just because we are a new service?	• Although frustrating to 'believers', it is true that intermediate care is considered a new service without demonstrated - or accepted - benefit
• Is health gain even measurable?	• Yes - it is difficult but important to measure
Politics	
• Who wants the data? Who is it for?	• This will influence the selection of outcome, but sometimes it is possible to influence the process or to design an evaluation that serves multiple masters
• How does one weight the many potentially different effects of intermediate care - for example, conclusions drawn based upon professional cultures, organisational politics, NHS policy demands or patient diversity (e.g. ethnic issues)?	• There are formal and informal techniques, from cost consequence analysis[25] to Delphi or nominal group technique[26] to qualitative treatment of any observed diversity to explicit acknowledgment that one particular perspective dominates your evaluation
Data	
• Do medical records data provide accurate information, or are they idealised? • Are medical records data pertinent to patient-centred outcomes?	• Data reliability is a serious issue and accuracy of records must be checked • Some data will be patient-centred - e.g. post-operative complications or discharge destination - but other important information, such as satisfaction, quality of life or advanced functional capacity, may be missed out
Methods	
• What methods exist for outcome evaluation aside from the RCT? • Do we have to use RCTs? What about matched pairs instead? • What if we evaluate intermediate outcomes as well as final outcomes? • Has an outcome evaluation been performed if organisational/human resource effectiveness, rather than clinical effectiveness, is measured?	• A fair alternative is the observational study • These are not mutually exclusive; matched pairs can be acceptable. • See section 5.6.2 • See section 5.6.3

[25] Drummond M. (1994) *Economic Analysis alongside Clinical Trials.* Leeds: Department of Health.
[26] Bingham R.D., Ethridge M.E. (Eds.) (1982) *Reaching Decisions in Public Policy and Administration: Methods and Applications*, New York and London, Longman Inc.

Table 5.2, continued

Analysis	
• Can single-site evaluations be linked, either to gain power by increasing sample size or to gain insights from findings across potentially diverse settings?	• Yes, although much depends on how each study was designed and implemented. It is strongly recommended that, depending on whether you want to gain predictive power or develop qualitative themes, a statistician experienced in meta-analysis or a social scientist experienced in qualitative research be consulted
• Is it possible to make a summative judgment about intermediate care when the outcomes conflict - if, for example, there is functional decline but increased subjective quality of life?	• The same issue is raised under 'politics'; most quantitative analyses select a 'main effect' on which to base a central analysis, then augment it with analyses of 'secondary effects'. See section 5.6.4.
• Whose perspective should matter more, the patients' or the professionals'?	• Probably this will be a subjective decision; note that the question assumes conflict between these, but is this assumption always justified?
Evaluator expertise	
• How can local evaluators access needed expertise?	• Contacting your NHS Research & Development Support Unit (RDSU) may prove very useful

There was a perception among many that the impetus to conduct outcome evaluation nearly always came from a source external to the intermediate care service itself, and further, that the commissioner's success criterion was either inappropriate to providers' concept of the service or, worse, might conflict with it. For example, some purchasers were willing to contract for intermediate care only if it reduced total length of stay, whereas the manager of the service believed that somewhat longer lengths of stay would lead to a more self-sufficient patient at discharge and that this result constituted a highly desirable outcome. Such conflicts need resolution. Those who have gained experience with intermediate care may wish to share ideas with others as to which outcomes are desirable to assess, and why. The objective would not be to dissuade anyone from conducting outcome-based evaluation, but to more appropriately define what would constitute success.

5.6 Questions about outcome evaluation

Other questions raised by participants, and noted in Table 5.2, refer to study design, data, methodology, and analytic approach. The questions about study design have been touched upon already; to reiterate, RCTs are a highly desirable form of outcome evaluation but may be difficult to arrange, while pragmatic trials with appropriate comparison groups and strong statistical adjustment for differences between groups are an attractive alternative.

5.6.1 Data

Questions about data focused on the reliability and validity of the patient record. Those in the field understood that data frequently were missing, and that even recorded information was sometimes suspect (for example, on discharge planning forms, having every item ticked 'yes' when it was clear on the wards that there was rarely time to complete the assessment and consultation process prompted by those forms) and often idiosyncratic (some doctors noting all laboratory tests ordered, along with results, other doctors noting only those results which required follow-up). Possibly more important, however, was the concern that even a perfectly completed patient record was too biomedical in focus and thus not patient-centred.

In the community district nurses, physiotherapists, and social workers from care managers to home care assistants tend to hold their own records. This makes the data extremely diffuse and time-consuming to collect. It is particularly difficult to gather information from the social care sector. Participants also commented that cross-trust data were virtually impossible to acquire. On the plus side, however, community-based practitioners tend to remember their patients/clients very clearly; if the evaluation team can arrange to collect data about contacts with a patient quickly enough, the probability of its being reliable and comprehensive may be greater than in the acute hospital.

When thinking about data collection, it is helpful to anchor it first in the research questions and evaluation objectives you have set. Objectives should be 'SMART **s**pecific, **m**easurable, **a**chievable, **r**ealistic, and **t**imely.[27] It becomes very tempting to collect all sorts of information, once one starts, but the temptation needs to be avoided. If data requirements are linked to evaluation questions, it then becomes possible to consider where one might locate the desired information (see Table 3.3, in Process Evaluation). Data sources need to be piloted, in order to confirm that what is

[27] See, for example, Evans D., Steiner A. *op cit.*, footnote 1.

supposed to be recorded, has been. In addition, when abstracting data from a patient record, it is good practice to test the inter-rater reliability of the process by having two or more people abstract the same items from a common set of records (say, 5-10% of the whole) and to determine the agreement rate. If it is not very high - 95% or better - there is genuine cause for concern. Equally, once data have been abstracted, it is best practice to double-enter it into the computer, in order to discourage typographical errors (it is very likely that someone will make a mistake, but highly unlikely that they will make the same mistake in exactly the same way).

Box 5.2 lists types of data that workshop participants identified as available, although irregularly so and with variable relevance, or desirable to obtain.

Box 5.2 Available or desired data in outcome evaluation of intermediate care

Easily available:	Hospital length of stay Hospital readmissions Mortality rates A&E statistics Discharge destination, including admissions to residential care
Desired, not always available:	Social services data, such as agency use Primary care use, including night use Patients' physical function, quality of life, self-control, cognitive status, social networks, rehabilitation or recovery potential Carers' stress, quality of life
Usually available, not always desired:	Barthel Index; seen as too limited in that only basic physical functioning is stressed

5.6.2 Intermediate outcomes

An intermediate outcome is a measure which, though not itself the desired endpoint, is known to be associated with that endpoint. A popular example is to take smoking cessation as an appropriate intermediate outcome, when the goal is to reduce rates of lung cancer, because the link between cigarette smoking and lung cancer is well established. In the case of intermediate care, if - for example - the goal were to improve patients' health outcomes after orthopaedic trauma, but it was not possible to monitor patients after hospital discharge, appropriate intermediate outcomes could

include the time between fracture and corrective operation, the time between operation and mobilisation, and the level of mobility at discharge. The first two factors have been demonstrated to be significantly associated with patient mortality at three months[28] ; the last factor, with patient function.[29] The intermediate outcomes are, in effect, process measures and are sometimes known as 'process outcomes' (just to really confuse things!).

The advantage of intermediate outcomes is that they are more amenable to short-term evaluation and offer a pragmatic alternative to summative evaluation. However, the usefulness of intermediate outcomes is contingent on there being *known* links to the desired objectives of the programme. Ideally, specification of intermediate outcomes should be evidence-based; in the absence of evidence, some experts believe that 'knowledge-based' criteria using expert professional consensus is a fair substitute.[30] Each evaluation team will have to set its own standards for acceptable associations between particular intermediate and summative outcomes; it is good practice, however, to clearly justify the choice that is made.

5.6.3 The relation between organisational effectiveness measures and outcome evaluation

Although most providers of intermediate care have patient needs at the centre of their concerns, patient outcomes are notoriously difficult - and certainly quite costly - to measure. Moreover, in at least some cases, patient results are frankly less important to the sponsors of an intermediate care programme than certain organisational outcomes, such as reduced costs, shorter lengths of stay, or increased skills among nurses and PAMS. Table 5.3 lists other potential outcomes, as they link to hypothetical programme objectives.

If an intermediate care service's primary objectives are to improve organisational effectiveness, then it is perfectly acceptable to have an outcomes evaluation to judge those effects. In contrast, the appropriateness of using organisational outcomes as intermediate measures of clinical effectiveness is contingent on the link between some aspect of organisational performance to patient outcome having been well-established. It is often assumed that an efficient service equates with a good quality service, but

---

[28] Todd C.J., Freeman C.J., Camilleri-Ferrante C., Palmer C.R., Hyder A., Laxton C.E., Parker M.J., Payne B.V., Rushton N. (1995) Differences in mortality after fracture of hip: the East Anglia audit. *British Medical Journal*, 310:904-908.

[29] Steiner, A. (1995) unpublished data.

[30] McGlynn, E.A. (1997) Six challenges in measuring the quality of health care. *Health Policy*, 16:4-21.

that has not been proven. However, in the US at least, comprehensive geriatric assessment (CGA) by a multidisciplinary team has been clearly associated with improved patient outcomes;[31] so the proportion of patients to receive CGA would be a legitimate intermediate outcome at the organisational level, for an older patient group. As with any outcome, it is essential that you know in advance how you will measure it, and that the measure you choose is reliable, valid, and responsive to change. That is, it will be able to reflect progress towards the programme's objectives if progress has, indeed, been made.

Table 5.3
Possible organisational outcome measures

Potential intermediate care objectives	Possible outcome measures
To reduce costs by reducing hospital days	• Length of stay, readmission days within 30 days of discharge
To increase responsiveness to local need	• Innovations to organisation and delivery of care resulting from local consultation
To develop a multidisciplinary team	• Formation of a team • Proportion of patients treated according to team approach • Satisfaction measures for patients (/users) and professionals
To create an assessment service	• Existence of staff trained in comprehensive needs assessment • Proportion of patients (/users) assessed • Number of problems identified and treated
To shift budgets from acute to community trust	• Changes to contracts • Unified budget

(Adapted from Evans D, Steiner A. *op cit.*, footnote 1)

5.6.4 Analysis of multiple outcomes

Most quantitative analyses depend upon a simplifying study design in which a single outcome is accepted as having the top priority, and all other measures are considered secondary to it.[32] Although one would hope that the 'main effect' - as the jargon goes - would equate to the outcome agreed to be most important, often an outcome is

[31] Stuck A.E., Siu A.L., Wieland G.D., Adams J., Rubenstein L.Z. (1993) Comprehensive geriatric assessment: a meta-analysis of controlled trials. *The Lancet*, 342:1032-1036.

[32] A set of main effect analyses can be planned, each with a different outcome; in this case, the outcome requiring the largest sample would determine sample size, with other analyses able to obtain greater precision.

chosen either because it has high policy relevance (for example, length of stay) or is easy to measure (for example, length of stay!). Seminar participants were frustrated by this approach to evaluation, which they viewed as simplistic and inappropriate.

They were more interested to learn how to combine complex information on multiple outcomes, implying a preference for measuring performance from several perspectives. Box 5.3 offers some ideas, including formal and informal techniques.

Box 5.3 How to weight different, or conflicting, findings

- Select only one outcome to focus on, and acknowledge that the study is limited. (All studies are.) This is an implicit weighting of that outcome against all others.
- Hold findings separate and qualitatively describe the apparent impacts of each. A more quantitative version of this is a cost consequence analysis, to be discussed in Section 6.
- Hold findings separate, but report them differently to different audiences, according to which outcomes you believe will matter most to a particular group. (This is a somewhat risky approach.)
- Use a formal group process such as the Delphi, modified Delphi, or nominal group technique[33] to establish consensus on which factors matter most. These techniques vary, but the general idea of the exercise is that participants first identify, privately, how important they think various factors are. Then they either discuss it in a group or are shown what other people thought before having a second chance to value the outcomes of interest. These final judgments are aggregated into a ranking.
- Conduct a study to elicit utilities from relevant players - whether professionals or service users. 'Utility' is defined as the amount of 'happiness' to be derived from something; it can be quantified, but in a subjective way - e.g. 'on a scale from 1 to 10' and used to place value on diverse outcomes such as changes in function, mobility, living arrangement, capacity for self-care, efficiency of bed use, professional development or cost of hospital treatment. The process involves imagining trade-offs until a point of 'indifference' is reached ('I think that saving two days in hospital is about as good as increasing a person's well-being score by 20 points"). If you are interested in utility-based analysis, you will want to consult an economist.
- Sometimes people simply list the benefits and losses in two columns. Although it might seem that the longer column 'wins', often what happens is that people realise how much they care about different factors - perhaps the only thing in the 'benefit' column is improved satisfaction with care, but those reviewing the list come to understand that user satisfaction is extremely important to them.

[33] Bingham R.D., Ethridge M.E. (Eds.) (1982) *op.cit.* footnote 26

Although one hopes that multiple measures will produce an image of clear dominance for or against the intermediate care intervention being evaluated, this is not always the case. Should outcomes conflict - for example, lower costs but less acceptable to staff - the interpretation of data can be challenging. Ultimately, the decision about how to manage it may be a product of practical and political constraints. Ideally, however, it would be the time, budget, and expertise available which would determine the sophistication of the evaluators' analytic approach, along with their judgment of whether or not it is essential, in their local situation, either to weight (/rank order) different outcomes or amalgamate them into the single outcome called 'failure or success'.

5.7 Outcome measurement

The final area of concern raised by seminar participants - outcome measurement - merits a separate section. Entire books and journals have been devoted to the subject (see Appendix 3); what is intended here is a brief discussion of the issues which participants identified as most central to their efforts to evaluate intermediate care (see Table 5.4).

5.7.1 Which outcomes?

By virtue of its whole-person approach, intermediate care is a treatment with multiple relevant outcomes. Because it is one element in a larger health care system, there are organisational and professional outcomes to consider as well. The glib answer to the question, 'which outcomes are most important to measure?' is 'all of them'. Obviously, this strategy rarely makes sense, so in the absence of examining all possible outcomes and integrating them into a single metric, evaluators must - MUST - narrow their focus and live with their choices. How, then, to do this?
First, it is useful to target outcomes for which reliable, valid and responsive measures exist. Appendix 1 provides a brief compendium of such measures, with comments regarding their appropriateness for assessing intermediate care.

Essentially, there are several types of health outcome measures, all of which could apply under the right circumstances:

- ***generic measures***, which cover overall health or well-being; for example, the Nottingham Health Profile or the SF-36. The advantage of generic measures is that they allow comparison across diverse health programmes or interventions. The disadvantage is that they can easily miss important changes. For example, there may be only one or two questions about entire domains of health, such as 'mental well-being' or 'pain'.

- ***condition- or disease-specific measures***, which target one illness or set of symptoms; for example, the Arthritis Impact Measurement Scale or the Diabetes Quality of Life Measure (see appendix 1). These will be better able to detect change than the generics, provided the measure is appropriate to the patient or intervention.
- ***dimension-specific measures***, which target one aspect of health, such as physical functioning or quality of life. Again, these have a better potential to accurately capture particular elements of a person's health, as well as changes in health; however, some of these domains are themselves subjective, even vague, and quite difficult to measure.
- ***ad hoc clinical indicators***, such as blood pressure or insulin levels. In some intermediate care interventions, these may be appropriate. More often they will not be as they are too narrowly focused. In addition, they may require professional input, which will raise data collection or analysis costs.
- ***patient-generated outcome measures***, for which patients specify what effects and what level of effect would be meaningful to them personally.[34] These are then used to measure baseline and post-intervention status in individually relevant ways. This type of scale has been developed in recognition of the discrepancy between personal and professional assessments of well-being for people with chronic disease, mental illness or permanent disability. They are especially sensitive to change; however, it is virtually impossible to use them to draw comparisons from patient to patient, group to group, or intervention to intervention.

Second, the selected outcomes must satisfy the commissioners of the evaluation but should also serve the people who are providing the intermediate care. At times, these choices will overlap. However, judging from what seminar participants told us, often they will not.

In this respect, evaluators have some choices. They can simply conform to the requirements of those who commissioned the work. Alternatively, they can supplement that with one or more outcomes of high value to other stakeholders or try to persuade those commissioning the evaluation that the outcomes first stipulated are inappropriate. In any event, it is important to be clear about whose perspective is taken in defining success (see Section 4, Implementation Evaluation).

[34] Steiner, A. (1995) unpublished data.

Table 5.4
Questions about outcome measurement in intermediate care

Issue	Question asked
Defining appropriate outcomes	• Which outcomes are we interested in? • Should health gain be defined as improved function or improved coping (e.g. with pain, or by family)? • How can we unpack the concept of gain, so that it comprises outcomes that are clinical, social, cost, risk, value for money, access to other services, domino effect on other areas of life, burden on carers, readmissions, quality of life-related, and more?
Establishing levels of success	• Are we trying to establish better outcomes or equal outcomes, or possibly even a slower rate of decline? • What values should be used as indicators of success? • What about ceiling effects?
Considering who shall choose the outcome to be measured	• Who determines which outcomes are the most important? • What outcomes matter to recipients of intermediate care? • Are less tangible gains, e.g. feeling cared for, more real to patients than outcomes such as number of contacts with professionals or length of stay? • How do the outcomes of interest to the trust compare to outcomes of interest to the patient? Can an evaluation include both?

Finally, it is well worth taking the time needed to consider carefully where your particular intermediate care programme is most likely to show benefit, and to measure that. Indeed, the importance of this cannot be over-emphasised. If those developing a service have a feel for where their relative advantage lies, they are strongly advised to include that area as one of the primary outcomes for evaluation.

5.7.2 How good is good enough?

Two issues were raised under this general heading. The first is sometimes referred to as the question of ***'equivalence or gain'***. Is the goal of intermediate care to actually improve health or organisational outcomes, compared to conventional practice, or is it to maintain outcomes but possibly at a lower cost to at least someone, or with greater acceptability to at least someone? Study samples have to be larger when the goal is to

prove equivalence (i.e. no difference) than when the goal is to establish a difference of a given size (see 5.8.2.). However, the equivalence outcome may be more realistic to achieve and considered sufficient justification for maintaining an intermediate care option as part of the continuum of care.

Even the statement of the question, however, reveals some important insights. For one thing, it is '***equivalence or gain***' with reference to some particular outcome - functional status, mortality, number of admissions or readmissions to hospital, the proverbial length of stay. Other outcomes *will* be different, by definition - intermediate care is a change of practice - so there will always be some gain (or loss) inherent in the new service's implementation.

It is a two-part problem, then; the evaluation must be designed with reference to a primary outcome, which may be more likely to remain approximately constant between intervention and comparison groups (equivalence) or to show a difference between groups (gain). But, in addition, the evaluation should also be designed to measure those areas where a difference is known to exist. For example, in the case of a nurse-led unit on a satellite site, there are the costs of transferring patients, the savings of paying nurses less than doctors, the reactions to potential power shifts arising from a transfer of clinical responsibility from consultant to nurse, to mention but a few. Even if there were no observable effects on patients' health or satisfaction with care, nor on their length of stay or contacts with therapists (again, to mention but a few possible outcomes of interest), the fact of providing the new service carries with it changes that could be measured. There is no such thing as perfect equivalence.

For another thing, equivalent or improved health outcomes are only two of the possible outcomes for patients. Evaluations can also target the rate of decline, and investigate whether intermediate care slows the rate of decline. This is a different type of 'gain' which requires a slightly different type of analysis. ***Consulting a statistician is strongly recommended*** before settling on an analytic strategy.

The second issue is that of ***statistical vs. practical significance***. Outcome evaluation designs set out to test hypotheses about the relationship between predictor and outcome variables, in part by 'testing the significance' of a finding that they appear to be related. Is the strength of relationship greater than what could have been observed, simply by chance? If so, a finding is said to be statistically significant (usually with reference to probability values of $<.05$, that is 1 in 500 chance or reference to confidence intervals around an estimate of the effect). To a naive consumer of

quantitative analysis, statistical significance is usually taken as excellent news; however, this p-value is a function of both the actual relationship between variables (between, for example, intermediate care and use of post-hospital services) and the number of people in the study. A large enough study will find virtually all the predictor variables are statistically significantly associated with the outcome variable. It may mean nothing at all.

What is important about the association between process and outcome is not only the statistical significance, but the practical or real-world significance. Is the difference large enough to matter to the patient, or to the organisation of care? Does it matter if you find that transfer to an intermediate care unit is significantly associated with a reduction in total length of stay of .01 days, on average? Probably not. If, however, you find that it is significantly associated with a reduction of 1.2 days, that may well matter to the trust. If it cuts the total days by one third, that is clearly an important result.

It is somewhat more difficult to assess the import of changes in the scores on functional status or quality of life scales. What if the Barthel is 2 points lower, on average, for the intermediate care patient than the conventional care patient? It is worth reviewing the literature and discussing with expert colleagues how much of a difference is needed to translate to a visible, clinically important impact on patients.[35] Then, in reporting results of this type of evaluation, one option would be to focus only on those findings that were both statistically significant and of practical importance. Making sure that you have the right amount of data to bring statistical and practical significance into approximate agreement is the subject of section 5.8.2.

Finally, it is a good idea to avoid using measures where most people are expected to cluster at the top end of the scale, because it will be very difficult to demonstrate improvement. This is called a ceiling effect; functional status scales that focus on the basic, rather than intermediate or advanced, activities of daily living are notoriously vulnerable to this effect. Unless your sample is particularly frail, a measure like the Barthel Index will be more useful as a potential confounding factor than as a measure of effectiveness.

[35] See, for example, Spector W.D., Katz S., Murphy J.B., Fulton J.P.. (1987) The hierarchical relationship between activities of daily living and instrumental activities of daily living. Journal of Chronic Disability, 40:481-489.

5.7.3 Who decides?

Seminar participants were very sensitive to the potential difficulty of selecting an appropriate outcome for evaluation. In particular, they highlighted the importance of finding out what was important to the recipients of care and they reflected that the outcomes which they, or the trust, found most important might be less important to service users. Virtually no information exists in the literature that could shed light on this important question. It has not been common practice to ask patients what their goals for recovery are, and relatively little is known about it.[36]

It would be a great step forward to preface an outcome evaluation with information gleaned through focus groups with patients (or narrative analysis). Similarly, the patient-generated quality of life indices may be useful in this context (see Appendix 1).

5.8 Ways to strengthen an outcome evaluation in intermediate care

Thus far, we have tried to provide guidance in response to participants' concerns. In this section, we make three suggestions that are independent of the workshop discussions, but which would make a difference to the quality of any outcome evaluation.

5.8.1 Pilot first

The suggestion here is not to pilot the intermediate care project *per se* (it is assumed you will have done that; indeed, many intermediate care projects are piloted without evaluation, only to be dropped for no clear reason). As is widely recommended, we mean pilot the evaluation. There is nothing as valuable to a study as being able to do it twice. Often it is only by piloting that it is possible to discover that a sensible-sounding data collection strategy will not be feasible or - more important - that you have completely forgotten about an important variable or - still more important - that you have mis-specified the primary outcome, or even the comparison group. It is immensely useful to attempt a thorough run-through before beginning. At times, it is the pilot study which can provide the information needed for sample size estimates as well (see 5.8.2).

With a process or implementation evaluation, the first few interviews may constitute a sort of pilot, but the information gleaned can be incorporated into the analysis, and into the next round of questioning. There is no need to stage it. With an outcome

[36] An exception to this is: Roberts H., Philip I., Bray J. (1996) *A Study of Needs Assessment and Goalsetting among Elderly People in the Community*. Project report to the South and West Regional Health Authority. Southampton: Elderly Care Research Unit, Southampton General Hospital.

evaluation, however, planning is all. Summative evaluation follows an essentially linear course with no changes and piloting should be one of the early steps along the way.

5.8.2 Determine your sample size scientifically

An outcome evaluation with too few observations - that is, too few patients, or admissions, or procedures (whatever your unit of analysis is) - is a frustrating waste of resources. It is a fairly simple operation for a statistician to calculate the number of people you will need in your study to have a good chance to detect practically significant changes in the outcome you care about, if that benefit actually occurs. It is common - but poor - practice to determine sample size according to how long you want the study to run, or how much money you have to spend. It is a better use of resources to calculate in advance an appropriate sample size and, if your budgets will not allow it, to opt for a different methodology instead.

Note that **it takes more observations to demonstrate equivalence than gain.** For gain, you need enough people in the study to have the power to conclude, with high confidence, that if there has been improvement to a certain level, you will detect it. For equivalence, you need a certain number of people to have the power to conclude, with high confidence, that if you observe *no* difference between groups, it really is because there *are* no differences. A larger sample is also generally needed for dichotomous (either/or) outcome variables, such as 'whether re-admitted', than for continuous variables such as 'number of readmission days.'

Finally, there is the important issue of subgroup analysis, first alluded to in Section 2. Nobody knows yet which type of patient benefits most from intermediate care - however variously benefit may be defined - and the issue of targeting cries out for attention. This implies evaluation designs capable of subgroup analysis, and that requires much larger samples than are typically seen. Nonetheless, the resources required for appropriate data collection, though considerable, would be well invested at this stage of intermediate care service development.

5.8.3 Choose the right comparison group

For most of this discussion, casual reference has been made to 'conventional' or 'usual' care. However, the choice of comparison group - speaking abstractly now - is one of the most important issues to resolve in an outcome evaluation. For hospital-based intermediate care options, most outcome evaluations described in the literature compared the days in an intermediate care unit to days on an acute ward, once patients

had been assessed as appropriate for transfer.[37] In many cases - particularly if the intermediate care unit is the first service of its kind in the trust, or if there is a perception that patients with good recovery potential are lingering in acute beds because the staff have not had time to concentrate on their (relatively less urgent) needs, this will be a good choice.

However, there is an argument to be made that a more appropriate comparator may be other post-acute wards in the hospital, such as an elderly care or stroke rehabilitation unit. These services are also considered somewhat 'Cinderella', may be located off-site, and may subscribe to treatment patterns that accept longer lengths of stay as the norm. By using patients in these settings as comparison group subjects (provided there is enough overlap in their demographic and casemix characteristics), there may be an opportunity to get at some of the fundamental questions for which seminar participants most wanted answers - specifically, which models of intermediate care are best?

For community-based post-acute care, the comparison may be between home and community hospital care; or between supported early discharge and conventional outpatient rehabilitation. This was the model followed in the Bradford Community Stroke Trial and the domiciliary stroke rehabilitation study in Nottingham (DOMINO).[38] In another early discharge RCT, in this instance of assessment followed by intensive home support, a factorial design was used with four groups - assessment and support, assessment and no support, no assessment but support, no assessment and no support.[39] The purpose of this arrangement was to control for the possibility that assessment was itself therapeutic. (The primary intervention to be evaluated was the home-based therapy; assessment, then, was a potential confounding factor. With reference to numbers, note that this design required a sample size of over 900 people!)

---

[37] Pearson A., Punton S., Durant I. (1992) *Nursing Beds: An Evaluation of the Effects of Therapeutic Nursing*. Royal College of Nursing Research Series. Harrow: Scutari Press.

[38] Gladman J., Forster A., Young J. (1995) Hospital- and home-based rehabilitation after discharge from hospital for stroke patients: analysis of two trials. *Age and Ageing*, 24:49-53.

[39] Townsend J., Piper M., Frank A.O., Dyer S., North W.R.S., Meade T.W. (1988) Reduction in hospital re-admission stay of elderly patients by a community based hospital discharge scheme: a randomised controlled trial. *British Medical Journal*, 297:544-547.

Finally, for admission avoidance schemes, an Oxford study from 1986 - before 'admission avoidance' entered the jargon - inferred success by comparing district general hospital admissions from areas with, and without, community hospitals.[40] A similar, geographically driven design has been proposed for evaluating rapid response teams.

At the heart of selecting a comparison group is the question of 'substitute vs. complement' (see Table 3.2 and Section 6). Sometimes intermediate care can be a complement - an add-on - to other services. This is the notion that 'it would be nice to have *all* patients transferred to a home-like environment for an extra week of nurturing' or 'it would be nice to intervene early *whenever* the slightest problem arose in the community' but 'we can't afford it' (this may prove true). To test the effectiveness of the intermediate care model in terms of saving on particular resources, however, what is wanted is a comparison group design that examines substitution.

5.9 Limitations of outcome evaluation

In this section, we have tried to demonstrate the immense power of outcome-based evaluation to confirm or reject claims of effectiveness. As with other forms of evaluation, there are numerous challenges to doing it well, but these have been reviewed with an eye to managing these difficulties almost as technicalities. However, and again as with other forms of evaluation, there will always be a limit to what an outcome-based evaluation can achieve.

Outcome evaluation is not an efficient strategy when the outcome of interest is essentially a rare event, because too many people will be required for the study. If, for example, an intermediate care unit admitted only those patients who were in good health before the trauma that brought them into hospital, and who had a positive prognosis of recovery, then evaluating the effect of the intermediate care on mortality or on long-term care placement would be an ineffective strategy, no matter how undesirable these outcomes would be.

Similarly, outcome evaluation does not work well for long-term outcomes, because it is not possible to control for intervening experiences that could also influence those outcomes. If what is most important about an intermediate care project is to promote health in such a way that the sorts of declines associated with high utilisation of health

[40] Baker J.E., Goldacre M., Gray J.A.M. (1986) Community hospitals in Oxfordshire: their effect on the use of specialist inpatient services. *Journal of Clinical Epidemiology and Community Health*, 40:117-120.

and social services are prevented, it may be difficult to manage a study that goes beyond a couple of years at most. In many cases it will then become difficult to establish causality, because the effects of the intervention will become attenuated.

Most important, however, outcome evaluation cannot open the black box and tell us what has gone on. As long as an intervention is complex - and as long as the focus of attention is intermediate care, the intervention will be complex - a predictor variable called 'intermediate care or not' is both reductionist and opaque. If the overall goal of an evaluation is to know whether to develop or replicate a programme, then even the best summative evaluation will have to be accompanied by a process evaluation that can describe what it is that works.

Finally, good outcome evaluation requires both time and expertise. Data collection can be time consuming and populations under examination may hinder the ability to gather adequate numbers for the required sample size. At the end of day if adequate resourcing is not available this approach may be better left to others.

Section Six: Economic Evaluation

Box 6.1

Highlights of this section

- Economic evaluation is a subset of outcome evaluation which focuses on both cost and outcome analysis.
- It can aid judgement by positioning services along a high low cost and a high low effectiveness axis.
- There are several different approaches to economic evaluation including
  - ◊ cost benefit analysis - when all benefits can be converted into monetary units
  - ◊ cost effectiveness analysis - where the benefit is not related directly to money, e.g. better functional status or less pain
  - ◊ cost utility analysis- where a subjective weighting is placed on a benefit, which may be years of life gained, freedom from pain or magnitude of acceptability (NB this approach can be related to QALYs)
  - ◊ cost consequences analysis - where costs and outcomes are measured and listed separately
- Economic analysis is linked to the concept of efficiency. The concern is not always 'is it cheaper' but 'does it help us to do more with the same resources?'.
- Marginal costs are concerned with identifying which factors cause cost to rise or decline and at what point they have an impact. Thus marginal costs will lessen as more is produced. This has implications for the timing of evaluation and the need to differentiate between set up and service costs.
- Opportunity costs refer to the 'what if' factor, where an assessment is made of how the resources (whether staff time, use of beds or patient pathways) would be used if the intermediate care services were not offered.
- Cost, in this sense, refers to **all** types of resource input including factors which are readily turned into monetary units and those which are not, such as 'psychic cost' (e.g. burnout). Carer costs must also be considered, though some attempts to equate these with monetary value have used subjective and in some instances questionable methods.
- Making use of data which is already collected will save time and effort but it may be necessary to ensure that the units in which they are expressed are clearly identified and can be compared.

Box 6.1 cont.

- Costing may be taken from a 'top down' view where whole costs are identified and divided proportionally between services (e.g. the intermediate care services may have 0.5 of a physiotherapist's time). Alternatively a 'bottom up' calculation can be made where the inputs for a specific patient are calculated, as may happen in an insurance-driven assessment of cost.

- Care must be taken to ensure that the people who receive intermediate care are not penalised on QALY assessments owing to age or frailty which is likely to lead them to having relatively low scores.

- As with all forms of evaluation, it is important to make clear the perspective from which the assessment is being made. Although, in principle, it should be as broad as possible, in practice it may be linked to the interests of those who have commissioned the work.

- As with any outcome evaluation, it is important to pilot data collection and analysis and to ensure an adequate sample size and appropriate comparison group.

- And finally - remember that economic analysis is essential when commissioning or providing services but it is equally essential to consider both cost and benefits in order to make sound judgements for the future.

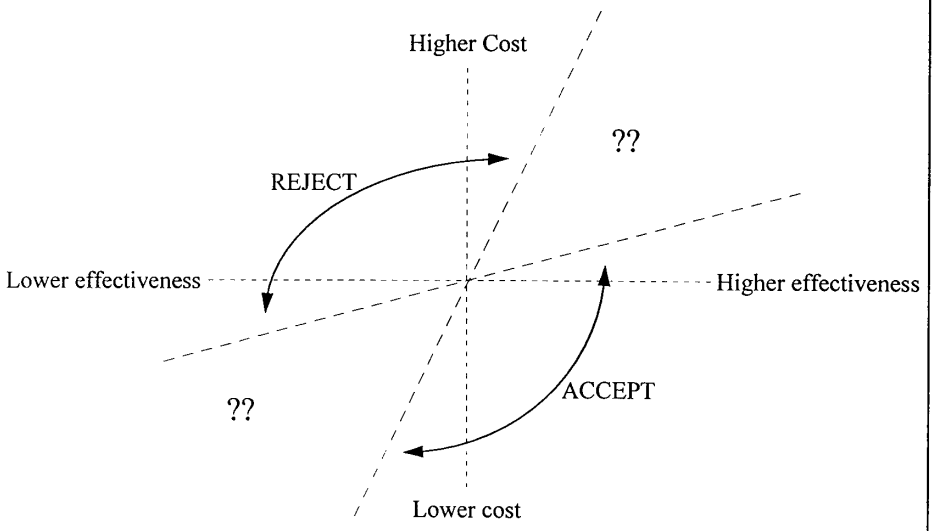
6.1 What is meant by economic evaluation?

Just as implementation evaluation can be viewed as a special case of process evaluation, so economic evaluation can be seen as a special case of outcome evaluation - one which incorporates costs into the judgment. Often taken simply to mean a review of 'how much does it cost or save?', this type of evaluation is actually ***as carefully focused on benefits as on costs***. Although, in practice, it still tends to be treated as an add-on, economic evaluation really should be considered part and parcel of any summative judgement of a programme's effectiveness.

6.1.1 The cost-effectiveness plane

Figure 6.1 provides a simple schematic of how economic evaluation can aid judgments about whether or not to promote the new 'technology' of intermediate care. After assessing cost and effectiveness, there are some 'zones' in which the decision to accept or reject a programme is straightforward. In the other zones, where both cost and effectiveness increase or where both decline, values other than rationality must determine the decision. It could go either way.

Figure 6.1
The cost-effectiveness plane



(Adapted from Laupacis A., Feeny D., Detsky A., Tugwell P. (1992) How attractive does a new technology have to be to warrant adoption and utilisation? *Canadian Medical Association Journal*, 146:473-481, with thanks to Alastair Gray).

6.1.2 Types of economic evaluation

The decision about how best to approach economic evaluation depends upon a number of factors (see Box 6.2). Options range from fairly basic costing exercises to quite complex cost-utility analyses. Note that the decision about method must be taken in the context of establishing effectiveness. If there is no information about effectiveness, the best thing to do is to get some! If, however, there appears to be no difference between intermediate care and the comparison condition, then the less expensive option is to be preferred. All that is required is to thoroughly identify the costs of each service (see section 6.3.1).

When there *are* differences between treatment approaches, other options become possible. **Cost-benefit analysis** is appropriate when it is possible to convert all the benefits into pounds and pence; **cost-effectiveness analysis** when it is possible to convert all the benefits into a single outcome that is non-monetary. For example, if the most important outcomes are length of stay, hospital readmission days, and ward-based staff costs, then a cost-benefit analysis would be appropriate. If, instead, the most important outcome is physical functioning, then a cost-effectiveness analysis to determine the cost for a given change in a functional status score would be preferred. Both these approaches have the advantage of producing a single, simple verdict in the

form of a ratio. Cost-benefit analysis in particular can be used to compare the value of services across any number of treatment categories. Cost-effectiveness analysis can help clarify the price to be paid for improving quality of care, and also allows for comparison across interventions, provided the selected outcome is relevant.

Two other options are finding favour in many health services research settings, one technical and state of the art, the other pragmatic and rather simple. The first is the ***cost-utility analysis***. In this approach, patient benefits are valued in terms of utility, or "the subjective assessment of the well-being gained from alternative interventions"[41]. There are numerous formal approaches to quantifying these qualitative judgments; they include *rating scales* which ask people to place different conditions on a scale from 0 to 100, where 0 represents death and 100 perfect health; *magnitude estimation*, where people must say how many times worse one condition is than another; *time trade-off*, where people are asked how many years of life they would forfeit to be free of a given condition; and *standard gamble*, where people choose the odds they would need to risk a choice (for example, a coin toss) between perfect health and death, rather than stay with a particular chronic condition forever. Two well-known indices that build on utility assessments such as these are the Rosser Index[42] and the EuroQol.[43]

Having created a utility measure (generically called a health related quality of life measure, or HRQoL), it would then be converted into a single metric such as the quality adjusted life year (QALY). If a health condition scored 70 percent in the HRQoL, then one year of life in this state would be worth 0.7 QALYs. Five years in that state would be worth 3.5 QALYs, less than having four years of perfect health (100% x 4) but equal to seven years with a HRQoL of only 50%. It then becomes possible to determine the cost per QALY associated with the intermediate care programme under evaluation.

If cost-utility methods are of interest, it is essential to consult a health economist with experience in using them. Although some interesting work has been done using cost-utility analysis to help with priority-setting,[44] there are numerous methodological

[41] Watson K. (1997) Economic evaluation of health care. In Jenkinson, C. (Ed.), *Assessment and Evaluation of Health and Medical Care: A methods text*, Buckingham: Open University Press.
[42] Rosser R., Kind P. (1978) A scale of valuations of states of illness: is there a social consensus? *International Journal of Epidemiology*, 7:247-258.
[43] EuroQol Group. (1990) EuroQol - a new facility for the measurement of health related quality of life. *Health Policy*, 16:199-208.
[44] Cohen D. (1994) Marginal analysis in practice: an alternative to needs assessment for contracting health care. *British Medical Journal*, 309:781-785.

challenges. For example, it is controversial to set death as the zero point on a rating scale; some people might find a life of intense pain worse than death. There may also be a problem with imagining how one would value different health states, compared to the reality of living them. People with disabilities may take a very different view from those contemplating limitations they do not currently experience.[45]

Box 6.2 Types of economic evaluation: a guide

Is there good evidence on the effectiveness of the intermediate care programme?
- If NO, best practice is to seek such evidence (see section 5)
- If YES, move to the next query

Does effectiveness differ significantly between intervention and comparison groups?
- If NO, do a cost minimisation analysis
- If YES, move to the next query

Can all outcomes be valued in monetary terms?
- If YES, do a cost benefit analysis
- If NO, move to the next query

Can outcomes be valued as a single non-monetary outcome?
- If YES, do a cost effectiveness analysis
- If NO, move to the next query

Can outcomes be measured as quality adjusted life years? (This is a patient-level measure that incorporates the value people place on different possible health states, including death, life with pain, life with disability, and other possibilities extending to perfect health.)
- If YES, do a cost-utility analysis
- If NO, do a cost consequence analysis

(Adapted from Gray, A. 'Cost indicators', King's Fund Intermediate Care Evaluation Seminar 29.10.97)

The second, admittedly softer, option is the ***cost consequence analysis***.[46] It has been recommended for application in settings where the effects are numerous and difficult to resolve into a single metric. For example, if the benefits of an intermediate care programme are at both patient and organisational levels, then cost consequence analysis may be the only feasible approach. It is not possible to create a summative verdict; however, by leaving that judgment to the stakeholders and decision-makers, it may reflect more realistically the mix of rational and political processes that go into

[45] Bauby J.D. (1997) *The Diving Bell and the Butterfly.* (translated by J. Leggatt) London: Fourth Estate.
[46] Drummond M. (1994) *op. cit.* Footnote 25.

developing a health service. In essence, by using the results to make a decision, the utility assessment is performed retrospectively rather than hypothetically, without the requirement to make explicit the implicit valuation of multiple outcomes. In all these methods, ***the objective is not to determine the cost of a service, but to determine its value***.

6.2 What questions are best answered by economic evaluation ?

As Table 6.1 indicates, there is an obvious set of questions about intermediate care that economic evaluation is best suited to answer. Does it cost less? Does it save money? If care alternatives are assumed to be of substantially equivalent effectiveness - or if all that matters is cost reduction - then a cost analysis or cost minimisation analysis may be the only evaluation required. In the first instance, the question answered is 'how much does this service cost?' In the second, the question is 'does it cost less than what we are doing now?'. In the latter case, an outcome-style

Table 6.1
Key Questions about Intermediate Care that Economic Evaluation Can Answer

Area of Interest	Key Questions
Cost savings	• Can intermediate care reduce costs? • Can intermediate care reduce length of stay in hospital? • Can intermediate care reduce the number of patients who need care?
Cost drivers	• What are the cost drivers in intermediate care; what makes the most difference?
Efficiency	• Can intermediate care increase throughput, i.e. the number of patients treated in acute wards? • Is intermediate care a cost-effective way to treat patients?
Trade-offs	• What are the opportunity costs of intermediate care? • What are the costs to informal carers? • Is intermediate care the best use of resources? • To what extent are cost savings in one sector offset by increased demand in another sector?
Value	• Is intermediate care worth supporting?

evaluation is implied, where the outcome (i.e. the dependent variable) *is* the cost. Another set of questions have to do with value for money - not 'is it cheaper?' but 'does it help us to do more with the resources we have?'. This is the concept of

efficiency. Only by examining both sides of the equation - inputs and outputs, what is invested (such as staff or beds) and what produced (such as number of patients successfully treated) - can judgments of efficiency be made. An outcome evaluation can report that patients get better with intermediate care, and do so more frequently or to a greater extent than patients without intermediate care; but it cannot say whether it has cost more to produce that outcome, or achieved more with a level or even lower investment. An economic evaluation can.

A third set of questions focused on the cost drivers, or leverage points. Which aspects of the new service make a difference? What is it that causes costs to rise or decline? It is not necessary to measure every cost, only those that are different or new, since what one wants to know is rarely 'how much?'. It is, more often, 'how much more?' or 'how much less?'. The focus of the evaluation is on the impact of a change.

In this regard, a central concept is that of *marginal cost*. This refers to the *extra cost associated with additional activity*; for instance, the cost of one more physiotherapist in the service, one more bed, or one more ward. Marginal costs will lessen, as more is produced. It may be more expensive to increase intermediate care beds from zero to one, than it is to increase the number from 12 to 13, even though the absolute unit of change is the same in both cases (see Box 6.3).

Box 6.3 Marginal costs and benefits*

Early on, in a service's development, it may cost more to treat a patient under the new system - to transfer the patient to another unit, to staff that unit, to organise additional therapeutic inputs - than it benefits the referring ward to have that bed 'unblocked.' The marginal cost / marginal benefit equation will not favour intermediate care. Later, however, as more patients are treated, the marginal costs of treatment decrease, because the transportation to the satellite site is already in place, the staff are there whether it is for one patient or twelve, and so on. At this point, marginal costs are lower than the marginal benefits of unblocking beds, and the cost/benefit equation begins to look more favourable. Indeed, one seminar participant reported that in her trust, it was only when the intermediate care service could treat 20-30 patients at any given moment that it became cost-effective, because it could then be equated with closing an alternative service (e.g. another hospital ward).

*with thanks to Margaret Stockham

The last set of questions refers to another important idea in economics, that of ***opportunity cost***. It refers to the value of what is *not* being done because resources are invested instead in the intermediate care service. This is sometimes called *'the*

value of the next-best thing', although that is something of a misnomer. Really, the idea is analogous to the notion of a counterfactual in outcome evaluation. Where would the nurses be, if not on the nurse-led unit? What would the beds be used for, if not for intermediate care? How would the community health team spend their time, if not on rapid response? By comparing the value of what has been sacrificed with the value of what is supported, it is possible to decide whether the current service is the best use of 'scarce resources' (more professional jargon, but easily understood).

Ideally, all programmes should be costed with reference to the opportunity costs of the service. In practice, opportunity costs are usually captured by taking market prices (for example, an occupational therapist's wage rate) as a fair proxy. Opportunity costs can also be identified qualitatively, through interviews with clinicians and/or therapists who either provide intermediate care or who refer their patients to the service.

6.3 Questions about economic evaluation of intermediate care

As might be expected, seminar participants had a clearer idea of the questions economic evaluation could resolve, at least in principle, than of how precisely to do it. Table 6.2 lists some of the issues they raised during discussions about incorporating the economic perspective into their evaluations.

The first question, 'is it necessary?', reflects an underlying resistance to accepting automatically the assumption that every evaluation has to examine costs, and to outcome-focused evaluation designs in general. If a service is going to be offered no matter what (for example, because there is the political will to do it) and if, moreover, the financial and health risks of the service are accepted as low, then an economic evaluation probably *isn't* necessary. The second question, 'what should be assessed?', is simply resolved as well; see Box 6.2.

In addition, the third question - about the expertise required - is one that can also be answered fairly readily. Local NHS Research and Development Support Units (RDSUs) may have the needed skills in-house, or be able to locate them. There are books and correspondence courses for the self-motivated.[47] Increasingly, universities around the country are sponsoring short seminars. Education of a rough and ready

[47] Drummond M.F., Stoddart G.L., Torrance G.W. (1987) *Methods for the Economic Evaluation of Health Care Programmes.* Oxford: Oxford University Press. For information about short courses, contact The Health Economics Research Group, Brunel University, Uxbridge. For information about distance learning, contact the Health Economics Research Unit, University of Aberdeen.

sort may be highly valuable, because there is a scarcity of competent consultants, relative to the demand for them. Accountants can be a good source of skilled support for costings. However, for an economic evaluation, the advice of a health economist would be preferred since, from an economic perspective, a cost is not necessarily monetary.

Table 6.2
Questions about economic evaluation of intermediate care

Issue	Question asked
Justification	• Is it really necessary?
Objective	• Should the focus be cost savings or cost effectiveness?
Expertise required	• Whose skills are needed to carry out this type of evaluation, an accountant's or an economist's? Given the scarcity of expert support, how can local evaluators learn enough to do the work on their own?
Methods of measuring costs	• Should costs be analysed as total costs during some period of time, or as cost per patient? • What is an alternative to traditional methods of estimating hospital costs as per-day averages, since the costs of intermediate care should be less than those of acute care (wherever delivered)? • Can diminishing marginal costs be specified? Is there a step costing procedure, and where are the breakpoints at which intermediate care becomes cost-effective?
Methods of measuring benefit	• How can intermediate care's benefits be established, given that utility approaches discriminate against old age and many intermediate care users are elderly?
Perspective	• How wide a perspective should the evaluation take? • Whose perspective should it take? • Who defines costs? How do different sectors define their costs?
Data	• Is it possible to obtain cost data, especially from the community health and social care sectors? • What is the best way to deal with the variations in how cost data is characterised (e.g. average cost, weighted cost, incremental cost, total cost)? • What is the most useful denominator for costings?
Study design	• At what point should economic evaluation be undertaken?

6.3.1 Costs

In economic terms, the term 'cost' refers to ***all types of resource inputs***, and not just financial outlays. It will include elements such as labour, drugs and consumables,

capital investments such as new premises or equipment, overheads, patients' or carers' travel costs, administrative requirements, and even 'psychic costs', such as finding one way of working more stressful than another, or one type of treatment more painful than another.[48] Psychic costs are not easily measured but are worth including qualitatively in any analysis of costs and benefits.

To measure costs, either a bottom-up or top-down approach is taken. With *bottom-up costing*, the actual use of services (for example, a 30-minute session between physiotherapist and patient) is recorded, and the price per 30-minute session is multiplied by the number of sessions that patient had. It is a reliable but time-consuming and data-intensive approach. With *top-down costing*, the total expense of each input (for example, one physiotherapist's salary) is identified, and then divided by your best estimate of the proportion of time spent by the physio on the intermediate care service.

Workshop participants noted that difficulties with cost measurement ranged from the basic challenge of obtaining any data at all to the form in which data were produced, even when information *could* be made available. With regard to obtaining cost information, there may be issues of gaining the trust of those who possess what they see as confidential data. More commonly, according to those at the seminar, community health care trusts and social services tend to not even know what the unit costs of care *are*. They don't have the information, even if they wanted to share it.

This perception is likely to be an exaggeration of the truth; many social service departments keep at least some records of units of utilisation (for example, hours of home care or day care) which can be translated into cost data. Certainly, hospital finance departments retain detailed information about their cost centres, from staff costs to laundry or catering. Most primary care surgeries maintain records of contacts, and it is possible to estimate, at least, how much time was spent per contact. In the absence of local data, however, there is a useful source of community care costings which may be applied as an approximation. It is a publication that lists virtually all types of community services, along with survey-based estimates of unit costs.[49] Provided that the economic evaluation follows a comparison group design - as it should - and that you have carefully collected data on utilisation, then even if the unit costs are not strictly accurate in your locality, the *comparison* between the costs of intermediate care and of usual treatment will be broadly accurate. Seminar

[48] From Evans D., Steiner A. (1997) *op. cit.*, footnote 1.
[49] Netten A. (1994) *Unit Costs of Community Care.* Kent: Personal & Social Services Research Unit.

participants also commented that they had trouble with the forms in which cost data were made available. For example, they have found that hospital finance departments vary in the way they characterise their costs. Some use average costs, some total costs (by week, month, or quarter), some incremental (i.e., marginal) costs, and others costs weighted to reflect overheads, casemix or some other factor. Some costings are ward-specific, while others are not. Some examined costs per episode, while others identified costs per patient. Indeed, there are experts who assert that, however challenging it is to measure benefits, it is more challenging to measure costs![50] Further, these variations pose special problems when the evaluation requires intersectoral analyses.

The best way forward - and it is not very satisfactory - is to be aware, in advance, that it will be necessary to clarify exactly how costs are expressed, and to check whether cost data from different sources will be comparable. If they seem not to be, it would be useful to consult a health economist - possibly at a university, research institute, RDSU, or health authority - to obtain the current thinking on how to translate multi-form cost data into a single currency.

6.3.2 The cost of informal care

Obviously, for some inputs - for example, the time given by family members to caring for somebody at home - there is no expense record in the central office files. If it is of interest to evaluate that cost shift from formal to informal care provision, the information will have to be collected independently. There are several methods of estimating the cost of informal care. One is called ***replacement cost***; if you had to hire people to do what the family or friend is doing, how much would it cost? Replacement costs are often thought to over-estimate the cost of informal care, because many people would want to aid the recovery of someone precious to them and would do some portion of the work, no matter what.

Another method is to estimate the ***opportunity cost***; what is the value, in the marketplace, of the person who is providing informal care? This too is a controversial approach, because it places different monetary value on different people, by virtue of their employment (see Table 6.3, compiled by a health economist). According to this system, a woman who takes leave of absence from her job as a solicitor, in order to take care of her recuperating mother, appears to have a far greater worth than a retired man who is caring for his recuperating wife. Hypothetically, the daughter may have

[50] Gray A., 'Cost Indicators' paper given at the King's Fund, 29.10.97.

been approaching burn-out and would not place a high value on her employment, whereas the husband greatly valued his leisure time; but the convention of assigning opportunity costs according to market prices would quite probably not take that into account.

6.3.3 Methods of measuring benefit

In section 6.1.2, the notion of QALYs was introduced. Some of the difficulties with the cost-utility method - which is, after all, still in its infancy - were presented. One limitation that was not mentioned, however, was its purpose; namely, to inform decisions about allocation of scarce resources among different types of health services. It is here that the use of QALYs to value intermediate care outcomes must especially be questioned, because many of the patients who are most likely to be recipients of intermediate care will receive relatively low HRQoL scores, regardless of how effective the intermediate care service may be. These patient groups include older adults, people of all ages with chronic disease, and possibly those with physical disability or mental health problems who require extra support to recover from a medical trauma. Even if the intermediate care were extremely successful in preventing functional decline or helping people meet their own quality of life objectives, for example, it would not be able to reverse time or undo permanent disability, and therefore could never dominate a utility-based list of health technologies.

It is very important that those who undertake economic evaluation of intermediate care understand in advance what its purpose is. If the purpose is to choose between several models of care for these patient populations, then any method may be applied (per Box 6.2). If, however, it is to help providers or purchasers choose between a service for these populations and a service for younger or abler people, then it may become a significant element in the evaluation plan to agree on a metric which will not disadvantage the intermediate care programme by virtue of who its patients are.

6.3.4 When to begin the economic evaluation

Seminar participants wanted to know the best time to conduct an economic evaluation. In particular, they were sensitive to the idea that, if undertaken at an early stage, the evaluation would reflect start-up costs which might decline once the service became embedded in the organisation of care, and could therefore present an inaccurate and overly negative picture.

Some evaluations are based on a staged study design, so that the initial investment and the subsequent returns can both be measured, and the trajectory of when savings begin to accrue can be mapped. It may be feasible to measure the costs of an intermediate care unit in parallel with a process-focused or implementation evaluation, and to follow these - once the model is clear and the service adequately implemented - with a well-specified outcome evaluation that incorporates cost measurement into the basic research design. Even then, however, costing should be comprehensive and include any special items, for example the cost of training staff to the intermediate care model.

6.3.5 Perspective

It may seem that every section in this report has contained a sub-section about perspective. In the economic context, this refers to the important question of whose perspective to take when assessing the extent to which intermediate care saves money, costs money, or provides a different set of benefits for a level investment. Cost savings to whom? It is a truism of economic evaluation that the perspective should be as wide as possible. It is also the case that the widest perspective - that of 'Society' - would be unreasonable. But how far from the most narrow perspective (for example, that of the clinical directorate within a single hospital) should one go?

Table 6.3 offers three different perspectives for intermediate care evaluation and the types of costs which would be included under each. These perspectives are of the health sector, patients and their families, and 'other'. But even within the health sector, there are choices to be made. An economic evaluation can take the perspective of the acute trust in which the intermediate care service is introduced; or the perspective of the NHS, which would include hospital costs but also costs of community health services and primary care; or the NHS and private sector, which would include all of the above but also any out of pocket expenses for private nurses, medications, equipment or use of private consultants or clinics.

Many intermediate care programmes rely not only upon health services, but also on social services; thus, a good economic evaluation will examine both health and social care costs, especially as hospitals may reduce costs only by shifting them to care managers and home helps in the community. However, it can be very difficult - time-consuming, expensive, cumbersome, and perhaps not possible - to collect social services data. As for the important area of costs to users, few economic evaluations of health care interventions have successfully tackled these, except as footnotes to a health service-based assessment.

Table 6.3
Possible perspectives for intermediate care evaluation and types of costs associated with each

Health sector	Patients and families	'Other' sectors
All (revenue and capital) costs of programme being considered This would include: • cost of initial programme • costs incurred or saved as a result of the programme	Out of pocket expenses such as over-the-counter medications, equipment, transportation to doctors Time spent by patient and/or carers, visiting, waiting for, or travelling to doctors and therapists This could be work time or leisure time, which would affect its valuation	Cost implications for social services, housing authority, nursing homes, or other agencies. Note that these could be part of the intermediate care programme or simply be affected by it.

(Gray, A., 'Cost Indicators' paper, King's Fund, 29.10.97)

When planning an economic evaluation, it is essential to identify the perspective which will be applied. It is also a good idea to consider how feasible it may be to take a broader perspective. One approach would be to test the impact of different perspectives on the cost-effectiveness judgment by incorporating multiple sectors' costs into a sensitivity analysis (see section 6.4).

6.4 Ways to strengthen an economic evaluation of intermediate care

The advice regarding economic evaluation is not unlike that offered in the last section, as economic evaluations centre on rigorously designed outcome studies. Their validity will depend on the way that effectiveness is judged, as much as on how costs are measured. Thus, the suggestion holds, to pilot as many data collection and analysis methods as possible first. Similarly, it will be essential to identify a study sample that is representative of the population of interest and to perform power calculations so that an adequate sample size is ensured. It is good practice to understand clearly why the comparison group you have selected is appropriate, and which outcomes are most important. Remember that if you are aware (or suspect) that the intermediate service offers a particular type of benefit, be sure to include it as one of the primary outcomes.

What has not been mentioned yet is the value of conducting a ***sensitivity analysis*** as part of the evaluation. Because the cost estimates you obtain may be spurious, it is a

good idea to calculate a *set* of cost-benefit or cost-effectiveness ratios based upon different costs from the ones you have. You could test, for example, the result if true costs are twice, or half, what your data say, if the costs are not only from the NHS perspective but also include social services, primary care, or patients' costs. For these last costs, it may be possible to obtain information for a subset of study subjects, and to use these to get a feel for how similar, or different, the overall picture becomes when they are included. The object of sensitivity analysis is to demonstrate how robust the findings of the economic evaluation are. If they appear to be vulnerable to slight variations, then the accurate specification of costs is clearly critical, and it would be worth investing extra resources to collect the best information possible. If, instead, the basic judgment of worth seems to be stable when cost estimates are varied, then the inferences drawn can probably be taken as read.

6.5 Conclusions

Most people think 'money' when they hear 'economics' and certainly money is part of it. For this reason, seminar participants felt that the drivers to conduct economic evaluation were obvious. "Everybody cares about costs," one said. Others noted that, with increasing competition for limited resources and possibly increased market power arising from primary care innovations, it was necessary to choose among care alternatives. To make the choices, both purchasers and providers needed information about the resources involved in producing and providing care.

In this section, we have tried to provide both an introduction to the terms and techniques of economic evaluation and to offer caveats regarding the points where classical economic evaluation may deviate from, or even undermine, a fair portrayal of the costs and benefits of intermediate care. In the next, final section, we discuss how to make this approach work with the other styles of evaluation in order to fit resources to needs, and in order to build a body of evidence about models of intermediate care.

Section Seven: Bringing it all Together

Box 7.1

Highlights of this section

- Both those who commission and those who do evaluation must agree where the focus of their attention should be directed.
- Pluralistic approaches, taking into account multiple concerns and methods are recommended. Similarly, phased or longitudinal studies should be considered.
- Matching methods to questions, skills and resources is critical. It is both possible and important to match evaluation questions to the appropriate method for answering them.
- There is no hierarchy of worth among different approaches to evaluation but there may be a hierarchy in terms of data collection and resources utilisation.
- Results from all sources of data should be presented in an integrated but non judgmental way, allowing those who commissioned the evaluation to make their own decisions about the relative merit or weighting of different results.
- The possibility of undertaking a meta analysis, integrating results for various sources should be considered. Similarly undertaking a systematic review of other studies can add to local understanding.
- Some units may find action research a useful approach to consider, remembering that it requires the same rigour as any other research method.
- It may be helpful to pose a series of questions about the type of local service, the stakeholders and the resources before developing your evaluation strategy.
- Remember, you can only achieve what you have the time, skills and resources to do so plan with care!

7.1 Choosing between options

In this section, we consider how to integrate the information presented previously in order to make the best choices for evaluation of intermediate care services, and we comment briefly on how the various options may be brought together in a realistic manner. We also suggest directions for further development, including the possibilities of meta-analysis for well-developed services that have completed some local evaluation and an action research agenda for developing new services in the area of intermediate care.

In previous sections, an overview was given of some of the different approaches to evaluation which can be considered within a whole strategy for examining intermediate care services. Each approach offers a specific perspective, focusing on one subset of a range of questions which need to be addressed in order to gain a complete picture. It must be stressed, however, that there is no single way in which a 'whole story' can be told, and it is up to those responsible for both commissioning and undertaking evaluation to decide where the focus of their attention should be placed. Hence the emphasis throughout has been on considering a pluralistic approach governed by the concerns and interests of the commissioners, the resources available, timing and population size, and the desired impact or application of the evaluation.

While the importance of gaining greater understanding of both the nature and efficacy of intermediate care services is widely acknowledged, there is no doubt that evaluation is complex. This relates not only to the methods used but also the context in which developments are occurring. For example service innovations of this nature are taking place against a background of enormous turbulence. Changes are occurring so rapidly that it is hard to hold anything still for sufficient time to see what is actually happening. The people who may benefit from intermediate care are not easily defined as they do not fit into a neat diagnostic category and their needs arise from a myriad of different clinical diagnoses. Clinical teams alter over time, and experience affects their confidence and ability to manage various client groups across a different range of settings.

Thus any approach to evaluation must be tightly defined, clear in its purpose and realistic in its aims. While it can be argued that *some* level of evaluation should form an essential component of the development work involved in introducing new services of this nature, it is equally important that the methods used are matched with the skills and resources available.

7.1.1 Reviewing the choices

To summarise the main observations made previously, we have suggested that, at its simplest, process evaluation determines what is done, implementation evaluation determines how it is done, outcome evaluation determines what happens as a result, and economic evaluation determines whether it was worth doing. As this characterisation implies, it is clear that evaluation can serve multiple purposes (see Table 7.1). Not only is it possible to match one's evaluation design to what one wishes, or needs, to find out; it is highly useful to plan accordingly and to reject certain questions if it is not feasible to implement the appropriate design.

Table 7.1

Types of evaluation and the questions they can answer

Type of evaluation	Types of questions it can answer
Process	Patterns of care - Organisational models - Assessment and referral - User views
Implementation	Impetus for change - Power balances - Stakeholders - Drivers and inhibitors
Outcome	Health outcomes - Organisational effects - Best models
Economic	Cost savings - Cost drivers - Efficiency - Trade-offs - Value

It does not make sense to ask the wrong question of a particular approach, simply because it 'feels' like the preferred way of doing things. Some people are drawn to qualitative open-ended interviewing or working with narrative storytelling techniques, while others prefer outcome measurement and standardised pre-validated questionnaires. But this is not sufficient reason for undertaking a particular sort of evaluation. Questions must match methods for, as the old saying goes 'If one only has a hammer, the whole world looks like a nail'. This hammer-and-nail mentality should be strenuously avoided!

7.1.2 Matching evaluation strategies to available resources

At the same time it is essential to 'cut your coat to suit your cloth'. Budget concerns cannot be ignored. If the budget will allow an implementation evaluation but not an outcome evaluation, then it is important to be clear from the outset about the strengths and limitations of the information gained and the manner in which it will be used. For example, in this situation it would not be possible to assure purchasers or the public of specific clinical effects, but it would be possible to demonstrate whether the service can be successfully introduced and accepted (or not, as the case may be). Whatever approach you choose, it will help immensely for commissioner and evaluator to come to a shared understanding about both the strengths - what you will be able to say - and the limitations - what you will not be able to say - indicated by what ever approach to evaluation you may choose.

It is also critical that the skills of those with responsibility for evaluation are matched to the aims of the study. Thus it is inappropriate, and some would say unethical, to ask those involved with service implementation to undertake rigorous and complex evaluation when they have neither the methodological research training nor the financial resources available. While it may be entirely reasonable, and indeed

desirable, to agree on a form of local evaluation which is confinable within these practical perimeters, it must be stressed that more ambitious studies need dedicated funding and expertise if the results are to be of value. A partnership between practitioners and evaluators would provide the ideal mix, bringing together the expertise of both in order to ensure that both the right question and the right method are chosen. Such negotiations are the ideal starting point for planning an appropriate evaluation.

7.2 Phased or staged evaluation

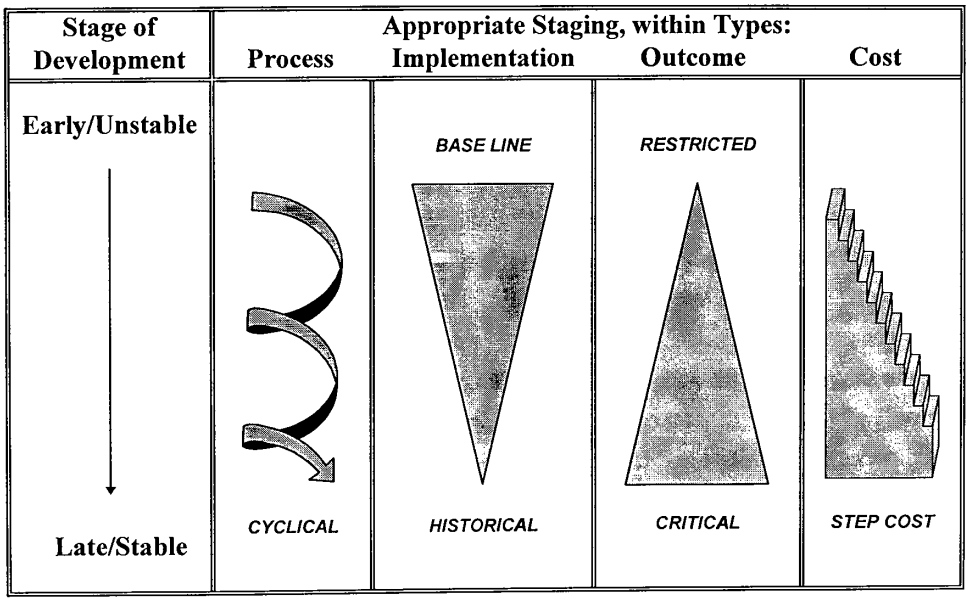
It is realistic to acknowledge that many intermediate care initiatives are at an early stage of development. Hence the context in which the service is offered will continuously be changing as new systems are established, teams learn to work in new ways and referral patterns alter as colleagues grow not only to know about the service but to gain confidence in its efficacy. Thus the areas under investigation are themselves undergoing change as new approaches to service delivery are introduced and traditional role boundaries are challenged.

Even if, in principle, an outcomes study may be your ideal type of evaluation, in this kind of situation it may not be wise, or even possible, to establish a study seeking generalisable cause and effect relationships. By the very nature of this type of inquiry, it would be critical that agreed variables could be held constant for long enough to gather sufficient data to give meaningful results. In reality, the increasing confidence of the clinical team, or shifts over time in clinical decision-making with resultant changes in referral patterns, cannot be held constant at this stage of implementation and meaningful results are unlikely to be forthcoming. In this situation it may be better to explore process and implementation questions initially and move to outcomes work when the service is more stable.

In contrast, other intermediate care providers have been working with short-term funding for several seasons and have by now created service delivery systems that appear to work well. For these services, it may be time - if it has not already been done - to consider a summative evaluation where the services in place are compared to more conventional treatments, perhaps in a neighboring locality that lacks intermediate care options or offers a different set of services. The implementation dynamics may be sufficiently well understood, and the care delivery process sufficiently well established, to justify bypassing those evaluation approaches and moving on to a comparative review of patients' clinical, functional, and quality of life outcomes.

We are strongly in favour of conceptualising evaluation in a phased or staged manner. (Figure 7.1 highlights this progressive approach.) Each service will be able to identify its own stage of development and to determine at what level to enter an evaluation plan. We would stress, however, that the recommendation to consider phased evaluation plans is not intended to promote a hierarchy of worth, but rather a hierarchy in terms of data collection and other resource needs. Whatever the evaluative method, and wherever it sits along the hierarchy of complexity to undertake, finding appropriate expertise will be a crucial element.

Fig. 7.1 Optimal approach to phased or staged evaluation, by type



The concept of phased or staged evaluation does not only apply to using the different approaches in a sequential or integrated manner. Even within evaluative approaches, there are phases to their application. As the schematic in Fig 7.1 depicts, process evaluation works best when conducted at regular intervals. Processes of care, and of organising, tend to develop and change. A sound evaluative strategy will take this into account and will continually update organisational understanding of what intermediate care means in terms of local operation. In contrast, implementation evaluation is of tremendous use in the early stages of developing intermediate care programmes. Later on, once the intervention is well embedded, such an approach would be less relevant to current practice, although a look backwards at the development process might offer the lessons of hindsight to planners of other interventions. The opposite is true for outcome evaluation. There is little scope for meaningful outcome assessment in the early stages

of a new programme, but it becomes crucial to measure effectiveness in terms of actual outcomes of care, once the model is up and running. Finally, with regard to costs, it is helpful to take a stepped approach in relation to the growth in the size of the programme.

In many cases, a process or implementation evaluation may be precisely what is wanted to answer the questions of most interest; stopping there does not necessarily signify a partial approach. As Parlett and Hamilton[51] have suggested, when speaking of those who focus only on outcomes, "...one has to look not only at the manuscript but also at the performance; that is, at the interpretation of the play by the director and actors. It is this that is registered by the audience and appraised by the critics". In fact, there is a strong argument to be made in favour of longitudinal process evaluation. For example, process evaluation can map referral patterns in order to identify variations over time. From an economic perspective, too, it may be necessary to build a cost model over a period of some months or even several years. For example, economies of scale will paint a very different picture from set-up and pilot costs to later days, when nursing-led post-acute hospital units are filled or the impact of admission avoidance on a particular GP's practice begins to be felt.

Timing is also a critical feature. Reviewing previous evaluations that studied the impact of primary nursing on clinical outcomes, it is evident that much of this work was considered flawed and lacking in consistency.[52] However, this was largely owing to a premature review of the impact of the changes brought about by primary nursing and too short a study period to gain comprehensive insight into effectiveness.

7.3 Integrating the findings

For those who are able to conduct a multi-method evaluation, whether in stages or in parallel, there is the important question of how to interpret the results in a coordinated way. In particular, there is the issue of which results should dominate. It is to be hoped that there will be convergence, and that a good process, successfully implemented, will be associated with increased cost-effectiveness and overall improvement in outcomes. That is very much an ideal scenario, however. More probably, services will have both strengths and limitations. The most successful programmes in terms of outcomes may depend on infusions of resources, which make

[51] Parlett M., Hamilton D. (1977) in Hamilton D. (Ed) *Beyond the Numbers Game: A Reader in Educational Evaluation,* Basingstoke: Macmillan.
[52] Giovannetti P. (1986) Evaluation of primary nursing in *Annual Review of Nursing research* vol. 4 pp 127 151 (Weilty H H, Fitzpatrick J J, and Taunton R L eds.) New York: Springer, Louis, CV Mosby.

it impossible to demonstrate economic savings. A satisfying process, in terms of staff morale or patient satisfaction, may not produce the much desired health outcomes such as improved physical function or reduced numbers of readmission to hospital. What to do then?

The evaluators' role must be to clearly describe the findings of each method, and to use those methods to draw out comparisons and trade-offs, as hinted above. In such a way, the investments and paybacks can be clarified, as can the unanticipated benefits or barriers to development. We would suggest that it is not the evaluators' role to produce a bottom-line value judgment unless all indicators point in the same direction. Short of that, it must be left to those who provide and/or have commissioned the service to decide, for example, which is more important, user views or consultants'; whether a higher cost programme with better outcomes is to be supported; whether a service with neutral costs and clinical outcomes, relative to conventional care, significant start-up costs (both financial and cultural) and greatly improved patient satisfaction is a net gain. To some extent this will be a political process. Again, what the evaluator must do is draw out clearly the benefits and drawbacks of a service, illuminated from various sources.

7.3.1 When is meta-analysis appropriate?

One approach to synthesising research findings is the meta-analysis. This is a systematic, quantitative form of integrating results from various sources. The original data are actually merged, in order to increase the sample size and therefore the power to detect changes if they have occurred. It would apply only to services that are well established, and that have conducted carefully designed evaluations using validated measures so that there is enough consistency to allow data to be merged with confidence. Few sites will be in a position to lead or participate in a meta-analysis, powerful as they can be. If you are, however, it is crucial to work with a statistician who has appropriate expertise from the start.

The most successful form of meta-analysis for complex interventions may be the US Frailty and Injuries: Cooperative Studies of Intervention Techniques (FICSIT) research.[53] This was a pre-planned meta-analysis to test a variety of interventions to prevent falls and/or fractures in older people in which each study committed in advance to taking a core set of measurements, in addition to others of specific relevance locally.

[53] NHS Centre for Reviews and Dissemination (University of York) and Nuffield Institute for Health (University of Leeds) (1996) *Effective Health Care Bulletin : preventing falls and subsequent injury in older people.* Published in association with Churchill Livingstone Vol. 2, No 4. 1-16.

These could be merged later on in order to discover the thrust of the evidence on effectiveness of this broad type of intervention. The potential for a FICSIT approach to be applied to the area of intermediate care is probably significant.

A less quantitative, but also useful, approach to combining information in order to integrate a body of evidence from various sources is the systematic review.[54] Essentially, a survey of studies is taken. Results can be analysed quantitatively or qualitatively, and it is good practice to assess not only the final outcome (did it 'work' or not?) but also to report on the quality of the intervention itself (process evaluation again). What makes a review systematic is that a fairly specific protocol is followed from the start, regarding which studies are to be included, why they are selected, and how the findings will be interpreted. Less formally, there may be considerable scope for providers or commissioners of intermediate care simply to compare notes - for example in workshop format - in order to tease out which models seem most effective, or which elements of intermediate care work easily and which are more difficult to arrange.

7.3.2 Action research - another way forward

In contrast to meta-analysis, which will only apply to a small group of readers, action research may provide a way for some to proceed.[55] It provides a framework which may be particularly pertinent as it offers a structure in which a phased evaluation, which is sensitive to local context, can be developed. Indeed, a number of seminar participants raised this as a way in which they would like to proceed, acknowledging its strength as both as an instrument of change and an evaluative tool.

Action research lies in the field of 'critical social theory' which has attracted interest in recent times, partly as a way of overcoming the dichotomy between positivist research, which is largely quantitative, and naturalistic research which is primarily qualitative. The value of action research lies in the way in which it takes note of the people who are involved in the development "..with immediate relevance to pressing social issues.." .[56] As Carr and Kemmis[57] suggest, technical and interpretative data can be brought together while using reflection on action to direct change. Thus theory and

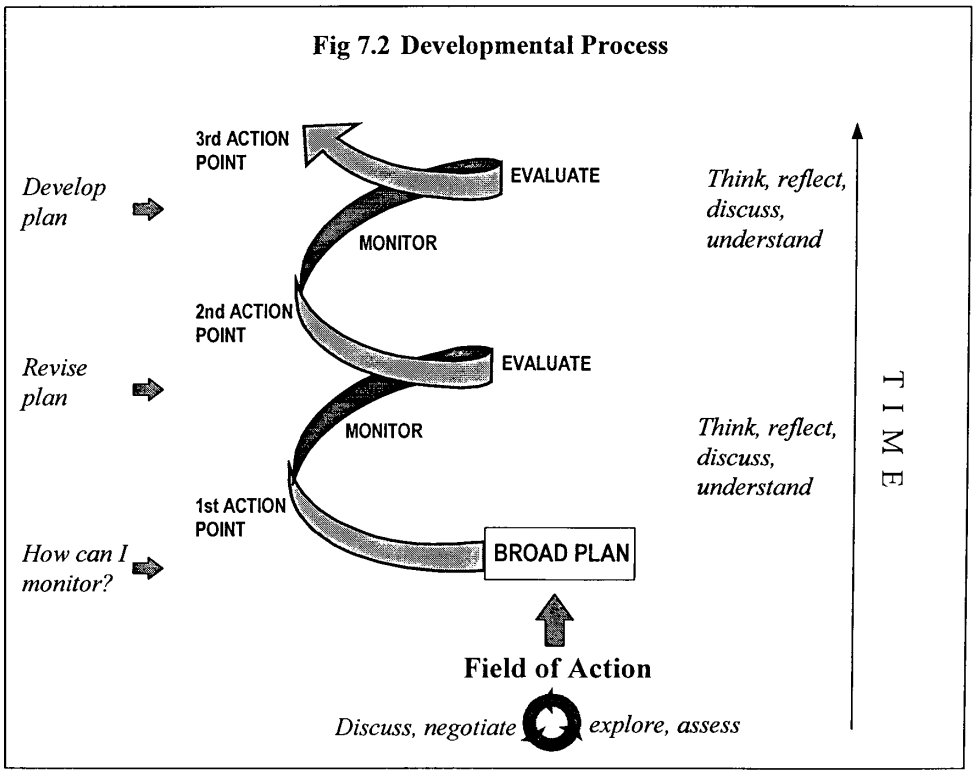
---

[54] Chalmers I., Altman D.G., Eds. (Eds.) (1995) *Systematic Reviews* London BMJ Publishing Group.
[55] Hart E., Bond N. (1995) *Action Research for Health & Social Care: a guide to practice*, Buckingham, Open University Press.
[56] McTaggart R. (1991) *Action Research : a short modern history,* Geelong, Deakin University Press.
[57] Carr W., Kemmis S. (1984) *Becoming Critical: Knowing Through Action Research,* Victoria, Deakin University Press.

action can develop together creating change "not in artificial settings where effects can be studied but in the real world of social practice".[58]

Throughout this text we have stressed the importance of being aware of the needs and wishes of the people concerned, recognising that these may range according to discipline and interest. Action research offers a framework in which such views can be brought together in order to move forward in harmony. In essence it involves a cyclical process, identifying an issue which needs to be addressed, gathering relevant data and feeding it back into the system to inform the next step of development. For example, in this situation it may be critical to know how many people would benefit from an intermediate care service, taking into account the views of doctors and nurses in community and acute settings. Addressing this problem as a team will not only lead to greater understanding *within* the team but provide hard factual information which will guide the next steps. Thus it is possible to combine a number of different approaches to data collection which may satisfy the variety of interested parties while aiding the developmental process (see Figure7.2).

Fig 7.2 Developmental Process



[58] McTaggart R. (1991) *op.cit* footnote 55.

Action research has attracted increasing attention over the past twenty years as a means of evaluating complex innovation and has provided valuable insights into development.[59] [60] However it is not without its critics and a word of caution should be raised about the dangers of sloppy use. The questions raised through the action research process must be as rigorously considered and formulated as in any other research design, as must the manner in which data are handled and analysed. The difference in this approach is primarily a philosophical one of partnership rather than distance between researcher and subject, with an intention to actively contribute to development. Thus it deserves both the time and resources of any other approach and cannot be used as a 'cheap option'.

7.4 Getting going - setting evaluation in a realistic context

To date we have focused on the type of questions which different approaches to research may answer. To conclude this report we turn to the essential issue of 'What will I do in my patch?'. We have set out below a series of questions, clustered around different issues which are likely to drive your decision making. They are intentionally broad ranging, seeking to raise the scope and breadth of questions which it is helpful to consider at the outset, in order to ensure that the efforts entailed in evaluation are of value.

7.4.1 The service

The shape and size of the service, the length of time it has been established, the setting in which it is offered and the number of patients seen now and planned for the future will all influence what it is possible or reasonable to evaluate. Asking the following questions may help you to decide where to focus.

- *How long has the service been established?* If it is still in very early days then it is unlikely to have settled into an established pattern of care. People may not know about it or may be cautious of transferring people with a degree of need which has traditionally been dealt with in an acute health care setting. Alternatively, however, if the service is well established then the range of choices may veer towards evaluation designs which require a more controlled environment.
- *How many patients are seen each week/month? Is this likely to alter?* Bear in mind that it is important to estimate the numbers needed to apply any form of statistical analysis and you need to know this beforehand. You could still be collecting data five years on!

---

[59] Towell D., Harries C. (Eds) (1979), *Innovation in Patient Care*, Croom Helm Ltd, London.
[60] Titchen A., Binnie A. (1992) Research Partnerships: Collaborative Action Research in Nursing, Institute of Nursing, Oxford.

- *What is the source of referral for your patients? Is this likely to widen or become more narrow with time?* In some instances the pressure on acute care beds is so high that providers do not feel able to comply with or indeed tolerate randomisation, as their priority is to create space for the next wave of patients. Similarly there may be a moral drive to prevent admission through the use of an intermediate care scheme when one is available locally. Thus the researcher may be pushed towards comparative matched samples rather than randomisation from a single potential cohort.
- *For how long is the service funded?* If it is being funded on 'soft money' for a short period of time, then you will be limited by that factor when considering what you can achieve.

7.4.2 The stakeholders or commissioners

Inevitably various stakeholders in a new initiative will have a view of what needs to be assessed which reflects the manner in which they work. For example, commissioners of services have a remit to ensure that public money is well spent in matching patient need to service provision and they have to make critical judgments in terms of prioritising. Alternatively clinicians may be interested in hard outcomes but also in quality issues related to patient or user views. In this area it may be useful to ask yourself the following series of questions.

- *Who has commissioned this work? Have the commissioners made clear the type of data they want? If you have been asked to design a study how much can you negotiate with the commissioners? Are their expectations realistic? Do you know their purpose in asking for the work?* Questions may be raised about the rationale behind the commissioners interest. It may be useful to consider, for example, the external pressures which they may be facing, such as a need to prioritise or a time scale in which decisions must be made. Early negotiation about who is responsible for the actual study design and agreement on what will be produced can aid in understanding the intended purpose of commissioned evaluation.
- *Has additional funding been provided for evaluation?* If not, then it is critical to come to some agreement about the amount of time which can be devoted to the evaluation and the implications this will have for everyday service delivery. If funding is available similar questions must be asked with a general rule of thumb that a small well executed study may be far more valuable than a broad but superficial overview.

- *If you are seeking funds can you judge/legitimise to potential funders why you are undertaking the work and the importance (against other demands) both locally and (potentially) nationally? Can you link your work to a wider strategic or political agenda which may increase its relevance to others?* While the rationale for undertaking your proposed study may be very clear from a local perspective, if funding is being sought then the arguments you present must be robust enough to be judged against other bids. Again it may be helpful to try to judge the starting point in terms of knowledge and interest of potential funders in order that an application can be tailored towards their interests.

7.4.3 The resources

We have already mentioned resources but cannot stress enough the importance of matching effort to outcome. At the end of the day resources are always limited in some way and it is critical that this issue is addressed realistically at the outset of a piece of work. We have grouped resources under three broad areas; namely skills, time and money.

- *What are the skills that will be needed to undertake this work well (bearing in mind that research and evaluation are specialist subjects in their own right)?* Designing a questionnaire or analysing quantitative or qualitative data requires skills that are not common among many practitioners, whatever their disciplines. It can be dangerous to over estimate knowledge of this subject and while it is not our intention to put people off it is always wise to seek help at the outset of a study rather than half-way through.
- *What evaluation skills are available within the team?* It is important to review the skills which are available within the team in order to develop a realistic plan which builds on strengths already available without raising unrealistic expectations.
- *Is it possible to seek advice from the local health authority or local academic institution?* Would this have cost implications? In the light of the comments made above advice may well be available through a number of different local agencies, even if it is only at the level of critiquing your proposal. There are many people who would be willing to offer help if asked, especially as there is currently a high interest in this type of work.
- *Is there any capacity within the team for time to be dedicated to evaluation?* Remember that however good the intention may be, this is often the area of work which goes to the bottom of the pile when life gets busy!

- *Is it possible to negotiate with another team to support evaluation (e.g. the health authority, region or academic organisation)?* It may be that a shared interest could lead to pooling of resources in everyone's best interest.
- *In real terms, how much time needs to be devoted to this study?* This issue merits serious attention. Most people seriously underestimate just how long both data collection and analysis can take. On the whole it is better to overestimate the amount of time needed.
- *How much time is available to do the work (with a rule of thumb which goes 'one third set up time, one third data collection and one third analysis')?* We suspect that you have gleaned our concern about time management by now!
- *Have you been able to do a realistic costing of the evaluation?* This relates back to issues of both time and skills which we have raised above.
- *Is there any funding which is dedicated?* If separate funding has been identified, it may be helpful to ringfence the money in order that it does not get pulled back into a service provision budget.

These questions are by no means finite, but are offered as a starting point for you to plan your own evaluation. There could well be other contextual issues which you feel must be addressed at a local level. These contextual concerns may impact on your final decision and merit careful consideration from the start. You will find, as your thinking develops, that what may seem like a multitude of concerns raised in this document will narrow to a manageable set as your requirements and plans come into focus. Good luck and do evaluate. It is only through reflective practice that intermediate care services can develop in a useful way.

Appendix 1: Suggested Evaluation Instruments

Purpose
This appendix offers a selection of measurement tools which may be appropriate for evaluation of intermediate care services. They have been chosen based upon the following selection criteria:

- conceptual relevance to intermediate care
- documented reliability and validity
- sensitivity to incremental change (either improvement or decline) over the short and medium term
- lack of 'ceiling effects'
- accessibility to the researcher.

Method
Compendia of rating scales and questionnaires were scrutinised for tools which , in the judgment of the reviewer, met the criteria[59]. Tools were sought which addressed patient outcomes under the following rubrics: condition specific measures, functional status measures, quality of life indices and patient generated measures. A small number of tools are described here which could be of use. Readers thinking about using other measures are advised to consult the compendia listed in the bibliography.

CONDITION SPECIFIC MEASURES

Guyatt's McMaster Chronic Respiratory Questionnaire (Guyatt, 1987)
Purpose: provides patient generated assessment of their well being related to respiratory difficulties
Description: patients select five most important areas, and rank on 7 point Likert scale. Also questions about dyspnoea, fatigue and emotional functioning. Interview takes 15-25 minutes.
Reliability and Validity: good test retest and correlational validity with clinical assessments.
Sensitivity to Change: not reported.
Relevance to IC: one of the more comprehensive disease specific measures. Patient generated items add conceptual relevance.

The Arthritis Impact Measurement Scales (AIMS)
Purpose: outcome measure of care, addresses physical, social and emotional well-being of arthritic patients. One of the most widely used outcome measures in arthritis research.
Description: 45 items grouped into 9 scales assessing mobility, physical activity, dexterity, household activity, social activities, ADL's, pain, depression, and anxiety. Self administered, takes about 15 minutes, most questions refer to problems during the past month.

---

[59] McDowell I., Newell C. (1996) Measuring Health: A guide to rating scales and questionnaires, New York: Oxford University Press.

Reliability and Validity: acceptable.
Sensitivity to change: acceptable to superior correlation with other instruments, and after therapeutic interventions such as total hip replacement.
Relevance to IC: focuses on how patient feels and functions, in aspects which active therapeutic input should be able to affect. Not validated for other clinical diagnoses.
Contact: John H. Mason PhD, Research and Evaluation Support Unit, Boston University Arthritis Center, 80 East Concord Street, Boston, MA, USA 02118-2394.

Diabetes Quality of Life Measure (Diabetes Control and Complications Trial Research Group, 1988)
Purpose: aims to measure diabetic patient's quality of life on dimensions of satisfaction, impact of disease on life, worry about diabetes, social and vocational worries.
Description: 41-57 item self administered questionnaire.
Reliability and Validity: internal consistency and test retest reliability adequate. Does not tap all relevant parameters of quality of life, could be used with generic quality of life measure.
Sensitivity to Change: unreported.
Relevance to IC: useful in conjunction with other relevant measures.

FUNCTIONAL STATUS MEASURES

Barthel Index
Purpose: measures functional independence in personal care and mobility. Originally intended for long term hospitalised patients and inpatient rehabilitation. Unproven as an outcome measure.
Description: Depending on version, 10-15 item rating scale completed by health care professional from medical records or direct observation. Completion in 2-5 minutes. Little consensus over which of several versions is definitive, although Modified BAI most commonly used in clinical assessment in Britain. Self report versions correlate poorly with health care professional assessment.
Reliability and Validity: Inter rater reliability excellent. Correlation with other instruments measuring function in stroke patients adequate.
Sensitivity to Change: Significant ceiling effect, which makes its use in IC questionable. Scoring categories are coarse, and fail to detect incremental change. Cannot be used to calculate changes over time, as scale is not continuous.
Relevance to IC: Widely respected as a good ADL scale, but narrow in scope, and does not detect low levels of disability. Where the emphasis in IC is on helping the patient to meet their own goals, a self report measure would be more conceptually acceptable. Despite its popularity as an assessment tool, the Barthel Index may not be sufficiently sensitive as an outcome measure in evaluation of intermediate care services. It could be used as a casemix variable, that is, to control for basic functional status when trying to predict more advanced functional outcomes.
Source: RCP and British Geriatric Society Standardised Assessment Scales for Elderly People, 1992.

Advanced Activities of Daily Living (AADL) (Reuben et al., 1990)
Purpose: to determine whether patients are capable of undertaking more advanced activities, which are part of their normal life.
Description: patient generated measure identifying activities important to them. Uses Guttman scaling technique.
Reliability and Validity: highly reproducible, strong construct and predictive validity reported.
Sensitivity to Change: greater than basic and advanced ADL scales.
Relevance to IC: useful in areas where function and roles are affected by illness.

GENERAL HEALTH STATUS AND QUALITY OF LIFE MEASURES

Short-Form-36 Health Survey (Rand Corporation and Ware, 1990)
Purpose: designed as a generic indicator of health status for use in population surveys and evaluations of health policy. Also used as an outcome measure in clinical practice and research, especially in conjunction with disease-specific measures.
Description: includes multi-item scales on eight dimensions:

- physical functioning
- role limitations due to physical health
- bodily pain
- social functioning
- psychological distress and well-being
- role limitations due to emotional problems
- vitality, energy or fatigue
- general health perceptions.

It may be self administered or used in face to face or telephone interviews. The SF 36 usually takes 5-10 minutes to complete; elderly respondents may need up to 15 minutes.
Reliability and Validity: extensively tested, strong internal consistency. Sufficient for comparing groups and individuals. Discriminates between types and severity of physical and psychological conditions. A British-English version of the SF 36, with altered phrasing, has been developed and tested against British population norms.
Sensitivity to Change: Change in health status can be assessed over the previous 4 weeks, or in the Acute form of the SF 36, in the past week.
Relevance to IC: addresses concepts of interest, amenable to therapeutic input from IC. Probably most useful in conjunction with disease-specific measures, where such are available.
Access: Permission to use the SF 36, free of charge, can be obtained from the Medical Outcomes Trust, PO Box 1917, Boston, MA, USA 02205. This organisation supports the development and distribution of standardised outcome measures, and produces updates on SF 36 administration, scoring and interpretation.

Quality of Well-Being Scale
Purpose: a health index which summarises a patient's current symptoms and disability in a single number that represents the social undesirability of the problem, expressed in QALYS. Intended as an outcome indicator, and in estimating present and future need for care. Applicable to individuals or populations, and any type of disease.
Description: a three stage process -

- assesses functional status through recording symptoms and medical problems in each of the previous eight days and classify the respondent's level of functioning;
- scales the responses according to preference weightings for the relative importance members of society assign to each function, giving an overall score, representing the person's well being at a point in time;
- estimates prognosis in terms of 'Well-Life Expectancy'(adequate for groups of patients, rather than individuals.)

Reliability and Validity: Strong correlation found in repeated preference rating exercises, stable over time. Content validity demonstrated through consideration of mortality, symptoms, problems and functional levels - all components of the concept of health. Correlates adequately with other functional and psychological distress rating scales.
Sensitivity to Change: Despite relatively coarse function indicators, sensitivity to change has been demonstrated in a wide variety of studies of treatment effects e.g. COAD, AIDS, diabetes, and arthritis.
Relevance to IC: useful for programme evaluations which include economic evaluation.
Access: Interview schedule and manual: RM Kaplan PhD, Professor and Chief, Division of Health Care Sciences 0622, School of Medicine, University of California, San Diego, La Jolla, California, USA 92093-0622. Self report version was, as of 1996, under development, and may now be available.

PATIENT GENERATED MEASURES

Schedule for the Evaluation of Individual Quality of Life (SEIQoL)
O'Boyle et al. (1992)
Purpose: measures patient's level of functioning in five self-nominated facets of life and the relative weight or importance attached to these, in decisions about quality of life.
Description: a structured interview elicits five areas of the patient's life that they judge to be most important to their overall quality of life. Subjects then rate their current status against each area using a visual analogue scale labeled 'as good as can possibly be' and 'as bad as could possibly be'. They then give a global rating of QoL an a single horizontal visual analogue scale. An overall QoL score is then generated.
Reliability and Validity: test-retest reliability after treatment very good.
Sensitivity to Change: sensitive to treatment effects if health is an important factor to the patient, but if patient selects factors other than health as important to their quality of life, may not demonstrate change over time.

Relevance to IC: given the emphasis in intermediate care upon attainment of patient's therapeutic goals, patient generated evaluations of the impact on quality of life are conceptually attractive. Moreover, these measures are more likely to be responsive to small changes. A drawback is that, since a standardised tool is not being employed, comparisons cannot be drawn across populations. It is most useful linked with disease-specific and global health status measures.

Reintegration to Normal Living Scale (Wood-Dauphinee et al, 1988)
Purpose: proposed as a proxy measure for quality of life. Focuses on functional capacity and compensation for loss, having been developed for use in younger people with a disability from which full recovery is not anticipated.
Description: 11 item Likert scale asking patients to rate the extent to which each item describes their situation. Addresses mobility in and out of the home, meeting self care needs, occupation, recreation, social activity, family role attainment, satisfaction with self and personal relationships, and capacity to deal with life events. Can be used in interviews, or as self report.
Reliability and Validity: adequate for use in group comparisons.
Sensitivity to Change: no data available. However, Likert format offers the possibility of measuring change over time.
Relevance to IC: Can be used regardless of clinical diagnosis, to address transition to self care and personal independence. Would need to be used with other measures.

Appendix 2: Glossary of Terms

Casemix	A measure of how severely ill a person is, and/or a measure of the underlying health status prior to illness. Evaluations should adjust for casemix, to avoid making unfair judgements of a programme's worth (e.g. if a ward with not very sick people is compared to a ward with extremely sick people, and found to have better outcomes, that is not a fair comparison unless the casemix has been measured and included in the analysis).
Complement	An add-on to existing arrangements (see **substitute**).
Condition-specific measure	A measure that assesses health for one particular illness or set of symptoms (also called **disease-specific**).
Cost	Any type of resource input.
Cost benefit analysis	A form of economic evaluation that describes all benefits in monetary terms, so that a ratio of costs to benefits produces a figure in terms of pounds; the programme or intervention with the lowest cost per benefit is selected.
Cost consequence analysis	A form of economic evaluation that arrays costs next to the range of different effects so that decision makers can place their own weights, or value, on those effects.
Cost effectiveness analysis	A form of economic evaluation that selects one measure of effectiveness (e.g. 10-point gain on a quality of life scale) and produces a ratio of the cost per 10-point gain; the lowest cost per quality improvement is selected.
Cost minimisation analysis	A form of economic evaluation where benefits are assumed to be equal under different interventions, so that only costs need to be measured; the lowest cost programme is chosen.
Cost utility analysis	A form of economic evaluation that attempts to capture mortality and judgements about quality of life under different health or disability

	circumstances in a single score (e.g. QALYs) and then uses the score in a ratio of cost per utility unit; the programme with the lowest cost per QALY is chosen.
Counterfactual	The visualising of what a study subject would have experienced if he or she had received a treatment other than the one actually received, all other things being equal (e.g., if he or she had received early intervention for a medical problem instead of being admitted to hospital farther along). Estimation of the counterfactual is at the heart of quantitative analysis (see **summative evaluation**).
Dimension-specific measure	A measure that assesses one particular aspect of health, such as physical function, psychological well-being, mobility, quality of life, endurance, etc.
Economic evaluation	A judgement of a programme's or intervention's value that incorporates both its costs and outcomes. (See **cost-benefit analysis, cost-effectiveness analysis, cost consequence analysis, cost minimisation analysis**, and **cost-utility analysis**.)
Focus group	A form of group interview that explicitly builds on interaction between participants in order to generate data. The method is used to help people to explore and clarify their views in ways that would be difficult in one-to-one interviews.
Generic measure	A global summary measure of health status (e.g. SF-36).
Implementation evaluation	A form of process evaluation that focuses particularly on the dynamics of introducing a change.
Intermediate outcome	A measure which, though not an endpoint in itself, is known to be associated with that endpoint; thus, if the intermediate outcome can be achieved, one can be confident that the longer-term outcome is likely to be achieved as well.

Inter-rater reliability	A method to ensure that information would be collected in the same way by different people; tested by having two or more people independently abstract data from the same source and comparing their results (which should be identical; if not, more training or a new approach is needed).
Marginal cost	The cost resulting from an additional unit of activity or output
Maximum variety sampling	An approach used in qualitative research to selecting people to evaluate; although a small group is chosen, an effort is made to include a mix where each person is different in some way from the one before; in that manner, the variety of perspectives is as great as possible. (See **representative sampling**)
Narrative analysis	The telling of stories and analysis of content is accepted as a legitimate way of accessing individuals perceptions of life events. The endeavour is said to arise from the work of Flannagan in developing training programmes for pilots in the late 1940's and lies within the field of naturalistic enquiry.
Observational study	A form of treatment/comparison group design in which the researchers do not determine which subjects are in the control group and which are in the treatment group. Many issues can only be studied observationally, and all observational studies must deal with problems of possible confounding factors (for example, sickness may be associated with whether a patient receives intermediate care or not, and also associated with the outcome of care).
Opportunity cost	The value of the next-best thing; the value of resources which could have been used for other purposes.
Outcome	A measure of health used specifically as an endpoint or a dependent variable.
Outcome evaluation	An approach to judging the value of a programme or intervention that focuses on the scientific measurement of association between

	the intervention and a desired endpoint (also called **summative evaluation**, because it sums up the effects of a change; in contrast, see **process evaluation**).
Patient-generated outcome	A measure that elicits individual patients' definitions and gradations of health change which they find personally meaningful.
Power calculations	The use of statistical formulas to determine the number of patients to be included in a quantitative study, in order to have a reasonable (80-90%) chance of detecting a difference of pre-determined size (see **significance**).
Process evaluation	Can have two meanings: (1) The measurement or description of what exactly constitutes a particular intervention (in intermediate care, what exactly is done, by whom, under what circumstances) - in order that the model can be replicated. (2) An approach to judging the value of a programme or intervention that uses the evaluation itself to help develop and improve the service as it goes (also called **formative evaluation**, because it attempts to form the new service or model, in order to produce a change; in contrast, see **outcome evaluation**).
Randomised controlled trial	The gold standard of **outcome evaluation**, because it enables clear estimation of the **counterfactual**. The essence of it is to use a comparison group that has been chosen at random.
Repeated measures analysis	A form of statistical modelling that is useful in longitudinal studies.
Representative sampling	An approach used in quantitative research to selecting people to evaluate.
Significance	A statistical term referring to the probability that the result observed in your study could have occurred by chance; if that probability is low (conventionally, less than 5% or less than 1%) then the finding is said to be statistically significant. However, it is good practice to consider whether a finding is also significant in practical terms as well (e.g. if a Barthel score

	changes by one point, on average, that might be statistically significant but is unlikely to affect actual health or function).
Substitute	A replacement for existing arrangements (see **complement**).
Triangulation	A qualitative method of analysis which collects data from a variety of sources and using a variety of techniques, in an effort to check the validity of one against the other and to identify a common 'truth' through consensus.
Utility	A quantification of the value people place on something. In economics, utility is sometimes defined as 'happiness' in that one tries to increase (or 'maximise') their utility.

Appendix 3: Bibliography

Baker J.E., Goldacre M., Gray J.A.M. (1986) Community hospitals in Oxfordshire: their effect on the use of specialist inpatient services. *Journal of Clinical Epidemiology and Community Health*, 40:117-120.

Bauby J.D. (1997) *The Diving Bell and the Butterfly.* (translated by J. Leggatt) London: Fourth Estate.

Bingham R.D., Ethridge M.E. (Eds.) (1982) *Reaching Decisions in Public Policy and Administration: Methods and Applications*, New York and London: Longman Inc.

Bruner J. (1987) Life as Narrative *Social Research* 54 (1) 11-32.

Carr-Hill R., Dixon P., Thompson. A. (1989) Too Simple for Words *Health Services Journal* 99 (1) 728-729.

Chalmers I., Altman D. (Eds.) (1995) *Systematic Reviews*, London: BMJ Publishing Group.

Cohen D. (1994) Marginal analysis in practice: an alternative to needs assessment for contracting health care. *British Medical Journal*, 309:781-785.

Connell N.A.D., Lees, D., Lees, P., Powell, P., Stafford, P., Sutcliffe, C. (1996) Costing and Contracting in the NHS: A Decision Support Approach, In Bourn M., Sutcliffe, C. (Eds.), *Management Accounting in Healthcare.* London: CIMA, pp. 31-40.

Drummond M. (1994) *Economic Analysis alongside Clinical Trials.* Leeds: Department of Health.

Drummond M.F., Stoddart G.L., Torrance G.W. (1987) *Methods for the Economic Evaluation of Health Care Programmes.* Oxford: Oxford University Press.

EuroQol Group. (1990) EuroQol - a new facility for the measurement of health related quality of life *Health Policy*, 16:199-208.

Evans, D., Steiner, A. (1997) A guide for local evaluation of PMS pilots. In McKeon, A., *Personal Medical Services Pilots under the NHS (Primary Care) Act 1997: A guide to local evaluation.* NHS Executive, HMSO, catalogue no. 97PP0130, 11-47.

Fielding N., Fielding J. (1986) *Linking Data, Qualitative Research Methods Series No. 4.* London: Sage Publications.

Flanagan J. (1954) The Critical Incident Technique, *Psychological Bulletin* 51 327-358.

Giovannetti P. (1986) Evaluation of primary nursing in *Annual Review of Nursing research* vol. 4 pp 127 151 (Weilty H H, Fitzpatrick J J, and Taunton R L eds.) New York: Springer, Louis, CV Mosby.

Gladman J., Forster A., Young J. (1995) Hospital- and home-based rehabilitation after discharge from hospital for stroke patients: analysis of two trials. *Age and Ageing*, 24:49-53.

Guba E., Lincoln Y. (1981) Effective Evaluation, San Francisco: Jossey-Bass.

Guyatt G., Berman L., Townsend M. et al. (1987) A measure of quality of life for clinical trials in chronic lung disease. Thorax, 42:773-778.

Hart E. Bond N. (1995) *Action Research for Health & Social Care: a guide to practice*, Buckingham: Open University Press.

Hornbrook M.C. Goodman M.J. (1991) Managed care: Penalties, autonomy, risk and integration. In Hibberd H., Nutting P.A., Grady M.L.(Eds.) *Primary Care Research: Theory and Methods* Washington, DC: US Dept. of Health & Human Services, Public Health Service, AHCPR,107-126.

Kemmis S. (1982) 'Action research in retrospect and prospect' in Henry C., Hook C., Kemmis S., McTaggart R. (eds) (1982) *The Action Research Reader* Geelong Deakin University.

King's Fund London Commission. (1997)*Transforming Health in London*. London: King's Fund.

Laupacis A., Feeny D., Detsky A., Tugwell P. (1992) How attractive does a new technology have to be to warrant adoption and utilisation? *Canadian Medical Association Journal*, 146:473-481.

Lewin K. (1958) *The Group Decision and Social Change in*: Macoby E (Ed) *Readings in Social Psychology,* London: Holt, Rinehart and Winston.

Lindblom C.E. (1959) The science of muddling through. *Public Administration Review*, 19:79-88.

McDowell I., Newell C. (1996) Measuring Health: A guide to rating scales and questionnaires, New York: Oxford University Press.

McGlynn, E.A. (1997) Six challenges in measuring the quality of health care, *Health Policy*, 16:4-21.

McTaggart R. (1991) *Action Research : a short modern history,* Geelong, Deakin University Press.

Nelson E.C., Wasson, J.H., Johnson D.J. and Hays R.D. (1996) Dartmouth COOP functional assessment charts: brief measures for clinical practice, in B Spilker (Ed.) *Quality of Life and Pharmacoeconomics in Clinical Trials (2nd edition)*. Philadelphia: Lippincott-Raven.

Netten A. (1994) *Unit Costs of Community Care*. Kent: Personal & Social Services Research Unit.

NHS Centre for Reviews and Dissemination (University of York) and Nuffield Institute for Health (University of Leeds) (1996) *Effective Health Care Bulletin : preventing falls and subsequent injury in older people*. Published in association with Churchill Livingstone Vol. 2, No 4. 1-16.

Norman I., Redfern S., Tomlin D., Oliver S. (1992) Developing Flanagan's Critical Incident Technique to Elicit indicators of High and Low Quality nursing Care from Patients and their Nurses, *Journal of Advanced Nursing* 17 590-600.

Parlett M., Hamiliton D. (1977). In Hamilton D. (Ed) *Beyond the Numbers Game: A Reader in Educational Evaluation,* Basingstoke: Macmillan.

Payne S., Betts P. (1995) Explanatory models of diabetes. Project funded by the Nuffield Foundation, as described in Steiner A. (Ed.) *Health in Social Sciences*, Special issue of *Research News*, Southampton: University of Southampton.

Pearson A., Punton S., Durant I. (1992) *Nursing Beds: An Evaluation of the Effects of Therapeutic Nursing*. Royal College of Nursing Research Series. Harrow: Scutari Press.

Pearson A., Punton S., Durant I. (1992) *Nursing Beds: An Evaluation of the Effects of Therapeutic Nursing*. Royal College of Nursing Research Series. Harrow: Scutari Press.

Pettigrew A. (1985) *The Awakening Giant, continuity and change in ICI* Oxford: Blackwell.

Roberts H., Philip I., Bray J. (1996) *A Study of Needs Assessment and Goalsetting among Elderly People in the Community*. Project report to the South and West Regional Health Authority. Southampton: Elderly Care Research Unit, Southampton General Hospital.

Robinson R. (1993) Economic evaluation and health care: costs and cost minimisation analysis. *British Medical Journal*, 307:726-728.

Robinson R. (1993) Economic evaluation and health care: Cost-benefit analysis. *BMJ*, 307:924-926.

Robinson R. (1993) Economic evaluation and health care: Cost-effectiveness analysis. *BMJ*, 307:793-795.

Robinson R. (1993) Economic evaluation and health care: Costs and cost-minimisation analysis. *BMJ*, 307:726-728.

Robinson R. (1993) Economic evaluation and health care: Cost-utility analysis. *BMJ*, 307:859-862.

Robinson R. (1993) Economic evaluation and health care: The policy context. *BMJ*, 307:994-996.

Robinson R. (1993) Economic evaluation and health care: What does it mean? *BMJ*, 307:670-673.

Rosser R., Kind P. (1978) A scale of valuations of states of illness: is there a social consensus? *International Journal of Epidemiology*, 7:247-258.

Shepperd S., Doll H., Jenkinson C. (1997) Randomised controlled trials. In C Jenkinson (Ed.), *Assessment and Evaluation of Health and Medical Care: A methods text*. Buckingham: Open University Press, p.6.

Spector W.D., Katz S., Murphy J.B., Fulton J.P. (1987) The hierarchical relationship between activities of daily living and instrumental activities of daily living. Journal of Chronic Disability, 40:481-489.

Steiner A. (1997) *Intermediate Care: A conceptual framework and review of the literature*. London: King's Fund.

Steiner A., Vaughan B. (1997) *Intermediate Care: A discussion paper arising from the King's Fund seminar held on 30th October 1996.* London: King's Fund.

Stokey E., Zeckhauser R. (1978) *A Primer for Policy Analysis*. New York & London: W.W. Norton & Company.

Stuck A.E., Siu A.L., Wieland G.D., Adams J., Rubenstein L.Z. (1993) Comprehensive geriatric assessment: a meta-analysis of controlled trials. *The Lancet*, 342:1032-1036.

Titchen A., Binnie A. (1992) Research Partnerships: Collaborative Action Research in Nursing, Insititute of Nursing, Oxford.

Todd C.J., Freeman C.J., Camilleri-Ferrante C., Palmer C.R., Hyder A., Laxton C.E., Parker M.J., Payne B.V., Rushton N. (1995) Differences in mortality after fracture of hip: the East Anglia audit. *British Medical Journal*, 310:904-908.

Towell D., Harries C. (Eds) (1979), *Innovation in Patient Care*, London : Croom Helm Ltd.

Townsend J., Piper M., Frank A.O., Dyer S., North W.R.S., Meade T.W. (1988) Reduction in hospital re-admission stay of elderly patients by a community based hospital discharge scheme: a randomised controlled trial. *British Medical Journal*, 297:544-547.

Wandelt M., Ager J. (1974) *Quality of Patient Care Scale,* New York: Appleton Century Croft.

Watson K. (1997) Economic evaluation of health care. In C Jenkinson (Ed.), *Assessment and Evaluation of Health and Medical Care: A methods text.* Buckingham: Open University Press.

Williams G. (1984) The Genesis of Chronic Illness: Narrative Reconstruction, *Sociology of Health and Illness* 6 (2) 175-200.

Wood-Dauphinee S., Williams J.I. (1987) Reintegration to normal living as a proxy to quality of life. *Journal of Chronic Disability*, 40:491-499.

Appendix 4: Evaluation Workshop 29/10/97

Participant List

Ms Lynne Barr	Director of Occupational Therapy South Tees Acute Hospital NHS Trust
Ms Trish Bennet	Nurse Practitioner Coordinator, Sir Alfred Jones Memorial Hospital
Mrs Pat Cantrill	Assistant Chief Nursing Officer, Department of Health
Dr Angela Coulter	Director, Policy and Development, King's Fund
Ms Ami David	Clinical Director, Community Health Service, Ravensbourne Trust
Dr Edward Dickinson	Associate Director, Research Unit, Royal College of Physicians
Dr Sue Dopson	Fellow, Templeton College, Oxford University
Ms Sue Dowling	Consultant Senior Lecturer, Bristol University
Mr Mark Etheridge	Researcher, Greenwich University
Ms Nora Flannagan	Director of Nursing, St Albans & Hemel Hempstead NHS Trust
Ms Mo Flynn	E.M.D.O.C., Orpington Hospital, Kent
Ms Louise Forward	Researcher, Health and Community Care Research Unit, University of Liverpool
Dr Sarah Furlong	Project Officer, Exploring New Roles in Practice, King's Fund
Dr Steve Gillam	Director, Primary Care Programme, King's Fund
Ms Jane Gooch	Senior Nurse, South Manchester University Hospitals NHS Trust
Ms Pippa Gough	Assistant Director of Policy, Royal College of Nursing
Ms Pat Gordon	Director, London and Northern Health Partnership, King's Fund
Dr Alastair Gray	Fellow, Wolfson College, Oxford
Ms Jenny Greer	Mancunian Community Trust
Mr Philip Hadridge	Service Development Manager, Anglia & Oxford NHS Executive
Ms Nancy Hallet	Director of Nursing, Homerton Hospital NHS Trust
Mrs Linda Hanford	Project Manager, Intermediate Care, King's Fund
Mr Neil Jessop	Assistant Director of Planning and Purchasing Enfield & Haringey Health Authority
Ms Fiona Johnstone	Liverpool Health Authority, Liverpool Health Authority
Dr Philip Leech	Principle Medical Officer, Quarry House
Ms Sue Last	Continuing Care Manager, Liverpool Health Authoirty
Mr Peter Lees	Director of Research and Development, Southampton General Hospital
Ms Tessa Lomax	Manager, Primary Health Care Projects, Plymouth Social Services
Ms Lisa MacFarlane	Research and Development, Southampton General Hospital

Ms Liz Mars	Development Manager, Rehabilitation Pilot Project Northenden Health Centre
Ms Mary Morrison	Health Advisor, Community Care Alarm Service, Newcastle upon Tyne
Dr Penny Newman	Consultant in Public Health/Primary Care, Havering Hospitals Trust
Mr Bamber Postance	ICB Research Co-ordinator, University of Greenwich
Ms Jane Price	Greenlalgh & Co. Ltd
Mrs Janice Robinson	Director, Community Care Programme, King's Fund
Ms Thelma Sackman	Nursing Directorate, Department of Health
Ms Judy Sanderson	Head of Policy for Disabled People, Department of Health
Ms Sasha Sheppard	Health Services Research Unit, University of Oxford
Ms Maureen Silcott	Clinical Audit Department, Wythenshawe Hospital
Dr Andrea Steiner	Institute for Health Policy Studies, Southampton
Ms Margaret Stockham	Deputy Chief Executive, Bedford and Shires Health Care NHS Trust
Ms Barbara Vaughan	Director, Nursing Developments, King's Fund
Ms Bronagh Walsh	School of Nursing and Midwifery, Southampton