Developing a mixed methods framework for process evaluations of complex interventions: the case of the National Exercise Referral Scheme policy trial in Wales

Graham Moore

Cardiff Institute of Society and Health, School of Social Sciences, Cardiff University

Developing a mixed methods framework for process evaluations of complex interventions: the case of the National Exercise Referral Scheme policy trial in Wales

Graham Moore

Cardiff Institute of Society and Health, School of Social Sciences, Cardiff University

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

CANDIDATE'S ID NUMBER	0635565
CANDIDATE'S SURNAME	MOORE
CANDIDATE'S FULL FORENAMES	GRAHAM FRANCIS

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ..... (candidate) Date ...3/8/10.....

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed ..... (candidate) Date ...3/8/10.....

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references[1].

Signed ..... (candidate) Date ...3/8/10.....

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate) Date ...3/8/10.....

[1] The thesis is linked to the pragmatic randomised controlled trial of the National Exercise Referral Scheme in Wales, conducted between 2006 and 2010. The trial was led by my supervisors Dr Simon Murphy and Professor Laurence Moore, with Dr Lawrence Raisanen acting as trial manager. I was not involved in the design of the trial or of the nested cost effectiveness evaluation, whose findings will be reported elsewhere. I however led the design, conduct and analysis of the process evaluation under the guidance of my supervisors, including all data components reported in this thesis, with the following exceptions. The thesis includes secondary data analysis of a routine monitoring database whose development was led by the trial manager and baseline questionnaire data collected as part of the trial. Though I had no input into the design and collection of these data sources, the analyses reported in this thesis were developed and conducted by myself. In addition, qualitative data from an interview with policy representatives incorporated questions developed by myself, though these questions were incorporated into an interview conducted by another member of the evaluation team, in order to avoid the need to approach these individuals for 2 separate interviews. Finally, qualitative interviews with patients were led by myself, though a second interviewer was present at all interviews.

Acknowledgements

I would like to thank my supervisors, Dr Simon Murphy and Professor Laurence Moore, for their support in the development and conduct of this thesis. I am also grateful to the NERS evaluation team, the scheme's national coordinators Elaine McNish and Jeannie Wyatt-Williams, and the research contract manager Janine Hale within the Welsh Assembly Government for allowing me to link my thesis to the wider evaluation, as well as to the professionals, coordinators and patients who provided data for the study. I would also like to thank Dr Claire Lane for advising on motivational interviewing fidelity monitoring procedures and for providing training materials, Professor Stephen Rollnick for comments on drafts of findings relating to motivational interviewing, Eleri Sian Jones for acting as second coder of recordings, Dr Nafees Din, Dr Pat Linck and Hayley Collicott for data collection assistance. As part of my PhD research training, I spent 2 months on an Overseas University Visit in the University of Melbourne, observing and contributing to the work of the McCaughey Centre, whilst continuing with writing up this thesis. I would therefore also like to thank Professor Elizabeth Waters for making this opportunity possible and Britt Johnson, Elisha Riggs and Rachel Clark for their hospitality and support. Finally, I would like to thank the Economic and Social Research Council for funding my studentship.

Articles published from the NERS evaluation

Murphy S, Raisanen L, Moore G, Edwards R, Linck P, Williams N, Ud Din N et al. (2010) A pragmatic randomised controlled trial of the Welsh National Exercise Referral Scheme: protocol for trial and integrated economic and process evaluation. BMC Public Health 10: 352.

Summary

Where possible, policies to improve public health should be evidence-based. Where political pressures and shortage of evidence force action in advance of evidence, effectiveness can be evaluated during policy rollout. Because the aetiology of public health issues is complex, successful policies will likely be complex in their design, their implementation and their interaction with their contexts and target audiences. Process evaluation is therefore crucial in order to inform consistent implementation, and alongside outcomes evaluation, in order to understand how outcomes are produced. However, limited methodological guidance exists for process evaluation.

This thesis develops a mixed-method framework exploring programme theory, diffusion, implementation, participant experiences and reach, which is applied to the evaluation of the Welsh National Exercise Referral Scheme (NERS). A logic model is developed via discussions with policy representatives. Diffusion is explored via qualitative interviews with policy representatives and local coordinators. Implementation checks draw on routine data, observation and self-report. Participant experiences are explored via qualitative interviews. Social patterning in reach is explored using routine monitoring data.

The study identifies challenges diffusing NERS into local practice, in relation to communication structures, support, training provision and the mutual adaptation of the scheme and its contexts. Implementation checks indicate a common core of discounted, supervised, group-based exercise, though some divergence from programme theory emerged, with unfamiliar activities such as motivational interviewing and patient follow-up protocols delivered poorly. Nevertheless, relatively high adherence rates were achieved. Key perceived active ingredients in practice included professional supervision, enabling patients to build confidence and learn to exercise safely, and the patient-only environment, seen as providing an empathic context and realistic role models. However, lower uptake emerged amongst non-car owners, with higher adherence amongst patients already moderately active at baseline, older patients and non-mental health patients. Implications for ERS implementation, outcomes interpretation and process evaluation methodology are discussed.

Table of contents

Table of tables

Table of figures

1 Introduction

1.1 Background and aims

Attention is increasingly turning to the need for evidence-based public health policy in order to use public money efficiently and avoid exposing target audiences to unexpected harms (Wanless et al., 2007; House of Commons Health Committee, 2009). Given the complex aetiology of problems targeted by public health, successful policies will often be complex and multi-faceted. Increasing complexity engenders significant challenges in developing evidence of what works. Recent Medical Research Council (MRC) guidelines therefore present a model to guide the development and evaluation of complex interventions (Craig et al., 2008a; Craig et al., 2008b), beginning with development of theory, before moving to feasibility testing ahead of definitive evaluation and promotion of wider usage.

Whilst randomised controlled trials (RCTs) are advocated by the MRC as the most robust means of establishing effectiveness within definitive evaluation stages where feasible, RCTs have been criticised for their inability to offer insights into the complex nature of causality within complex interventions. Implementation of complex interventions occurs in a somewhat fluid manner, whilst outcomes are not directly caused by the intervention, but produced in the interaction of target audiences with the intervention in context (Pawson and Tilley 1997). MRC guidelines respond to such criticism by arguing that process evaluation ought to be conducted within RCTs, in order to retain strengths in terms of internal validity, whilst also understanding the implementation and functioning of the intervention in context.

Whilst process evaluation is crucial within any definitive evaluation, challenges faced by process evaluation are likely amplified in policy evaluation settings, where evaluation is embedded into policy rollout and it is therefore not possible to adhere strictly to the development and evaluation process described above. The MRC argue that where limited theory development and piloting work has taken place, the risk of weak implementation is heightened, complicating the challenges associated with understanding programme implementation and functioning. As the need to embed

evaluation into policy rollouts becomes increasingly recognised therefore (Wanless 2007; House of Commons Health Select Committee 2009), the need for high quality process evaluation perhaps becomes increasingly critical.

Process evaluation is however a poorly demarcated field, led by few guiding principles. Whilst a number of frameworks propose aims for process evaluation (Steckler and Linnan 2002), these typically stop short of offering comprehensive methodological guidance. The principle aim of this thesis is therefore to develop and implement a comprehensive framework for process evaluation of complex interventions, and to reflect on its usefulness for understanding the implementation, and functioning of complex interventions. The context for this will be the policy trial of the National Exercise Referral Scheme (NERS) in Wales. NERS was implemented in Wales in 2007 in an attempt to standardise exercise referral practice throughout Wales, with a pragmatic RCT and nested economic and process evaluation, built into its first year.

Effectiveness and cost-effectiveness data will be reported elsewhere and hence process evaluation is presented largely as a stand-alone activity. An understanding of implementation is useful as an end in itself, informing efforts to ensure that implementation remains coupled to the theoretical perspectives informing intervention development. Process evaluation however adds value to, rather than acting as a substitute for, robust effectiveness evaluation, answering questions which must be first addressed if effectiveness data are to be subsequently interpreted. Oakley et al. (2006) recommend analysis and interpretation of process evaluation in advance of outcomes analysis in order to minimise tendencies for post-hoc rationalisation; a model adhered to in this thesis. Hence, process evaluation is treated as a precursor to outcomes evaluation, with implications of process data for subsequent analyses of trial outcomes and directions for integration of process data into analysis of outcomes discussed throughout.

1.2 Research design

Whilst early MRC guidelines focused on formative functions of process evaluation in theory development (Campbell et al. 2000), more recent guidance has largely centred

around implementation, starting from the assumption that a clear theory is in place prior to process evaluation (Steckler and Linnan 2002). This thesis combines both perspectives in order to understand how NERS is conceived, delivered and experienced in practice. A mixed-methods design is adopted which rotates between qualitative exploration of causal processes and quantification of intermediate outcomes. Methods include discussions with policy representatives to elicit programme theory, semi-structured interviews with national and local policy representatives to explore diffusion, a combination of structured observation, implementer self-report and routine monitoring data to assess implementation, semi-structured interviews with patients and exercise professionals to explore patient experiences and use of routine monitoring data and baseline trial data to explore patterning in programme reach.

1.3 Overview of chapters

Chapters 2 to 4 are the literature review chapters. These begin in Chapter 2 with an introduction to MRC guidelines, and introduction to the need for process evaluation and challenges in understanding the implementation and functioning of complex interventions in policy evaluation settings. Chapter 3 then goes onto develop the framework used in this study, through a review of current practice and discussion of potential modifications to existing frameworks. Chapter 4 focuses specifically on the example of exercise referral schemes, using themes developed in Chapter 3 to identify key weaknesses in their evaluation and illustrate how systematic process evaluation may help to move this field of research forward. The methodological approach of this thesis is then described in detail in Chapter 5.

Chapters 6 to 8 present findings. Given the complexity of the research design, each chapter is accompanied by a recap of the data sources used, and comprehensive discussion of findings and their implications for implementation and trial outcomes. Chapter 6 first describes the development of a logic model to define NERS programme theory. The chapter then turns to exploration of how the intervention was diffused into practice, using qualitative data from national and local policy representatives. This offers key insights into how implementation structures, contextual factors and implementer agency shape local adoption of the intervention,

as well as offering insights into challenges implementing the NERS model and the processes through which the scheme becomes routinised into its context. Finally, the chapter presents quantitative data, which ultimately describes the outcomes of diffusion processes, evaluating consistency of the delivered intervention with programme theory.

Chapter 7 then pursues two issues uncovered by implementation checks. Two planned components of NERS, motivational interviewing and goal setting, were delivered poorly. In the case of MI, this was predicted when protocol documents revealed that only one hour of training had been provided. Concerns were communicated to policy representatives, leading to implementation of training courses. The first half of the chapter involves a mixed method sub-evaluation, combining longitudinal monitoring of consultations before and after training with qualitative interview data from exercise professionals and the MI training provider to explore for whom and under what circumstances practice begins to move closer to an approach consistent with MI. In the case of goal setting, concerns were expressed by the trial manager that routine monitoring data indicated that goal setting was poor. Policy representatives communicated the need for setting of measurable and time-bound goals to coordinators. The second half of the chapter involves secondary analysis of routine monitoring data, exploring whether goal setting quality improved over time and whether higher quality goal setting processes were linked to programme adherence.

Having defined the intervention and explored the impacts of efforts to keep implementation on track, Chapter 8 focuses on understanding how the delivered intervention is experienced by patients and implementers in order to understand how it is perceived to be working and the emergence of social patterning in programme reach. The chapter uses qualitative interviews with patients and exercise professionals and quantitative analysis of routine monitoring data exploring patterning in uptake and adherence by baseline characteristics. The thesis concludes in Chapter 9 with an overview of the implications of findings reported for the implementation of ERS and for interpretation of outcome effects, and methodological reflections on the implementation of the process evaluation framework.

2 Evaluating complex interventions

2.1 Chapter overview

This chapter aims to explore key challenges in evaluating complex interventions, and to introduce the need for comprehensive process evaluation in order to understand the implementation of complex interventions and guide interpretation of their outcomes. The chapter begins by introducing the Medical Research Council (MRC) framework for development and evaluation of complex interventions, describing the proposed development and evaluation process, aims and methods for evaluation and functions of process evaluation within definitive evaluation stages. It will then be argued that MRC guidelines represent a useful ideal model, but that in some cases, it is necessary to integrate evaluation into policy rollouts, likely involving evaluating interventions which have been introduced into a wide range of contexts with limited formative work. The extent to which this approach increases the challenges associated with understanding the implementation and functioning of complex interventions, and the increasingly critical role of process evaluation as the value of embedding evaluation into policy rollout becomes increasingly recognised, will then be discussed.

2.2 The development and evaluation of complex interventions: MRC guidelines

Health promotion and public health are inherently political activities, driven by a desire to improve quality of life and reduce healthcare costs (Foresight, 2007). Many well-intentioned interventions have however proven unsuccessful or counterproductive, such as a bicycle education programme which increased accidents amongst children, presumably through increasing confidence and risk exposure (Carlin, Taylor and Nolan, 1998). Basing policies on good intentions and common sense may therefore prove fiscally wasteful, whilst harming those they are intended to help. Following the evidence-sceptical 'conviction politics' of Thatcherism, discourses of evidence-based policy have increasingly been emphasised since election of the 1997 Labour government (The Cabinet Office, 1998), with the need both to base decisions on evidence and to develop evidence as policies are rolled out increasingly recognised (House of Commons Health Committee, 2009).

Given the complex aetiology of outcomes targeted by public health interventions, recognition is growing that effective interventions are likely to be complex. A complex intervention is defined by the MRC as comprising multiple components whose actions interact to produce change (Craig et al., 2008a; Craig et al., 2008b). Whilst earlier MRC guidelines focused on complexity in intervention design, updated guidelines recognise that further dimensions of complexity include their implementation and their interaction with their settings. Understanding a complex intervention has been compared to the science of chemistry where interest lies not only in the independent actions of programme elements, but in the actions of the compounds produced by their combination in context (Hawe, Shiell and Riley, 2009). This section provides an overview of MRC guidelines, describing their proposed development and evaluation process, before focusing on aims and methods for evaluation and challenges for understanding the implementation and functioning of interventions arising from their complexity.

2.2.1 The development and evaluation process

Recognition of the need for consistent good practice in developing and evaluating complex interventions led to publication of initial MRC guidelines in 2000 (Campbell et al., 2000). In response to criticisms including the overly linear nature of the proposed development and evaluation process, excessive focus on models from pharmaceutical research and limited consideration of context, guidance was updated in 2008 (Craig et al., 2008a; Craig et al., 2008b). New guidance proposed an iterative development and evaluation process focusing upon: i) development, ii) feasibility and piloting, iii) evaluation and iv) implementation.

According to these guidelines, development should involve consideration of relevant evidence bases, identification and development of appropriate theory, and modelling of change processes and outcomes. At this stage, programme developers must ask themselves if they can unambiguously describe the intervention and articulate a coherent theory for how it will work. If not, further development is needed. Once developed, feasibility is established through piloting, exploring issues such as acceptability to the target audience and barriers and facilitators of delivery or

participation, as well as estimating variability in target outcomes for sample size calculations. This may involve small-scale studies combining assessment of outcomes with formative process evaluation. Refinements may then be made to intervention theory, with key components or delivery mechanisms further developed as it becomes apparent that some are infeasible, unacceptable or require additional development to overcome contextual barriers.

The fully developed intervention, whose feasibility is now tentatively established, should then be evaluated. Consideration should be given to practicalities of wider implementation throughout the development and evaluation process, to avoid wasting resources through conducting robust evaluations of interventions which cannot be implemented. However, efforts to encourage wider adoption and implementation should ideally follow the definitive evaluation stage.

2.2.2 Definitive evaluation: measuring outcomes and understanding processes

2.2.2.1 Measuring outcomes

Given that complex public health interventions involve commitment of substantial public money, quantification of outcomes is crucial. This is not always straightforward, as outcomes will likely be multiple, whilst different weight may be given to different outcomes by different stakeholders. Indeed, some authors have pointed to a tendency for health impacts of interventions whose primary objective is not health-related, such as income supplementation, to be ignored (Thomson et al., 2004). Hence, MRC guidelines highlight the need for anticipated outcomes to be clearly defined prior to evaluation, most likely during theory development, given that this stage will be concerned with understanding how outcomes will be produced.

In evaluating outcomes in medical settings, randomised controlled trials have been crucial in exposing weaknesses of observational studies. For example, whilst once thought to reduce various chronic diseases, trials of Hormone Replacement Therapy found that it increased many of these outcomes (Humphries and Gill, 2003). Due to its persuasive ability to eliminate confounding factors and establish internally valid

causal inferences, the RCT has become nested at the top of the hierarchy of evidence within medical research (Davies, Nutley and Smith, 2000). Following the New Labour drive for evidence-based policy, increasing attempts have been made to apply principles of evidence-based medicine to public health (Davies et al., 2000), with a recent report by the House of Commons Health Select Committee (2009) arguing that 'having a control group or project to which an intervention or treatment can be compared is a fundamental tenet of good research, so is randomisation' (p32). Hence, whilst traditionally used to establish 'efficacy' under ideal conditions, RCTs have been extended to real-world 'effectiveness' (Roland and Torgerson, 1998) and are recommended by the MRC as the most internally valid means of estimating outcomes (Craig et al., 2008a; Craig et al., 2008b).

There are many challenges in applying RCTs to complex interventions above and beyond those experienced in evaluating pharmaceutical interventions. Challenges include inability to blind participants and implementers to treatment conditions (Sibbald and Roland, 1998), reintroduction of biases due to patient attrition following receipt of an unfavoured treatment allocation (Gidlow et al. 2008), contamination arising from control group members awareness of the intervention elsewhere (Nutbeam et al., 1993; Stephenson and Imrie, 1998; Moore et al., 2007; Sanson-Fisher et al., 2007), difficulties negotiating ethical and political objections to randomisation (Isaacs et al. 2007) and ensuring that the trial and its participants offer an externally valid reflection of real world practice (McKee et al., 1999; Godwin et al., 2003; Rothwell, 2005). MRC guidelines therefore warn against methodological dogmatism and whilst critical of tendencies for randomisation to be too easily written off, explicitly acknowledge that RCTs cannot always be conducted. Impacts of smoke-free legislation for example could neither feasibly nor ethically be evaluated using an RCT, and hence relied upon observational methods (Holliday, Moore and Moore, 2009).

Where researchers can demonstrate that an RCT is feasible and ethical, it is the position of this thesis that it is the most suitable approach for answering effectiveness questions. The role of the RCT however remains a source of significant polarisation within the social sciences, with some insisting that it is the gold standard (House of

Commons Health Committee, 2009), whilst others argue that it is an unsuitable means of evaluating complex interventions due to its limited ability to illuminate the complex nature of causality (Pawson and Tilley, 1997; Tones, 1997; Green and Tones, 1999; Freeman, 2009; Mackenzie et al., 2010). As recognised within new MRC guidelines however, the significant energy invested in arguing against RCTs could perhaps be more usefully invested in understanding how concurrent research activities may allow evaluators to retain its strengths whilst also addressing its shortcomings. The need for process evaluation nested within RCTs in order to understand the implementation of complex interventions and how outcomes are produced will now be discussed.

2.2.2.2 Understanding causal processes, implementation and context: the need for process evaluaton

As described above, RCTs attempt to isolate the amount of change in outcomes of interest which is attributable to the introduction of an intervention. Effect sizes are derived from comparisons between two groups between whom the only meaningful difference is whether or not the intervention was offered. This analytical focus has however been described by critics as implicitly supporting an oversimplified 'if x then y' view of causality, drawing upon Humean notions of constant conjunction. According to critics, where sole focus is placed on aggregate impacts, interventions are presented as having absolute causal power to determine outcomes, and as acting in an undifferentiated manner among passive recipients (Pawson and Tilley, 1997; Clark, MacIntyre and Cruickshank, 2007; Berwick, 2008).

In arguing for a more nuanced understanding of causality, Pawson and Tilley's (1997) Realistic Evaluation describes an epistemologically flexible approach underpinned by critical realism (Sayer, 2000) which focuses not on whether interventions 'work', but on 'what works, for whom and under what circumstances'. To illustrate the generative view of causality underpinning Realistic Evaluation, it is perhaps helpful to take an example from the natural sciences. If one lit a match in a gas-filled room, one might expect an explosion. However, on most occasions, striking a match does not cause an explosion. Hence, the action does not deterministically cause the outcome. Instead, the action (striking the match) introduces a mechanism

(the spark), causing an outcome (the explosion) which is contingent upon the context (the gas-filled room). Applying this view of causality to complex interventions, Pawson and Tilley argue that these 'work' through introducing ideas and opportunities which are sufficiently suited to their settings and target populations to activate important mechanisms of change (Pawson and Tilley 1997). Interventions do not have absolute causal power to produce change, but change is produced in the interaction of the target audience with the intervention, an interaction impacted by the implementation of the intervention, participant characteristics and the context in which the intervention is experienced.

Understanding 'what works' therefore first requires clear definition of the intervention with which participants interact. Defining a complex intervention is however not always straightforward, even if one has a clear understanding of the intended intervention. In trialling a drug, one might be confident that it will not morph into something different as it is transported across sites, given that it is an inanimate object with no theories of change or conscious desires to help its recipients. Hence, evaluators may be concerned simply with issues such as compliance and dose. However, given that interactions between implementers and participants typically form the basis of complex interventions (Pawson and Tilley 1997; Berwick 2008), implementation will likely be somewhat fluid. As evidenced by a tendency for better outcomes in sites achieving fuller implementation (Saunders et al., 2006; Strange et al., 2006; Thyrian et al., 2007), once an intervention is developed, the causal chain linking it to impacts begins with activities to ensure effective delivery (Miller and Rose, 2009). Evaluation must therefore pay significant attention to understanding how implementation is achieved and clearly describing what was actually delivered.

Rather than being passive recipients of the delivered intervention, participants' backgrounds, values and circumstances will influence their interactions with it (Pawson and Tilley 1997). Though in RCTs, participant characteristics should not vary between groups, they will vary within groups, causing patterning which is ignored where RCTs only present aggregate analyses. Hence, a priori defined sub-group analyses may offer a useful means of examining emerging patterning and understanding 'for whom' the intervention works. Indeed, sub-group analysis is often

seen as a crucial phase in Realistic Evaluation, though should be accompanied by examination of the mechanisms through which patterning emerges (Connelly, 2002; Kazi, 2003; Clark et al., 2007).

Rather than being implemented in closed laboratory settings where the intervention can be isolated from other extraneous influences, complex interventions are delivered in open systems (Sayer 2000), where contextual factors will constrain and facilitate the actions both of implementers and of participants. Hence, both implementation and the manner in which the target audience interacts with the intervention will vary across contexts. Project Northland for example, a community based intervention to reduce alcohol misuse in children, achieved highly promising impacts in rural settings (Perry et al., 2002), though impacts were weaker in urban settings, with intended actions of the intervention perhaps drowned out by the multitude of pro-alcohol stimuli in the urban environment (Komro et al., 2008). Hence, an understanding of 'under what circumstances' the intervention can be most effectively implemented, and will be most likely to result in positive impacts, is crucial (Pawson and Tilley 1997). Like participant characteristics, contextual factors should be comparable between trial arms in RCTs, though will vary within trial arms, particularly in multi-site trials (Wolff, 2001).

It is unfair to suggest that experimentalists have relied fully upon the stereotyped view of causality that critical realists have often attributed to them. Indeed, Baert (2005) has argued that critical realists have commonly built a role for themselves through attacking straw-man caricatures of causal assumptions no longer embraced within the social sciences. Many advocates of RCTs do argue for a need to explain how outcomes are produced (Rychetnik et al., 2002), whilst recent years have seen increasing attention to the need to develop techniques such as multi-level modelling to capture variability in responses between patients, contexts and over time (Brown et al., 2008).

However, in many trials to date, sole emphasis has been placed upon estimating aggregate effects, with little attention paid to understanding how outcomes were produced. Indeed, in early MRC guidance (Campbell et al., 2000), no emphasis was

placed on the need to understand implementation during the definitive trial or to explore how target audiences interacted with it in context. The role of process evaluation activities was limited to intervention development phases, centring around defining the intervention and modelling its causal processes. Hence, complexity was conceived somewhat narrowly as relating solely to illuminating the actions of a 'black-box' intervention package, perhaps under the assumption that if well designed, this could be replicated in new contexts and would produce the same outcomes.

High profile recognition of the need to understand implementation amongst advocates of RCTs is however increasingly evidenced in documents such as the updated CONsolidated Standards of Reporting Trials (CONSORT) guidelines (Schulz et al. 2010), which state that 'sufficient details to allow replication, including how and when they were actually administered' should be included. Earlier guidelines made no mention (Altman, 1996) or less explicit mention of the need for detailed assessment of implementation (Moher, Schulz and Altman, 2001). Furthermore, updated MRC guidelines reflect the increasing influence of approaches such as Realistic Evaluation, describing a need for process evaluation to serve functions including assessment of implementation quality, as well as clarification of causal mechanisms and identification of contextual influences on implementation and functioning. Indeed, these guidelines acknowledge that only through incorporating process evaluation and understanding how the intervention is implemented and how it produces change in context is it possible to 'build a cumulative understanding of causal mechanisms, design more effective interventions and apply them appropriately across groups and settings' (Craig et al., 2008a).

Whilst debates surrounding the perceived oversimplification of causality within RCTs have been used to argue against RCTs and for approaches focusing on theories of change and implementation however (Mackenzie et al., 2010), process evaluation is seen by the MRC as a vital accompaniment to, but not a substitute for robust effectiveness evaluation (Craig et al. 2008a). Process evaluation may serve valuable stand-alone functions, such as informing efforts to ensure that implementation of complex programmes maintains consistency with programme theory, and providing generalisable guidance on how to ensure that evidence based principles are applied in

practice. However, just as outcomes evaluation cannot fully stand without process evaluation, the contribution of process evaluation is significantly enhanced where conducted as part of a wider evaluation package including robust assessment of effectiveness. Process evaluation ought to be seen therefore both as a means of informing high quality implementation of complex interventions and as a precursor to understanding outcomes, with process evaluation data analysed and interpreted in advance of outcomes analyses in order to avoid post-hoc rationalisation and provide a priori guidance for subsequent outcomes interpretation (Oakley et al. 2006).

2.2.3 Summary

In summary, MRC guidelines recommend an iterative development and evaluation process, consisting of development, feasibility and pilot testing, evaluation and implementation, with each phase feeding back into the previous and forward into the next. Guidelines recommend a comprehensive approach to evaluation focusing on effectiveness, implementation and cost-effectiveness. RCTs are recommended for outcomes evaluation where feasible, but due to the contingency of outcomes on human agency and context, must be accompanied by concurrent process evaluation if quality of implementation is to be ensured and findings interpreted. Process evaluation will likely serve important stand-alone functions in guiding high quality implementation of complex interventions. However, when nested within RCTs, process evaluation offers a means to retain the strengths of the RCT in terms of understanding effectiveness, whilst addressing its limitations in terms of understanding implementation and causality.

Whilst thus far, discussion has focused on challenges in developing and evaluating complex interventions, assuming researchers to have a degree of control over both development and evaluation, increasing attention is turning to the need for evaluation to be embedded into policy rollouts. The role of policy evaluation in developing evidence for public health interventions, and challenges in understanding the implementation and functioning of policies as they are implemented will now be discussed.

2.3 Challenges in applying MRC guidelines to evaluations embedded within policy rollout

MRC guidelines recognise that there may be situations in which evaluators are forced to go straight to definitive evaluation, or to evaluate a scheme whose implementation has begun, such as where commissioned to evaluate an existing or imminently planned policy. Indeed, embedment of evaluation into policy rollout is a crucial means of developing inadequate evidence bases and informing future policy development. This section discusses the need for evaluation to be built into policy rollouts, before going on to discuss challenges for understanding implementation and functioning of complex interventions which may arise in such circumstances.

2.3.1 Evidence based policy or policy based evidence?

Whilst earlier MRC guidance (Campbell et al., 2000) was criticised for presenting development and evaluation as an unrealistically linear process, guidance issued in 2008 recognised the iterative nature of this process (Craig et al., 2008a; Craig et al., 2008b). However, this framework still represents a somewhat ordered process, with evaluation preceding full implementation and preceded by significant modelling, development and piloting. This is perhaps consistent with idealised models of evidence-based policy in which policymaking begins with identification of a problem, followed by goal setting, formulation of objectives and selection of the best strategy from a number of well-researched alternatives, followed by evaluation (Becker and Bryman, 2004), with long-term policy commitments not made until evidence is available. From this perspective, the duty of programme developers and evaluators is to ensure that enough evidence-based options are available, and that these are accessible to policymakers. In the USA for example, databases are held of evidence-based public health interventions, with funding often contingent on selecting evidence-based approaches (http://www.nrepp.samhsa.gov/find.asp).

Despite increased focus upon evidence-based policy in the UK however, it has often been seen as the exception rather than the rule (Lindblom, 1959; Leicester, 1999). Coote et al. (2004) describe evidence as one factor amongst many in UK policymaking, with desire to win voters, resources, pressure from lobby groups, media exposure, international power relations, personal ideologies, desire to innovate

and need for quick results seen as key influences. In addition, there is often insufficient evidence to meet the policy needs of the day, and as new problems are identified, new solutions may be needed. Whilst, in the USA development and piloting of policy ideas prior to wider implementation has become a common part of policymaking (Jowell, 2003), such a model has yet to become commonplace in the United Kingdom. Jowell (2003) suggests that a key reason for this is the UK's centralised government structure, with the state structure of the USA allowing policies to be piloted in a few states before gaining momentum in media or party political debates.

Developing and evaluating interventions may be a long and slow process, and the need to act may in many cases supersede the desire to wait for appropriate evidence (Coote et al. 2004). However, Macintyre (2003) argues that although lack of evidence should not be an excuse for inaction, and doing 'something' which is not evidence-based may be better than doing nothing, the impacts of these 'somethings' must be carefully assessed. Rather than obsessing about whether policy is evidence-based, a more pragmatic model for partnership between academia and policymaking where evidence is not available, may involve policymakers developing a policy, before commissioning an evaluation to be embedded into its rollout (Coote et al. 2004). Indeed, given that interventions of the scale or complexity likely needed to address complex social problems cannot be funded or implemented solely within academic structures, embedding evaluation into policy rollout may be the only realistic means of developing such evidence, providing valuable information to guide decisions on long-term continuation of funding or future policy development (Creegan and Hedges, 2007). Evaluation may take the form of natural experiments (Petticrew et al., 2005). Where viable, randomisation may be built into policy roll-outs to facilitate integrated policy trials (House of Commons Health Committee, 2009), such as in the case of the Welsh Assembly Government's Primary School Free Breakfast Initiative (PSFBI; Moore et al. 2007). Rather than evidence-based policy in its purest form, this model can perhaps be better described as policy-based evidence, with evidence sought to justify or refine policy implementation.

2.3.2 Political and methodological challenges in embedding evaluation into policy roll-outs

Though the need to embed evaluation into policy rollout is increasingly being recognised, this has yet to become common practice. A recent report by the House of Commons Health Select Committee (2009) stated that whilst 'it is crucial for policy and plans for evaluation to be designed thoughtfully and in conjunction with one another...this rarely happens'. Policies are described as often being introduced with no clear aims, and with structures changing before the intervention has chance to settle, weakening impacts and reducing evaluability. In the absence of good quality evaluation, patient satisfaction or descriptions of practice come to be accepted as evidence of good practice. Wanless (2007) argues that much of the multi-billion pound investment received by the NHS in recent years has ultimately been wasted, with expensive programmes such as new pay deals and policies on the management of long-term conditions introduced without evaluation built into their rollout.

Integration of evaluation into policy rollout engenders substantial political and methodological challenges. In some instances, the policy may have featured as a major political commitment and will likely be subjected to intense media scrutiny, with publicity making it difficult to conduct pilots without substantial contamination (Sanson-Fisher et al., 2007). Whilst policy representatives will likely have used media attention to persuade the public of the policy's worth, commissioning an evaluation is an admission of uncertainty, with this contradiction likely producing political tensions. For example, Rutter (2006) argues that the certainty with which Sure Start was portrayed as a solution to child poverty created perceptions that an RCT was unethical, leading commissioners to rule out the most robust option.

As argued by Gorard (2002), political investment in positive findings will likely lead funders to apply pressure for palatable conclusions. The greater likelihood of positive results in studies funded by interested parties is well documented with respect to pharmaceutical industry funded trials (Davidson, 1986; Friedberg et al., 1999). Ultimately, the policy evaluator is placed in a position comparable to that of a researcher funded to test a drug by a pharmaceutical company which stands to benefit from evidence of its effectiveness. Pressures may not be explicit, but there will likely

be a tendency towards less critical acceptance of findings which support the policy than of critical findings. A pilot conducted in advance of commitment to implementation, and which reveals that a policy is flawed or ineffective may save the political embarrassment of pursuing a worthless policy, and evaluations ought to be explicitly designed with the aim of drawing out weaknesses so that policy might be strengthened (Jowell, 2003). However, where committed to a policy prior to testing, finding that the policy is ineffective or delivery mechanisms are weak is dire political news.

Whilst critical findings will likely be more unpalatable where evaluation accompanies high profile policy rollouts, they may also be more likely to emerge than in many forms of research, due to compromises in research timetables. As described above, the MRC suggest that interventions should be carefully modelled, with piloting and feasibility testing preceding definitive evaluation (Craig et al., 2008a; Craig et al., 2008b). This approach likely leads to identification and removal of redundant components and integration of activities to address implementation challenges prior to definitive evaluation. In policy evaluation however, the limited time available prior to wider implementation may mean that development and evaluation are conducted by separate professional groups, whilst feasibility testing may be bypassed. As MRC guidelines state:

> Too strong a focus on the main evaluation, to the neglect of adequate development and piloting work, or proper consideration of the practical issues of implementation, will result in weaker interventions that are harder to evaluate and are less likely to be worth implementing (Craig et al. 2008b p4).

Challenges in understanding the implementation and functioning of the intervention therefore likely become more extensive than where formative evaluation has ironed out many initial kinks, perhaps requiring more in-depth exploration of the issues which should be explored by process evaluation in any definitive trial (Craig et al., 2008a; Craig et al., 2008b).

Programme theory will likely not be fully articulated. Challenges associated with maintaining consistency with programme theory as the policy moves rapidly from

one context to the next will likely emerge throughout the definitive evaluation. Furthermore, feasibility issues in the implementation of complex new structures will likely come to light only after implementation has begun. These issues must be understood within any definitive evaluation, with the transition from a small pilot to a full-scale definitive trial likely not without its challenges. However, in policy evaluation settings, exploration of these issues is perhaps more likely to identify extensive unanticipated issues which significantly impact the implementation and functioning of the intervention. Policy trials therefore offer a highly useful context in which to develop comprehensive frameworks for process evaluation, which might contribute to an understanding of how efficient large scale implementation might be achieved, whilst also aiding interpretation of trial outcomes.

2.3.3 Summary

In summary, whilst the need for evidence in policymaking is increasingly recognised, evidence represents one factor amongst many influencing policymaking, whilst evidence-based solutions to policy needs are not always available. Whilst the development and evaluation of policy pilots prior to wider implementation may be desirable, a useful though underused compromise may be to integrate evaluation into policy roll out. This however engenders significant challenges in negotiating political investments in positive outcomes, whilst also meaning that key stages of MRC guidelines may be bypassed. In policy trials in which development and evaluation stages are conducted by different parties, challenges in understanding of implementation and functioning of the intervention are perhaps heightened, with policy trials therefore offering a highly useful context in which to develop comprehensive process evaluation frameworks.

2.4 Chapter summary

In recent years, recognition has grown of the need to use evidence to inform policy in order to avoid proliferation of fiscally wasteful and potentially harmful policies. Increasing recognition that the aetiology of social problems targeted by public health policy is complex has engendered recognition that interventions to reverse or prevent their aetiology must also be complex. As interventions become increasingly complex, the task of gathering evidence for their effectiveness becomes complicated. The need

for development of high quality evidence has led to emergence of MRC guidelines for the development and evaluation of complex interventions. This guidance, as well as recommending a development and evaluation process which involves significant theory development and piloting ahead of definitive evaluation, acknowledges the need for comprehensive process evaluation within definitive evaluation stages, in order to understand the implementation and functioning of the intervention.

Whilst MRC guidelines focus predominantly on development and evaluation processes for researcher-led programmes, the embedment of evaluation into policy rollouts represents a vital means of learning from policy where evidence is insufficient. Whilst a need for process evaluation is common to all definitive evaluations, challenges in understanding implementation and the functioning of complex interventions are perhaps amplified in such settings, given that evaluators have limited influence over programme development, whilst limited piloting work may have been conducted. As recognition grows of the need to embed evaluation into policy rollouts, the need for high quality process evaluation becomes increasingly critical. As will now be explored throughout Chapter 3 however, whilst recognition of the need for process evaluation is increasing, limited methodological guidance exists for process evaluation, with significant ambiguity remaining to be resolved over how process evaluation might serve the functions described above.

3 Process evaluation in public health research

3.1 Chapter overview

A need for process evaluation has been described throughout Chapter 2 as arising from complexity in the design and implementation of complex interventions and their interactions with their settings and target audiences. However, although calls for process evaluation are increasing, these have typically not been accompanied by attempts to provide guidance for its conduct. MRC guidelines for example simply state that 'process evaluations should be conducted to the same high methodological standards and reported just as thoroughly as evaluation of outcomes' (Craig et al., 2008a). The aims of this chapter are therefore to explore key areas in need of development if process evaluation is to achieve the functions described by the MRC of understanding implementation, causal processes and contextual factors impacting intervention functioning. The chapter begins with a brief overview of the history of process evaluation and emergence of guiding frameworks within public health research, before reviewing current practice in the conduct of process evaluation. Section 3.3 will then discuss key priorities for a process evaluation framework suitable for complex interventions. Whilst the review is not limited to policy evaluation settings, aspects of process evaluation which may need greater emphasis in this context than in smaller scale or researcher-initiated studies are discussed throughout.

3.2 Process evaluation in public health research: history and recent practice

3.2.1 A brief history of process evaluation

Whilst the term process evaluation can be found in the public health literature as early as the 1960s (Steckler and Linnan, 2002), formal recognition of its value only gathered pace during the late 1980s after a key article described the need to avoid Type 3 error (Basch et al., 1985) (i.e. concluding that a programme is ineffective when it was implemented incorrectly). Throughout the following decade, the study of programme implementation was advanced in a series of preventive programmes (Finnegan et al., 1989; McGraw et al., 1989), focusing on understanding cost

effectiveness and outcomes, quality and quantity of delivery, the amount of intervention received, intermediate outcomes and causal pathways (Pirie et al., 1994).

Use of process evaluation was subsequently advanced largely by researchers involved in these early projects. Quantitative monitoring systems were used to examine key components within the Community Intervention Trial for Smoking Cessation (Corbett et al., 1990), whilst researchers in the Child and Adolescent Trial for Cardiovascular Health described measuring participation of schools, teachers and children, dose and fidelity and the compatibility of the intervention with the school context (McGraw et al., 1994; Perry et al., 1997).

As the complexity of public health interventions grew, so did the aims of process evaluation, triggering recognition of the need for frameworks to guide process evaluation development. These have mostly taken the form of lists of key targets for assessment by process evaluation. For example, Baranowski and Stables (2000) identified 11 aims: i) recruitment, ii) maintenance, iii) context, iv) resources, v) implementation, vi) reach, vii) barriers, viii) exposure, ix) initial use, x) continued use and xi) contamination. Around this same time, the RE-AIM framework (Glasgow, Vogt and Boles, 1999; Glasgow et al., 2001) emphasised a need to understand: i) reach ii) effectiveness/efficacy, iii) adoption (i.e. the percentage of eligible settings where the intervention is adopted), iv) implementation and v) maintenance. Shortly afterwards, a further framework was advanced by Steckler and Linnan (2002), which described process evaluation as complementary to outcomes evaluation, and as a means of establishing the internal validity of an outcomes evaluation as a test of programme theory. This framework recommended focus upon i) context (local factors which impact implementation), ii) fidelity (the extent to which the intervention is delivered as conceived), iii) dose delivered (the amount or number of units of intervention offered to participants), iv) dose received (the extent of participants active engagement in the scheme), v) reach and vi) recruitment practices.

Whilst leaving substantial scope for variability in the operationalisation of key constructs, where operationalised effectively, the now widely cited framework by Steckler and Linnan (2002) may offer a useful starting point for addressing the

priorities identified within Realistic Evaluation (Pawson and Tilley 1997). These include a need to understand what is implemented (fidelity and dose), the emergence of outcomes through interaction of participants with the intervention (dose received and reach) and contextual influences on implementation and outcomes (context). As described throughout Chapter 2, these priorities are largely reflected in the proposed functions of process evaluation within recent MRC guidance, which focus on the need to understand implementation, causal processes and contextual factors.

Whilst outlining key aims however, frameworks have stopped short of describing how to best conduct a process evaluation. Saunders and colleagues (2005) have since attempted to flesh out the Steckler and Linnan (2002) framework, describing a fictional example of process evaluation development. The authors describe a development process involving i) describing the programme and its underlying theory, ii) defining complete and acceptable delivery (i.e. core components), iii) development of research questions and methods, and consideration of resources, context and programme characteristics and iv) finalising a process evaluation plan. Advances include more explicit focus on describing and linking implementation checks to programme theory, a commitment to mixing methods, and an emphasis on use of data for both formative and summative functions within the definitive evaluation. However, limited consideration is offered of challenges linking methods to research questions, how methods are to be combined, or to exploring the linkage of process evaluation components to one another and to scheme outcomes.

In meeting the increasing calls for process evaluation within complex interventions, further development of guidance for the development and conduct of process evaluation is therefore of paramount importance. Before proposing a comprehensive framework for process evaluation, this section will review practice in process evaluation in the years since the emergence of the aforementioned frameworks. A Medline search using the terms 'process evaluation' and 'health promotion or public health or health behaviour or physical activity or diet or smoking or alcohol' (i.e. the 'holy four' health behaviours described by McQueen (1987) as impacting health outcomes) revealed 167 articles between 2003 and the time of writing. Abstracts indicated that 107 reported empirical data from a process evaluation. Full texts of 73

articles published in journals subscribed to by Cardiff University were examined for context, aims and methods. Given the focus of this thesis on process evaluation within an RCT, particular emphasis will be placed on aims and methods of process evaluations within RCTs.

3.2.2 Contexts and aims of process evaluation articles

Contexts and aims of process evaluation articles are presented in Table 1. Most formed part of a wider evaluation including assessment of effectiveness. In 24 (32.9%) cases, studies were embedded within an RCT, whilst in 30 (41.1%) these were combined with before-and-after or quasi-experimental evaluations. The remaining 19 (26.0%) were stand-alone process evaluations or feasibility studies.

Notably, whilst MRC guidance focuses on the need for process evaluation to assess implementation, most studies did not measure implementation. Quality of implementation (ie. fidelity) was assessed in only 28 (38.4%) studies, quantity (i.e. dose delivered) in 19 (23.9%), with both quality and quantity assessed only in 12 (16.4%). Many therefore bypassed consideration of what the intervention was in practice, moving straight to assessment of participants' experience of an undefined intervention. The most common aim overall was assessment of dose received or participant experience (n=41; 56.2%). Whilst 21 studies (28.8%) examined contextual issues or barriers and facilitators to implementation, in most cases little emphasis was placed on understanding how consistency or divergence in delivery emerged, or the interaction of the intervention with its contexts. Fewer studies assessed reach (n=18; 24.7%), whilst 4 (5.5%) assessed recruitment practices, 17 (23.3%) assessed acceptability to implementers, and 5 (8.2%) explored implementers' perceptions of causal processes.

In process evaluations conducted within RCTs, approximately two-thirds assessed either fidelity or dose delivered (14 out of 24 studies), with studies most likely to focus on dose received or participant experience (20 studies). Contextual issues or barriers and facilitators of implementation were explored in only 7 studies. Hence, recent practice does not facilitate the wider functions of process evaluation within RCTs recommended by MRC guidance (Craig et al. 2008a), of understanding

implementation, causal processes and the interaction of the intervention with its context.

Most studies assess a combination of aims, with virtually no two studies selecting the same combinations. Notably, some report data from a process evaluation rather than a complete study and data on other aspects were likely collected though not published. Griffin et al. (2010) for example report that qualitative data were collected but will be reported in future articles. Justifications for prioritisation of particular aspects of implementation for study or publication are rarely forthcoming. Linkage between aims (e.g. how implementation impacts patient experience) is rarely explicitly explored.

In addition, theoretical or empirical linkage to outcomes is often absent. Whilst typically offering a brief description of the planned intervention, few present a theoretical model for how programme activities were anticipated to produce desired outcomes. Indeed, one article describes negotiating a logic model to guide identification of key components, though chooses not to present this logic model (Griffin et al. 2010). Hence, implications of data for the functioning of the programme are often unclear. Direct linkage of implementation to outcomes, was conducted in a small number of studies (n=6; 8.2%) through correlating outcomes with delivery of key components. Others report use of process evaluation for post-hoc explanation of intervention success (Audrey, Holliday and Campbell, 2006a) or failure (Bleijlevens et al., 2008).

Table 1 Context and aims of identified process evaluation studies

	RCT	Other outcomes evaluation	Formative evaluation or stand alone	Fidelity	Dose delivered	Dose received	Reach	Context/ barriers and facilitators	Recruitment	Acceptability/ feasibility to implementers	Integration with outcomes	Mechanisms / theories of change
(Johnson et al., 2010)	X			X	X	X	X	X				
(Griffin et al., 2010)		X		X	X							
(Pearson et al., 2010)	X			X	X	X	X					
(Singh et al., 2009)	X			X		X	X	X		X		
(Liu et al., 2009)		X		X								
(Butler et al., 2009)			X			X						
(McConnon, Kirk and Ransley, 2009)	X					X						
(Neumark-Sztainer et al., 2009)	X					X						
(Williams et al., 2009)		X			X					X		
(Rabiei et al., 2009)		X		X	X			X		X		
(Bertens et al., 2009)		X		X		X			X			
(Heim, Stang and Ireland, 2009)		X				X						
(Petrunoff, 2009)			X	X		X		X				
(Kratz, Ponce and Yancey, 2008)			X					X		X		
(Young et al., 2008)	X			X	X		X	X				
(Wegner et al., 2008)			X	X				X				
(Jancey et al., 2008)		X				X	X					
(Escoffery, Glanz and Elliott, 2008)	X			X				X				
(Bleijlevens et al., 2008)	X			X		X				X		
(Sy and Glanz, 2008)			X		X			X				
(Shek, 2008)		X		X								
(Stein et al., 2008)	X					X				X		
(Gnich et al., 2008)		X					X		X			
(St. Louis et al., 2008)		X			X							

	RCT	Other outcomes evaluation	Formative evaluation or stand alone PE	Fidelity	Dose delivered	Dose received	Reach	Context/ barriers and facilitators	Recruitment	Acceptability/ feasibility to implementers	Integration with outcomes	Mechanisms / theories of change
(Phongsavan, 2008)												
(Macniven, 2008)			X			X						
(Khun and Manderson, 2007)			X									X
(Stafstrom and Larsson, 2007)		X										X
(Danielzik, 2007)		X								X		X
(Hunt et al., 2007)	X					X	X	X				
(Feathers et al., 2007)			X	X		X	X					
(Steele, Mummery and Dwyer, 2007)	X					X						
(Meier, Stock and Kramer, 2007)		X				X						
(Maiorana et al., 2007)		X		X						X		
(Inchley, Muldoon and Currie, 2007)		X						X				
(Nicolaides-Bouman et al., 2007))	X			X		X				X		
(Nishikido et al., 2007)			X							X		
(Shek, Lee and Sun, 2007)		X		X								
(Donnelly et al., 2007)	X					X	X				X	X
(Campbell et al., 2007)	X				X	X	X	X			X	X
(Rosecrans et al., 2007)		X		X	X	X	X	X		X		
(Jalleh, 2007)			X		X							
(Rankin et al., 2006)			X					X				
(Huston et al., 2006)	X				X	X				X		
(Papadaki and Scott, 2006)		X				X						
(Martens et al., 2006)	X			X		X				X	X	
(Kolt et al., 2006)	X					X						
(Kloek et al., 2006)		X		X	X	X	X	X				
(Jurg et al., 2006)	X				X	X						
(Bolam et al., 2006)			X			X						

	RCT	Other outcomes evaluation	Formative evaluation or stand alone PE	Fidelity	Dose delivered	Dose received	Reach	Context/ barriers and facilitators	Recruitment	Acceptability/ feasibility to implementers	Integration with outcomes	Mechanisms / theories of change
(Audrey et al., 2006a)	X											X
(Orton, 2006)			X			X		X	X	X		
(Sannibale et al., 2005)		X		X			X					
(Hunt et al., 2005)	X				X	X	X				X	
(Mavimbe, Braa and Bjune, 2005)			X									
(Pearlman et al., 2005)		X										
(Curran et al., 2005)		X		X	X			X				
(Salmon et al., 2005)	X			X	X	X						
(Merom, 2005)			X			X	X					
(Marino, 2005)		X				X						
(Ganley et al., 2004)		X				X						
(Lowe et al., 2004)			X			X						
(Ritchie et al., 2004)		X							X			X
(Lobb et al., 2004)	X			X	X	X	X					
(Ronda et al., 2004)		X		X		X		X				
(Power et al., 2004)			X					X				
(van Sluijs et al., 2004)			X					X		X		
(Gerald et al., 2003)			X									
(Steckler et al., 2003)	X			X	X	X	X					
(Simmons and Voyle, 2003)		X										
(Forsetlund et al., 2003)		X		X								
(Neumark-Sztainer et al., 2003)	X					X		X		X		

3.2.3 Selection and combination of methods

Process evaluation methods are summarised in Table 2. Fifty-four (74.0%) articles report quantitative methods, 44 (60.0%) report qualitative methods, whilst 27 (40.3%) report both. Studies evaluating fidelity, dose delivered or reach typically used quantitative assessment methods, including structured observations used by 16 (21.9%) or routine data sources, such as attendance lists and database records of intervention activities (n=28; 38.4%). Qualitative observation of implementation activities was used by 9 (12.3%) studies.

The most common tool was structured questionnaires, used in 37 (50.7%) studies, occasionally for assessment of fidelity or contextual factors, though most commonly to provide measures of exposure or satisfaction. Indeed, some process evaluations comprise nothing more than a satisfaction questionnaire (e.g. (Heim et al., 2009; McConnon et al., 2009; Neumark-Sztainer et al., 2009; Williams et al., 2009), occasionally accompanied by open questions (6 studies; 8.2%) which were either quantified or analysed qualitatively. Studies were approximately equally divided as to whether participant views of the intervention were explored simply via measures of exposure or satisfaction, or via qualitative interviews. Indeed, the second most common method was qualitative interviews, including semi-structured interviews and focus groups in order to explore perceptions of the programme and its causal mechanisms, used in 33 (45.2%) studies. Qualitative interviews were also commonly used for the purposes of exploring barriers and facilitators of implementation and acceptability to implementers. Within RCTs, dominant methods were structured questionnaires (18 out of 24 studies), programme monitoring data (14 studies), qualitative interviews (9 studies) and structured observation (7 studies).

Table 2 Methods used in identified process evaluations

Reference	Quantitative data	Qualitative data	Structured observation	Qualitative observation	Structured Questionnaire	Open ended questionnaire	Programme monitoring data	Other routinely available data source	Structured interviews	Qualitative interviews or focus groups	Correspondence or meeting notes	Implementer	Participant
(Johnson et al., 2010)	X				X		X					X	
(Griffin et al., 2010)	X						X					X	
(Pearson et al., 2010)	X				X								X
(Singh et al., 2009)	X				X		X					X	X
(Liu et al., 2009)	X						X					X	
(Butler et al. 2009)	X				X								X
(McConnon et al. 2009)	X				X								X
(Neumark-Sztainer et al. 2009)	X				X								X
(Williams et al. 2009)	X				X							X	
(Rabiei et al. 2009)	X	X	X		X					X		X	
(Bertens et al. 2009)		X					X			X		X	X
(Heim et al. 2009)	X	X			X	X							X
(Petrunoff 2009)	X		X		X							X	X
(Kratz et al. 2008)	X				X							X	
(Young et al. 2008)	X	X	X				X					X	
(Wegner et al. 2008)		X		X									X
(Jancey et al. 2008)	X	X					X			X			X
(Escoffery et al. 2008)	X	X	X				X		X	X		X	X
(Bleijlevens et al. 2008)	X	X			X	X			X		X	X	X
(Shek 2008)	X		X									X	
(Stein et al. 2008)		X		X						X		X	
(Gnich et al. 2008)		X					X			X		X	X
(St. Louis et al. 2008)	X												

	Quantitative data	Qualitative data	Structured observation	Qualitative observation	Structured questionnaire	Open ended questionnaire	Programme monitoring data	Other routinely available data source	Structured interviews	Qualitative interviews or focus groups	Correspondence or meeting notes	Implementer	Participant
(Phongsavan 2008)	X					X	X						X
(Macniven 2008)	X	X			X	X							X
(Sy and Glanz 2008)	X	X			X		X			X		X	X
(Khun and Manderson, 2007)		X		X						X		X	X
(Stafstrom and Larson 2007)		X								X	X	X	
(Feathers et al. 2007)	X	X	X			X				X		X	X
(Danielzik et al. 2007)	X				X							X	X
(Steele et al. 2007)	X	X			X					X			X
(Meier et al. 2007)	X				X					X			X
(Maiorana et al. 2007)		X	X	X	X							X	
(Inchley et al. 2007)		X		X						X		X	
(Nicolaides-Bouman et al. 2007)	X						X		X			X	X
(Nishikido et al. 2007)		X					X				X	X	
(Campbell et al. 2007)	X	X			X	X				X		X	X
(Shek et al. 2007)	X		X									X	
(Donnelly et al. 2007)	X						X		X				
(Rosecrans et al. 2007)	X	X	X		X		X			X		X	X
(Jalleh 2007)	X							X					
Hunt et al. (2006)	X	X		X	X		X					X	X
(Rankin et al. 2006)		X		X						X	X	X	
(Huston et al. 2006)	X	X					X					X	X
(Papadaki and Scott 2006)	X	X			X		X			X			
(Martens et al. 2006)	X	X	X		X		X			X		X	X
(Kolt et al. 2006)	X	X			X	X							
(Kloek et al. 2006)	X				X			X				X	X

	Quantitative data	Qualitative data	Structured observation	Qualitative observation	Structured questionnaire	Open ended questionnaire	Programme monitoring data	Other routinely available data source	Structured interviews	Qualitative interviews or focus groups	Correspondence or meeting notes	Implementer	Participant
(Jurg et al. 2006)	X	X	X		X				X			X	
(Bolam et al. 2006)		X								X			
(Bere et al. 2006)	X				X							X	X
(Audrey et al. 2006a)	X	X	X		X					X		X	X
(Orton 2006)	X	X			X		X			X		X	X
Sannibale et al. (2005)	X												X
(Hunt et al. 2005)	X						X					X	
(Mavimbe et al. 2005)	X	X	X				X			X		X	
(Pearlman et al. 2005)		X		X			X					X	
(Curran et al. 2005)	X		X		X			X				X	X
(Salmon et al. 2005)	X		X		X							X	X
(Merom 2005)	X				X		X					X	X
(Marino 2005)		X								X			X
(Ganley et al. 2004)	X				X								X
(Lowe et al. 2004)	X	X			X					X		X	X
(Ritchie et al. 2004)		X								X		X	
(Lobb et al. 2004)	X				X		X					X	X
(Ronda et al. 2004)	X	X			X		X		X	X		X	x
(Power et al. 2004)		X								X		X	X
(van Sluijs et al. 2004)		X								X		X	
(Gerald et al. 2003)		X								X		X	
(Steckler et al. 2003)	X	X	X		X		X			X		X	X
(Simmons and Voyle 2003)		X		X						X		X	
(Forsetlund et al. 2003)	X	X					X			X	X	X	X
(Neumark-Sztainer et al. 2003)	X	X			X			X		X		X	X

A key challenge for process evaluation given this diversity of its aims and methods is how these multiple methods are combined and presented. However, it is often near impossible to critique the conduct of many available studies due to the opaqueness in their reporting. Multiple data sources are often pooled in a manner which makes it unclear which methods resulted in which findings. This opaqueness in design and execution is perhaps most evident in the 27 studies which combine quantitative and qualitative methods (see Table 3), of which few explicitly state a justification for combining methods. Where justifications are offered, these typically include triangulation (Sy and Glanz, 2008), added depth (Steele et al., 2007) or addressing separate questions (Rosecrans et al., 2007).

It is often however unclear how these aims were achieved, with some simply presenting quantitative data followed by a sentence or two stating that qualitative data concurred with reported findings (Orton, 2006). In most cases (69.6%) a heavier emphasis was given to quantitative data. Some incorporated open-ended questions into structured questionnaires, simply summarised with little analysis (Kolt et al., 2006) or quantified (Huston et al., 2006). In many studies, only a brief summary of qualitative data is given after descriptive quantitative data, with no quotations used to illustrate informants responses (Sy and Glanz, 2008). In some cases, qualitative data collection was described in methods, yet was not mentioned in results (Jurg et al., 2006). It is therefore often difficult to ascertain what has been gained through mixing methods. Where an approximately equal emphasis is given to both approaches, methods in some instances answer different questions, such as a study which reports quantitative data on fidelity, dose and reach, before presenting qualitative data on the acceptability of the intervention (Rosecrans et al., 2007). In other cases, qualitative and quantitative analyses are more closely integrated, with analysis subdivided into themes and quantitative and qualitative data relating to each of these themes presented alongside one another (Audrey et al., 2006a).

Table 3 Justifications for combining quantitative and qualitative methods and approaches to integration of methods in process evaluations

Reference	Main emphasis	Reason for combining	Integration of qual and quant data
(Rabiei et al. 2009)	Approximately equal	None given	Analytical approach not described
(Hunt et al. 2006)	Approximately equal	Address separate questions	Quantitative descriptive findings presented, followed by qualitative findings.
(Feathers et al. 2007)	Approximately equal	None given	Quantitative descriptive findings presented, followed by qualitative findings.
(Campbell et al. 2007)	Approximately equal	None given	Quantitative descriptive findings presented, followed by qualitative findings.
(Heim et al. 2009)	Quantitative	None given	Qualitative data analysis simply a summary of responses to open ended quotations
(Young et al. 2008)	Quantitative	None given	No qualitative data presented. Linkage between data sources and findings unclear
(Jancey et al. 2008)	Quantitative	None given	Analytical approach not described
(Escoffery et al. 2008)	Quantitative	None given	Very brief (two paragraph) summary of qualitative data on practitioner opinions, with no quotations used
(Bleijlevens et al. 2008)	Quantitative	None given	Very brief (two paragraph) summary of qualitative data on practitioner opinions, with no quotations used
(Sy and Glanz 2008)	Quantitative	Triangulation	Very brief (one paragraph) summary of qualitative data on practitioner opinions, with no quotations used. Agreement or discrepancy between data sources not clearly examined.
(Steele et al. 2007)	Approximately equal	Triangulation, added depth	Qualitative data used to confirm quant findings (simply stated that this had been the case) and to answer additional separate research questions
(Huston et al. 2006)	Quantitative	None given	Qualitative data quantified and summarised quantitatively
(Papadaki and Scott 2006)	Approximately equal	Triangulation	Quantitative descriptive findings presented, followed by qualitative findings.
(Martens et al. 2006)	Approximately equal	None given	Results rotate back and forth between quantitative and qualitative findings, presenting both under same thematic subheadings.
(Kolt et al. 2006)	Quantitative	None given	Minimal qualitative data (open ended questions included within structured questionnaire) quantified and summarised.
(Jurg et al. 2006)	Quantitative	Address separate questions	Qualitative data collection described in methods, but no analysis presented
(Audrey et al. 2006a)*	Approximately equal	None given	Results rotate back and forth between quantitative and qualitative findings, presenting both under same thematic subheadings.
(Orton 2006)	Quantitative	None given	Agreement of focus groups with quantitative data described at end of some paragraphs
(Mavimbe et al. 2005)	Quantitative	None given	Some disjuncture between descriptions of themes in methods and results. One paragraph summary of quant fiindings
(Lowe et al. 2004)	Quantitative	None given	Very brief (one paragraph) summary of qualitative data incorporated within one theme also containing quantitative data.
(Steckler et al. 2003)	Quantitative	None given	
(Forsetlund et al. 2003)	Qualitative	None given	One table of quantitative description. Main focus on qualitative data
(Ronda et al. 2004)	Approximately equal	Triangulation	Results rotate back and forth between quantitative and qualitative findings, presenting both under same thematic subheadings.
(Neumark-Sztainer et al. 2003)	Quantitative	None given	Most qualitative data simply quantified. One paragraph of thematic description
(Rosecrans et al. 2007)	Approximately equal	Address separate questions	Quantitative data on fidelity dose and reach followed by qualitative analysis of acceptability
(Macniven 2008)	Quantitative	None given	Very brief (one paragraph) summary of qualitative data on practitioner opinions, with no quotations used. Agreement or discrepancy between data sources not clearly examined.

In essence therefore, process evaluation appears to be a field of research where anything goes. Anything from a mixed-method study combining systematic quantitative monitoring of implementation with qualitative exploration of the experiences of implementers and participants (Escoffery et al., 2008; Rabiei et al., 2009), to administration of a few items on programme satisfaction (Kolt et al., 2006; Heim et al., 2009; McConnon et al., 2009; Neumark-Sztainer et al., 2009; Williams et al., 2009) can call itself a process evaluation.

3.2.4 Summary and conclusions

In summary, process evaluation is driven by diverse and often poorly defined aims. Limited attention to programme theory commonly hampers an understanding of the implications of implementation variability for programme functioning, whilst assessment of implementation is rarely accompanied by exploration of how the intervention came to be more or less successfully implemented. In fact, most process evaluations do not assess implementation, with many simply comprising satisfaction questionnaires or some exploration of patient responses to the intervention. Limited attention has been paid to linkage between evaluation aims and how these contribute to achieving broader functions of informing improved implementation and understanding how outcomes were produced. Hence, recent practice in process evaluation does not typically enable it to serve the functions described within MRC guidelines of understanding implementation, causal processes and the interaction of the intervention with its context, and within evaluation of effectiveness, does little to illuminate how outcomes are produced.

Methodologically, process evaluation appears to be a field where almost anything goes, and it is near impossible to critique the conduct of many available studies. Such studies perhaps include too many aims in the same article, and hence address none of these particularly well. Given the diversity of aims, multiple articles with clear linkage which demonstrates their contribution to the wider evaluation may be a more appropriate means of presenting process evaluation findings than forcing everything into a single article. Integration of multiple methods is often poorly structured, making it difficult to ascertain which data comes from which source, and how authors have reached their conclusions (Williams et al., 2009). As described by Oakley at al. (2006), the apparent bias towards quantitative methods is somewhat surprising given

that the term 'process' is often presumed within the social sciences to be synonymous with qualitative methods (Denzin and Lincoln 2000).

Whilst the quantity of process evaluations is growing, there is therefore a clear need for development of frameworks which go beyond describing potential aims, and move towards providing systematic guidance on the selection and combination of methods to achieve the diverse aims of process evaluation, whilst also paying attention to systematic linkage between aims and methods and the role of each component in contributing to wider goals of informing implementation and aiding interpretation of outcomes.

3.3 A new framework for process evaluation

As described in Section 3.1, the widely cited Steckler and Linnan (2002) framework offers a potential starting point to guide the development of process evaluations which aim to serve the functions identified within MRC guidance of understanding implementation, causal processes and the relationship of the intervention with its context. As also discussed, this framework has been extended by Saunders and colleagues (2005) to include more explicit initial emphasis on programme theory in order to guide implementation checks. However, substantial room remains for variation in the operationalisation of key process evaluation components, for further incorporation of methodological guidance and for development of systematic linkage between programme aims. The aims of this section are therefore to propose a framework which considers theoretical and methodological issues in operationalising process evaluation components, provides methodological guidance on how to address process evaluation aims and offers a structure to link individual components to one another and to outcomes, viewing each component of process evaluation as contributing incrementally to wider aims of informing improved implementation and interpretation of outcomes.

In order to provide this linkage, aims will be discussed in a temporal manner, beginning with describing the planned intervention, moving towards understanding how it is delivered and finally towards how the delivered intervention works. Targets for assessment within process evaluation will therefore not be treated in an atomistic manner, but the implications of one phase for the next discussed. Discussion will

begin by focusing on the need to describe the programme in terms of its key components and underlying theory as described by Saunders and colleagues (2005), followed by the need to understand the processes through which the planned intervention takes shape across contexts. The need to measure the consistency of implementation with programme theory will then be discussed, before the need to explore how the delivered intervention is experienced, and how and for whom it produces change. The manner in which these various aspects of process evaluation might be linked to provide a comprehensive understanding of implementation and programme outcomes is discussed throughout and brought together explicitly at the end of this section.

3.3.1 The planned intervention: programme theory

Attempting to understand a complex intervention is something of a daunting prospect without a clear understanding of what the programme is and how it is expected to work (Saunders et al., 2005). Hence, as stated within MRC guidance for evaluating complex interventions, as a starting point for evaluation, 'a good theoretical understanding is needed of how the intervention causes change' (Craig et al., 2008a; Craig et al., 2008b). As will be illustrated in Chapter 4 however in relation to exercise referral schemes, evaluators have often simply examined the outcomes of a black-box intervention package, ignoring the actions of individual components and their synergy, leading to conclusions that nothing works and more research is needed, or that everything works and the intervention must be reproduced in full (Pawson and Tilley 1997). Hence, all we can conclude is that an often poorly described package did or did not produce intended aggregate outcomes in the settings and groups to which it was delivered (Michie et al., 2009).

Programmes can only ever be as good as the theories built into them (Tilley, 2004) and it is commonly recommended that public health interventions should be based on formal behaviour change theory (Riddoch, Puig-Ribera and Cooper, 1997; Department of Health, 2001). However, many formal theories are rather weak predictors of behaviour (Bartholomew, Parcel and Kok, 1998; Shepperd et al., 2009) and interventions based on formal theory are not always more effective than those which are not (Eakin, Glasgow and Riley, 2000; Adams and White, 20032005). However, whether based upon formal evidence, experience or 'common sense', interventions are

always 'theories incarnate' (Pawson and Tilley, 1997), in that they represent a manifestation of theoretical beliefs that action will lead to desired outcomes. Policy development draws on a range of latent theory drawn from a range of sources and evaluating a policy involves testing an external party's theories of change; theories which will likely have not been made explicit.

Whilst Steckler and Linnan (2002), amongst others (McGraw et al., 1989; Carol, 1997; Pawson and Tilley, 1997; Weiss, 1997; Bartholomew et al., 1998; Saunders et al., 2005), highlight the value of referring to programme theory in the early stages of designing process evaluation, their framework largely assumes that a clear programme theory is already be in place, and therefore offers no guidance for how theory is to be developed or elicited. By contrast, as described in Chapter 2, early MRC guidance focused on process evaluation for purely formative purposes, largely ignoring implementation (Campbell et al., 2000). However, as described in Chapter 2, process evaluation must do both. In a researcher-controlled development and evaluation process, articulation and development of programme theory will take place prior to definitive evaluation, and hence this stage will involve describing theory which has been developed.

However, given that policy trials will often involve evaluating schemes where limited formative process evaluation has been conducted, these activities are forced to take place alongside the definitive trial. Process evaluation within policy trials must therefore start by eliciting a clear understanding of programme theory. Articulating programme theory will likely involve discussions with policy representatives, with programme theory usefully summarised in the form of a logic model illustrating key intended programme components, hypothesised intermediate causal processes and longer term outcomes (Kellogg Foundation, 2004). Reviews of relevant literature will likely be useful in exploring the plausibility of programme theory.

Though this late elicitation of programme theory may reveal weak links or theoretically conflicting activities which it is too late to modify, it will provide valuable guidance in understanding challenges in implementation and in interpreting outcomes, with transparency allowing internal and external critique of the plausibility of intervention theory. Furthermore, measurement of implementation quality easily

becomes a subjective notion (Dusenbury et al., 2003) unless there is a clear understanding of the core spirit of the planned intervention. As will be described below, variation in delivery is to be expected as an intervention moves from one context to the next, and may serve a useful role in ensuring sustainability (Goodman and Steckler, 1989), so long as adaptations remain consistent with programme theory (Hawe, Shiell and Riley, 2004a). Hence, eliciting programme theory serves a key role in defining complete and acceptable delivery (Saunders et al., 2005) in terms of the irreducible core of the intervention (Greenhalgh et al., 2004), with variability considered low fidelity only where involving departure from programme theory.

3.3.2 Diffusion into local practice

As the first component of Steckler and Linnan's framework for process evaluation, the authors recommend considering contextual influences on implementation. However, as described in Section 3.2, most process evaluations have bypassed consideration of challenges in ensuring implementation, beginning by exploring whether or not the intervention was delivered as intended. In small-scale evaluations, the intervention is perhaps delivered in a controlled manner by small numbers of motivated implementers, or even by researchers themselves. Hence, challenges associated with wider implementation may not always need in-depth exploration until after evaluation. Indeed MRC frameworks recommend that whilst consideration be given to the feasibility of large-scale implementation throughout the development and evaluation process, efforts to promote and facilitate this ought to begin only after the programme has been shown to be effective (Craig et al., 2008a; Craig et al., 2008b). Hence, initial focus on delivery and receipt of the intervention, supplemented simply with descriptive contextual information to avoid over-generalisation to other settings, is perhaps justified.

However, in evaluations of multi-site interventions such as policy trials, this becomes more problematic. Large scale trials rely on the intervention being effectively implemented across multiple sites in order to evaluate a version of the intervention which resembles that which would be happening if the trial was not taking place, with delays representing a key challenge to policy trials. In the evaluation of the Primary School Free Breakfast Initiative (Moore et al. 2007), whilst short-term follow up was planned for 4 months after baseline, most schools had not implemented the scheme by

this time. Even at 12 month follow-up, some were yet to implement. Rapid and uncontrolled adoption can however also be deeply problematic. As will be described in Chapter 4, rapid adoption of motivational interviewing has led to often extremely poor implementation (Miller and Rollnick, 2009). Hence, delays in implementation will likely mean that policy trials evaluate an intervention which is not fully up and running, whilst excessively rapid uptake may enhance the risk of weak implementation. As the range of settings into which the policy will be delivered and the number of stakeholders in whom behaviour change is required increases, so do challenges ensuring effective implementation (Craig et al. 2008a).

Contextual influences on implementation clearly need to be understood. However, Steckler and Linnan's framework is somewhat limited by its focus on context as the sole dimension of complexity anticipated to impact the implementation of complex interventions. Implementation is a human activity, and whilst contextual factors may mediate whether actions produce desired results, contextual influences cannot be understood in isolation from human actions (Spillane, Reiser and Reimer, 2002). The roles of agency and context in shaping implementation are dynamic and reciprocal, and local implementers do not passively accept contextual challenges and give up, but respond through modifying implementation activities or changing the local context (Hawe et al., 2004a). A more appropriate starting point for understanding how implementation is achieved in multi-site policy trials is therefore to explore implementation activities, local factors impacting implementation, and implementers' responses to contextual challenges.

A small number of process evaluation studies have attempted to describe the process of successful implementation. Inchley and colleagues (2007) for example examined the process of becoming a healthy school, highlighting how a sense of shared ownership and empowerment was achieved through needs assessments, allowing schools flexibility in interpreting the aims of the programme and its activities and involvement of community members in decision making. Collaboration was achieved through creation of multi-agency steering groups, whilst compatibility with current practices was seen as a key determinant of whether the scheme would achieve routinisation.

Whilst such exemplars within studies identifying as process evaluations are relatively rare, the diffusion literature is clearly of relevance (Rogers, 2003; Greenhalgh et al., 2004). This literature is, for want of a better term, highly diffuse, and notoriously difficult to review (Greenhalgh et al., 2005). This section does not provide a comprehensive review, but introduces key concepts from the diffusion literature and their potential value in examining the process of ensuring the adoption and implementation of new interventions across multiple trial sites. The most widely known framework within this literature, Diffusion of Innovations theory (DoI) defines diffusion as a process through which an innovation is communicated, through certain channels, over time, among members of a social system (Rogers, 2003). In health promotion research, diffusion studies have predominantly focused on social marketing based innovations, linking communication channels to reach and uptake in target populations (Greenhalgh, Robert and Bate, 2003). However, exploring diffusion processes offers substantial potential as a means of understanding how interventions take shape at the local level.

According to DoI, diffusion into organisations begins with efforts to initiate adoption by relevant stakeholders, followed by implementation. Where successful, implementation ultimately ends with the innovation becoming routinised, losing its identity as a separate bounded entity, and becoming an integrated part of the organisation (Rogers, 1995). The final three components of the aforementioned RE-AIM framework (Adoption, Implementation and Maintenance), can therefore be conceived as representing phases of the diffusion process. A planned intervention may fail to be adopted widely enough to produce impacts, may be adopted though not implemented in a manner which reflects the core intent of the planned intervention, or may fail to become institutionalised, becoming dispensable after initial funding cycles (McLeroy et al., 1988; Goodman and Steckler, 1989; Rogers, 1995).

Much diffusion research has focused on characteristics of successfully diffused innovations, including relative advantage, compatibility, low complexity, trialability and observability (Rogers, 1995; Greenhalgh et al., 2003). However, whilst often portrayed as stable innovation 'properties', these are 'neither stable features of the innovation nor sure determinants of their adoption. Rather, it is the interaction amongst the innovation, the intended adopters and a particular context that determines

adoption' (Greenhalgh et al., 2004) (p598). National standardisation may for example involve introducing a new innovation which is similar to but distinct from current local practice. Judgments of relative advantage, compatibility and complexity will therefore likely be made not solely through rational comparison with existing schemes, but will be linked to investments in previous practice and emotional responses to instruction to change (Brehm, 1966). Qualitative research focusing upon perceptions of the innovation in context (Fitzgerald et al., 2002; Spillane et al., 2002) may therefore offer a deeper understanding of the roles of these perceptions in shaping implementation.

Another key focus of diffusion literature has been the effectiveness of communication structures. Policy innovations commonly involve multiple levels of diffusion. In diffusing a new exercise referral scheme into practice for example, a first stage may involve persuading local coordinators to change their practice. Subsequently, coordinators will likely be tasked with securing adoption by leisure centres' and health professionals; key gatekeepers in determining whether the scheme will ultimately diffuse to patients (Bartholomew et al., 2006). Hence, the policy reaching its target audience is contingent on multiple levels of effective communication.

Communicating new innovations, particularly in centralised systems, commonly relies upon change agents who occupy a midway position between adopters and the agency aiming to produce change. Change agents are tasked with establishing a perceived need for change in intended adopters, building rapport, opening up information exchange, triggering intention to change and supporting translation into action (Rogers, 1995). Whilst communications are typically strongest between individuals or groups with a high degree of homophily (i.e. similarity) (Fitzgerald et al., 2002), new ideas are typically introduced from outside adopters' social networks, involving communications with heterophilous others (Rogers, 1995). Change agents will therefore typically be most successful in communicating new innovations where sufficiently homophilous with intended implementers and where oriented towards needs of intended implementers more closely than to those of the change agency (Rogers, 1995). This is perhaps a challenge in policy innovations, which are often introduced by civil servants who are both of dissimilar status to intended adopters and intrinsically linked to the change agency.

The implementation phase of the diffusion process is characterised by uncertainty and tendencies to seek information and technical support from change agents (Rogers, 2003). Whilst public health innovations are increasingly characterised by complexity (McLeroy et al., 1988), perceived complexity is typically inversely associated with adoption and implementation success. Judgments of complexity are intimately linked to adopters' understandings of the innovation and perceptions of whether they have the skills and resources to implement it, and hence will at least in part be impacted by communications. An implementer may perceive an innovation to be simple having misunderstood it, leading to inappropriate implementation (Miller and Mount, 2001; Miller and Rollnick, 2009). However, the implementer may understand the innovation, but lack the skills or resources to use it (Spillane et al., 2002). Perceptions of complexity are therefore likely linked to the adequacy of support, training and resources accompanying the policy signal (Spillane et al., 2002; Greenhalgh et al., 2004), with innovations more likely to be successfully implemented where augmented with sufficient training and support (Greenhalgh et al., 2004). Underestimation of training and support has been cited as a common pitfall in complex schemes, which typically require acquisition of new skills or new applications of skills (Bartholomew et al., 2000; Owen et al., 2006).

As described, from the perspective of change agencies, the most successful diffusion outcome is routinisation. However, whilst authority innovation-decisions (i.e. decisions made for the entire social system by a few individuals in positions of power) typically achieve adoption rapidly, they may take longer to achieve routinisation (Rogers, 2003; Greenhalgh et al., 2004). As discussed, implementation is influenced by understandings of the policy and agreement with its core intent; likely greatest where the innovation has emerged through consensus amongst members of a social system. In addition, whilst local ownership and empowerment have been linked to successful implementation (Bartholomew et al., 2006; Inchley et al., 2007), authority-innovation decisions may reduce ownership and control. Furthermore, sustainable implementation involves mutual adaptation or reinvention, as the organisation adapts to the innovation, whilst the innovation adapts to the organisation (Ringwalt et al., 2004). However, within top-down implementation models, focus on clear protocols and elimination of ambiguity, including safeguards against reinvention to ensure

quality (Rogers, 1995), commonly increases conflict through limiting autonomy and placing onus on the organisation to adapt to the innovation (Matland, 1995). Hence, in efforts to implement interventions across contexts, tensions between the need for standardisation and for local reinvention need to be understood.

Given the complexity and openness of systems into which policies aim to become a part (Hawe et al., 2009), understanding how diffusion is achieved across contexts necessitates a flexible qualitative approach (Hawe et al., 2004b). These data will likely offer substantial insights into the roles of issues such as implementers' perceptions of the innovation, communication structures, perceptions of training and support and adaptive local tailoring in shaping the form of the programme which is ultimately implemented in practice. The timing of collection of this qualitative data perhaps needs to be considered carefully, with issues such as the perceptions of the innovation likely changing over time, though where multiple data collection points are not possible, retrospective interviews some time after initiation may provide opportunity for the greatest degree of reflection on the diffusion process. Quantitative implementation checks are then needed to assess the outcomes of the diffusion process, in terms of the extent to which congruence with programme theory is achieved.

3.3.3 Implementation: fidelity and dose delivered

Protocols will likely stipulate what activities will be delivered, and the amount of each component to be delivered, in terms of duration, volume or intensity. Hence, Steckler and Linnan (2002) break implementation down into fidelity and dose delivered, with fidelity defined as the integrity which the programme is delivered, whilst dose delivered is defined as the amount of intervention delivered (Steckler and Linnan, 2002). Essentially, these measures represent assessments of the internal validity of the outcomes evaluation as a test of programme theory (Steckler et al., 2003). Where implementation is not assessed, findings are based on the naïve assumption that intervention as conceived and as delivered are one and the same (Basch et al., 1985), potentially leading to dismissal of sound theories due to poor implementation, or inspiring false confidence in weak ideas which were strengthened at the local level during implementation.

As described, variation in delivery is a normal and adaptive part of the diffusion process (Goodman and Steckler, 1989), and hence should only be considered poor fidelity where representing divergence from programme theory. However, the theory driving intervention design will often not be the theory that is eventually tested (Tilley, 2004). An understanding of the process of diffusion will have likely begun to clarify which components are being delivered in a manner consistent with programme theory and in reality, implementation will rarely be viewable dichotomously as a success or failure (Carroll et al., 2007). Likely outcomes are that some activities will be delivered, others may not, some may be delivered in some areas though not others, or with variable quality or dose. Additional components will likely have been introduced locally to address perceived shortcomings. Where implementation checks reveal that the intervention substantially departs from programme theory, implications for the proposed causal chain should be considered. A modified logic model may in some instances become necessary prior to examination of participant experiences or intervention outcomes (Oakley et al., 2006).

As will be described in Chapter 5, checks ought to be in place at the earliest opportunity, and where possible should cover the whole trial period or multiple time points in order to understand how implementation changes over time. The integration of routine monitoring structures into the intervention offers a means of facilitating long term evaluation of implementation and negating concerns regarding Hawthorne effects, where monitoring artificially changes practitioners behaviour. However, process evaluators may need to identify key components not covered by monitoring structures, and integrate additional monitoring activities, using a range of observation and self report, the respective merits of which are discussed in Chapter 5.

Given the high risk of divergent implementation in policy trials, consideration needs to be given to how the evaluation will respond to emerging shortcomings. Where key components are undelivered, theoretical contributions to outcomes remain untestable and attention ought to turn to understanding how delivery might be improved if the policy is to continue. Indeed, emerging shortcomings may lead to efforts to improve delivery during the trial, with impacts of these efforts for delivery and outcomes needing consideration. In a four year trial of the 'Active by Choice Today' intervention, observations of fidelity and dose delivered led to extended training and

clarification of protocol documents (Wilson et al., 2009). The authors argue that this did not represent changing the intervention mid-trial, but rather ensured that implementers were able to deliver it better, with changes in programme reach monitored to examine impacts of these improvements. Whilst taking place alongside the trial, process evaluation within researcher initiated trials, and perhaps to a greater extent in policy trials may serve formative functions, keeping implementation on track, rather than purely summative functions (Saunders et al. 2005).

3.3.4 Responses to the intervention: participant experiences

As described in Section 3.2, exploring participant responses to an intervention, through measures of exposure, satisfaction questionnaires or qualitative interviews, has become the most common focus of process evaluation. However, this has all too often taken place in the absence of an understanding of implementation. Responding to participant views on how well the intervention meets their needs and modifying the intervention will clearly be contingent upon knowing what was delivered. Hence, in addition to establishing the internal validity of outcomes assessment as a test of programme theory, implementation checks serve a valuable function in defining the programme, prior to exploring patients' experiences in order to gain insights into how the delivered intervention worked.

Steckler and Linnan (2002) describe a need to assess how the intervention is experienced by its target audience in terms of 'dose received' (Steckler and Linnan, 2002). Whilst sometimes operationalised as 'exposure' (Baranowski and Stables, 2000), Steckler and Linnan (2002) define dose received as 'active engagement', rendering the passive tone of the term 'received' something of a misnomer. Active engagement however is a difficult construct to operationalise as interventions become more complex. For example, whilst Steckler and Linnan (2002) argue for a focus upon how much of a set of educational materials a patient can recall, it is difficult to see how an equivalent measure might be developed for complex interventions. Satisfaction questionnaires have therefore become a common proxy for engagement.

Where Steckler and Linnan's operationalisation of dose received is perhaps most problematic is that it is explicitly defined as a 'characteristic of the target audience', inviting an interpretation that if an intervention is delivered as intended but does not

produce a positive impact this is due to patient non-compliance. Hence, whilst Baranowski and Stables' (2000) relatively passive definition risks viewing the intervention as acting upon passive recipients (Pawson and Tilley, 1997), Steckler and Linnan's (2002) more active definition risks over-privileging agency. Inequalities arising from patterning in responses to an intervention can for example be blamed on poorer patients who failed to engage in the opportunities made available to them. Patient experiences are neither a characteristic of the patient, nor of the intervention, but are a product of the interaction between patient and intervention in context (Pawson and Tilley, 1997; Clark et al., 2007).

One of the most substantial overall weaknesses of many process evaluation frameworks including that of Steckler and Linnan (2002) is that whilst acknowledging the value of mixing methods, terminology and key constructs overemphasise quantification. The process evaluation literature, as described in Section 3.2, is approximately equally divided as to whether 'dose received', or participant experiences, have been explored through satisfaction questionnaires or qualitative interviews. Satisfaction questionnaires can provide an insight into affective responses to key components, but can do little to illuminate the complex processes linking activities to outcomes. Exploring these issues calls for largely qualitative exploration of the interaction of patients with the delivered intervention (Curry, Nembhard and Bradley, 2009). The term 'dose received', which is excessively passive and biased towards quantity rather than quality of patients' interaction with a scheme may be better replaced with a term such as 'participant experiences'.

3.3.5 Measuring reach and social patterning

Reach is defined within the RE-AIM framework, and within Steckler and Linnan's framework as the proportion of the target audience who participate in the intervention. Abrams et al. (1996) argue that public health impact is the product of reach and efficacy, with significant impact dependent on achieving positive outcomes and reaching a large proportion of the target audience, whilst equitable interventions aim to reach all patient sub-groups to a similar extent (Glasgow et al., 1999; Glasgow et al., 2001). Reach is typically assessed via attendance records (Steckler and Linnan, 2002).

In community-based interventions, it can be difficult to determine an appropriate denominator to calculate the percentage of the target audience participating in the intervention. An intervention delivered via a closed social unit such as a school may identify the percentage of schools taking up the intervention, and the percentage of teachers or pupils participating in the intervention (Young et al., 2008). In very small scale physical activity promotion schemes using only one or two practices, some have identified the proportions of patients within that practice who would have been eligible, calculating response rates within this denominator (Stevens et al., 1998). However, in larger schemes using a wider range of entry routes, obtaining accurate estimates of the size of the target audience is challenging.

Quantitative evaluation of reach where the denominator is more difficult to define may perhaps focus most productively on understanding which sub-groups are most likely to come into contact with and participate in the intervention. Identifying patterning in attendance and responses by patient characteristics is often seen as a crucial initial stage in realist approaches to evaluation (Kazi, 2003; Clark et al., 2007). As with participant experience however, reach is not a characteristic of the target audience but an outcome of the interaction between the intervention and the target audience. Combining quantitative profiling of programme reach with qualitative data on patient experiences may offer a useful means of understanding how social patterning emerges in responses to the programme.

3.3.6 Linking study components: a hierarchical framework for process evaluation

As described in Section 3.2, there is often little clear justification for which aspects of implementation or patient response to an intervention have been selected for evaluation, with linkage between components and how these combine to address wider aims of improving implementation and interpreting outcomes often unclear. Hence, drawing on the themes described above, a hierarchical framework for process evaluation is presented below in Table 4, in which each stage explicitly builds upon findings of the previous stage and feeds into the next, building incrementally towards a comprehensive understanding of implementation and intervention functioning. This begins by clearly defining what the programme is intended to be and how it is intended to work, drawing upon relevant literature and communications with

programme developers. Subsequently, attention turns to diffusion, focusing upon activities to diffuse the intervention into local practice and the dynamic roles of context and human agency in the process of implementing the scheme across contexts, likely relying heavily upon qualitative methods. Whilst qualitative data will offer insights into the process of implementation, quantification of the outcomes of these processes in terms of the consistency of what is implemented with programme theory, allows clear definition of the programme with which patients' interact, validating the trial as a test of programme theory.

Clear definition of the programme as delivered then facilitates exploration of how patients interact with programme activities and how these interactions produce change. The need to understand causal processes linking delivery to outcomes at this stage likely lends itself best to qualitative methods focusing upon the interaction between participants and the delivered intervention. Qualitative methods will likely provide significant insights into emerging social patterning in responses to the scheme, with subsequent attention to measuring patterning in response to the programme by patient characteristics interpretable in light of these understandings. Hence, the framework taken as a whole is anticipated to be greater than the sum of its parts, with focus on one element such as patient experience in isolation from an understanding of implementation significantly weakening the insights gained.

Table 4 A process evaluation framework for complex interventions

Component	Methodological and analytical focus
Programme theory	Clear elicitation of programme theory, in terms of key planned programme components, causal pathways and intended outcomes
Diffusion	Qualitative exploration of how diffusion activities, contextual factors and actions of implementers shape delivery across local contexts
Implementation	Quantitative measurement of the consistency of delivery with programme theory and quantity of intervention delivered
Participant experience	Qualitative exploration of how the intervention is experienced by patients, and the causal processes through which change is promoted in context
Reach	Quantitative measurement of patterning in programme reach by patient characteristics and variability in programme delivery

As recommended by Oakley and colleagues (2006), these analyses should be completed prior to analysis of outcomes, in order to provide guidance for the interpretation of outcome effects. A subsequent stage likely after outcomes analysis, and which will not be included in this thesis for time and word count reasons, is integration. At this stage, implementation data may be integrated into outcomes

analyses in order to examine patterning in scheme outcomes according to implementation variation. Directions for integrated analysis will likely emerge throughout analysis of process data, as insights emerge into perceived active ingredients of the programme, their causal processes and the settings in which they facilitate change.

3.4 Chapter summary

In summary, process evaluation is a poorly demarcated and inconsistent field of research. Recent practice in process evaluation is insufficient to serve the functions described in MRC guidance, of understanding implementation, causal processes and context, with most such studies not assessing implementation or context, whilst exploration of causal processes is often understood simply in terms of exposure or satisfaction. Methods are typically mixed in an opaque manner, and linkage between study components or between process evaluation and outcomes unclear. A framework has been suggested in this chapter which focuses upon theory, diffusion, implementation, participant experience and reach, rotating between qualitative exploration of causal processes and quantitative assessment of intermediate outcomes. Chapter 4 will further develop these ideas by exploring the themes making up this framework in relation to exercise referral schemes and how their evaluation might be advanced by adopting this framework.

4 Exercise referral schemes: the case for comprehensive process evaluation

4.1 Chapter overview

Despite a literature base spanning almost 2 decades, little is known about what works, for whom and under what circumstances in exercise referral schemes. As will be argued throughout this chapter, this is in large part due to the simplistic manner in which these increasingly complex interventions have typically been evaluated. The aims of this chapter are to explore key gaps in the ERS literature using themes developed in Chapter 3, and to highlight the potential roles of process evaluation in developing this literature. Following a brief overview of the emergence of ERS, it will be argued that although a developing concept, underlying theory is rarely made explicit, whilst efforts to integrate theory into evaluation typically simply add one or two links to the causal chain rather than articulating a comprehensive programme theory. A number of latent hypotheses underpinning ERS and their theoretical basis will then be explored. The chapter will then move on to discuss the need to examine diffusion and implementation of ERS across contexts before exploring how and for whom ERS work, through exploration of patient experiences and social patterning in reach within the definitive trial.

4.2 Exercise referral schemes in the United Kingdom: emergence and effectiveness

Exercise referral schemes (ERS) aim to promote long term change in physical activity, typically through providing a minimum of health professional referral to an exercise facility, where the patient receives a discounted exercise programme for 10-12 weeks. Since the 1990s, ERS have proliferated rapidly throughout the UK amidst widespread belief in effectiveness. However, in the mid-1990s, concerns began to surface regarding the need for evidence to justify their use (Iliffe et al., 1994). Reporting data collected in 1994, Fox and colleagues (1997) identified 35 established ERS and 71 pilot schemes throughout England, most of which were insufficiently resourced to allow evaluation. Months later, Riddoch and colleagues' (1997) reviewed primary-care physical activity initiatives including case studies of several ERS. These were

seen by stakeholders as having positive effects including emergence of social support and growing self-confidence, with individualised exercise and supervision highlighted as key mechanisms for supporting change, though effects were seen as short-term unless a continuing strategy was available.

Between the mid 1990s and 2001, a small number of trials (Taylor et al. 1998; Stevens et al. 1998; Harland et al. 1999) and observational evaluations (Lord and Green, 1995; Cochrane and Davey, 1998; Martin and Woolf, 1999; Damush et al., 2001; Greater Glasgow Health Board, 2004; Dugdill, Graham and McNair, 2005; Dinan et al., 2006) consistently demonstrated only short-term impacts on physical activity. Only two studies attempted to examine very long-term outcomes, using surveys 3-5 years after completion, with one finding that two-thirds of patients were still active, though achieving a response rate of only 40% (Day and Nettleton, 2001), whilst the other found no long term differences in physical activity between adherers and non-adherers (Jackson et al., 1998). Nevertheless, ERS received endorsement by the Department of Health in 2001 National Quality Assurance Frameworks stating that:

> The health service has a key role to play in giving people not only advice, but also the support they need in making changes to improve their health. Referral schemes can form an important option. (Department of Health, 2001) iii).

By 2006, a number of additional trials (Lamb et al., 2002; Harrison, Roberts and Elton, 2005b), observational studies (Damush et al., 2001; Greater Glasgow Health Board, 2004; Dugdill et al., 2005; Dinan et al., 2006) and systematic reviews (Morgan, 2005; NICE, 2006c), continued to provide evidence only of small or short-term impacts on physical activity, whilst another showed impacts on quality of life among over 65s, though did not measure physical activity (Munro et al., 2004). Hence, attempts were made to reign in the enthusiasm with which ERS were being endorsed, with National Institute for Health and Clinical Excellence guidance stating that they should be endorsed only as part of a properly designed and controlled effectiveness study (NICE, 2006a). Concerns that this was being interpreted as instruction to disinvest in existing services led to revised recommendations, advising continuation of high quality schemes to address clinical conditions, although ERS 'solely for the purpose of promoting physical activity should only be commissioned or endorsed ...

when they are part of a properly designed and controlled research study to determine effectiveness' (Department of Health, 2007).

Two further UK trials published this same year (Isaacs et al., 2007; Mutrie et al., 2007) offered higher intensity programmes than observed previously, involving fully supervised patient-only classes, both showing short-term effects on physical activity which were not maintained to 6 months. Interestingly, both also showed benefits for psychological well being, maintained despite lack of impact on physical activity, whilst Mutrie and colleagues also found significant impacts on cardio-respiratory fitness. A systematic review also published in 2007 (Williams et al., 2007) included meta-analysis of five UK trials (Harrison et al. 2005b; Isaacs et al. 2007; Lamb et al. 2002; Stevens et al. 1998; Taylor et al. 1998), concluding that for 1 person to become moderately active, 17 would need to be referred. The review also discussed findings from non-controlled studies, indicating limited uptake and adherence (Lord and Green, 1995; Cochrane and Davey, 1998; Martin and Woolf, 1999; Damush et al., 2001; Greater Glasgow Health Board, 2004; Dugdill et al., 2005; Dinan et al., 2006). On average, 1 in 3 referred patients did not attend a first appointment, whilst completion ranged from 12-42%. A more recent observational study reported completion of 52% (Lee, Griffin and Simmons, 2009) .

Despite significant questions remaining regarding whether ERS can be effective, Sowden and Raine (2008) have argued that widespread proliferation and unshakeable belief in effectiveness has made future conduct of rigorous effectiveness studies difficult if not impossible in England, whilst Simmons and colleagues (2009) argue that RCTs are no longer viable where services are already in place. Interestingly, in making this point, Simmons and colleagues interpret findings of small impacts in the meta-analysis by Williams and colleagues as demonstrating the absence of equipoise; findings interpreted by the lead author of the review (Williams 2009) as supporting the need for further trials. Though UK evaluative literature has been limited to England or Scotland, ERS have proliferated to a similar extent in Wales, with most areas running local schemes prior to 2007 until the randomised rollout of the National Exercise Referral Scheme, as will be described below.

As described above, significant emphasis has been placed on the need for effectiveness studies. However, numerous trials and observational studies have brought us little closer to understanding which models of delivery might be effective, for whom and under what circumstances. Whilst additional effectiveness studies are needed, evaluation must move away from simply examining the aggregate effects of a package of intervention, and towards an approach which allows incremental development of theory to inform future development. Key strengths and weaknesses of the ERS literature will be explored throughout this chapter, using themes developed in Chapter 3 of theory, diffusion and implementation, patient experiences and reach. No previous studies have examined all of these issues simultaneously, resulting in a somewhat piecemeal set of insights into functioning of often poorly defined programmes. First, a brief overview of the National Exercise Referrals Scheme (NERS), whose evaluation forms the context of this study will be provided.

4.3 The case of the National Exercise Referral Scheme in Wales

In the early half of last decade, most local health board areas in Wales operated local ERS. Their heterogeneity amidst limited evidence of effectiveness led the Welsh Assembly Government to identify improving standards of delivery as a key action area (Welsh Assembly Government, 20032005ab). Intent to learn from local schemes throughout Wales in order to inform development of training and infrastructure for a Wales-wide scheme was first signalled in 2005 (Welsh Assembly Government, 2005a). Publication of guidelines for delivery of high quality ERS followed (Welsh Assembly Government, 2006), with protocols for the National Exercise Referral Scheme (NERS) developed through reviews of practice and consultations with local coordinators.

NERS protocols described a scheme in which patients were recruited via a wide range of health professionals and referred to a local authority leisure centre. During a first appointment, patients were to receive a health check as well as being invited to discuss behavioural change using motivational interviewing, before agreeing goals with the exercise professional and entering a 16 week exercise programme, including patient-only classes and one-to-one gym-based instruction, fully supervised by a level 3

qualified exercise professional. Contact was to be made at 4 weeks to discuss progress through the scheme and further health checks were to be conducted at 16 weeks, before patients were signposted to other exercise opportunities. Contact was to be maintained through further telephone calls at 8 months and a further one-to-one consultation at 12 months.

An evaluation was commissioned in late 2006, with NERS protocols due to be implemented in 2007. The method chosen for evaluating the impacts of NERS was a pragmatic randomised controlled trial. Consistent with MRC guidance, the trial was accompanied by a nested economic evaluation to establish cost effectiveness. Design and conduct of the process evaluation formed the basis of this thesis. The thesis will report some secondary analyses, conducted by the author, of a routine monitoring database developed and collected by the trial manager, and will also use some baseline data collected for the trial[2]. Development and collection of all remaining data components were however led by the author, as were their analysis.

4.4 The role of theory in the development and evaluation of exercise referral schemes

As described in Chapter 2, MRC guidelines recommend development of interventions through consideration of relevant evidence and theory. ERS evaluations have sometimes involved researcher-initiated programmes (Harland et al. 1999; Mutrie et al. 2007), though others have evaluated schemes over whose development evaluators have had little control, such as the trial by Isaacs et al. (2007) which temporarily suspended entry to an existing scheme other than via trial participation. This section discusses the use of programme theory in the development of ERS and post-hoc applications of theory to its evaluation, before describing some of the latent hypotheses linking active ingredients to intended outcomes. It is beyond the scope of this thesis to test all of the hypotheses discussed. However, as described throughout, understanding what these hypotheses are provides the basis for understanding implementation.

[2] An overview of the design of the full evaluation is available in the following article: Murphy S, Raisanen L, Moore G, Edwards R, Linck P, Williams N, Ud Din N et al. (2010) A pragmatic randomised controlled trial of the Welsh National Exercise Referral Scheme: protocol for trial and integrated economic and process evaluation. *Bmc Public Health* 10: 352.

Study	Referral process	Baseline consultation	Scheme exit consultation	Referral criteria	Tailored exercise programme	Motivational communication strategies	Goal setting approach	Supervision by exercise professional	Group-based / individual	Cost and duration
Taylor et al. 1998	Written invite from GP or research team	lifestyle assessment, discussion about exercise perceptions and goals, physiological assessments, and advice on use machines	n/a	Patients aged 40–70, identified as smokers, hypertensive or overweight on medical records	Yes	None described other than advice during assessment	Unspecified discussion of goals	Available on request	Unspecified	Subsidised for 10 weeks
Stevens et al. 1998	Assessment questionnaire and letter sent by GP to identify patients	Physiological measures, medical assessments, lifestyle questionnaires and introduction to PA diary.	Review of progress	Sedentary (i.e. below recommended activity level) men and women aged 45-74	Yes	None described other than advice during assessment	'Do your best' approach. Physical activity diary provided.	None described	Unspecified	Unspecified. Lasted 10 weeks
Harland et al. 1999	Approached by researcher at general practice	Motivational interviewing	n/a	Sedentary (i.e. active <3 times per week)	No	Motivational interviewing	None described	None	Unspecified	Free. Thirty vouchers - no expiry date stated
Lamb et al. 2002	Unspecified	Advice and information (same as in control group)	n/a	Sedentary patients (i.e. taking less than 120 minutes) aged between 40 and 70 years	No	None described	Aim to meet current guidelines for PA.	All walks led by a walk leader	Group based and individual walk options	Free. Unlimited duration
Munro et al. 2004	Recruitment letter to eligible patients on practice register	none	None	Patients over 65 not in the most active quintile	No	No	None	Yes		Free for up to two years
Harrison et al. 2005	Referred by GP. Referral forms faxed to research team	Advice and information on increasing activity	Review of progress and exit route information	Sedentary patients (< 90 mins weekly activity) with CHD risk factors.	Yes	None described other than advice during assessment	None described	None described	Unspecified	Subsidised for 12 weeks

Isaacs et al. 2007*	Referred by GPs, practice nurses, and other health professionals	Physiological measures, fitness testing and 45 minutes worth of lifestyle questionnaires etc.	Physiological measures, fitness testing and 45 mins lifestyle questionnaires	Sedentary patients, aged 40 to 74 years, with CHD risk factors:	No	None described	None described	Activities involved supervised classes and walks	Group based	Free or subsidised for 10 weeks
Lord and Green 1999	Referred by GP	Assessment and discussion of exercise options	Questionnaire completed at 10-week consultation	Sedentary patients (undefined) aged 18-65, with no contraindications to exercise and CHD risk factors	Unspecified	None described	None described	Unspecified	Group-based and individual	Standard cost of a prescription for 12 weeks
Greater Glasgow Health Board (2004)	Referral by GP	Counselling session and agreement of goals	Follow up at 3, 6, 12 months. Content undescribed	Sedentary patients (undefined)	No	'Following recommended guidelines' (unspecified)	Negotiated with patients	None (mainstream services)	Group-based and individual	Subisdised for 3 months
Mutrie et al. (2007)	Recruited by research staff when attending outpatient appointments	None	None	Women with diagnoses of early stage breast cancer	No	None	Group discussions of goal setting after exercise classes	Full supervision	Patient only group based activity, with discussion afterwards on themes such as self efficacy and goal setting	Unspecified for 12 weeks
Jolly et al. (2009)	Referral by GP or nurse of self-referred	One hour person-centred baseline interview, including discussion of pros and cons of changing, perceived barriers and resources, implementation intentions and goal setting	Focus explicitly upon reinforcing internalisation of physical activity involvement	Judged by health professional as sufficiently motivated and likely to benefit from activity	No	Self determination theory	Goals focus upon behaviour change (how, what, where)	Unspecified	Unspecified	Unspecified cost for 12 weeks

Table 5. Characteristics of exercise referral schemes evaluated in UK settings

4.4.1 Use of theory in the development and evaluation of ERS

Table 5 presents an overview of the content of each UK effectiveness study where details are specified. ERS have evolved significantly during the last 15 years, with each scheme representing a manifestation of different programme theories. At present, there is no 'best practice' model, and further development is needed. However, in common with many complex interventions (Michie et al., 2009), theoretical assumptions are rarely articulated, regardless of whether researcher-initiated or focused on existing schemes.

Few have explicitly based delivery on a formal behaviour change theory. Exceptions include the study by Harland et al. (1999) which uses motivational interviewing, though gets its theoretical basis wrong, claiming that MI is based on the transtheoretical model. Another describes basing group discussions on topics guided by an integrated model combining self-efficacy theory and the transtheoretical model (Mutrie et al., 2007), though this is the only component for which a theoretical basis is cited. Whilst citing use of a small scale pilot, this largely assessed the feasibility of the intervention, paying limited attention to refining programme theory (Campbell et al., 2005). A forthcoming trial explicitly tests the hypothesis that autonomy-promotive consultations will enhance effectiveness of an existing ERS, through promoting internal motivation (Jolly et al., 2009).

Whilst formal theory has been rarely explicitly used in developing ERS, a range of behaviour change theory has been used in evaluation, to identify potential mediators. Some studies have examined impacts on stage of change; part of the transtheoretical model (Prochaska and Diclemente 1984). This model assumes that prior to making permanent changes, patients move back and forth between; i) pre-contemplation, ii) contemplation, iii) preparation, iv) action and v) maintenance, with movement to a higher stage increasing the likelihood of stable change. Application of the trans-theoretical model to physical activity has typically been of limited success, perhaps due to poor operationalisation, limited applicability or the fact that stage progress may only be only one step towards behavioural change (Adams and White 2003, 2005). In one ERS, between group differences in stage of change and physical activity were observed to 6 though not 12 months (Lamb et al. 2002), whilst in the other,

intervention groups increased their stage of change to a greater extent than controls at 6 months, though no between group difference in physical activity was observed (Isaacs et al. 2007). In one observational study (Greater Glasgow Health Board 2004), although less than 1 in 3 patients returned for 12-month assessment, increases in stage of change and physical activity were observed in this subsample.

Two trials have assessed changes in intrinsic and extrinsic barriers, assuming that removal of key barriers will make it easier for patients to align behaviour with their true preferences (Taylor, Doust and Webborn, 1998; Isaacs et al., 2007). Analysis of 9-month follow up data in one (Taylor, 1997) demonstrated impacts on barriers including perceptions of having someone to exercise with, as well as perceptions of having the right clothes, having the energy to exercise and feelings of embarrassment associated with exercising, though no impact on time-based barriers (Taylor, 1997). Changes were linked to improvements in anthropometric measures and adherence to the 10-week exercise programme (Taylor and Fox, 2005). However, no significant impacts were observed on long-term physical activity, perhaps indicating that these factors did not impact behavioural change, or that they too had subsided by 12-month follow up. In the other, barriers reduced equally in intervention and advice only control groups (Isaacs et al. 2007).

A small number of studies have tested impacts on constructs from self-efficacy theory (Bandura, 1977) or self-determination theory (Deci and Ryan, 1985). Self-efficacy theory argues that behaviour change is likely where seen as an effective means of achieving desired outcomes, and within the individual's capabilities. Self-efficacy has consistently emerged as a key predictor of physical activity behaviour change, with one review finding significant correlations with physical activity in all descriptive studies and impacts of self-efficacy promoting interventions on behavioural change (Keller et al., 1999). A more recent review of psychological mediators in youth physical activity programmes also found self efficacy to be the most consistent mediator of change (Lubans, Foster and Biddle, 2008).

Self-determination theory focuses on internalisation of motivation, arguing that stable change is most likely where internally motivated (i.e. where intrinsically enjoyable, or linked to personally valued outcomes). Internalisation of motivation is more likely

where psychological needs of autonomy, competence and relatedness are supported. Perhaps the most common use of self-determination theory has been assessments of autonomy-promotive practices in school physical education, shown to increase participation and activity levels in children (Standage, Duda and Ntoumanis, 2006; Spruijt-Metz et al., 2008; Chatzisarantis and Hagger, 2009). Studies with adults include a weight management intervention focusing on promoting autonomous behavioural regulation, which found significant impacts on weight loss and exercise (Silva et al.) and delivery of fitness classes to female attendees of leisure centres (Edmunds, Ntoumanis and Duda, 2008) leading to greater need satisfaction, positive affect and attendance. There is substantial conceptual similarity between constructs of 'self-efficacy' within self-efficacy theory and 'competence' which is conceived as central to internal motivation within self-determination theory. Hence, self efficacy and motivation typically rise and fall in parallel (Marshall and Biddle, 2001).

In one ERS, Jones et al. (2005) found that self-efficacy improved amongst completers of a 12-week gym-based programme, though worsened among non-adherers, indicating potential beneficial impacts, though adverse effects of trying but failing to change. Edmunds et al. (2007) found that patients for whom motivations become more internalised were more likely to adhere than those whose participation remained externally motivated. Morton et al. (2008) found that patients who adhered to an ERS demonstrated greater self-determined motivation than those who did not, though differences were already apparent in baseline data, with patients whose behaviour was more internally motivated at baseline more likely to subsequently adhere. These studies have not to date examined whether changes led to behavioural change, though as described, a trial is underway to assess the impacts of an autonomy-promotive approach to ERS delivery on internal motivation and physical activity (Jolly et al. 2009).

In summary, whilst formal theory has rarely been used in developing ERS, attempts have been made to apply theoretical models to evaluation, through testing impacts on theory-based change processes. However, these have drawn equivocal conclusions, indicating impacts on psychological mediators which decline by long-term follow up, whose growth has not been accompanied by behavioural change, or whose associations with long-term change have not been evaluated.

Identification of mediating mechanisms represents an important step in understanding how ERS work. However, understanding outcomes requires more than adding one or two links to the causal chain. Even if we know that a scheme impacted internal motivation for some patients, it is important to consider how this might have occurred and for whom. Each intervention involves a package of components, with each component included because developers believe that it will strengthen the whole and serve a function in facilitating change. However, little effort has been made to articulate the proposed active ingredients of ERS, or make explicit the hypotheses regarding how each contributes towards outcomes. Consideration will now be given to the roles of activities to recruit patients, promote scheme adherence and support long term change.

4.4.2 How might ERS work? Latent theories underpinning ERS

4.4.2.1 Recruitment processes and scheme uptake

ERS attempt to identify individuals likely to benefit from change and support them in behaviour change efforts. Given this individualistic focus and relative cost, ERS offer greatest potential as a targeted intervention, rather than as an intervention to produce population based change (Harrison et al. 2005a). To date inclusion criteria have varied across studies, sometimes including all 'sedentary patients' (Harrison et al. 2005b), though other times including only one clinical group (Mutrie et al. 2007). Criteria are typically interpreted and applied by health professionals, perhaps skewing uptake towards those most likely to contact a health professional, such as women and older patients (Green and Pope, 1999), whilst some evidence suggests that referral criteria are often not systematically applied (Graham, Dugdill and Cable, 2005). In one countywide scheme, a central referral mechanism to identify inappropriate referrals resulted in removal of 1 in 6 patients on the basis of medical unsuitabilility, readiness to change, psychosocial factors or already being active (Johnston et al., 2005). Furthermore, patients' likelihood of receiving advice from their GP appears variable according to practitioner characteristics, such as their own activity levels (McKenna, Naylor and McDowell, 1998).

Criticisms are often directed towards ERS trials for bypassing health professional advice, with researchers recruiting patients directly (Department of Health, 2001; Gidlow et al., 2008). Hence, rather than just a means to identify patients, health professional advice is often seen as an active ingredient, with a latent hypothesis being that advice from an esteemed figure may trigger commitment to behavioural change through raising patients' awareness of the need to change. A number of reviews have examined impacts of advice on physical activity (Ashenden, Silagy and Weller, 1997; Riddoch et al., 1997; Eaton and Menard, 1998; Eakin et al., 2000; Lawlor and Hanratty, 2001; Smith, 2004), with a review of four of these reviews concluding that brief advice supported by written materials produced modest though short-term effects in some patients (Hillsdon et al., 2004).

The short-term nature of change, typically amongst a minority of patients, is perhaps unsurprising, given that advice-based interventions assume that patients are rational decision-makers who if presented with the correct information, will act on it. Such assumptions are perhaps most consistent with the Health Belief Model (Strecher and Rosenstock, 1997), which argues that change is predicted by perceptions of personal susceptibility to and seriousness of health consequences of not acting and perceived efficacy of behaviour change in preventing these consequences, with advice from an esteemed source acting as a cue to action. Whilst this model recognises the importance of perceived barriers in restricting change efforts, it emerged as a means of understanding one-off preventive actions such as screening or immunisation, for which there may be few barriers. Utility is limited in explaining attempts to promote complex long-term changes (Strecher and Rosenstock, 1997).

Tulloch and colleagues (2006), whose review found larger effects of advice and counselling by allied health professionals rather than or as well as by general practitioners argue that there may be value in an interdisciplinary model in which physicians provide referrals to physical activity counsellors, in order to both raise awareness of the need to change and support the development of strategies to overcome barriers; an approach perhaps consistent with ERS which adopt behavioural counselling and support for changing behaviour into their structures.

4.4.2.2 Activities to support adherence

Entry on the basis of advice by an authority figure will likely mean that change is somewhat externally motivated (Markland and Tobin, 2010). Key roles of the exercise programme may therefore include supporting internalisation of motivation in order to assist development of an exercise habit. However, adherence to ERS is typically poor. Hence, before impacts can be realised, achieving higher levels of attendance is a priority (Williams et al. 2007). Key activities to support or incentivise attendance typically include patient consultations and financial incentives, roles of which will now be discussed.

4.4.2.2.1 Patient consultations

Most ERS include consultations on scheme entry, and sometimes again throughout the scheme or on scheme exit. Where described (see Table 5), consultations typically begin with health or lifestyle assessments. Though primary care health checks might motivate short-term change for some patients (Amoroso et al., 2009), patients have likely already been made aware of their conditions, and health checks principally serve a function of ensuring that exercise is not contraindicated. Content beyond assessment is often poorly specified, sometimes including agreement of an exercise programme, advice, or some (usually unspecified) form of goal setting.

In recognition of the typically poor adherence to ERS, Department of Health (2001) guidelines recommend explicit use of baseline consultations for motivational purposes through integration of theory-based motivational communication strategies such as motivational interviewing and goal setting. Two UK trials have attempted to integrate counselling strategies based on self-determination theory (Jolly et al., 2009) or motivational interviewing (Harland et al. 1999) into consultations. Another observational study also describes delivering counselling based on recommendations, though the recommendations in question are unspecified (Greater Glasgow Health Board 2004). A substantial movement towards integration of motivational counselling into ERS has emerged in overseas evaluations (Hosper, Deutekom and Stronks, 2008; Kallings et al., 2008; Lawton et al., 2008; Sorensen et al., 2008; Leijon et al., 2009; Romé et al., 2009). Observational studies typically show improvements in physical activity where ERS are combined with motivational counselling, whilst RCTs show impacts to be no greater than motivational counselling alone. However, one RCT of a

scheme underpinned by motivational interviewing demonstrated impacts on physical activity to 24 months among women aged 40-74 (Lawton et al. 2008). Potential roles of motivational interviewing and goal setting in supporting adherence will now be discussed.

4.4.2.2.1.1 Motivational interviewing

4.4.2.2.1.1.1 What is MI and how does it work?

Motivational interviewing (MI) is 'a collaborative, person-centred form of guiding to elicit and strengthen motivation for change' (Miller and Rollnick, 2009). Whilst owing substantial debt to Rogerian client-centred counselling (Rogers, 1959), and linked to concepts such as cognitive dissonance (Festinger, 1957), and self-efficacy (Bandura, 1977), motivational interviewing was not explicitly derived from any formal theory. It was developed by William Miller who, having become aware of emerging disjuncture between his approaches to substance abuse counselling and those taught in formal training (Miller and Rose, 2009), made explicit his approach in a concept piece, published as a now seminal paper (Miller, 1983).

A burgeoning number of trials and reviews have since demonstrated efficacy in a range of patient groups, contexts and behavioural domains (Dunn, DeRoo and Rivara, 2001; Burke, Arkowitz and Menchola, 2003; Britt, Hudson and Blampied, 2004; VanWormer and Boucher, 2004; Hettema, Steele and Mller, 2005; Rubak et al., 2005; Resnicow, Davis and Rollnick, 2006; Van Dorsten, 2007; Martins and McNeil, 2009). However, development of evidence for its efficacy has outpaced the quest to explain how MI works (Miller, 1996). Although parallels exist, the growing myth that MI was based on the trans-theoretical model led Miller and Rollnick to remove many references to this model from the 2nd edition of their key text (Miller and Rollnick, 2002). More recently, parallels have been drawn with self-determination theory (Markland et al., 2005; Britton, Williams and Conner, 2008), including explicit focus on autonomy-promotion and the social contexts necessary to support internal motivation.

In 2009, Miller and Rose (2009) identified two active ingredients of MI; a relational component focusing upon its interpersonal spirit, and a technical component focusing upon elicitation of commitment-related change talk, whilst limiting sustain talk.

Consistent with Rogers' (1959) concept of accurate empathy, an empathic, non-judgmental environment is considered a necessary context for clients to explore their ambivalence; a hypothesis supported by significant correlations between practitioner empathy and client behavioural change (Miller, Benefield and Tonigan, 1993; Moyers and Martin, 2006; Gaume, Gmel and Daeppen, 2008), and the failure of a trial using technical aspects of MI in a more authoritarian style to promote change (Kuchipudi et al., 1990).

Actions of more technical ingredients were linked initially to cognitive dissonance (Festinger, 1957), a state of psychological discomfort caused by holding two contradictory ideas, though cognitive dissonance has been replaced within MI by the concept of ambivalence, a pre-decisional phenomena which is not inherently troubling until the point of having to choose (van Harreveld et al., 2009). A negative link between sustain talk and behavioural change emerged in early studies (Miller et al., 1993) and has been replicated more recently (Baer et al., 2008). Although evidence of a link between change talk and behavioural change was not immediately forthcoming (Miller et al., 1993; Miller, Yahne and Tonigan, 2003), when focus moved to the type of change talk and its patterning throughout the consultation, behavioural change was predicted directly by the strength of commitment language elicited (Amrhein et al., 2003; Moyers et al., 2007; Hodgins, 2009).

In healthcare settings, sustain talk is commonly elicited by imposing ideas and values through directive advice giving (Lillian Southwick and Rui, 1991; Fogarty, 1997), presenting complex debates as rational choices. Patients will likely know that physical activity would be of health benefit, though change may conflict with core values such as spending time with family, with this ambivalence maintaining the status quo. Heavily advocating only one side of patients' ambivalence will likely be experienced as a threat to autonomy, triggering the emotional response of reactance (Brehm, 1966; Fogarty, 1997) whereby the patient is forced to reassert autonomy and justify their current behaviour through articulating arguments against change (Rollnick et al., 2005). Hence, the practitioner may achieve the opposite of their intention, directing the patient towards talking themselves out of change.

By contrast, in MI-based consultations, to minimise sustain talk, confrontation is avoided through responding empathically to resistance, and resisting impulses to present rational solutions to patients' problems, removing the need for the patient to defend their behaviour. Advice is provided only after establishing that the patient wants advice, ensuring that control remains with the patient. The practitioner elicits change talk through use of open questions and reflective listening, allowing the patient to explore views on how important change is to them and why, barriers to changing, and the discrepancy between their current behaviour and their deeply held values, essentially offering patients the opportunity to talk themselves into change within an empathic and non-judgemental interpersonal context (Rollnick, Miller and Butler, 2008).

4.4.2.2.1.1.2 Applications to physical activity and potential role within ERS

Application of MI to physical activity is a relatively new field. A 2009 review (Martins and McNeil, 2009) identified 9 trials using MI to promote physical activity, with 7 conducted in the 2 years prior to the review. Two trials overlooked by this review were the aforementioned New Zealand based ERS (Lawton et al., 2008) and one comparing a single session of brief negotiation (a reduced form of MI) delivered by a health promotion practitioner against direct advice (Hillsdon et al., 2002).

No significant impacts on physical activity were reported in four trials (Perry et al., 2007; Befort et al., 2008; Bennett et al., 2008; Greaves et al., 2008), though one reported near significant impacts on physical fitness ($p=0.06$) (Perry et al., 2007), whilst another impacted exercise self-efficacy (Bennett et al., 2008). Six reported short-term improvement in physical activity (Harland et al., 1999; Brodie and Inoue, 2005; Bennett et al., 2007b; Carels et al., 2007; Hardcastle et al., 2008; Lawton et al., 2008). Only three reported outcomes beyond 6 months, with impacts not maintained to 12 months in one (Harland et al., 1999), whilst in a second, impacts were observed at 12 months in per-protocol analysis though not intention-to-treat analysis (Hillsdon et al., 2002) and in the third, impacts were maintained to 12 and 24 months (Lawton et al., 2008).

'Successful' interventions were delivered by a range of providers including nurses (Brodie and Inoue, 2005; Bennett et al., 2007b), physical activity specialists

(Hardcastle et al., 2008) and doctoral students (Carels et al., 2007), and to a range of patient groups including long-term cancer survivors (Bennett et al., 2007b), overweight patients (Carels et al., 2007; Hardcastle et al., 2008), and patients aged 65+ with chronic heart failure (Brodie and Inoue, 2005). In some cases these were a stand-alone treatment (Brodie and Inoue, 2005; Bennett et al., 2007b; Hardcastle et al., 2008), whereas in others these were an adjunct to a weight management programme (Carels et al., 2007) or an ERS (Harland et al., 1999; Lawton et al., 2008). Hence, whilst long-term efficacy remains to be consistently demonstrated, evidence is beginning to indicate that for some patients, MI may offer utility for promoting physical activity.

In considering the potential role of MI within the causal chain for ERS, it is worth considering what patients will have experienced prior to attending. ERS typically begin with either health professional advice or a self-determined decision to seek referral. Referral seekers will perhaps have begun the process of linking change to their personal values, triggering a decision to seek support. For patients entering on advice, how this is delivered likely influences whether it triggers reactance or begins to trigger commitment to change. A recent UK study of interactions between doctors and diabetes patients found that where patient-centred communication strategies were used, these led to more positive affective responses from patients, though also observed that their use was rare (Moran, Bekker and Latchford, 2008).

Whilst health professionals' advice-giving styles are beyond the control of ERS implementers, a key role for MI may be to actively avoid further entrenching resistance arising from advice to change. Patients may enter having been told that they must change, and hence having spent time cognitively defending their position, and may therefore attend prepared with arguments against change. Delivering consultations in a manner which allows exploration and resolution of ambivalence and avoids further entrenching these arguments may help the patient engage with support offered. However, whilst details on content of consultations are rarely provided, principle emphasis is typically placed on assessment and advice-giving. Hence, the patient may be immediately rendered passive by a process of assessment and closed questioning (Rollnick et al., 2008), wresting control away at the very beginning of the therapeutic relationship rather than promoting autonomy.

There is plausibly therefore a key role for patient-centred approaches such as motivational interviewing in the causal chain for ERS between health professional advice and entry to the exercise programme, supporting patient autonomy, avoiding entrenchment of arguments against change, and strengthening and eliciting motivation for change prior to entering a programme which offers the support to achieve this change. Given its emphasis on allowing patients to link behaviour change to their values, this process may also lead to identification of specific, personally valued goals to motivate attendance.

4.4.2.2.2 Goal setting

4.4.2.2.2.1.1 Goal setting theory and applicability to health behaviour change

A goal is 'that which one wants to accomplish; it concerns a valued, future end state' (Lee, Locke and Lantham, 1989). Goal setting theory assumes that conscious desires influence action, that intentional behaviour is implicitly goal-directed and that making goals explicit enhances motivation through focusing attention upon how desired outcomes will be achieved (Locke and Latham, 2002). Whilst emerging within industrial psychology, Locke and Latham (2006) argue that goal setting can be applied to any outcome over which the individual has some control. Strecher and colleagues (1995) have however expressed concerns regarding tendencies for goal setting theory to permeate health behaviour interventions with limited attention to which principles might be transferable.

Two central assumptions are that goals must be difficult and specific (Locke and Latham, 2002). However, though difficulty may ensure productivity in industrial settings, among clinical populations, any behavioural change may be seen as difficult and whilst goals should perhaps be challenging enough for their achievement to have an impact on self-efficacy, achievability is crucial (Bovend'Eerdt, Botell and Wade, 2009). Goal success will likely breed higher motivation and further success, whilst failure may breed demotivation (Locke and Latham, 2002). Hence, Bodenheimer and Handley (2009) report that goal setting interventions in primary care typically include assessment of patients' confidence, with the practitioner playing a key role in encouraging patients to aim for goals which are challenging, yet achievable. Locke and Latham (2006) argue that goal setting is likely to prove most effective in

triggering change in synergy with intervention components which support self-efficacy. In the context of an ERS for example, goal setting accompanied by the support of an exercise professional may allow for achievement of goals which the patient is not confident achieving alone, enhancing self-efficacy and motivation to set and achieve further goals.

Locke and Latham (2002) argue that specificity improves effectiveness through reducing ambiguity. Setting measurable and time-bound goals facilitates self-monitoring of progress and enables provision of feedback, with feedback on achievement enhancing self efficacy (Locke and Latham, 2002). Bodenheimer and Handley (2009) argue that in primary care settings, a tendency to set proximal goals appears to have been widely endorsed, and that in practice, proximity leads to greater specificity. For example, a patient focusing upon losing 4 lbs in the following month as a means of progressing towards a distal goal of losing 3 stone, may be more able to break this goal down further to specific action plans.

Locke and Latham (2002) argue that in industrial settings, goals tend to be equally motivating whether assigned or self-set. However, this is described by Bodenheimer and Handley (2009) as the least transferable element of industrial goal setting theory, with change more likely when patients participate in decisions (Heisler et al., 2002). Locke and Latham (2002) argue that assignment of difficult goals may enhance self-efficacy where interpreted as an expression of confidence. However, whilst the goal-setter in industrial settings is perhaps able to reinforce success or punish failure, or may be an esteemed individual whom the employee wishes to impress, an exercise professional with no pre-existing relationship with the patient, assigning goals seen as excessively difficult, will likely trigger doubts about professional judgement rather than increasing patient self-efficacy.

Though imposing goals may be counter-productive, one recent study found that following feedback on current physical activity, participants could articulate the extent to which they wished to change, though were unable to arrive at concrete goals unguided (Saini and Lacroix, 2009). Hence, Bodenheimer and Handley (2009) argue that collaborative goal-setting has become the accepted norm; as reflected in recommendations within the American College of Sports Medicine's guide to exercise

on prescription for clinicians, which describes a need to negotiate goals with patients (Jonas and Phillips, 2009). These recommendations state that the most useful exercise prescriptions provide 'Specific, Measurable, Achievable, Realistic and Timely (SMART) goals' (p58). There are a multitude of meanings assigned to the acronmym SMART, the most common being 'Specific, Measurable, Attainable, Relevant and Timebound' (Rubin, 2002). ACSM recommendations are a slight deviation from this definition, which arguably offers the closest fit with goal setting theory (Rubin, 2002).

4.4.2.2.2.1.2 Goal setting and physical activity

A 2004 review of interventions promoting physical activity through goal setting identified 8 studies among adults, of which 6 offered positive results (Shilts, Horowitz and Townsend, 2004). More recently, Dishman and colleagues (2009a) conducted an RCT of a 12-week workplace intervention combining organizational action and personal and team goal-setting, finding positive impacts on physical activity, with greater increases amongst participants setting more challenging goals (Dishman et al., 2009b). Two recent trials have examined uses of technologies for setting, monitoring and providing feedback on goals, both finding impacts on physical activity immediately after intervention (King et al., 2008; Block et al., 2009). Others have integrated goal setting into interventions involving group discussions of behaviour change and goal progress (Redfern et al., 2008; Thompson et al., 2008; Xue, Yao and Lewin, 2008). Of these, only one study which focused on multiple behaviours did not impact physical activity, though achieved aggregate improvements in fruit and vegetable consumption (Thompson et al., 2008).

Hence, whilst outcomes have typically been measured in the short term, and it is difficult to disentangle its role from other intervention components, the balance of research findings to date indicates that goal setting may be a useful intervention component within physical activity interventions. Within ERS the potential synergy between goal setting and support for self efficacy provided by professional support has been described. In addition, ERS have increasingly begun to include substantial social support elements, consistent with a number of aforementioned trials which combine goal setting with opportunities for participants to discuss goals with other patients, with motivation and confidence perhaps enhanced by social support.

However, understanding the quality of goal setting is critical if their potential role in the causal chain is to be realised, with previous uses of goal setting in ERS typically unspecified or contravening goal setting theory. Several simply state that goals were discussed (Taylor et al., 1998), or negotiated with patients (Greater Glasgow Health Board, 2004). Only two offer insights into a specific approach. In the first, whilst participants were made aware of public health recommendations, they were simply encouraged to do more than they were currently doing (Stevens et al., 1998); a 'do your best' approach which runs counter to the fundamental principle that goals should be specific. In the second, a predetermined goal of 120 minutes per week was imposed (Lamb et al., 2002), which violates assumptions that goals are more effective when relevant to the patient, whilst also representing a large and potentially unachievable increase for sedentary patients.

4.4.2.2.3 Financial incentives

An additional component included in almost all ERS to date in order to enhance adherence involves offering subsidising access to an exercise programme. Hence, two common latent hypotheses are that: i) reduction of a cost barrier will make it easier for patients to become more active, and ii) by the end of the programme, the new behaviour will be sufficiently habituated for removal of the discount not to cause discontinuation of the behaviour. However, the function of financial incentives in motivating adherence has been studied infrequently.

Financial incentives can involve rewards, whereby patients are paid to change behaviours (Marteau, Ashcroft and Oliver, 2009), or discounting, whereby a cost is temporarily reduced. Whilst one ERS has used rewards (Hosper et al. 2008), discounting is more common. The former approach has been criticised by proponents of self-determination theory, given that laboratory studies typically indicate a tendency for extrinsic rewards to undermine intrinsic motivation (Deci, Koestner and Ryan, 1999). Use of financial incentives has been seen by some as coercion, which undermines internal motivation through constraining autonomy (Claassen, 2007).

However, this presumes that the incentive is being used to cajole patients into an unwanted behaviour change. Patients will often want to change, though this desire may conflict with other core values, such as spending time with family. Hence, desire

for change is thwarted by more immediate rewards associated with the less healthy behaviour or by the punitive nature of behavioural change. Rather than constraining autonomy therefore, financial incentives may help patients align behaviours with their true preferences, through making the desired behaviour more immediately rewarding or less punitive (Marteau et al., 2009). Whilst increasingly used, impacts on motivation have yet to be widely studied in the health behaviour domain (Marteau et al., 2009). Furthermore, impacts of financial discounting have been less extensively studied than rewards, despite likely having somewhat different impacts on motivation. Whilst rewards may temporarily motivate an undesired change, a discount for something which the patient has no interest in doing is likely little incentive.

Separating impacts of discounts from other ERS components has not always been possible. One exception was a study by Harland et al. (1999), who offered 30 free exercise vouchers. Short term change was observed in a condition combining vouchers with intensive motivational interviewing, though not in conditions providing only the discount or only behavioural counselling suggesting a synergistic action, though even these changes were not maintained to 12 months. A US trial found that adding financial incentives to a supervised exercise programme did not increase adherence (Wing et al., 1996). However, a more recent American study randomising patients to receive full or partial subsidy following GP referral found a 25% increase in sessions attended in the full subsidy condition (Shepich, Slowiak and Keniston, 2007). Hence, there is equivocal evidence that financial discounting might in some circumstances increase attendance, though less evidence that impacts last after withdrawal of the discount.

4.4.2.3 Promoting behavioural change: the exercise programme

The purposes of financial incentives and patient consultations are largely to enhance adherence. However, it is primarily through attending an exercise programme that the patient is anticipated to develop an exercise habit. The nature of exercise programmes offered by ERS has been highly variable. In early trials such as that by Taylor et al. (1998) professional supervision was available only on request, whilst others specify contact with professionals only in consultations, offering discounted access to mainstream services (Harland et al. 1999; Stevens et al. 1998; Harrison et al. 2005b). By contrast, recent schemes provide fully supervised classes (Mutrie et al. 2007;

Isaacs et al. 2007). Hence, early schemes latently theorised that health professional advice accompanied by a consultation to plan an exercise programme would be enough to facilitate development of an exercise habit, whilst a theoretical shift has been observed towards emphasising roles of intensive support and supervision.

Studies are now beginning to explore the roles of professional support within ERS. One involved 133 female completers of a 10 week scheme, finding higher identified motivation (i.e. acting due to a sense of the behaviour as personally important) amongst patients who perceived that the professional was supportive of autonomy and competence (Markland and Tobin, 2010), as well as amongst those who reported higher levels of social assimilation into the exercise environment and higher levels of relatedness to others within that environment. It is not however clear whether this refers to support during consultations, or whether this scheme offered a supervised programme. As described above, an ongoing study by Jolly et al. (2009) is using a randomised controlled trial to assess whether autonomy-promotive practices lead to enhanced internal motivation and physical activity. Again, this focuses on consultations rather than the exercise programme itself.

As will be elaborated in Section 4.5, qualitative studies highlight the value of supervision by a knowledgeable and supportive professional and of explicit opportunity for social interaction in supporting change. The role of emerging social networks in supporting change has been largely ignored within quantitative research (Thurston and Green, 2004), and along with professional supervision, appears to have very recently begun to permeate the theories of programme developers, with early trials placing little or no emphasis on provision of explicit opportunities for social interaction, whilst more recent trials commonly include patient-only exercise classes (Mutrie et al. 2007; Isaacs et al. 2007; Hosper et al. 2008).

4.4.2.4 Supporting long term maintenance of change

Although Riddoch and colleagues argued in 1997 that ERS were unlikely to be successful in the long-term unless supported by strategies to maintain changes beyond the programme, participants are typically expected to become independent exercisers after a short-term programme. If post-intervention strategies are offered in evaluated interventions, this is typically not evident within evaluation reports. Hence, this

implicitly hypothesises that sufficient psychological change will have been achieved for patients to independently maintain this new habit.

The typically short-term nature of changes observed however suggests that this may be a naïve assumption. Indeed, from a socio-ecological perspective, which views behaviour as constrained and facilitated by multiple levels of influence (McLeroy et al., 1988) including intrapersonal factors, social and cultural context and physical environments, decaying effects of interventions which focus solely on intrapersonal factors is unsurprising. Increasing sedentary behaviour in modern society is not a result of successive generations becoming lazier, but arises from changing social structures which have socialised activity out of daily living (McElroy, 2002). Individualistic interventions aim to support patients in developing strategies to overcome higher level impediments, but these remain largely unchanged (Stokols, 1996).

Department of Health (2001) guidelines recommend that in addition to consultations on entry to and exit of an ERS, schemes should maintain contact with patients who have not been attending the scheme to try and encourage them back, and that ongoing telephone contact maintained after leaving the scheme. This recommendation is supported by the only ERS to date to demonstrate impacts on physical activity beyond 12 months, in which the main component was monthly telephone support from an exercise specialist for a 9-month period (Lawton et al., 2008). Indeed one recent review of telephone interventions for physical activity and dietary change found that 11 out of 16 studies using telephone contact as a primary method of supporting change in physical activity produced positive outcomes (Eakin et al., 2007). Maintaining dialogue throughout the programme and beyond, via face to face contact where the patient is still attending the scheme and via medium such as the telephone where the patient is not, may serve a function in reducing relapse (Iliffe et al., 2008).

4.4.3 Summary

In summary, ERS are highly heterogeneous and are driven by varying, complex and typically unarticulated programme theories. A crucial stage in evaluating an intervention is to understand its underlying theory, through considering what each component is, how it might work, and synergy between components. However theory

has rarely been explicitly used in the development ERS, whilst efforts to use theory to guide evaluation have typically involved simply adding one or two links to the causal chain. Hence, it is typically not clear what is being evaluated, or how it is intended to work. Patients advised by health professionals to enter an ERS may be encouraged by this advice to attend, though alternatively, patients whose awareness of the scheme has been raised by other means may make a more self-determined decision to seek referral. Given that behaviour change will likely be externally motivated for many patients, a key challenge for ERS is to encourage internalisation of exercise motivations. Hence, use of baseline consultations for motivational purposes, through motivational interviewing in order to elicit change talk followed by harnessing of this change talk into specific, personally valued goals may allow patients to enter an exercise programme with attention focused on the need for change and the actions required to achieve it. Professional supervision may provide patients with the support to achieve these goals, enhancing self efficacy and a sense of autonomy, and increasing internal motivation, whilst group-based ERS may also offer the opportunities for formation of social networks supportive of change. Discounted exercise may remove a barrier to behaviour change, helping patients to align behaviour with their true preferences during the period when an exercise habit is being formed. Whilst developers often assume that sufficient intrapersonal change will occur for independent behavioural change to be achievable after withdrawal of the programme, pathways out of ERS and impacts of post-scheme contact on maintenance of changes in physical activity have received limited exploration.

4.5 Diffusion and implementation

4.5.1 Diffusion and implementation of ERS

As described above, whilst questions remain regarding the actions of key components within ERS, there is theoretical reason to believe that high quality schemes which incorporate evidence-based principles might be successful in facilitating change for some patients. However, a plausible model means nothing unless the intervention remains coupled to its underlying theory. Careful attention is therefore needed to understanding how the planned intervention diffuses into practice, and the consistency of implementation with this intended model. In ERS evaluations however, little attention has been paid to understanding challenges associated with diffusion or to

examining the quality of their implementation, under the somewhat naïve assumption that intervention as planned and intervention as delivered are one and the same.

The lack of emphasis on diffusion is perhaps in part due to the fact that evaluation studies have often involved interventions delivered by a small number of implementers within limited geographical regions, often involving one or two GP practices, and few leisure centres and exercise professionals. Hence, challenges associated with diffusing protocols into multiple settings have perhaps not been encountered. As described in Chapter 2, MRC guidelines recommend that efforts to understand and facilitate wider implementation should follow demonstration of effectiveness; a stage not yet reached within ERS research. However, the reality is that evidence-based or not, ERS have diffused widely throughout the UK, whilst their highly variable quality is perhaps evidenced by their widely variable uptake and adherence levels. Whilst attempts have been made to provide guidance and advice for the process of diffusing ERS into practice, the lack of research into these issues means that these have had limited grounding in evidence (Welsh Assembly Government, 2006).

There is good reason to believe that attempts to diffuse new protocols for ERS into multiple sites will encounter significant challenges. As argued by Sowden and Raine (2008), a key barrier to the conduct of rigorous evaluations of ERS, and to the diffusion of new protocols, is that these will need to be implemented in contexts where existing schemes are believed to be effective. Instruction to standardise practice represents an authority-innovation decision, which as described in Chapter 3, likely leads to faster adoption, though greater risk of weak implementation than where programmes emerge through consensus within the local system. Changes to existing practice will likely involve introduction of new and unfamiliar activities, with inclusion of complex activities such as motivational counselling enhancing the risk of partial adoption or poor implementation (Department of Health 2001). As the scale of delivery is increased, the number of stakeholders whose adoption will be required for successful implementation increases complexity further, requiring effective communications between heterophilious professional groups. Furthermore, a scheme developed at a national level will require tailoring to achieve compatibility with local settings and populations. Understanding how the scheme and the organisation

mutually adapt to one another is crucial in understanding how the scheme achieves routinisation, and how local tailoring is conducted to potentiate mechanisms of change at the local level (Rogers 2003; Hawe et al. 2004).

The only aspect of diffusion to receive significant attention in the ERS literature to date is their adoption into health professionals' practices. Harrison et al. (2005a) found that in a scheme involving 125 practices almost half referred 10 or less patients over a 5 year period, whilst 17% referred more than 100 per year. In one qualitative study, most patients reported asking to be referred, with health professionals rarely mentioning the scheme unless prompted by patients (Wormald and Ingle, 2004). Whilst the authors attribute this to limited communication of the scheme to health professionals, studies with health professionals have raised concerns regarding their endorsement of ERS. In one, most cited smoking cessation as their most commonly used referral service due to greater confidence in the smoking-health evidence base, perceptions of smoking cessation as a simpler behaviour change and medico-legal considerations (Graham et al. 2005). In another, referrers described limited confidence in exercise professionals' knowledge of clinical conditions, leading many to refer only the lowest risk patients (Wiles et al., 2008). Challenges securing adoption of new practices amongst leisure service partners or exercise instructors have received little attention.

The training and support requirements for full implementation of ERS have also received limited attention. One case study of a countywide scheme in Somerset describes developing a consultancy and support service with expertise including health psychology and exercise physiology, in order to mediate between health professionals and leisure centre staff, provide regular workshops and enable service providers to meet a countywide recognised standard of expertise (Crone, Johnston and Grant, 2006). Most evaluations however provide little detail on the training and competencies of implementers, perhaps in part because ERS have traditionally involved activities perceived as simple, and which implementers are well qualified to deliver, such as tailored exercise programmes or discounted exercise cards. However, even where incorporating complex and unfamiliar activities such as motivational counselling (Greater Glasgow Health Board, 2004; Hosper et al., 2008; Kallings et al., 2008; Lawton et al., 2008; Sorensen et al., 2008; Leijon et al., 2009; Romé et al., 2009),

diffusion into practice has almost always been presented as entirely unproblematic with no description offered of training, support and monitoring structures. Only one intervention, incorporating consultations based on self-determination theory, has appeared to acknowledge the challenges associated with changing practitioners behaviour, providing 2-day courses prior to delivery, a 2-month pilot phase to allow skills to be practiced and provision of top-up training and support during the trial (Jolly et al., 2009).

In addition to understanding how the programme is diffused into practice, Department of Health (2001) guidelines for ERS recommend monitoring implementation for quality assurance; a purpose overlapping significantly with assessment of fidelity and dose in process evaluation. However, to date, whilst few studies explore diffusion, none present data on fidelity or dose. Given the limited exemplars relating directly to ERS, and the fact that incorporation of approaches such as motivational interviewing represents perhaps the most fundamental recent change in practice in ERS delivery, discussion will now focus on a growing literature highlighting the importance of understanding the diffusion of motivational interviewing into practice and the consistency of implementation with underlying theory.

4.5.2 Diffusion and implementation of motivational interviewing

In recent years, Miller and Rollnick have attempted to draw attention to the challenges associated with effectively delivering MI, arguing that interventions purporting to use MI often violate its core spirit (Rollnick and Miller, 1995; Moyers et al., 2005). In 2009, the authors published a paper listing key theoretical misrepresentations of MI, arguing that it had become conflated with a number of theories and approaches with which it had conceptual overlaps, but important distinctions. Notably, highlighting what MI *is not* has been dependent on clear definition of exactly what MI *is*, a stage not reached within ERS research due to poor description of programmes and their underlying theory.

Key misrepresentations of MI included the aforementioned tendency to erroneously see MI as based on the trans-theoretical model (Miller and Rollnick 2009). The authors also expressed concerns that it had become seen by many as a way of tricking people into agreeing to a course of action. Furthermore, whilst decisional balance, or

listing of pros and cons for change was discussed by Miller and Rollnick (2002) as a potential means of gaining leverage to elicit change talk amongst more ambivalent patients, the authors argue that many practitioners have reduced 'MI' to this non-prescriptive technique, which is in fact recommended only for the most ambivalent patients (Miller and Rollnick 2009). MI has also been conflated with cognitive behavioural therapy, despite being humanistic rather than behaviourist, eliciting motivations already present rather than installing new skills. Studies therefore often test theories which diverge substantially from MI, contaminating existing evidence bases and hampering their development. Unfounded claims to be using MI in routine practice likely lead practitioners to take comfort in misguided belief that their practice is evidence-based, whilst failing to replicate benefits observed in controlled studies.

In 2009, Miller and Rose (2009) proposed a causal chain for MI which began with understanding diffusion and implementation, explicitly recognising that the first step in changing patients' behaviour is changing practitioners' behaviour. Explicit emphasis was placed on training practitioners and supporting them in becoming competent practitioners; a process described as having clear parallels to helping patients change their behaviour. Similarly, Madson and colleagues have recently highlighted the need to consider issues such as practitioner self-efficacy, intention to use MI as well as attitudes toward MI, in addition to providing knowledge of MI through training (Madson, Loignon and Lane, 2009).

Professionals such as fitness instructors tasked with delivering MI within an ERS will almost by definition be more accustomed to practitioner-centred, directive styles of promoting change. Movement towards MI therefore requires abandonment of learned professional practices and adoption of new practices, requiring the practitioner to accept the congruence between MI and their goals and values. Adopting the role of behaviour change counsellor may be a move which professionals are not well motivated to support. For example, whilst studies have commonly reported positive reactions amongst trainees (Lane et al., 2003; Broers et al., 2005), in one study, health care providers saw MI as taking too long to deliver, raising concerns that the relatively passive role of the practitioner clashed with their beliefs that provider input should be less limited and that patients should be provided with more information (Miller and

Beech, 2009). Hence, practitioners saw their role as to provide solutions, whilst MI sees solutions as being strongest when elicited from within.

Where ready to adopt MI, practitioners must be provided with the skills, confidence and opportunity to implement MI. Two key difficulties in ensuring fidelity have however been misconceptions that MI is easy to learn, and tendencies to overemphasise similarities with current practice. These challenges likely arise from a tendency for practitioners to make sense of new ideas through using prior practice as a frame of reference, particularly where the new information is not communicated clearly and differences emphasised, resulting in it being seen as little different to current practice and hence easy to adopt (Rogers 2003). Such misconceptions perhaps underlie the limited description of training or implementation in ERS claiming to use MI (Harland et al., 1999; Lawton et al., 2008) and are by no means limited to ERS, with Miller and Rollnick (2009) describing commonly receiving requests to teach MI in one or two hour workshops; an approach comparable to attempting to teach a complex sport or musical instrument over lunch.

An increasing body of research is attempting to inform best practice in training practitioners to use MI. Two-day courses are a common approach (Lane, Hood and Rollnick, 2008), although consensus is building that these are inadequate to facilitate proficient use without ongoing support (Wahab, Menon and Szalacha, 2008). One study found that although a two-day training course convinced clinicians that they had learned MI, practice samples showed insufficient change to benefit patients (Miller and Mount, 2001). As Miller and Rollnick argue 'initial 2-day training can provide a certain head start, but real skill and comfort grow through disciplined practice with feedback and coaching from a knowledgeable guide' (Miller and Rollnick, 2009). Hence, research is increasingly attempting to understand the support needed to reinforce introductions to MI and ensure that skills can be practiced (Miller, 2004; Bennett et al., 2007a; Smith JL, 2007; Mitcheson, Bhavsar and McCambridge, 2009). In one such study, dieticians randomly allocated to receive a tailored two-day MI course, followed by a one day top-up course and on demand advice and feedback, exhibited significant changes in counselling style, leading to improved patient outcomes (Brug et al., 2007). In a similar study, a 1.5 day course for GPs followed by 2 half-day top-up sessions influenced their counselling style, and patient attitudes to

behaviour change (Rubak et al., 2009). In one recent study, following training and ongoing feedback, skills of smoking cessation practitioners continuously improved throughout a 2 and a half year period and only at the end of this period did all practitioners reach minimum competence standards (Forsberg et al., 2010).

In evaluating a scheme claiming to adopt MI, transparency over structures to facilitate its implementation is essential. Given that the issue of how to best promote use of MI in routine practice is far from resolved, understanding practitioners' reactions to introductions to MI, and experiences of attempting to use its principles in practice may provide valuable insights into the successes and failures of integration efforts. As well as describing how implementation was supported, evidence should be provided that it was delivered. Self-reports of MI proficiency may bear little resemblance to expert ratings (Miller and Mount, 2001; Miller, 2004), and hence Miller (2001) advises assessing fidelity through direct monitoring of practice samples.

Development of measures to quantify MI fidelity has gathered pace in recent years (Moyers et al., 2005; Madson and Campbell, 2006) and the Motivational Interviewing Treatment Integrity (MITI) scale (Moyers et al., 2005) has emerged as the gold standard (Pierson et al., 2007), though requires intensive training, and has not been widely adopted. The Behaviour Change Counselling Index (BECCI; (Lane, 2005) offers a simple alternative, providing a single global score and has been validated for use by coders with a small amount of training. Hence, well validated and simple methods exist for quantifying the consistency of practice with an MI-based approach. Whilst not to date adopted within the ERS literature, moves toward assessment of fidelity are observed in other physical activity studies. Several recent trials describe quality control measures such as reviews with implementers to discuss tape-recorded consultations (Bennett et al., 2007b; Befort et al., 2008; Hardcastle et al., 2008), whilst only three have directly quantified fidelity (Carels et al., 2007; Bennett et al., 2008; Greaves et al., 2008). Whilst equivocal findings of the utility of MI for promoting physical activity may indicate that MI is more useful for some patient groups than others, in particular treatment contexts or intensities, differing outcomes may stem largely from failures to deliver MI which have not been identified through process evaluation.

4.5.3 Summary

In summary, whilst ERS have diffused rapidly into practice, limited attention has been paid to understanding how high quality implementation might be achieved or to scrutinising the quality of implementation. Quality has perhaps been somewhat difficult to define due to the absence of coherent theoretical models of intended delivery. However, following a clear explication of programme theory, attention must turn to understanding the consistency of delivery with programme theory. Whilst complexity in design has increased, methods for evaluation have remained simplistic, with challenges associated with incorporation of complex and unfamiliar activities ignored. Where poor implementation goes unnoticed, erroneous conclusions will be drawn, whilst interventions may fail to replicate the effects on patients' behaviour observed in studies from which theory is drawn.

4.6 Participant experiences of ERS

Gidlow and colleagues describe an important but often undervalued role for qualitative research in understanding patients' journey through ERS, offering insights into how ERS might work and for whom in context (Gidlow et al., 2008). Where timetables allow piloting to take place, qualitative research may play a substantial formative role in developmental pilot phases. For example, in one Dutch trial (Hosper et al., 2008) preliminary qualitative data indicated concerns regarding the dependence of patients on programme structures which informed changes to protocols prior to full trial, including provision of motivational counselling on programme exit and regular post-programme opportunities to try out free activity options.

Whilst useful in formative stages, qualitative data exploring patient experiences is also crucial within definitive evaluations. After all, what is delivered in practice may diverge from that which was intended, likely rendering much prior theory redundant. Hence, once the intervention as delivered is clearly defined, the focus of process evaluation should turn to exploring how the intervention is experienced. However, only a minority of outcomes evaluations have to date also published qualitative analyses, including two trials (Mutrie et al. 2007; Taylor et al. 1998) and two observational studies (Lord and Green 1995; Martin and Woolf 1999). A minimal qualitative element was also reported alongside the trial by Isaacs et al. (2007), which presents a 'selection of quotes' from open questions in list form with no analyses. A

number of stand-alone qualitative studies exploring patient experiences of ERS have also emerged in recent years (Crone et al. 2005; Singh 1997; Stathi et al. 2004; Wormald and Ingle 2004). This section will offer a brief overview of key findings from qualitative studies, before examining insights into outcomes where both quantitative and qualitative data are analysed, and discussing key limitations of the qualitative evidence base to date.

4.6.1 Findings from patient experience studies

The primary focus of qualitative ERS research has typically been the functioning of the exercise programme, although some studies have also touched on the role of other programme components, including discounts and health professional advice. Whilst in some studies, health professional advice has been seen as providing a legitimate justification for setting aside time to participate in exercise (Singh, 1997; Schmidt et al., 2008), in other cases, health professional advice has been seen as playing little role in patients journey through the scheme, with most patients asking for referral rather than being advised to enter (Wormald and Ingle 2004). Schmidt et al. (2008) highlight the perceived importance of discounts to patients in a deprived district, though also report that more than half of patients stated that they would not continue to exercise once the discount was removed.

As touched upon in Section 4.4.2.3, consistently emerging themes in relation to the exercise programme have included roles of the instructor in educating patients and supporting confidence and motivation (Stathi, McKenna and Fox, 2004; Wormald and Ingle, 2004; Hardcastle and Taylor, 2005). Patients in one qualitative study reported that the structure and supervision available within the leisure centre encouraged activity that they would not have engaged in at home, valuing guidance to navigate unfamiliar equipment (Wormald and Ingle, 2004). In another, patients reported positive contrasts between expectations of an authoritarian and forceful instructor and experience of a professional who demonstrated non-judgmental empathy (Wormald et al., 2006). Hence, whilst studies highlight a perceived need for education and advice, it appears important that this takes place within an empathic, non-judgmental relational context.

Another key theme emerging in almost all qualitative ERS studies has centred around the value of interaction with other patients. Interviewees in almost all qualitative studies have identified 'collective exercise with people in similar circumstances' (Singh, 1997) as a key mechanism for supporting adherence (Lord and Green, 1995; Singh, 1997; Wormald and Ingle, 2004; Crone, Smith and Gough, 2005; Hardcastle and Taylor, 2005). It is often not clear how much opportunity for social interaction was offered. Indeed in some studies, patients have expressed views that the exercise environment offered disappointingly limited opportunity to interact with other patients (Stathi et al., 2004; Wormald and Ingle, 2004). The benefits of group-based exercise have however emerged as a key component in at least two studies which were based on programmes involving patient only classes, both of which were limited to women who highlighted the value of a female only exercise context (Emslie et al., 2007; Schmidt et al., 2008). Whilst it may not always be acceptable to offer services only to one sex, other dimensions of similarity among patients are likely important. For example, participating in a scheme alongside other patients facing similar health challenges may facilitate social assimilation (Crone et al., 2005; Schmidt et al., 2008).

4.6.2 Combining qualitative methods with quantitative data

Where combined with quantitative evaluation, patient interviews have offered insights into how compatibility between activities important to participants and those offered have shaped adherence or outcomes. In the Stockport ERS evaluated by Lord and Green (1995), interviewees highlighted the importance of an ongoing relationship with a single trusted instructor, with discontinuity of instructors cited as a key reason for non-adherence, likely contributing to an overall adherence rate of only 18%. In a scheme achieving similarly poor adherence (12%), Martin and Woolf (1999) found that non-adherers were more dependent on others in the exercise environment, perhaps indicating that insufficient support or opportunity for social interaction was offered to meet the needs of most patients. Interviewees in one longitudinal qualitative study (Hardcastle and Taylor, 2005) focused upon the role of the interpersonal skills of the exercise professional in supporting autonomy, competence and internal motivation. However, in the trial reported by Taylor et al. (1998), completers typically reported having been more active than minimum public health recommendations prior to intervention at baseline, with the scheme perhaps only enhancing motivation among

already highly motivated patients. Professional supervision was available only on request and the scheme achieved no long term impacts on physical activity.

In one qualitative study among women undergoing treatment for breast cancer (Emslie et al., 2007), the patient-only group environment was seen as helping minimise anxieties about changes in appearance and allowing patients to develop valued social support networks. However whilst trial outcomes demonstrated positive impacts on quality of life and positive affect to 6 months (Mutrie et al., 2007), no impacts were observed on physical activity. Hence, whilst social support was perhaps beneficial in maximising participation and promoting psychological well-being, it appeared less effective in supporting long term change in physical activity, with mental health benefit perhaps mediated by social processes rather than by exercise. This finding of impact on psychological wellbeing in the absence of greater behavioural change than a control group was observed in another group activity based ERS the same year, which demonstrated impacts on depression (Isaacs et al. 2007). Of a list of 15 quotes from participants in this latter trial, 4 focus on disappointment at being 'dropped' after the programme, and difficulty maintaining change after the scheme.

4.6.3 Key limitations of qualitative evidence

In addition to limitations in relation to poor definition of the programmes under investigation and often limited accompaniment with effectiveness data, key limitations of this evidence base include a narrow focus on the exercise programme itself. As argued throughout Chapter 3, interventions do not deterministically cause change. Rather, changes occur in the interaction between patients and the intervention in context. Hence, a crucial starting point in understanding how patients' interact with an intervention is to understand patients' reasons for attending, and how well the scheme meets their wants and needs. Whilst some studies conducted during participation have emphasised patients' reasons for attendance, these have been limited to older patients (Stathi et al., 2004) or older women (Hardcastle and Taylor, 2001), highlighting issues such as functional fitness and desire for increased social connectedness (Stathi et al., 2004). Other studies have involved retrospective interviews (Lord and Green, 1995; Martin and Woolf, 1999; Wormald and Ingle, 2004), reflecting back on the programme, though offering limited focus on motivations for entering the scheme. Whilst one asked patients to indicate why they were referred from a closed list, this

reflected health professionals' reasons for referring the patient rather than patients' own motivations (Martin and Woolf, 1999). Hence, although patients' motivations will likely be more influential than those of external parties such as health professionals (Miller and Rollnick, 2002), and although recommendations for ERS delivery are increasingly highlighting the need to tailor delivery to patients' motivations (Department of Health, 2001), these have often been seen as a secondary consideration after health professionals' reasons for referring the patient.

In addition, studies have offered little exploration of future intentions, or concerns about the prospect of becoming independent exercisers after the programme. Whilst some of these studies are conducted during participation, even retrospective studies have missed the opportunity to explore these issues. Two simply describe whether patients were still active (Martin and Woolf, 1999; Wormald and Ingle, 2004), with neither exploring experiences of attempting to remain active. The only exception is the aforementioned trial by Isaacs et al. (2007) who present a list quotes from patients, with many focusing on reasons for non-adherence such as illness and injury. Hence, there has perhaps been a tendency to focus somewhat too firmly on experiences of the exercise programme itself as an isolated unit, rather than as a fleeting event in patients' lives, which may or may not help patients to make long-term changes.

Finally, qualitative data exploring processes of change has almost exclusively involved patient interviews, with views of implementers overlooked. Given that patient interview studies typically exclusively or disproportionately represent the views of older women, their role in understanding social patterning in responses to ERS is limited. Exercise professionals by contrast may offer an additional perspective on the actions of the programme, whilst being better placed to offer insights into which patient groups appear to respond most positively to the intervention and why.

4.7 Patterning in scheme reach

As described in Chapter 3, a key aim of process evaluation is to understand the extent to which the scheme reaches its intended audience, with interventions having most impact where effective and achieving widespread reach (Glasgow et al., 1999; Glasgow et al., 2001). As was noted, estimating the size of target audiences is however often challenging in targeted community based interventions. ERS will typically only

reach a minority of their potential audience (Fox et al. 1997) and in one study of referral to an ERS in a district with 125 potential referring practices, only 4% of the sedentary population were referred over a 5 year period (Harrison et al. 2005a). Indeed, given their often highly inclusive referral criteria, were ERS to reach all eligible patients, they would likely not be resourced to cope.

Exploration of programme reach may however usefully focus on profiling the types of patients entering such schemes and patterning in the likelihood of completing the scheme. Whilst uptake and adherence have generally been poor, attention has recently begun to turn toward understanding social patterning in responses to ERS. This literature has largely drawn upon observational studies, perhaps because the small scale nature of many trials would render such analyses underpowered. In the trial by Taylor et al. (1998) for example, no demographic predictors of adherence were observed, though only approximately 40 intervention patients completed the study. Pragmatic policy trials, given their scale, offer an opportunity to move towards understanding patterning in adherence in real world settings, with concurrent qualitative data offering explanations for how patterning emerges. Adherence may also subsequently be linked to patterning in outcomes.

A systematic review in 2005 (Gidlow et al., 2005) found no demographic predictors of adherence, though concluded that most studies reported adherence poorly, or provided limited socio-demographic profiling. Findings have been progressed in the years since this review, largely by the authors of the review, with the bulk of such studies reporting data from a countywide service coordinated by the Somerset Physical Activity Group (Gidlow et al., 2007; Crone et al., 2008; James et al., 2008), involving biweekly professionally supervised exercise classes at a local leisure centre for 8-12 weeks. In terms of referral process, this scheme was unique in its use of a 'central referral mechanism' removing referrals deemed inappropriate prior to entering the scheme (Johnston et al., 2005). Another (Chinn et al. 2006) focused upon participation in the trial of the aforementioned Newcastle Exercise Project (Harland et al. 1999), comprising MI and discounted exercise. Another focused on an observational evaluation of the scheme trialled by Harrison et al. (2005a), whilst one pooled routine monitoring data from 6 London based schemes, whose content is undescribed

(Sowden et al., 2008), and another pooled data from 2 undescribed schemes in the North-West of England (Dugdill et al., 2005).

4.7.1 Demographic correlates

ERS have typically achieved higher uptake amongst females, with all leisure centre based UK RCTs attracting almost twice as many females than males, perhaps due to a tendency for these patients to be more likely to come into contact with a health professional in any given year. However, although more likely to enter ERS, women are often less likely to complete them (Dugdill et al., 2005; Gidlow et al., 2007; James et al., 2008). James and colleagues argue that this may arise from the multiple social roles of women leading to prioritisation of needs of other family members over their own, greater emphasis on the role of social support as a determinant of physical activity in females (McMunn et al., 2006), and greater impact of anxieties regarding the leisure centre environment. The former hypothesis is perhaps supported by findings from one female only qualitative study, in which women describe health professional advice as providing a legitimate justification for taking time out from competing family commitments (Schmidt et al., 2008). The latter two hypotheses are perhaps supported by two of the qualitative studies described above in which the women only environment helped to eliminate anxieties about the exercise environment (Emslie et al., 2007; Schmidt et al., 2008) as well as enhancing mental health outcomes through emerging social support (Emslie et al., 2007).

Older patients have been consistently shown to be more likely to complete ERS than younger patients (Gidlow et al., 2007; James et al., 2008; Sowden et al., 2008). This trend is likely attributable in part to additional time pressures amongst working patients (Dunbar, 1992). Indeed, qualitative studies have pointed to a tendency for ERS to offer activities only at times poorly suited to the needs of patients of working age (Wormald and Ingle, 2004).

4.7.2 Medical correlates

A tendency has emerged for lower adherence levels amongst mental health patients (Dugdill et al., 2005; Crone et al., 2008; James et al., 2008). Although in one study, attendance of a first appointment was highest amongst patients referred for 'fitness' or 'mental health', scheme completion was not reported (Harrison, McNair and Dugdill,

2005a). Patients referred for overweight were also more likely than those referred for cardiovascular reasons to drop out of the programme in one study (James et al. 2008).

These trends perhaps arise from a tendency for patients' motivation to be undermined by their conditions. This is a key challenge previously identified in relation to using exercise to treat mental health (Seime and Vickers, 2006). Similarly, significant weight-loss may not be easily achieved through ERS alone, undermining motivation amongst overweight patients. Mental health patients may also experience impaired social functioning, whilst mental health patients and overweight patients likely experience greater anxieties about exercising in front of others. Indeed, a tendency for mental health patients to struggle to assimilate into the exercise environment has been cited in one qualitative study (Wormald and Ingle, 2004). Given the aforementioned linkage between social assimilation, personal relatedness and internal motivation amongst ERS participants (Markland and Tobin, 2010), it is likely that challenges assimilating into the social environment damage the chances of motivation becoming internalised.

4.7.3 Psycho-social correlates

Three studies have investigated baseline psycho-social characteristics impacting scheme adherence. One found that non-completers were more reliant on support of others when attending the leisure centre (Martin and Woolf, 1999), although given that completion rates of only 12% were achieved, this perhaps indicates that only a very highly motivated subsample of patients benefitted from the scheme, which receives little description. A second reported that participants with modest expectations were more likely to achieve their goals and to adhere to a 12-week gym based programme (Jones et al., 2005) consistent with goal setting theory which argues that unrealistic expectation will likely result in goal failure and amotivation (Locke and Latham, 20022006). Although in Jones and colleagues' study, baseline stage of change and self-efficacy were not associated with adherence, a small scale study with 16 ERS patients reported higher internal motivations at baseline amongst patients who went onto complete an ERS (Morton et al., 2008), with the scheme more likely to be completed for patients who reported exercising for autonomously determined reasons.

4.7.4 Socioeconomic correlates

Findings on uptake and adherence by socioeconomic status are equivocal. Gidlow and colleagues (2007) reported that uptake was lower for participants in deprived areas as well as among rural-dwelling participants in Somerset, although in one study across 6 Greater London primary care trusts, referrals were higher in areas of deprivation, with attendance and completion comparable across areas (Sowden et al., 2008). Harrison and colleagues (2005a) found no patterning of uptake by deprivation levels, although did not examine adherence beyond a first appointment. Refusal to participate in the trial of the Newcastle Exercise project (Chinn et al., 2006) was higher amongst participants who lived in areas with high levels of deprivation or had lower levels of education. In addition, non-participants were more likely to report poorer health and being an adult carer, although were less likely to have children living at home. Although more non-participants considered that they were already sufficiently active to gain health benefit, these patients typically had less knowledge of the benefits of physical activity and attached less importance to it in maintaining health. Furthermore, non-participants were more likely to cite fear of leaving their home unattended to exercise, though less likely to cite having no one to exercise with as a barrier to exercise.

4.7.5 Summary

In summary, social patterning is beginning to emerge in quantitative profiling of ERS adherence by patient characteristics, highlighting the need to focus on tailoring programme delivery to the needs of population sub-groups. This represents a welcome move beyond asking simply whether ERS 'work' and towards asking for which types of patients needs are best met by current models of ERS delivery. Findings are however inconsistent and studies have often been accompanied by limited attempts to link patterning to the nature of the programme provided, hampering an understanding of whether certain patient groups respond better to certain types of programmes. Qualitative data from elsewhere offers a number of potential explanations for how patterning emerges, and mixed-method research would perhaps be useful in simultaneously examining how and for whom patterning emerges in responses to clearly defined intervention models.

4.8 Chapter summary

Whilst ERS have proliferated rapidly throughout the UK, there is limited evidence of their long-term effectiveness. The literature reviewed throughout this chapter points to a number of key shortcomings in the current evidence base which might be addressed through effective process evaluation alongside effectiveness studies. ERS have typically been poorly described, with underlying theory unarticulated, and to date evaluators and programme developers are left with insights into the aggregate impacts of poorly defined schemes, whose theoretical plausibility and reasons for ineffectiveness cannot be easily critiqued.

The impacts of ERS might be improved through careful consideration of programme theory and adoption of evidence-based principles into protocols, such as MI and goal setting. However, in diffusing new protocols into practice, challenges are to be expected in persuading implementers of the need to change, whilst introducing new and unfamiliar activities significantly heightens complexity and risk of weak implementation. As the scale of implementation grows, so does the diversity of contexts and populations targeted by the intervention. To date, no studies have explored the challenges involved in implementing ERS or how protocols achieve fit with varying local contexts and populations. Moves towards integration of complex and unfamiliar activities such as MI have been treated as entirely unproblematic.

There is a vital role for qualitative data in understanding how ERS work and for whom. Attention to implementation prior to analyses of qualitative data on patient experiences would however allow these to be understood in the light of exactly what was delivered, allowing incremental changes to be made to programme delivery. A number of studies have begun to explore patterning in reach. Whilst attempts to move beyond aggregate analyses and towards understanding emerging patterning in responses to ERS is encouraging, prior focus on implementation would likely offer insights into which types of programme produce social patterning, whilst combining quantitative profiling with qualitative data may offer insights into how patterning emerged in context.

In this thesis, it is anticipated that following the stages of process evaluation described in Chapter 3 will offer significant insights into the functioning of a National Exercise

Referral Scheme. Through developing an understanding of exactly what the intervention is, both in theory and in practice, as well as qualitative exploration of how the scheme is delivered and experienced by patients, insights will be offered into challenges diffusing national policy into local practice, how the delivered intervention works, and for whom.

5 Methodological issues in the NERS process evaluation

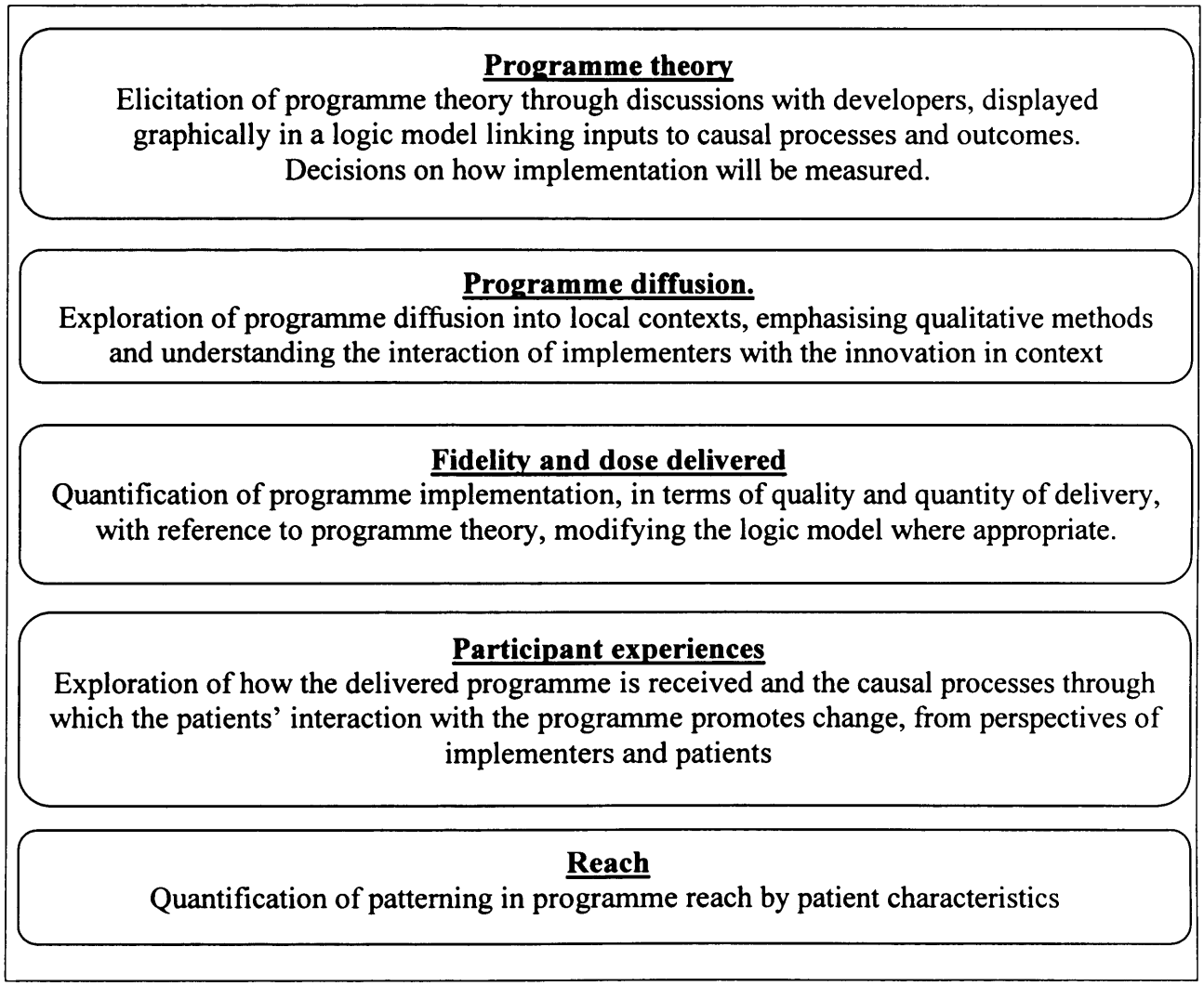
5.1 Aims and research questions

This thesis aims to provide rich detail on how NERS was implemented and experienced by implementers and patients, in order to provide insights into how implementation might be improved and in order to facilitate interpretation of outcomes effects. The study is rooted in beliefs that the causal chain for complex interventions begins with efforts to ensure effective delivery, and that interventions cannot be understood in isolation from the contexts in which they are implemented or the human interactions involved in their delivery and receipt (Berwick, 2008). Hence, functions of process evaluation are both to inform the high quality implementation of complex interventions and to understand the causal chain leading to the production of trial outcomes.

The study uses as a starting point the framework for process evaluation described by Steckler and Linnan (2002), which suggests focus on: i) context ii) fidelity, iii) dose delivered, iv) dose received, iv) reach and v) recruitment. However, as described in Chapter 3, this framework includes a number of potential areas for development. Firstly, eliciting programme theory is conceived as a key process evaluation activity, without which fidelity cannot be defined (Saunders et al. 2005). Secondly, the term 'context' is replaced with the term 'programme diffusion', allowing for a less deterministic interpretation of the role of context in shaping delivery, aligned with a critical realist understanding of causality (Sayer, 2000) which focuses on interactions of implementers with diffusion activities in context (Greenhalgh et al., 2004). Thirdly, whilst conceived as a separate construct within Steckler and Linnan's framework, recruitment of relevant stakeholders is conceived as part of the process of programme diffusion. Fourthly, within the phrase 'dose received', the term 'dose' is problematically biased towards quantification, whilst the term 'received' is problematically passive. The position of this thesis is that patients' interactions with an intervention are best understood in context via qualitative methods. Hence, the term 'dose received' is replaced with the term 'participant experience'. Finally, in relation to reach, emphasis will be on exploring emergence of social patterning in referral, uptake and adherence. There is also a need for more in-depth consideration of

methodological issues in the selection of process evaluation methods and for more explicit linkage between process evaluation components, as described throughout Chapter 3.

Figure 1. Process evaluation framework developed for the NERS process evaluation



An overview of the NERS process evaluation framework is provided in Figure 1. Essentially the framework rotates between qualitative exploration of causal processes and quantification of intermediate outcomes. Following explication of programme theory, qualitative exploration of programme diffusion examines causal processes through which the intervention takes shape across varying local contexts, whilst implementation checks quantify the intermediate outcomes of diffusion processes. Quantification of implementation in turn provides a framework for understanding participant experiences, through defining the intervention with which the patient interacts, with qualitative data on patient experiences guiding hypotheses regarding

causal pathways and patterning in reach. In order to minimise interpretation biases, analyses will take place prior to outcomes analyses (Oakley et al., 2006).

A primary aim of this thesis is to apply this evaluation framework and reflect on its usefulness for informing improved implementation and interpretation of outcomes within randomised controlled trials. Hence, findings chapters will address a number of key empirical questions in relation to NERS. Subsequently, the discussion chapter draws together empirical findings, discussing implications for exercise referral practice, for interpretation of trial outcomes and for conduct of process evaluation within policy trials. The thesis addresses the following broad empirical questions:

- RQ1: How is national policy for exercise referral diffused into local practice?
- RQ2: How consistent is the delivered intervention with programme theory?
- RQ3: How and for whom does the delivered intervention facilitate adherence and behavioural change?

In the early stages of implementation, additional questions emerged from examination of training manuals which revealed that only a one-hour course in motivational interviewing had been offered, leading the author to communicate concerns to policy representatives that this was not sufficient to facilitate practice. Though no recommendations were offered at this stage, two-day training courses were implemented in response to these concerns and became the subject of a mixed-method sub-evaluation, examining practitioners' views on the acceptability of training and the feasibility of putting it into practice, whilst quantitatively monitoring how fidelity was impacted by the course. This study essentially addressed the following emerging research question.

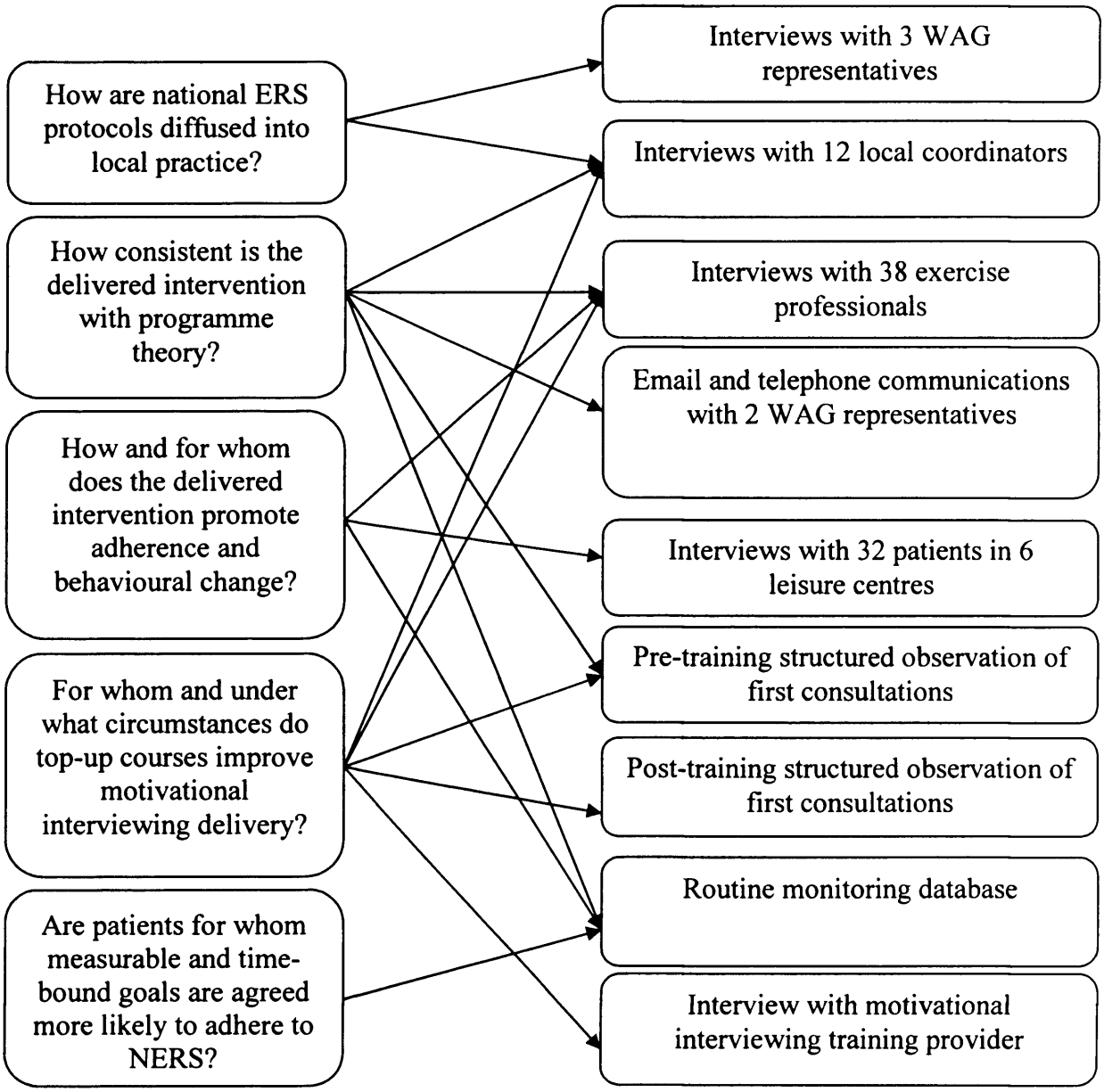
- RQ4: For whom and under what circumstances does a two-day training course in motivational interviewing lead to increased consistency with motivational interviewing?

The highly variable delivery of a related key component, goal setting, was also communicated to implementers, and provided an opportunity to explore whether goal setting quality increased over time and the role of baseline goal setting quality in predicting patient adherence, addressing the following primary research question.

- RQ5: Are patients for whom measurable and time-bound goals are agreed more likely to adhere to NERS?

An overview of linkage between research questions and data sources is presented in Figure 2. Given the complexity of study design and the number of methods and samples in each stage of the study, findings from each process evaluation component chapter will be preceded by a brief recap of data sources, and followed by a detailed discussion which draws out its implications for the evaluation as a whole and for the next stage of the study.

Figure 2. Linkage between data sources and research questions



This chapter however provides a detailed overview of the design of the process evaluation as a whole. First, an overview of key challenges in combining quantitative and qualitative methods and positions of this thesis on these issues will be discussed. Subsequently, the development of qualitative aspects of the evaluation will be described followed by an overview of quantitative components, concluding with a brief guide to empirical chapters.

5.2 Adopting a mixed methods approach

Public health and health promotion are positioned sometimes uncomfortably between medicine and social science; two fields traditionally with significant misgivings about one another's work. Whilst early medical dominance of public health ensured a hegemonic position for quantitative research (Baum, 1995), this began to be challenged in the mid 1990s. In a seminal paper published in the British Medical Journal (Pope and Mays, 1995), the authors argued for greater recognition of the role of qualitative research in 'reaching the parts other methods cannot reach' and against 'rigid demarcation of qualitative and quantitative research as opposing traditions'. The same year, Baum (1995) argued against maintenance of divisions within public health research, stating that:

> 'Some public health issues need quantification and ask questions such as: How many? How much? How often? What change? Others are more qualitative and require us to know things such as: Why did that happen in that context? Why do some participate and others not?....most public health researchers dealing with public health issues generated by policy-makers, communities or other public health practitioners will want to ask both types of questions.'

More recently, Curry and colleagues (2009) identified 5 circumstances under which one may wish to integrate qualitative data into an outcomes study. These are i) when investigating complex phenomena that are difficult to measure, ii) in order to generate data needed for a comprehensive understanding of a problem, iii) to gain an insight into potential causal mechanisms, iv) to develop sound quantitative measurement processes and instruments and v) to study special populations. Many of these justifications resonate with aims of process evaluation, which attempts to understand causal processes which are likely too numerous, complex and unpredictable to measure. Given the diversity of questions which any comprehensive evaluation must address, mixing methods is essential, with insistence on mono-method research

representing the biggest threat to advancement of knowledge (Onwuegbuzie and Leech, 2005).

However, despite increasing calls for mixed-methods approaches, the quantitative-qualitative debate has been inhibited by significant boundary maintenance efforts, with researchers often feeling the need to declare allegiance to one side or the other, as if extolling both traditions undermines one's academic credibility (Oakley, 2000). Divisions are maintained by the paradigmatic fashion in which textbooks and methods courses present the traditions (McEvoy and Richards, 2006). Indeed, Onwuegbuzie and Leech (2005) argue that breaking down divisions may require an overhaul of methods teaching, with quantitative and qualitative methodology modules replaced with courses which teach both simultaneously, promoting an approach of selecting methods appropriate to the research question (Gorard, 2004).

Substantial practical difficulties remain to be overcome. These include the degree of integration between methods and the extent to which methods are mixed or parallel (Creswell and Clark 2003), how findings from one approach illuminate the other and how to maintain clarity whilst describing increasingly complicated designs and findings (O'Cathain et al. 2008). This section discusses paradigmatic challenges in mixing methods and positions of this thesis, before moving on to discuss key practical challenges in mixed-methods research.

5.2.1 Paradigmatic challenges

A useful illustration of the paradigmatic distinctions often drawn up between quantitative and qualitative methods is observed in Denzin and Lincoln's (2000) Handbook of Qualitative Research. In this text, it is argued that 'qualitative researchers stress the socially constructed nature of reality, the intimate relationship between researcher and what is studied, and the situational constraints that shape inquiry...They seek answers to questions that stress *how* social experience is created and given meaning' (p 8). Subsequently, the authors claim that quantitative researchers rely on positivism, focus upon causal relationships rather than processes, and emphasise the notion of value-free inquiry (Denzin and Lincoln, 2000); p8). As argued by Sullivan (2002) many social science researchers have rejected quantitative methods, due to ideological objections with positivism and quantitative research,

fuelled by such definitions. However, whilst recognition of the value of qualitative research is crucial, given the need to also address questions which require quantification, movement towards a fully qualitative paradigm should be resisted.

The tendency for quantitative methods to be labelled as positivist in a climate of fierce anti-positivism is perhaps partly responsible for the critical decline in quantitative expertise within UK social sciences (Lynch et al., 2007), which hampers the development of a culture of rigorous policy evaluation (Creegan and Hedges, 2007). However, arguments against quantitative methods have typically involved caricaturing positivism (Bennett, 2005) and misrepresenting assumptions made by quantitative researchers in order to align them with these caricatures (Yu, 2006). 'Positivist' has been engineered into a term of abuse (Clark et al. 2007), hurled lazily at anyone who performs quantitative research, whereas in reality, positivism has been widely rejected for decades (Phillips and Burbules, 2001). As Clark (1998) argues, to make a case for one's approach through arguing against a perspective that no-one embraces does little to strengthen the arguments of the alternative.

Differences between approaches are commonly exaggerated and similarities systematically ignored (Onwuegbuzie and Leech 2005), with purists conflating method with epistemology (Baum 1995). Just because sometimes we wish to describe a phenomena, whilst other questions call for us to count its prevalence does not mean that we are assessing fundamentally different phenomena (Gorard and Taylor, 2004). Similarly, there is no reason why a strong social constructionist cannot use numbers, or why a positivist cannot describe phenomena in words. In reality, quantitative researchers make sense of their findings in a narrative manner, whilst qualitative researchers implicitly quantify, identifying themes based on how commonly they emerge, or commenting on the prevalence of different perspectives in their data (McEvoy and Richards, 2006).

Whilst Morgan (2007) argues for movement away from research paradigms, Johnson and Onwuegbuzie (2004) describe mixed-methods research as a third paradigm. Unlike quantitative and qualitative paradigms however, the mixed-methods paradigm does not reject research which does not mix methods, but accepts that some problems lend themselves to quantitative research, others to qualitative research, and others such

as comprehensive policy evaluation, must use both. Such an approach fits most naturally with moderate, anti-dogmatic worldviews such as critical realism (McEvoy and Edwards 2006). It is the position of this thesis that distinctions between quantitative and qualitative approaches are more constructed than real, and are sustainable only through endorsement of false dichotomies. A number of these dichotomies and the positions of this thesis on these issues will now be discussed.

5.2.1.1 Objective reality or subjective experience?

Ontologically underpinning debates surrounding the two opposing research paradigms are arguments surrounding the nature of truth claims. Those operating within a quantitative paradigm are often described as viewing the primary goal of research as discovery of truth. From this 'positivist' perspective, if researchers use valid and reliable methods, truths will become observable. In practice, few quantitative researchers subscribe to this naive realism. By contrast, taken in its strongest form, the social constructionist view of reality attributed to qualitative researchers by Denzin and Lincoln (2000) suggests that there is no truth beyond the minds of individuals, and that social phenomena are entirely constructed through inter-subjective processes. However, just as few quantitative researchers subscribe to naïve realism, few qualitative researchers subscribe to a relativist world-view (Coffey and Atkinson, 1996), which in its strongest form renders social research a futile enterprise, with findings representing nothing beyond the research process which produced them (Sayer 2000).

Knowledge is influenced both by the realities of our daily lives and by interpretations of those realities (Johnson and Onwuegbuzie 2004). Indeed, it is not always necessary or desirable to capture reality. As argued in Chapter 3 for example, the perceived advantages of a new innovation will likely play a significant role in shaping local acceptance. However, whether the innovation is an improvement according to some objective criteria becomes an irrelevance, as the constructed perceptions of its advantages amongst implementers in a given context will be what guides adoption decisions (Greenhalgh et al. 2004). Similarly, implementers' theories of change, whilst fallible, likely exert greater influence on implementation than any objective reality. Hence, exploring perceptions and experiences adds significant value to an

understanding of implementation and participation in complex programmes, regardless of truth value.

However, in some instances it is desirable to get close to understanding objective social realities. Complex interventions aim to influence real phenomena beyond the minds of individuals, and hence, an explicitly realist ontological position is necessary in order to understand their outcomes and the processes through which change is facilitated. Adopting a realist ontology and accepting the existence of mind-independent social realities however leads to epistemological questions regarding the extent to which social reality can be captured through social research. In quantitative research, standardisation is often seen as an essential mechanism to avoid extraneous influence and allow comparability, with attempts made to minimise influences of the researcher on the production of data. Qualitative researchers by contrast commonly accept that the researcher's immersion in the research process is necessary in order to gain in-depth data.

Claims to objectivity in quantitative research are commonly criticised through arguments that tools of measurement inherently represent the values of the dominant culture (Bourdieu and Passeron, 1977), with value-free enquiry impossible, given that researchers' beliefs and attitudes will play a key role in shaping research questions, how they are addressed and how findings are interpreted. The fallibility of measurement has been highlighted by feminist researchers in reference to the use of measures designed by men, which unfairly favour responses of males (Gilligan, 1982). Other examples include IQ tests, once thought to represent an objective measure of intelligence, but biased towards white middle-class respondents (Mensh and Mensh, 1991).

Qualitative purists often suggest that quantitative researchers do not acknowledge their subjective influence on the research process. However, whilst such researchers may aim for measures to provide meaningful representations of social realities, few would argue that their measures correspond perfectly to an objective reality. Indeed, any good quantitative researcher will acknowledge and discuss sources of bias in the data, for example in terms of social desirability bias or Hawthorne effects (Campbell, Maxey and Watson, 1995), and would acknowledge issues which compromise their

position, such as funding by an interested party. Indeed, safeguards built into modern RCTs, such as random allocation by an independent third party (Torgerson and Roberts, 1999), represent recognitions that our positions may cause conscious or unconscious subversion of the research process. As argued by Phillips and Burbles (2001), identifying examples where unacknowledged subjectivities have compromised objectivity is an intellectually weak means of arguing that the entire approach is fatally flawed. Whilst subjectivities may enter the research process, where exposed through opening up work to critique, attempts can be made to minimise these in future studies which examine whether observed trends remain or are an artefact of these biases. Critique will however likely take place most fervently where new knowledge challenges existing discourses, particularly where social science remains the reserve of privileged social groups (Sayer 2000).

Few would argue that subjective influences can be removed, although as argued by Pawson and Tilley (1997), we know that we can never be perfectly clean, yet we still wash. Rather than an achievable goal, objectivity can be seen as a useful regulative ideal (Johnson and Onwuegbuzie 2004). Whilst rejecting relativism, understandings of social realities will always be fallible, situated in discourse and limited by available means of interpretation (Sayer, 2000). Rather than a one-to-one correspondence with reality, the frames of reference and theoretical models used by researchers represent socially constructed means of representing, if not transparently, complex realities.

There is however a trade-off between the extent to which subjective influence can be minimised, and the depth of data which can be obtained. In quantitative research, through attempting to disengage, richness of data is compromised and data are limited to the superficial. Qualitative methods by contrast heighten the need to reflexively recognise the interactional nature of the production of data. In two articles reporting the same facts, the positions of the author will inevitably shape how the story is told, although the underlying story may remain the same. Knowing those positions will allow the reader to draw their own conclusions about how the underlying story is being represented. However, the need to move beyond the superficial and add depth to quantitative findings, only achievable through significant researcher immersion in the production and interpretation of data, remains one of the most compelling arguments

for mixing methods in evaluation research (Curry et al. 2009). As argued by Berwick (2008):

> 'methods that seek to eliminate bias can sacrifice local wisdom, since many designs intentionally remove knowledge of context and mechanisms. That is wasteful. Almost always, the individuals who are making changes in care systems know more about mechanisms and context than third-party evaluators can learn with randomized trials' (p 1183).

This thesis uses a range of quantitative and qualitative methods, all of which are subject to influences which prevent them from truly representing a transparent window into objective realities. Impacts of researcher position and challenges such as Hawthorne effects and social desirability biases are discussed throughout Sections 5.3 and 5.4, and in discussion of findings.

5.2.1.2 Inductive or deductive?

A second key dichotomy has centred around whether research should be inductive or deductive. Quantitative researchers commonly use methods such as forced choice questionnaires, often criticised for putting words into participants' mouths (May, 2001). Such methods essentially test preconceptions of the researcher, and one is unlikely to discover new findings unless actively looking for them. By contrast, with the popularity of approaches such as grounded-theory, which focuses upon 'discovery' of theory which 'emerges' from data (Glaser and Strauss, 1967), qualitative research has commonly been seen as inductive, allowing insights beyond the preconceived ideas of the researcher, though more deductive approaches such as framework analysis (Pope, Ziebland and Mays, 2000) are common in health research.

Endorsing this dichotomy however ignores the need for a combination of induction and deduction in any enquiry (Gilbert, 2006). The notion that qualitative researchers do not enter into research with ideas about what they might find is absurd, whilst formation of hypotheses for deductive testing by quantitative researchers relies on inductive reasoning, based upon hypotheses formulated from prior literature and experience. Hence, feedback loops emerge, with theory providing justifications for collecting data, which in turn contributes to modification of theory. Human reasoning rotates between inductive and deductive reasoning and all research is driven to

differing degrees by theory and data (Morgan, 2007). Methods may be usefully mixed for abductive purposes, with the more inductive results of qualitative research acting as inputs for the more deductive work of quantitative research and vice versa (Morgan, 2007). This movement back and forth between approaches is supported by recent researchers such as (Mendlinger and Cwikel, 2008) who describe 'spiralling between qualitative and quantitative data in attempting to understand women's health issues', and as described throughout Chapter 4, is crucial in understanding links between causal processes and their outcomes.

This thesis is neither fully data driven nor theory-driven. Instead, analysis at each stage is guided by a combination of prior theory and emerging themes. In policy evaluation where large amounts of data are needed, whilst a relatively short window of opportunity likely exists for their collection, pragmatic considerations will influence the order in which data are to be collected (Creswell and Clark, 2007). For example, monitoring structures for the purposes of implementation checks will likely need to be in place from the start of the scheme, whilst qualitative components may be more meaningful where collected after implementers have had time to experience the intervention in practice. However, whilst not always able to influence the data which are collected for subsequent phases, previous stages of analyses inform the theoretical frameworks used for analyses of subsequent components, with relationships between study components iterative and dynamic.

5.2.1.3 Context specificity or generalisability?

A further dichotomy stems from whether studies seek context-specific or generalisable information. Within quantitative research, efforts are commonly made to use sampling methods which result in groups of participants who represent a wider population, with the goal of generalising findings, whereas in qualitative research, more in-depth data are collected from smaller groups, commonly selected to represent a range of views and experiences rather than for the purposes of representativeness. However, it is difficult to imagine any finding which would be so unique as to have absolutely no applicability to any other context. This would render social research a voyeuristic enterprise, with the impossibility of replicating the same context twice precluding application of any finding to future programme development. However, similarly few trends will transcend all historical and cultural contexts, or translate to all patient

groups. The quest for generalisability may mask a multitude of complexity, with limited aggregate outcomes of an intervention in a nationally representative population masking significant benefits in some sectors, and harms amongst others (Pawson and Tilley, 1997; Sayer, 2000).

A more realistic view is that research findings are transferable from one patient group or context to another to varying degrees. Using the aforementioned example from the natural sciences, striking a match will likely cause an explosion only in very specific contexts such as a gas-filled room. There are perhaps however fewer contexts in which throwing a live hand-grenade into the room will not. Similarly, some forms of complex intervention will transcend contexts, or be adaptable across contexts, to a greater extent than others. Hence, evaluators perhaps need to focus on understanding how much of existing knowledge bases might be transferable to a new group or set of contextual circumstances (Morgan, 2007).

As a minimum, quantitative sub-group analyses and integration of mediators and moderators into analyses can offer a means of understanding transferability from one patient group or context to the next (Connelly, 2002). Furthermore, understanding the contextual contingencies required for the actions of implementers to produce positive effects represents a vital role for qualitative research within policy trials (Pawson and Tilley, 1997; Hawe et al., 2004b). Within this thesis, qualitative data will be used to explore patient experiences of the programme, as well as exploring the interaction of implementation activities with their local contexts.

5.2.2 Practical challenges

The value of mixing methods within intervention trials is increasingly being recognised, with largely qualitative process evaluations built into a number of recent trials (eg.,Moore et al. 2007; Starkey et al. 2005). However, whilst becoming increasingly common (O'Cathain, Murphy and Nicholl, 2007; O'Cathain, 2009), a number of practical difficulties remain to be resolved. Several reviews of mixed-methods research indicate that justifications for combining methods are often not forthcoming (Creswell, Fetters and Ivankova, 2004; Hanson et al., 2005; O'Cathain, Murphy and Nicholl, 2008). In addition, a review of 118 mixed-methods studies funded by the Department of Health between 1994 and 2004, (O'Cathain et al., 2008)

identified a marked lack of transparency in reporting of methods, particularly in qualitative components. As described in Chapter 3, in process evaluation, mixing methods is commonly undertaken without justification, with qualitative data often analysed in a superficial and subservient manner. The justifications for rotating between qualitative exploration of causal processes and quantification of implementation in this thesis have been described throughout Chapter 3.

Due to the aforementioned rarity of researchers equally skilled in both traditions, mixed-methods research commonly relies upon teamwork, comprising individuals with one or the other skill set working together. However, if made up only of specialists in both areas, questions need to be asked surrounding how integration will be managed. Hence, whilst there is a need for specialists in both fields, there is also a need for researchers who are able to use both methods effectively, and this thesis provides the opportunity for a previously quantitative researcher to become a mixed-methods researcher.

Perhaps arising from divisions within research teams, quantitative and qualitative study components can often be seen almost as parallel studies rather than components of one larger study (Bryman, 2007). To some extent, fragmentation is inevitable, given that publishing all findings of a complex study in one article likely requires an unacceptable level of parsimony in one or more elements, compromising the depth which was likely the reason for adopting mixed-methods. Publishing the entire NERS evaluation or even the process evaluation, in a single article would result in findings becoming superficial and methods opaque. More pragmatic approaches may be to publish components separately, with clear links between them, or with an additional paper which summarises key findings from all study components without presenting any new empirical data (Stange, Crabtree and Miller, 2006). Whilst Stange et al. (2006) recommend publishing sequential qualitative and quantitative articles in the same journal, it will not always be possible to negotiate this with journal editors, with journals often preferring to emphasise one methodology or the other (Bryman, 2006). In this thesis, whilst analyses will for the most part be presented in sequential phases rather than directly mixed, discussions will present these as part of a whole rather than a series of stand-alone studies, through reflecting on insights offered by previous and subsequent findings. Findings from each component will be followed by detailed

discussion to clarify what it adds to the unfolding story of the NERS evaluation, and its implications for analyses to follow.

5.2.3 Summary

In evaluating public health interventions, understanding outcomes through quantitative research, and causal processes through qualitative research is essential to offering insights into how, for whom and in what contexts schemes produce change. Mixing methods is relatively rare within the social sciences, due to significant investments in boundary maintenance efforts amongst methodological purists, and a lack of researchers skilled in both traditions. Mixed methods approaches are however consistent with critical realist interpretations of causality which focus on understanding outcomes and developing programme theory, viewing objectivity as a regulative ideal, rather than fully achievable, embracing the value of understanding objective realities and subjective interpretations and exploring the contextual contingencies for the transferability of findings to new settings and patient groups. This section will now move onto a detailed description of the methods used within this thesis, before the chapter concludes with a guide to empirical chapters.

5.3 Methods used in this thesis

This section describes the methods adopted within this thesis. The order in which findings are presented will reflect the framework presented in Figure 1. However, given that as indicated in Figure 2, some process evaluation components draw on multiple data sources, whilst some data sources are used in multiple process evaluation components, the development of methods is described under methodological themes rather than according to the process evaluation component to which they relate in order to avoid undue repetition of methodological considerations. The chapter then concludes with an overview of empirical chapters, describing how data sources will be used to address the research questions described in Section 5.1.

5.3.1 Eliciting programme theory

To elicit programme theory, the author consulted protocol documents and notes from planning meetings with policy representatives, and constructed a draft model of key inputs and the roles that each was anticipated to play in affecting long term behavioural change. Hypothesised change mechanisms drew upon the review of

theoretical literature presented in Chapter 4. This draft model was then sent to two representatives within the Welsh Assembly Government, who were asked to indicate the extent to which the model reflected their theories for how NERS would bring about change, and to identify areas of disagreement or additions or subtractions to the model. Representatives then telephoned the author to provided feedback, and the model was modified and re-circulated until no changes were required.

5.3.2 Qualitative study components

5.3.2.1 Aims

Qualitative components aimed to understand programme diffusion and explore participant experiences of the delivered intervention, from the perspectives of patients and exercise professionals. A further emerging use was to explore exercise professionals' responses to mid-trial MI training.

5.3.2.2 Selection of methods

5.3.2.2.1 Interview type

Given the experiential nature of the research questions requiring qualitative data, these study components were well suited to interview-based approaches (Silverman, 2005). Interviews have been defined by Berg (2004) p75) as 'simply a conversation with a purpose'. Britten (1995) describes three types of interview: i) 'structured interviews', typically used for quantitative research, ii) semi structured interviews 'conducted on the basis of a loose structure consisting of open ended questions that define the area to be explored...and from which the interviewer or interviewee may diverge in order to pursue an idea in more detail' (Britten 1995 p251), or iii) 'depth interviews', which may cover only one or two topics but explore these in depth with questions emerging as the conversation develops. The chosen degree of structure may be driven by the specificity of research questions and the extent to which a deductive or inductive approach is justified by what is already known. For this thesis, semi-structured interviews were selected for qualitative components, consisting of open questions and topics for discussion, which were neither prescriptive nor-exhaustive, allowing the interviewer to explore and probe key issues of interest arising from understandings of

the literature and earlier stages of analyses, whilst allowing practitioners and patients to identify and discuss issues not anticipated by the interviewer.

5.3.2.2.2 Group versus one to one interviews

Interviews may be conducted in a one-to-one or group format. Indeed, previous ERS studies have used group interviews (Wormald and Ingle, 2004; Wormald et al., 2006), individual interviews (Hardcastle and Taylor, 2001; Hardcastle and Taylor, 2005) or a combination (Stathi, Fox and McKenna, 2002; Stathi, McKenna and Fox, 2003; Stathi et al., 2004). Whilst a group setting perhaps means that not all patients will offer opinions on all topics, interaction can provide insights into consensus and conflict in views and experiences (Frey and Fontana, 1991), with discussion among participants likely drawing out details which interaction between one participant and an interviewer would not have elicited (Bloor et al., 2000). The group setting also offers opportunity to elicit a wider range of perspectives more quickly (Berg, 2004). However, group dynamics may lead participants to respond in a different manner to how they would were they asked the same question in a one-to-one interview, particularly where there is a hierarchy amongst participants. Lower status participants may be less likely to contribute or express disagreement with higher status individuals, leading to false consensus and overrepresentation of the views of higher status participants (Reed and Payton, 1997). Furthermore, group size may compromise the depth with which a topic may be explored.

Choices of methods were influenced by these issues as well as pragmatic considerations. In implementer interviews, it was crucial to allow frank and open critical reflection on experiences of NERS. Given that conducting interviews in a group format may intensify tendencies for positive appraisal, in order to avoid appearing critical of other implementers, one-to-one interviews were conducted for exercise professionals and their line managers. One exception was the interview conducted with policy representatives, which for pragmatic reasons took place in a group format, involving two civil servants involved in the transition to NERS and a representative appointed after its commencement. An interview was scheduled as part of the economic evaluation, and it was considered that including questions for process evaluation in this interview would be more viable than attempting to arrange further interviews, given the time constraints on civil servants. It is acknowledged that a

degree of hierarchy may have been present, with the recently appointed member of this group perhaps less likely to express disagreement.

As will be discussed below, patient interviews were conducted immediately after exercise classes. Initial conversations with implementers whilst these were being planned indicated that classes were typically attended by between 4 and 10 patients at a time. In most centres, all finished at the same time. Hence, conducting one-to-one interviews would involve either sampling two patients (one per researcher) or expecting patients to linger until first interviews were completed. Both were unsatisfactory, as the first reduced the number of patients whose views could be represented, whilst the second involved unacceptable inconvenience to participants. Given that the programme was experienced in a group format, group interviews were considered likely to offer insights into group processes and how the programme facilitated change. Hence, most interviews were conducted in a group format. In a minority of centres however, patients attended the gym in groups, but were given individual programmes which finished at different times by as much as half an hour, allowing patients with a higher level of fitness to exercise for longer. In these instances, a group interview would involve significant inconvenience to participants' who finished the sessions first, and one-to-one interviews were chosen. Six one-to-one interviews were conducted, as well as six group interviews including a total of 26 patients.

5.3.2.2.3 Face to face or telephone

Interviews may be conducted by telephone or in person. Whilst use of the telephone may make rapport building difficult, telephone interviews may also allow a greater sense of anonymity (Sturges and Hanrahan, 2004). Quantitative research comparing structured face-to-face or telephone-based interviews has commonly indicated similar findings between methods (Rohde, Lewinsohn and Seeley, 1997; Brustad et al., 2003). Whilst fewer studies have compared modes of qualitative interviewing, one study comparing themes from telephone and face-to-face interviews indicated similar results (Sturges and Hanrahan, 2004).

The choice between modes may however be influenced by resource implications. In evaluating a national scheme, face-to-face interviews may involve substantial

travelling, time and financial cost. This thesis was conducted in Cardiff; up to 250 miles from some trial sites. Hence, implementer interviews were conducted by telephone. Most implementers had already met the researcher during implementation meetings in mid-Wales, and hence establishing rapport via telephone was likely to be less difficult than where there was no prior contact between interviewer and interviewee. For patient interviews however, as these would involve participants who were perhaps less familiar with the ongoing evaluation, it was considered important to invite patients to participate in the study and to conduct interviews in person. Hence, a smaller number of case study centres were selected for the conduct of patient interviews, with researchers travelling to six sites.

5.3.2.3 Semi-structured interview topic guides

5.3.2.3.1 Welsh Assembly Government representatives

Data from WAG representatives was used primarily to explore programme diffusion. A topic guide therefore explored perceptions of how effectively the scheme had been implemented as well as perceived barriers and facilitators of implementation and how these had been overcome.

5.3.2.3.2 Area coordinators

Data from coordinators was also used primarily to understand programme diffusion. A topic guide was developed exploring coordinators opinions of the NERS protocols and the move to national standardisation, views on how well the protocols worked in their area and any issues which had made the scheme particularly easy or difficult to implement in their area.

5.3.2.3.3 Exercise professionals

Data from exercise professionals was primarily used to understand causal processes through which the delivered scheme produced change, and perceived patterning in patient responses to the programme. A topic guide was developed exploring professionals' opinion of the scheme, their role within the scheme, views on patients' readiness to change on scheme entry and perceived processes of change. For professionals who at the time of interview had recently attended MI training, an additional topic was included asking about perceptions of this training.

5.3.2.3.4 Patients

Data from exercise professionals was primarily used to understand causal processes through which the delivered scheme produced change. A topic guide was developed exploring participants' opinion of the scheme, reasons for attendance, perceived changes and processes of change, self efficacy for exercise, and barriers/facilitators of attendance.

5.3.2.3.5 MI training provider

As the MI training provider initiated contact with the evaluation team, asking to share his experiences, it was considered likely that a more structured schedule would be unnecessary. The interview therefore simply explored perceptions of patient responses to the course and perceived issues in attempting to implement the skills learned in practice.

5.3.2.4 Accessing participants

5.3.2.4.1 Welsh Assembly Government representatives

Three Welsh Assembly Government representatives involved in implementing the scheme were invited by economic evaluators to take part in a telephone interview. Process evaluation topics were integrated into this planned interview, conducted by the economic evaluator.

5.3.2.4.2 Area coordinators

Twelve local health board areas were involved in the NERS trial. Each area was headed by a local coordinator. Contact details were provided by the national coordinator, and coordinators were each invited by email to participate in a semi-structured interview. In two areas, coordinators were new in post at the time of interview. Hence, in one, the previous coordinator was invited to interview whereas in the other, the coordinators' line manager was interviewed. Coordinators in all 12 areas consented to participate.

5.3.2.4.3 Exercise professionals

Area coordinators acted as gatekeepers for professionals and were asked to provide email contacts for their professionals, so that professionals could also be invited to interview. Contact details were obtained for the 41 professionals in post at the time of

interviews in the 12 areas. Each professional was sent an email inviting them to participate in a telephone interview and asking them to provide suitable times for an interview lasting up to one hour. Professionals were sent up to 3 emails before being considered non-responders. In total, 38 professionals in 12 areas participated. One arranged two appointments, but at both times, could not be contacted at the arranged time. Another 2 did not reply to email invitations.

5.3.2.4.4 Patients

Exercise professionals acted as gatekeepers for patient interviews. Centres were sampled purposively so as to represent a range of geographic areas and levels of area level deprivation. Centres were divided into North, South and West Wales groups, with each group dichotomised around the median Welsh Index of Multiple Deprivation score. Random number generators were used to sample a case study centre from each group, each located within a different area. Within each centre, participants were selected opportunistically. The researcher first contacted the coordinator asking permission to visit, before asking permission of an exercise professional for two researchers to attend the centre. Professionals were asked to advise patients booked into the classes on the day of the visit that two researchers would be in attendance in order to invite patients to take part in an interview about their experiences of the scheme. Such an approach automatically excludes any referred patients who do not enter the programme, and hence cannot offer insights into reasons for non-entry. However, as NERS operated a rolling programme, with patients entering and leaving the scheme at different times, this meant that interviews would include patients at a variety of stages of the programme, with some patients interviewed on their first day, whilst others were in the closing stages of their programme. All patients attending the centre on the day of the case study visit were asked to participate in a group or one-to-one interview. Only two participants declined this invitation, in both cases due to having other prior commitments. Thirty-two patients took part, including 28 females and 4 males. Mean (and standard deviation) age of patients was 59.8(12.5) years, and ranged from 24 to 82

5.3.2.4.5 MI training provider

As part of the sub-study reported in Chapter 7, the training provider delivering the MI training course was interviewed. The training provider was informed by email that a

sub-study was being conducted to examine impacts of training on subsequent delivery of MI, and in reply, contacted the evaluation team offering to share his views on training delivery. Hence an interview was scheduled.

5.3.2.5 Data collection procedures

5.3.2.5.1 Implementer interviews

Telephone interviews were conducted at the time agreed with the interviewee. At the commencement of the interview, the interviewer reiterated that the interviewee was under no obligation to take part, could refuse to answer any question without being asked to give a reason and could terminate the interview at any time, before verbal consent was sought to switch on the digital voice recorder connected to the telephone line. In most cases, interviewees demonstrated a significant willingness to discuss their practice, with interviews often going on beyond the 1 hour that interviewees had been asked to set aside. Whilst interviewees had been made aware that they were free to leave at any time, in all cases when the interview overran, practitioners were asked if they wished to cut the interview short, although none took this option. These interviews were conducted shortly after the end of randomisation in October 2008.

5.3.2.5.2 Patient interviews

Patients were approached as a group at the beginning of their exercise class and the research project was explained to them. Patients were asked to participate in an interview immediately after the class. These were typically held in the café of the leisure centre, where one researcher bought patients a drink whilst the other gave out participant information sheets and consent forms, offering explanation of any issues where requested. After asking verbal consent to tape record the interview, the recorder was switched on and the interview conducted. These typically lasted approximately 30 minutes, although patients were informed that they were under no obligation to take part, could refuse to answer any question or leave the interview at any time without being asked to give a reason. Interviews were conducted approximately 6 months after randomisation began, in April 2008.

5.3.2.6 Researcher position

In any research method, the researcher's situation will play a role in shaping the understanding that is produced. However, the interactive nature of the interview

heightens involvement of the researcher in every stage of data generation and interpretation, influencing questions that are asked, which emerging avenues are pursued and the theoretical perspectives from which data are analysed. Hence, the researcher is an active co-constructor of shared understandings (Holstein and Gubrium, 2004). This does not mean that issues arising from qualitative research are not valid or do not represent anything beyond the research process (Miller and Glassner, 2002). However, interaction between the same participants and a different interviewer would likely elicit and emphasise different aspects of the same realities.

The position of the researcher was perhaps particularly important in interviews with scheme implementers. There is an inevitable power dynamic between an evaluator linked to a trial and implementers whose jobs may depend on positive outcomes (Kuper, Lingard and Levinson, 2008). In government funded research, the interviewer may be seen as more oriented towards the needs of the paying client than to patients and implementers, impacting the openness of communications with the researcher. Indeed, as described in Chapter 2, it may be challenging to avoid being influenced by the desire for palatable results among funders with substantial interest in findings. Linkage of the interviewer to the trial also perhaps made it less likely that implementers would be critical of their experiences of the trial itself.

However, where data begin to indicate that things have begun to go wrong in implementation, the evaluator is placed in a position of attempting to explain shortcomings, whilst avoiding unbalanced apportionment of blame or findings which allow interpretations which apportion blame unfairly. Non-delivered components could easily be blamed on exercise professionals not doing their jobs properly, and the author became keenly aware of a responsibility not to provide such data without attempting to understand how and why discrepancies had emerged, with discussions and analyses increasingly focusing on informing improved delivery of poorly delivered components and representing multiple perspectives on how shortcomings had emerged.

In an effort to minimise power dynamics and allow open and frank reflection, the interviewer attempted to position himself as one step removed from the trial and to emphasise his role as a mediator between key decision-makers and those

implementing the programme on the ground. This encouraged critical reflection on positive and negative experiences of the programme and its evaluation which could be anonymously fed back to decision makers in order to provide recommendations for improving delivery. The interviewer used open questions and allowed professionals to direct discussion onto other topics as they saw fit, listening reflectively and using empathy statements to put practitioners at ease when expressing criticisms, whilst being careful to avoid explicit agreement or disagreement. It was however crucial to avoid misleading participants into believing that the interviewer was entirely separate from the trial and to avoid overstating the influence that feedback on concerns would have on practice, through highlighting that issues could be fed back, but that this did not guarantee action. Whilst the role of the interviewer was clearly not passive, transcripts indicated that the vast majority of the talking was usually by implementers rather than the interviewer, with open questions and reflective listening eliciting in-depth data from participants' perspectives.

5.3.2.7 Analyses

A thematic approach to analysis was adopted for all qualitative data (Braun and Clarke, 2006). Whilst as described in Chapter 4, many previous process evaluations do not clearly identify their approach to analysis, most implicitly use thematic analysis, often describing organising data into themes with little description of how this was achieved. Braun and Clarke argue that thematic analysis is a commonly used, epistemologically flexible, but often poorly demarcated approach to qualitative analysis suitable for exploratory qualitative research, describing a 6-step process:

i) Familiarizing yourself with your data (transcribing data, reading and re-reading the data, and noting down initial ideas),

ii) Generating initial codes (coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code).

iii) Searching for themes (collating codes into potential themes, gathering all data relevant to each potential theme)

iv) Reviewing themes (Checking if the themes work in relation to the coded extracts (Level 1) and the entire data set (Level 2), generating a thematic 'map' of the analysis.)

v) Defining and naming themes (Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.

vi) Producing the report: (Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis).

Analysis was iterative and took place alongside data collection (Coffey and Atkinson 1996), beginning immediately after the interview, with the interviewer making notes on emerging issues and themes before data were transcribed. The approach to analysis was neither fully inductive nor deductive, but rotated between inductivism and deductivism (Gilbert, 2006). Given the volume of qualitative data, only about a quarter of interviews were transcribed by the author. Additional transcripts were received one-by-one as soon as they were transcribed and were read thoroughly on receipt whilst listening to the interview recording, and transcription errors corrected. During this process, initial notes were made of potential emerging themes. After checking, the transcript was coded using QSR Nvivo. Subsequent transcripts were coded using the codes already generated, with additional codes applied where appropriate. Once all transcripts were coded, these were combined into sets or themes of related codes. Themes were then reviewed firstly by reading through all data included under each theme and then comparing this against the entire dataset, before a thematic framework was arrived at and final themes and sub-themes identified and described, with quotations selected to illustrate descriptions.

At this stage, the author returned to the relevant literature in order to locate findings in relation to existing theory and empirical findings, and thematic frameworks were modified to take into account new aspects of the data drawn out following consideration of the existing literature. Drafts of qualitative analyses were sent to project supervisors, who commented on the thematic framework and suggested modifications where appropriate. A further researcher who had assisted with data collection for patient interviews was sent the uncoded data file and asked to comment on analyses. This was not possible for implementer interviews.

5.3.2.8 Ethical considerations

During patient interviews, patients were each given a consent form and information sheet and invited to ask any questions prior to providing consent. Verbal consent was sought prior to switching on the recording device for the interview. All interviews were downloaded onto a password-protected university network folder and deleted from the digital recorder at the earliest opportunity. All interviews were anonymised during transcription, with no interviewee identified at any stage. Interviewees were informed that every effort would be made to ensure that their data remained anonymous during reporting. In the cases of the MI training provider and policy representatives, this was not possible. Hence, these informants were assured that nothing would be reported without first seeking permission. The section of the thesis in which the training provider was cited was sent to the training provider for approval prior to being viewed by anyone outside of the evaluation team, whilst policy representatives saw and approved all analyses, including their data, prior to submission.

5.3.3 Quantitative study components

5.3.3.1 Aims

Quantitative components aimed to measure implementation in terms of fidelity and dose delivered, as well as the delivery of non-prescribed elements, the extent of programme reach and baseline patient characteristics.

5.3.3.2 Selection of methods

Targets for assessment were guided by the programme logic model which will be described in detail in Chapter 6. Key components forming the basis of assessments are presented below in Figure 3. As described in Chapter 4, previous process evaluations have typically quantified implementation through structured observation, self reports or secondary data sources such as routine monitoring data. All three approaches are used in this thesis for a range of purposes. Justifications and uses of each method will now be discussed.

Figure 3 Targets for quantitative assessment of NERS implementation

Targets for fidelity assessment

- Motivational interviewing, goal setting and health check, to be delivered in first consultations
- The conduct of subsequent consultations at 4 and 16 weeks, during which goal progression will be reviewed. Protocols stipulated that 4 week contact was to be made regardless of whether the patient was still attending the scheme, in order to encourage patients back into the scheme.
- The application of a £1 discounted rate of exercise for 16 weeks
- Delivery of supervised exercise classes, mostly following a group based structure but with opportunity for one to one classes where appropriate.
- Patients to be limited to professional only classes for the first 4 weeks.
- Contact with patients after leaving the scheme, at 8 and 12 months from scheme entry.

Targets for assessment of dose

- Length of consultations between patients and professionals
- Number of classes per week available for patients to attend at each centre
- Programme duration
- Duration of discount

Targets for assessment of reach

- Percentage uptake and adherence
- Social profiling

Non-prescribed programme components

- Activities offered during the exercise programme
- Exit route activities

5.3.3.2.1 Secondary data sources

The process evaluation used two data sources collected for another purpose. These were routine monitoring data collected as part of the programme, and baseline questionnaires collected for the trial. These data were collated by the trial manager, and fully anonymised before being received for analyses. Uses of these data sources will now be discussed, before moving on to discuss the development of primary data sources to measure aspects of implementation for which data were not incorporated into monitoring structures.

5.3.3.2.1.1 Routine monitoring data

5.3.3.2.1.1.1 Rationale for using routine monitoring data

The key secondary data source used for this thesis was a database developed between the trial manager and policy representatives, completed by professionals after each consultation. The quality of this data is clearly contingent upon honest and full completion by implementers. However, this is also true of self-reports (Young et al., 2008), whilst Hawthorne effects will in some instances lead to similar biases in direct observation (Campbell et al., 1995). The fact that this data forms an integral part of the programme helped to minimise additional research burden on implementers and to reduce concerns regarding Hawthorne effects (Campbell et al., 1995) where participants act in an unnatural manner because they are being monitored by a process which will be removed after the study (Audrey et al., 2006b). As the database was integrated into practice to facilitate long-term monitoring, were practice to be improved by awareness of being monitored, such effects would continue after the trial. The use of data collected for another purpose within process evaluation also reduces resource requirements, allowing a broad overview of the entire programme duration at little extra cost. In addition, this allows process evaluation to provide examples of how the data might be used to assess programme quality, informing longer term programme monitoring and evaluation, ensuring standards of delivery are maintained after completion of the trial.

5.3.3.2.1.1.2 Variables from routine monitoring data

A summary of the uses of routine monitoring data is presented below in Table 6. Professionals were expected to record dates that consultations were held, to record a range of quantitative health check and lifestyle data and to record the content of consultations, in terms of topics discussed, measures taken and goals set. Binary variables for the delivery of health checks (i.e. whether or not a health check was delivered on scheme entry and exit) and the delivery of goal setting (i.e. whether or not goals were recorded in the database) within consultations were created using this database. Furthermore, professionals had been advised that goals were to be specific, measurable, achievable, relevant and time-bounded. Although assessments of relevance and achievability were not possible from written records, goals were coded dichotomously according to whether or not they contained a measurable, time-bound element (e.g. 'to lose 4 pounds in 4 weeks' or 'to be able to walk up the hill next to my

house without feeling breathless in 4 weeks' as opposed to 'lose weight' or 'get fitter'). A 1 was awarded for each case where a goal contained both a measurable element and an endpoint at which goal achievement could be assessed and new goals set, and a 0 for all other cases.

Table 6. Uses of routine monitoring data for implementation assessments in NERS

Variable	Coding	Use
Health check delivery	At least one item of health check data – 1 No health check data present – 0	Fidelity
Goal setting delivery	Goals recorded in database – 1 No goal recorded in database – 0	Fidelity
Goal setting quality	Record contains measurable, time-bound element – 1 Record contains no measurable, time-bound element – 0	Fidelity
Attendance at entry / exit consultations	Health check or lifestyle data completed – 1 No health check or lifestyle data present – 0	Reach
Contact at /adherence to 4 weeks	Hand coded string variable - contacted – 1, not-contacted – 0 - still attending NERS – 1, left NERS - 0	Reach
Contact at 8 months	Hand coded string variable - attendance – 1, non-attendance – 0	Fidelity
Contact at 12 months	Health check or lifestyle data completed – 1 No health check or lifestyle data present – 0	Fidelity
Programme duration	Number of days between first and last consultation	Dose delivered
Referral reason	Binary variables – - Weight management - CHD risk factors (non-weight related) - Mental health	Social patterning in reach

Protocols also stipulated that 4 week contact should be made regardless of whether the patient was still attending in order to encourage leavers back to the programme, and that patients were to be contacted at 8 and 12 months to discuss progress since leaving the scheme, in order to maintain dialogue with patients about their activity outside of exercise classes. Hence examining percentages of i) non-attendees contacted at 4 weeks, ii) completers contacted at 8 months and iii) completers contacted at 12 months offered a useful indication of fidelity to patient reviewing protocols. In relation to dose, calculating time expired between scheme entry and exit allowed for programme duration to be quantified. In relation to reach, dates of first, mid-programme and end of programme consultations also gave an indication of the percentages of referred patients who entered the scheme, and percentages who subsequently dropped out

between 0-4 weeks, 4-16 weeks or completed the scheme. Attendance figures also provided the denominator for fidelity checks of the percentage of delivered consultations including each intended component.

Consistency checks however revealed that area coordinators were completing attendance date fields in at least 3 different ways, with the presence of a date sometimes only indicating that an appointment was booked, not that the patient attended. Hence, patients were considered to have attended entry, exit and 12-month consultations where data were entered from either health checks or lifestyle questionnaires scheduled for consultations. For 4 and 8 month contact, details entered into open string variable fields were coded by hand to ascertain whether the patient was or was not contacted or still attending. The final use of routine monitoring data was to identify reasons for referral for analyses of social patterning in reach.

5.3.3.2.1.2 Baseline trial data

An additional source of secondary quantitative data was baseline questionnaire data collected for the purposes of the trial, used in this thesis for the purposes of examining social patterning in uptake and adherence by baseline characteristics. Use of this data allowed quantification of patterning using measures which would subsequently also be available to assess patterning in trial outcomes.

Prior to randomisation, patients completed a short postal questionnaire including age in years, postcode, sex and car ownership (yes or no). Postcodes were linked to total deprivation scores from the Welsh Index of Multiple Deprivation. Patients also completed the General Practice Physical Activity Questionnaire *(GPPAQ)* (NICE, 2006b). The GPPAQ asks patients how many hours per week they participate in activities from 2 categories 'physical exercise, such as swimming, jogging, aerobics, football, tennis, gym, workout etc' or 'cycling, including cycling to work or during leisure time', with response options of none, less than 1 hour, 1 to 3 hours, and 3 hours plus. A patient activity index is assigned to indicate whether the respondent is inactive (no structured activity), moderately inactive (less than 1 hours structured activity per week), moderately active (1 to 3 hours structured activity per week from one category, or 0 to 1 hours from both categories), or active (3 or more hours of structured activity per week from one category, or 1 to 3 hours per week from both).

5.3.3.2.2 Variables not captured within secondary data sources

The routine monitoring database was constructed prior to commencement of this thesis, and the author had no input into its design. Hence additional data components were required where the database did not capture activities seen within the NERS model as central to outcomes. These were:

i) Delivery of motivational interviewing
ii) Duration of consultations
iii) Application of a £1 discounted rate of exercise for 16 weeks
iv) Delivery of exercise classes and exit routes.

In addition, discussions with implementers indicated that many extended programme dose through allowing access to NERS classes after official exit from the scheme, and that most areas had introduced discounts to replace the scheme discount at 16 weeks. Systematically capturing details on local variation in extended programme dose was crucial to understanding how much intervention patients were offered. These aspects of implementation were assessed via a combination of practitioner self report and observation of tape recorded practice samples.

5.3.3.2.2.1 Structured observations –motivational interviewing and consultation length

5.3.3.2.2.1.1 Rationale for using structured observations

Given the absence of structures to assess motivational interviewing fidelity, the relative merits of self-report or direct observation were considered. Direct observation involves generation of data through immersion in the situation under investigation (Spradley, 1980) and is often preferred over self-report as it helps avoid discrepancy between what implementers say they do, and what implementers actually do (Young et al., 2008). The usefulness of direct observation is however limited to situations in which delivery can be observed unobtrusively, with Hawthorne effects potentially leading to similar biases. In NERS, motivational interviewing was delivered in private one-to-one consultations between patient and professional. Hence, having a third person present would likely damage rapport, leading to the consultation becoming a stilted version of a real life consultation.

However, during training programmes, near-zero correlations have been reported between practitioner reports of motivational interviewing proficiency and tape recorded samples rated by experts (Miller and Mount, 2001; Miller, 2004). Asking

implementers to self-assess their own practice would likely lead to excessively positive appraisal (Audrey et al., 2006b), whilst ability to critically reflect on one's delivery of a complex skill will be impacted by one's understanding of that skill. A skilled practitioner may be more aware of their limitations and see themselves as having room for improvement (Miller, 2001), whilst less skilled implementers may view the approach as simple, having not understood it (Miller and Rollnick, 2009). Hence, as recommended by Miller (2001) implementers were asked to record consultations, to be rated for fidelity using a validated rating scale. Whilst not immune to Hawthorne effects, a small digital recorder was considered less intrusive and likely to have less impact on delivery than a researchers' physical presence.

The likelihood that recordings would reveal shortcomings became apparent during early readings of training manuals, which revealed that professionals had been offered only one hour of combined training in MI and goal setting. Indeed, the absence of structures to monitor MI delivery within the programme for quality control purposes was in itself seen as potentially problematic. Hence, a need to build in structures to monitor MI delivery was one of the first priorities emphasised by the author as the process evaluation was being designed. As this was being negotiated and written into a resubmission to the ethics committee, policy representatives implemented additional training courses in motivational interviewing. The first was delivered in August 2008 with only a few attendees. At this point, ethical approval was still pending. Further courses were planned for October 2008, January 2009 and April 2009. Ethical approval was received shortly before the October course, and hence recording equipment was distributed to areas whose professionals were due to attend this course, in order to obtain pre-training recordings, prior to recordings being collected from areas attending later courses. A planned assessment of current practice became a longitudinal evaluation of fidelity before and after training, embedded within a qualitative exploration of attempts to deliver MI since training, drawing upon qualitative interviews with professionals, coordinators and the MI training provider described above, with those attending the largest course in October asked to provide further recordings 6 months later.

5.3.3.2.2.1.2 Variables from structured observations

Recordings were coded at both time-points for fidelity to MI using the Behaviour Change Counselling Index (BECCI; Lane et al. 2005). Whilst the Motivational Interviewing Treatment Integrity scale is considered the gold-standard measure of fidelity (Moyers et al., 2005; Pierson et al., 2007), this requires substantial training to be used reliably, whereas training guidelines recommended within the manual for the Behaviour Change Counselling Index were more achievable given the resources available for the study, involving reading several chapters of key texts and watching a demonstration video. Coding was conducted independently by two psychology graduates, one of whom was the author, who had both attended a two-day training course in MI. The author contacted the lead developer of the scale, who provided advice on its use and supplied coder training materials (Lane, 2005). The measure demonstrated excellent inter-rater reliability (r=0.90), with mean scores of coders differing by less than 2%. In addition to providing a measure of fidelity to MI, recordings also allowed for consultation durations to be quantified.

5.3.3.2.2.2 Structured interviews – the NERS exercise programme and the £1 discount

5.3.3.2.2.2.1 Rationale for using structured interviews

For aspects of delivery which can be unambiguously described, do not involve self-evaluation of a skill, cannot be observed, or for which Hawthorne effects may be a greater threat to validity than social desirability bias, self-reports may offer a useful means of gaining an overview of how a component is delivered. Two main options are structured questionnaires or interviews. Structured interviews are similar to structured questionnaires, in that they involve closed questions, defined prior to the conduct of the interview (Berg, 2004). However, with a questionnaire, the participant reads the questions and records answers themselves, whereas during a structured interview, the interviewer reads questions and notes the interviewee's verbal response. Comparisons between data obtained through self-completion questionnaires or structured interviews have typically been limited to studies examining the usefulness of questionnaire data for diagnosis of psychiatric disorders, as a cheaper alternative to structured interviews, assuming structured interviews to be the more valid option against which self report measures are validated (Hyler et al., 1990; Fairburn and Beglin, 1994).

The key rationale for using structured interviews was that this removed the need for questionnaires to be posted to implementers and back to the evaluation team, reducing burden for evaluators and implementers, and potentially increasing response rates. Implementers were all invited to participate in semi-structured interviews as described above, and hence inclusion of a brief structured section meant collection of this data at little extra time cost. By contrast to routine monitoring data, which cover the entire scheme duration, these data are however limited by their cross-sectional nature, representing delivery at one point in time.

5.3.3.2.2.2.2 Variables from structured interviews

Decisions on how the £1 discount was applied and scheme exit routes were made at the local authority level. Coordinators were therefore asked to state how much patients were charged for group exercise classes, whether there were any exceptions (i.e. patients charged more or less), what services were available at this rate (e.g. supervised classes, supervised gym use, independent centre use or independent pool use) and for how long the discount was offered. In addition, coordinators were asked whether patients were limited to supervised classes for the first 4 weeks and to list the types of exit route options offered in their area. Coordinators were also asked whether post-scheme discounts had been introduced.

Structured interview schedules for exercise professionals aimed to describe the exercise programme, in terms of the number of classes offered in each centre, the types of classes offered and the range of available times. Professionals were therefore asked firstly to name all the leisure centres where they delivered the scheme, how many group-based exercise classes they themselves offered in each centre (including group based gym sessions), the types of exercise classes offered in each centre and whether these classes were exclusive to NERS patients. For centres served by multiple professionals, responses were summed to give a total number of classes per centre. Professionals were asked whether they ran classes alone or jointly with another professional, and in centres where classes were delivered by two professionals, the total was divided by two. Professionals were also asked to indicate the number of one-to-one sessions they offered per week, other than consultations or inductions, and whether weekend or evening classes were available at each centre. Professionals were

also asked whether they allowed patients to continue accessing their classes after leaving the scheme as well as whether discounts were offered.

5.3.3.3 Accessing participants

5.3.3.3.1 Structured interviews

Structured interviews were conducted alongside qualitative interviews, for which recruitment procedures are described in Section 5.2.

5.3.3.3.2 Consultation recordings

Coordinators acted as gatekeepers for consultation recordings. All 12 area coordinators agreed to distribute recording equipment to exercise professionals to record consultations for assessment of MI fidelity, though one responded too late to conduct recordings prior to additional MI training. Of the 35 professionals approached within the remaining 11 areas, 31 provided recordings. However, 4 had been unable to record any consultations until after receiving the 2-day course and were excluded. A further 4 provided unusable recordings (eg. inaudible recordings or incomplete paperwork). Hence, usable recordings were provided by 23 professionals within 10 areas. Three professionals from one area reported having been unable to obtain consent from patients and 1 reported having no initial consultations during the time equipment was available. Of these 23 professionals, 14 attended the October training course. Data collection procedures were repeated 6 months later for these professionals, at which point, 2 had left their post, whilst another had been unable to gain patient consent to record, with 11 providing follow up data.

5.3.3.4 Data collection procedures

5.3.3.4.1 Structured interviews

Interview procedures are described above in relation to qualitative interviews. Implementers were asked to provide yes/no or numeric responses to a small number of closed questions prior to conduct of semi-structured interviews.

5.3.3.4.2 Consultation recordings

For consultation recordings, the researcher arranged an appointment with the area coordinator to provide equipment and paperwork and discuss recording procedures.

Coordinators were asked to distribute equipment to professionals and to request that professionals record their next two to three consultations, after obtaining signed consent from patients. The researcher arranged a second meeting with the coordinator to collect equipment when recordings had been completed.

5.3.3.5 Analyses

Summary statistics were calculated using SPSS version 16 and Stata version 11. As will be described in Chapters 7 and 8, secondary analyses involved variability in adherence according to goal setting quality and socio-demographic profiling of scheme uptake and adherence. These involved regression analyses adjusted, conducted using STATA version 11. More detailed descriptions of specific analyses are provided in the relevant findings chapters.

5.3.3.6 Ethical considerations

Probably the most sensitive patient-level data were the recorded consultations provided by exercise professionals. Exercise coordinators were briefed prior to conduct of these recordings on the need to obtain patients written consent and to allow the patient to read the information sheet and ask questions, as well as to ensure that recordings were stored in a secure location where they could not be accessed by a third party until the researcher returned to collect recordings. In a small number of cases, recordings were provided without accompanying consent forms, in which cases these were destroyed immediately. Once entered onto a secure drive of the university network and password protected using encryption software, files were removed from the recorder. The routine monitoring database was fully anonymised by the trial manager through removal of patient names and contact details prior to being received by the author.

5.3.4 Ethical approval

Plans for a smaller scale process evaluation were written into the application to the NHS ethics committee for approval to conduct the trial in 2006, and ethical approval was received prior to the commencement of this thesis. However, at the commencement of this thesis, planned activities were significantly altered and extended. This coincided with changes to the main trial and was therefore written into a notification of major changes to the ethics committee, which received approval in

August 2008. Data collections for elements not covered by the original ethics application were delayed until after this time.

5.4 Guide to empirical chapters

5.4.1 Chapter 6: Programme theory, diffusion and implementation

Empirical chapters begin in Chapter 6, which presents the NERS theoretical model developed through discussions with policy representatives, before presenting analyses of qualitative interviews with policy representatives and coordinators, exploring how this intended model was diffused into local practice. Whilst data collections were not guided by an explicit theoretical framework, initial thematic data analyses indicated substantial overlap with key concepts from Diffusion of Innovations theory (Rogers 2003; see Chapter 3). This analysis offers insights into how divergences in programme delivery emerged, and is followed by data which quantifies the consistency of implementation with the NERS logic model, drawing upon structured observation, interviews and routine monitoring data. Whilst presented sequentially, data collections and analyses overlapped, with timing of implementation checks driven by pragmatic considerations, whilst interviews regarding programme diffusion were conducted a year after implementation began, in order to allow implementers to reflect on the scheme having been implementing for a year.

5.4.2 Chapter 7: Formative aspects of the NERS process evaluation - implementation of motivational interviewing and goal setting

The focus of Chapter 7 is upon informing improvement of two key programme components whose delivery was not optimal during the trial period. This illustrates the formative as well as summative nature of process evaluation in the context of policy trials. Feedback was provided to policy representatives partway through the trial on weaknesses in delivery of MI and goal setting, leading to action to correct weaknesses. The first half of the chapter explores mid-trial efforts to diffuse MI into practice. Professionals' responses to the 2-day training course in motivational interviewing and impacts on fidelity of practice are explored using semi-structured interviews with the MI training provider, area coordinators and exercise professionals combined, followed

by quantitative fidelity checks before and 6-months after the training course. Qualitative data are then directly linked to individual level quantitative change over time in order to provide an understanding of how and for whom changes occurred (Creswell and Clark, 2007). The second half of Chapter 7 explores variation in adherence by goal setting quality. This was a second component whose delivery was compromised. Coordinators were made aware of the limited quality of goal setting via policy representatives, though no further training was provided. This section focuses both on whether goal setting quality improved over time, and whether higher quality goal setting processes and goal types were associated with higher levels of scheme adherence.

5.4.3 Chapter 8: Patient experiences and patterning in programme reach

Chapter 8 presents quantitative and qualitative data, with methods combined essentially to explore different aspects of the same research question. The chapter begins with qualitative data from exercise professionals and patients, examining views on how NERS facilitates adherence, generating hypotheses regarding processes of change and social patterning which are then tested through quantitative profiling of programme reach. Analyses for qualitative study components were conducted concurrently, although analyses of exercise professional interviews were completed first, allowing participant interviews to be framed explicitly as building on findings from exercise professional interviews. The research design adopted by Chapter 8 is largely consistent with an exploratory design, whereby qualitative analyses are followed by quantitative analyses for the purposes of offering insights into the emergence of patterning prior to quantitatively testing emerging hypotheses (Creswell and Clark, 2007).

6 Programme theory, diffusion and implementation

6.1 Chapter aims

This chapter addresses the following two research questions:

- How is a national policy for exercise referral diffused into local practice?
- How consistent is the delivered intervention with programme theory?

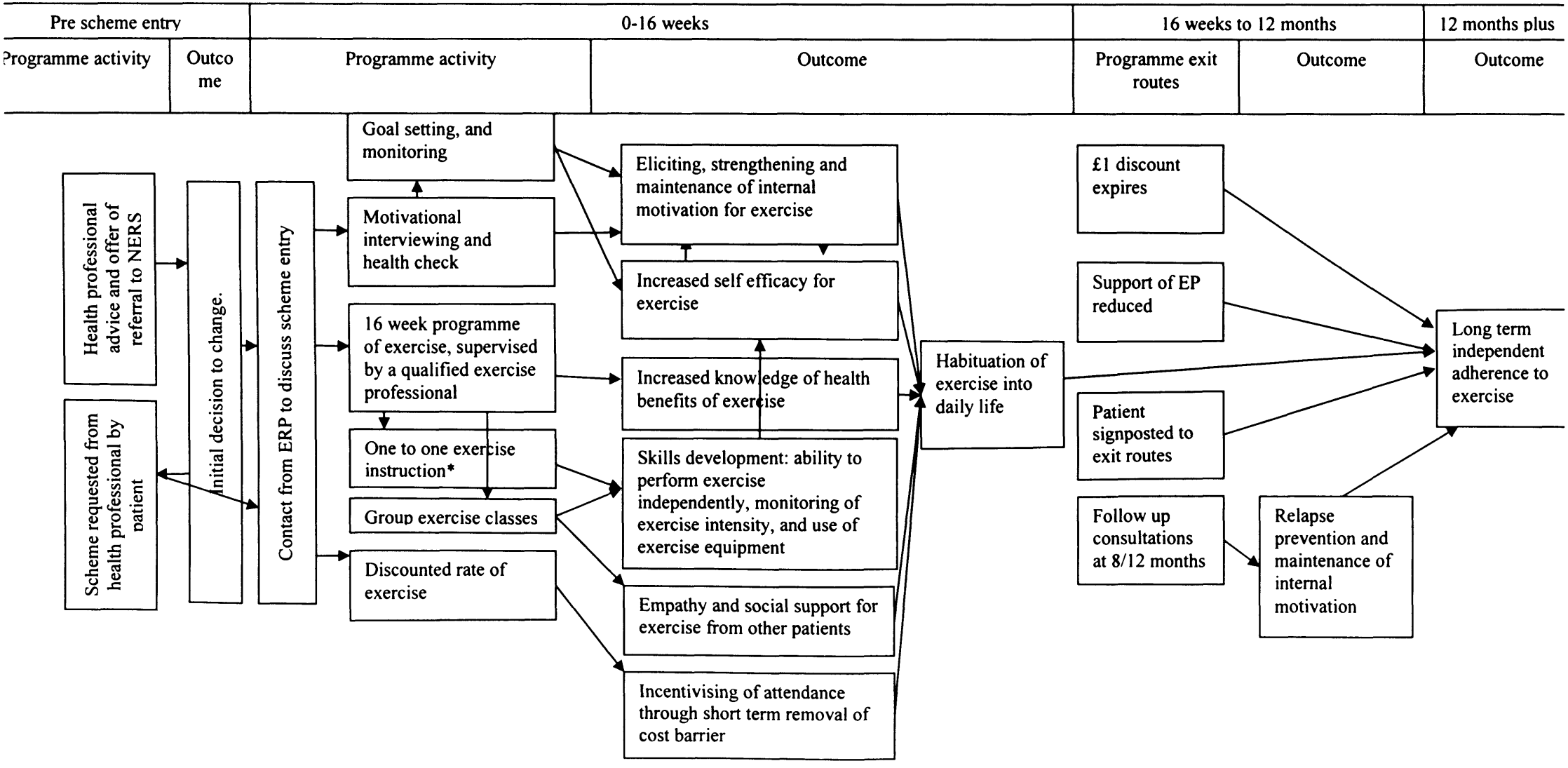
The chapter begins in Section 6.2 by presenting the NERS theoretical model, derived from discussions with policy representatives. Section 6.3 then explores perceived influences on the rate of adoption of NERS into local practice and barriers and facilitators to implementing the NERS model at the local level. Section 6.4 then presents quantitative implementation checks, describing the delivery of key components. Discussion of implications of implementation for programme functioning is guided by the NERS model.

6.2 Programme theory: the NERS theoretical model

A model of planned activities and hypothesised mechanisms of change is displayed below in Figure 4. The patient's decision to change is seen as either preceding a decision to seek referral or as triggered by health professional advice. After agreeing to be referred, the patient is contacted by an exercise professional and invited to a consultation. The professional conducts a health check in order to ensure that exercise is not contraindicated before inviting the patient to discuss behavioural change using motivational interviewing principles, eliciting change talk and guiding the patient towards linking change to their deeply held values, enhancing internal motivation. The change talk elicited during this process is harnessed into personally relevant, measurable and time-bound goals negotiated with the patient, further enhancing motivation through focusing patients' attention on desired changes. The patient then enters a 16-week programme of predominantly group based exercise, intended to provide an understanding of the health benefits of exercise and the skills to exercise safely and independently given the limits placed on them by their conditions. Group based classes are intended to provide an empathic environment and to foster social support for exercise. Patients are to be encouraged to attend 2 exercise sessions per week. Patients' goals are reviewed throughout the programme at approximately 4

weeks, in order to enhance self efficacy for exercise and maintain motivation. Protocols state that non-attendees are to be contacted at this stage to encourage them back to the scheme. In addition, a £1 per session discounted rate of exercise is intended to act as an incentive for attendance. It is anticipated that after 16 weeks, physical activity will be habituated so that support can be reduced and the discount withdrawn and once signposted to exit routes during an exit consultation, the patient will be able to maintain activity levels independently. Additional telephone contact at 8 months and a final consultation at12 months aim to maintain motivation and prevent relapse.

Figure 4 A theory of change for the National Exercise Referral Scheme, developed through discussion with national policy representatives



6.3 *Diffusion of the NERS model into local practice*

6.3.1 Context and aims

The trial sampled patients from the first two phases of implementation of the above NERS model. Each area was allocated a target number of referrals to hit, based on population size, before patients were sampled, allowing the protocols to be up and running when sampling began. The first Phase 1 area was ready for sampling in September 2007, though remaining areas took between 2 weeks and 3.5 months to catch up. Seven further areas were sampled for the second phase, with the first two areas ready for sampling in January 2008. Four further areas took between 1 and 4 further months to be ready for sampling, whilst the seventh was still not ready for sampling when target numbers had been reached and randomisation terminated in October 2008. As indicated in Table 7, population adjusted rates of referral to the trial varied between areas from 0.6 to 1.9 patients per 10,000 of the county population per annum in phase 1 (pro rata from September 15th 2007 to the date of the last referral) and from 0.9 to 2.4 in phase 2 (per annum pro rata from January 1st 2008 to the date of the last referral). In most areas, NERS was offered in more than 75% of local authority run leisure centres. One area however achieved relatively high referral rates, though concentrated resources on only 2 out of 9 centres. The number of classes per centre ranged from 1.8 to 7.5, and was highest in areas focusing resources on fewer centres. Scheme entry and completion rates varied within and between phases. The aims of this section are to explore how NERS was diffused into local practice, offering insights into how patterning in adoption rates emerged, perceived challenges maintaining consistency with programme theory, and how the scheme became routinised into local practice.

Table 7. Area level characteristics and levels of diffusion during trial period

Phase	WIMD median rank*	Months between sampling beginning for phase and for area	Referral rate (per 10,000 of county population)**	Number and percentage of leisure centres offering NERS***	Classes per centre per week***	Scheme entrants	Scheme completion
1	3	3	1.4	7 (100)	3.4	83.4	61.5
	6	1.5	1.6	2 (22.2)	7.3	80.8	46.5
	7	0.5	0.6	6 (100)	2.7	84.1	39.1
	9	0	1.9	6 (75.0)	6.2	85.0	10.7
	11	0.5	0.9	4 (66.7)	7.5	79.4	44.1
	12	3.5	0.9	4 (100)	1.8	91.4	42.9
	Phase one averages	1.5	1.2	4.8 (77.3)	4.8	84.0	40.9
2	1	1	1.7	4 (100)	5.0	92.9	54.1
	2	2	1.8	4 (100)	4.0	90.5	50.8
	4	0	2.3	5 (83.3)	2.5	88.6	54.3
	5	4	0.9	6 (60)	3.2	81.0	40.5
	8	1.5	2.4	7 (100)	2.3	84.8	29.4
	10	0	1.9	2 (50)	7.0	85.7	40.0
	Phase two averages	1.4	1.8	4.7 (82.2)	4.0	87.3	44.9

* Ranks based on median 2008 Welsh Index of Multiple Deprivation scores for districts within each county (Welsh Assembly Government, 2008)

** Number of patients referred to NERS trial per 12 months pro rata, divided by estimated population size in mid-2007 according to figures published by the Welsh Assembly Government (Welsh Assembly Government, 2009).

*** Based on reports provided by exercise professionals in October 2008

6.3.2 Data sources and analysis

Analyses draw upon qualitative interviews with 3 policy representatives and 12 local area coordinators. Findings are described under the following thematic headings:

i) Perceptions of the innovation before and after implementation: relative advantage, compatibility and complexity

ii) Communications between change agents and adopters.
   a. Coordinator as adopter: communications with policy representatives
   b. Coordinator as change agent: communications with health professionals
   c. Coordinator as change agent: communications with leisure centres

iii) Supporting movement from initiation to implementation: training and management support.

iv) Towards routinisation: restructuring the local context and reinventing the innovation.
   a. Changing local leisure systems to achieve fit with NERS
   b. Reinventing the innovation
   c. Sustainability of the referral criteria

Where discussed by both coordinators and policy representatives, findings begin with views of policy representatives, followed by views of coordinators. Qualitative responses are linked throughout to data on diffusion in order to generate hypotheses for how variation emerged. Fewer verbatim quotes are used than in subsequent qualitative phases due to greater difficulties in maintaining anonymity.

6.3.3 Findings

6.3.3.1 Perceptions of the innovation before and after implementation: relative advantage, compatibility and complexity

In reflecting on coordinators' perceptions of the new scheme, policy representatives focused primarily on issues experienced during the move to national standardisation, commenting that emotional attachments to previous schemes had limited coordinators willingness to implement NERS and that even though they had involved everyone who was involved in delivering a scheme in the development of the protocols, this had not

fully pacified these tendencies. Resistance to the scheme was seen as heightened by randomised roll out for the purposes of a trial, with objections to the trial attributed to a limited culture of evaluation in leisure services.

Whilst policy representatives focused largely on the transition coordinators reflected on NERS having been implementing it for approximately a year. NERS was seen by some as having involved small changes in practice, and as such, achieving compatibility relatively quickly, with one coordinator commenting that 'it's been fairly easy to implement the scheme. Obviously we had one initially, so it was just a question of tweaking it'. By contrast to reports of policy representatives, comparisons to previous schemes overwhelmingly favoured NERS. Some focused upon advantages such as increased patient monitoring and group-based structures, and their perceived benefits for patient retention. Others commented that national implementation had led to more cohesive links with other areas, with greater ability to share problems and solutions arising from adoption of common protocols. In two areas, coordinators commented that previous schemes had relied upon mainstream staff being released for a few hours a week to deal with referrals, with one commenting that 'if a facility manager had a member of staff go sick then exercise referral would suffer because they would juggle their staff in and it would be least priority.' Hence, having a team of professionals funded specifically to deliver exercise referral was welcomed.

In many areas, the move to national standardisation was described as raising standards of delivery, with some offering substantial critical reflection on previous practice. One coordinator for example commented that in the scheme she had set up previously 'there were no standards it was people just doing really exactly really what they wanted to'. In one area achieving above average referral and adherence, the coordinator expressed frustration at the nature of the previous scheme as 'a token gesture', appearing to value provision of structures to improve status of exercise referral locally and to 'get everyone in a positive frame of mind, in terms of exercise referral'.

One of the highest performing areas was however the area where no previous scheme existed. This area's coordinator commented that lack of investment in a previous service simplified implementation, with ERS seen as a valuable new innovation and as

'something that everybody was waiting for'. Despite taking longer to be ready for sampling than most other areas, this area overtook many earlier starters, achieving an average referral volume despite this late start, integrating the scheme into all centres and achieving among the highest uptake and completion rates. In addition, in one area achieving below average referral rates and the lowest uptake level, acceptance of NERS still appeared limited by local views of the previous scheme as better, with the coordinator describing a well-developed scheme as having been 'pulled right back so that the new ones that were coming up in other local authorities, we were all starting from the same line'. This was the only area where NERS was still described as inferior in many respects to previous practice.

6.3.3.2 Communications between change agents and adopters

6.3.3.2.1 Coordinator as adopter: communications with policy representatives

During planning and implementation, regular meetings were held between area coordinators and national policy representatives. However, despite these efforts to open up communication channels, policy representatives described difficulties communicating with coordinators. Whilst seen as arising in part from the initial lack of a dedicated national coordinator to follow up issues between meetings, these were seen as arising largely from reluctance of local coordinators to communicate with civil servants. One spoke of tensions arising due to the primary obligation of civil servants to support the minister responsible for the scheme, and limited time available to support coordinators. Communications were however described as improving following appointment of a national coordinator, with meetings now led by an individual whose experience of ERS delivery ensured trust and respect. The new national coordinator was described as being seen in a different light to how the civil servants initially coordinating NERS were seen by local coordinators, with coordinators communicating more openly as a result. One representative commented that it had made a huge difference having a national coordinator with experience of delivery on the ground and respect as a result of that experience.

Some local coordinators commented that initial lack of communication had contributed to divergent initial interpretations of protocols, with one commenting 'every single authority was doing something different, nobody knew what they were

doing because there was not a national coordinator in place...it was very much do it how you can you know, do whatever you can'. Whilst many spoke positively of communications with the civil servant initially coordinating the scheme, coordinators were unanimous in support for creation of a full time national coordinator role, seen as increasing the volume and promptness of communications. The appointment was seen as enhancing cohesion within the network, facilitating identification of solutions for common problems, leading to smoother implementation across areas and increasing empathy with challenges in implementation. One coordinator for example described the value of 'just having somebody who's actually been an instructor, who's worked on the coal face, who understands the problems. It just felt before that nobody understood exactly what it was like.'

6.3.3.2.2 Coordinator as change agent: communications with health professionals

Policy representatives commented that coordinators' limited initial perceptions of the need for change, had knock-on effects for communications with health professionals in the early stages of the scheme. In many areas, the scheme took substantially longer than anticipated to be adopted by health professionals. Policy representatives commented that whilst timely implementation had been hampered by administrative shortcomings in collating paperwork, some coordinators' reluctant acceptance of the role of change agent had resulted in the coming changeover not being communicated to referral partners until coordinators had the paperwork in their hands. In addition, policy representatives cited the paper-based referral pad as a key barrier to facilitating adoption by some referral partners, with one representative commenting that GPs were reluctant to fill the form in and often didn't do it particularly accurately. However, communication of these concerns via coordinators had led to efforts to find alternatives such as computerised referral. In addition, policy representatives described receiving a number of queries through the coordinators saying that their GPs wanted to get paid for referral, though took a decision not to financially incentivise adoption.

Policy representatives also described objections to the evaluation amongst referral partners, though one attributed these to the manner in which the evaluation was communicated by coordinators, with coordinators perceived as not always convinced of its justifications, and hence being tasked with communicating a process that they

did not fully support. Support for communicating the trial to health professionals was offered by policy representatives and the evaluation team, with some coordinators describing these as beneficial, whilst others declined these offers, commenting that health professionals had made up their mind.

Alternative explanations for delays in start-up were offered by local coordinators, who focused on the challenging nature of communications with health professionals for newly qualified coordinators, or difficulties arising from periods with no coordinator in post. One experienced coordinator for example commented that 'I know a lot of people were new into the coordinators post and to sit in front of 8 or 10 GPs is quite a daunting thing.' The coordinator of the last area ready for sampling to the trial argued that being new to the role she had sought support from a contact within the LHB on how to communicate the scheme to professionals, with this contact advising against switching over until referrers had been brought on board. Whilst delaying implementation, this coordinator argued that by 'going through the correct route, when we actually started the scheme it made it much, much easier and because of that then the surgeries, as I said before that wouldn't or weren't interested in coming, actually stepped on board.' The phase 1 area taking the longest to be ready for sampling had operated without a coordinator during much of the transition, with the previous coordinator leaving as NERS launched. The new coordinator in this area commented that there 'wasn't anyone working on the scheme to actually try and get the referrals from the health professionals'. The lowest referral rate was achieved by a further area operating without a coordinator for much of the trial.

Coordinators expressed a consensual view that communicating the trial to professionals had been difficult, with health professionals objecting to randomisation and sometimes refusing to refer until randomisation was complete. Most however commented that referral partners had been receptive to the scheme itself. In many cases, relationships developed during previous schemes were seen as facilitating communication of new information to health professionals. In accessing referral partners, some spoke of identifying appropriate personnel to communicate the innovation to on a practice-by-practice basis, whilst others spoke of always going through the practice manager, or targeting practice nurses who were then asked to communicate the innovation to doctors. Concerns about litigation, or perceptions that

promoting exercise was not a priority were cited by some, though these were seen as minority views. Although some health professionals were described as perceiving that local schemes did not need to be changed, these issues were seen as having been overcome through communicating advantages of the scheme. However, the coordinator of the aforementioned area where NERS was still seen as somewhat inferior to prior practice described negativity among referral partners and difficulties communicating with health professionals, arguing that this was 'because we'd had such a successful scheme. I think if we had not had a scheme before they would have embraced the idea.'

Four coordinators, including 3 achieving above average referral rates, commented that government backing had increased the esteem of the service and receptiveness of health professionals to communications. One for example commented that 'it's a far more professional service. GPs are far more comfortable with referring into a national scheme'. Maintenance of communications was often described as an ongoing process, with some describing maintaining relationships through systematic feedback on patients' progress. One coordinator, achieving an above average referral rate, attributed this largely to systematic structures for providing feedback. In one area achieving a low referral rate, the coordinator who came into post midway through the trial spoke of recently consulting with referrers on the feedback they would like, commenting that professionals 'wanted to know whether the patient has started the scheme, and then at the end of 16 weeks what they've done' with parsimonious feedback being introduced to boost referral volumes.

6.3.3.2.3 Coordinator as change agent: communications with leisure centres

Communications with leisure centres were often seen as relatively easy, due to involvements in previous schemes or relationships developed in previous roles. Most coordinators commented that leisure centres had been supportive of NERS, and that they had experienced no difficulties communicating the service. The coordinator of the area with the highest absolute referral volume commented that acceptance of the scheme had been reinforced by the increasing income generated as referrals gathered pace, commenting that centres were 'very happy with our group sessions and our numbers, because obviously they're financially driven and it looks good on their finance report.' However, 3 coordinators, described conflicts with centres' financial

priorities, and resentment at 'giving away cheap or free sessions to people that could possibly be paying full price'. One spoke of addressing concerns through communicating with leisure service managers and reaching compromise in terms of running classes at quieter times commenting that during 'the dead times in the gym, you know you might as well have your £1s coming in'.

Some commented that centres had begun to see financial benefits of the scheme as a feeder into mainstream services, with patients viewed as potential future members, in some cases citing promotion of internal exit routes above options beyond the centre in order to achieve congruence with financial priorities. By contrast, 3 cited difficulties exiting patients into mainstream services, with staff resisting patients' access due to concerns over litigation. One of these coordinators commented that 'a lot of the staff have dug their heels and said I'm not qualified to deal with so and so, and we're saying well yes you are, you know level 2 gym qualification which means yes you're qualified to deal with our clients because they're now mainstream. So there has been a battle there.' Overall, four coordinators described encountering resistance integrating the service into leisure centres. With one exception, such concerns were expressed in areas with low referral volumes. Whilst one area achieved a completion rate almost identical to the national average, the remaining areas achieved 3 of the 4 poorest adherence rates. Two offered the scheme in all centres despite these challenges, though offered a below average number of classes per site per week. The remaining two offered a large number of classes per centre, though concentrated on fewer potential centres.

Whilst most coordinators were based within leisure service departments, in one area, NERS operated through a physical activity and sport department, with this separation causing challenges negotiating access to facilities, with the coordinator having to pay to use leisure facilities. Rather than fully supporting the scheme, the local leisure trust was described as setting up a service in competition with NERS, offering introduction to activity classes supervised by level 3 qualified instructors, with the coordinator commenting that the local trust were 'able to deliver exactly what we deliver'. This was seen as particularly problematic during recruitment to the trial, with health professionals described as bypassing NERS and advising patients to access this alternative service where they would be guaranteed entry. Indeed, whilst one of the

first areas ready for sampling to the trial, referral rates during randomisation were amongst the lowest.

6.3.3.3 Support for implementation: management and training

According to policy representatives, transitioning from previous schemes conferred the advantage that the skills base for ERS delivery was largely in place. Implementation was however described as hampered by the limited management support for coordinators, seen initially as the responsibility of local authorities, with policy representatives describing having assumed that they would get proper management support, though that had not always been the case. Filling this unanticipated void was seen as impossible within the structures available at the beginning of the scheme, with NERS coordinated nationally by civil servants who had insufficient time to oversee coordinators' practice. The failure to create a national coordinator post from the development stage was seen as a key weakness in implementation planning, with policy representatives arguing that 'what was probably lacking in the initial context was someone who had the time to be able to go out and actually spend one on one time with some of the coordinators so that we really understood what some of them were doing'. Whilst commenting that aspects of new protocols such as the £1 rate for exercise classes had been followed, representatives commented that protocols such as those stating that patients were to be contacted at 4 weeks even if they had not attended classes were often not being adopted, leading to high drop-out rates, and that these divergences had gone largely unnoticed until her appointment. Structures were being put in place to provide additional resources where these deviations from protocols were put down to insufficient time.

As well as initial shortcomings in national coordination, challenges implementing NERS at the local level were described in some areas as arising from periods without a coordinator. As described, areas operating without a coordinator appeared to struggle to obtain referrals. In addition, one area had a coordinator in post during the transition to NERS, though the post became empty during the trial. Whilst achieving the highest population-adjusted referral rate for a phase 1 area, this area achieved a completion rate 4 times below the national average, and almost 3 times lower than the 2nd lowest performing area.

For the most part, coordinators commented that professionals' training had been excellent, although training volume was seen by many as causing problems, with ability to cope linked to how quickly NERS had been adopted by referral partners in the local area. However, several commented that limited attention had been paid to coordinator training, identifying a lack of training on administrative and managerial aspects of the role, often having been promoted from instructor roles and having limited managerial experience. One for example commented that 'a lot of people were promoted from exercise professional to coordinator, where yes you have the background and knowledge, but obviously budgeting skills and you know sort of putting together a presentation. We just have to do off our own backs'.

6.3.3.4 Towards routinisation: restructuring the local context and reinventing the innovation

6.3.3.4.1 Changing local leisure systems to achieve fit with NERS

NERS was often described as having potential to be effective, but as often offering limited fit with the local context. In particular, several coordinators identified a shortage of low intensity exercise opportunities, with one describing 'a total lack of classes for people of kind of low impact could actually attend'. In response, several areas attempted to change the local context through developing exercise opportunities within leisure centres or community centres, to be used as exit routes for NERS patients, but which would also be available to members of the public. In some areas, delivery of these complementary structures was pending, with funding having been granted. However, in others, these services were seen as beginning to transform the culture of leisure services, triggering movement towards offering more low intensity opportunities for individuals relatively new to exercise, with one coordinator commenting that 'the leisure centres' seen the success of those classes and they've taken on a bit of a mantle now of trying to expand the 4 walls of the leisure centre into the community'.

6.3.3.4.2 Reinventing the innovation

Where policy representatives identified local variation, this was typically in reference to threats to quality and standardisation. Coordinators however commonly commented on the need to reinvent, sometimes simply through issues such as retaining control

over branding in order to maintain a sense of local ownership. One coordinator for example commented that 'we don't particularly want to copy another authorities name ... we want to develop our own name, or brand if you like.'

More extensive reinvention included experimenting with linking the programme with other existing structures in order to foster links with health care systems, with one coordinator commenting on having tried 'an innovative programme in the initial side with partnership with hospitals the NHS, and having a physiotherapist working within the program'. This coordinator had also integrated nutrition classes into NERS. Another coordinator described supplementing professional training in order to increase nutritional support for the large volume of patients referred for weight management, and to offer a wider range of services to attract more patients, commenting that 'I wanted to make sure they were nutrition level 2, I wanted to make sure they were aqua qualified so they could teach aqua with, aqua based activity and I wanted to make sure they were all walk leader trained'. Coordinators who discussed experimenting with or adding to the protocols were typically experienced coordinators, who had strongly endorsed the value of national standardisation. Hence, rather than a means of subverting protocols, reinvention was seen as a means of building on the foundation provided by the protocols.

Reinvention was sometimes described as needed in order to communicate the scheme to local populations with varying needs or to mobilise the scheme to more isolated patients. Consistent with quantitative data indicating higher completion in more affluent counties (Table 6), several coordinators described the scheme as better received and easier to implement in more affluent communities, identifying a need to tailor communications to local patients' motivations and understandings. One coordinator commented that promotional materials were often aimed above people's literacy levels, and needed to be tailored locally, and that many local people were 'very poorly educated so the way you communicate has to be, you know designed so that people can actually take on that information. It is no good getting technical because people will just think I'm frightened I don't understand this information I'm not going to absorb it'. The coordinator of one of the poorest areas, made up of several isolated communities, spoke of attempting to make NERS more accessible through offering classes in school halls or community centres. However, success of these

efforts had at times been limited by shortages of facilities and exit routes, and limited local support.

6.3.3.4.3 Sustainability of referral criteria

Aside from concerns regarding whether funding would continue, coordinators cited one key threat to sustainability. At the time of interview, coordinators unanimously agreed that referral numbers were coming through at an increasing pace, indicating that referral to NERS was increasingly becoming integrated into health professionals practice. However, this volume of referrals was now beginning to be seen as problematic. Several coordinators expressed concern about waiting list build-ups, arguing that the scheme could not be sustained without reducing the patient to professional ratio. Some commented that the openness of criteria made the scheme available to patients who 'don't technically need this level of expertise' but that implementers 'can't not accept them because they're in the selection criteria', meaning that those in need were caught in long waiting lists. Notably, in both phase one and two, the areas which achieved the highest population-adjusted referral volumes simultaneously achieved the poorest scheme completion rates, with these area's coordinators describing a need for alterations to the innovation such as reviews of referral criteria and risk stratification. Others however described criteria as a vast improvement on what had been in place previously, with eligibility more explicitly defined, typically focusing upon a need for more staff and seeing narrowing of referral criteria as a retrogressive step given the efforts that had gone into getting health professionals to use the service.

6.3.4 Summary and implications

Whilst coordinators' adoption of NERS protocols was essentially an authority innovation-decision (Rogers, 1995), efforts to involve local coordinators in deciding the form of the scheme had been made through consulting local coordinators in developing protocols. However, these efforts were not seen as fully preventing resistance. Whilst relative advantage is typically one of the most important predictors of successful diffusion (Rogers, 1995; Greenhalgh et al., 2004), coordinators were described as perceiving previous schemes as better than NERS, with these perceptions attributed by policy representatives to emotional investments in previous practice and objections to evaluation. By contrast, sometime after adoption, comparisons between

previous schemes and NERS made by coordinators typically favoured NERS. According to Diffusion of Innovations theory, whilst initial change can be punishing and is often resisted, positive observable impacts often lead to positive reinforcement (Rogers, 1995; Bartholomew et al., 2006). Perceived impacts such as increased patient retention arising from scheme monitoring structures and benefits of being part of a national network of coordinators, likely helped to reinforce implementation. Whilst, initial negativity likely stemmed from meanings being framed around restriction of local ownership and autonomy (Bartholomew et al., 2006; Inchley et al., 2007), the innovation perhaps over time began to be reframed as an opportunity to improve practice.

Both parties commented that the civil servants initially coordinating the scheme had not had sufficient time available to provide sufficient communication and support. However, the quality of communication was likely also impacted by the structures of communication channels, with policy representatives and coordinators both citing limited communication from the other party. Key roles of change agents in communicating new innovations involve persuading intended implementers of the need for change, building rapport and opening up lines of communication (Rogers, 1995); roles typically performed most successfully where there is a high degree of homophily between change agents and intended implementers, and where the change agent is oriented towards the needs of intended implementers more closely than those of the change agency (Rogers, 1995). However, civil servants were perhaps not optimally positioned to achieve these goals, due to heterophily in professional status and close orientation to the agency imposing an unpopular move. In combination with the disempowerment arising from instruction to change, power imbalances in communications likely triggered defensiveness in coordinators, leading to expressions of frustration among civil servants at coordinators responses. Indeed, there are clear parallels with the classic behaviour change scenario, where an authority figure such as a doctor tells the patient to change, triggering defensiveness, followed by frustration in the doctor at patient non-compliance, further entrenching patient resistance (Fogarty, 1997). Altering communication dynamics through appointment of a highly respected peer to the role of change agent likely pacified these imbalances (Goodman and Steckler, 1989; Rogers, 2003), with the new national coordinator's experience of ERS delivery leading to higher credibility in the eyes of coordinators and ability to

empathise with challenges involved in delivering ERS. Given that diffusion was seen as slowed by objections to the evaluation in addition to objections to the move to national standardisation, communicating justifications for the trial via a respected peer may have also helped to limit this resistance.

Whilst some coordinators discussed divergences in implementation arising from limited communications, reports of other coordinators that NERS protocols were easy to implement were perhaps at odds with reports by national policy representatives of the low fidelity of some aspects of the intended model. It is likely however that the aforementioned communication difficulties led to the innovation being sometimes perceived as less complex than it was. Adopters typically use experience of similar innovations as their main frame of reference when making sense of new innovations, causing exaggeration of similarity and failures to identify key differences, with limited communications perhaps heightening a tendency to rely upon these frames of reference (Rogers, 1995; Spillane et al., 2002). Hence, aspects which represented more fundamental change in practice were perhaps overlooked or seen as peripheral to the core intervention, leading to perceptions of full implementation despite non-implementation of some aspects which were central to policy representatives' theories of change.

Challenges persuading coordinators of the need for change were also seen by policy representatives as leading to limitations in communication with health professionals, with coordinators seen as accepting the role of change agent somewhat reluctantly. Alternative explanations for slow adoption were however offered by coordinators, including the absence of a coordinator to fulfil this role in the initial stages of the scheme in one area, as well as promotion to referral partners being seen as a daunting task for newly promoted professionals. Whilst as discussed, successful change agents are commonly of similar social and professional status to adopters (Rogers, 1995), general practitioners were likely perceived as higher social status groups, making communication efforts intimidating. Indeed, this may have in part explained a tendency for many coordinators to target practice nurses and managers rather than directly targeting GPs, with these stakeholders perhaps seen as less high status than general practitioners and hence easier to communicate with. In addition, experienced

coordinators likely had existing relationships with referral partners which new coordinators may have needed to build.

Several coordinators commented that government backing was influential in securing adoption by health professionals'. Previous studies incorporating health professionals' views on using ERS have raised concerns regarding effectiveness (Graham et al., 2005), and ability of ERS staff to deal with patient conditions (Wiles et al., 2008). However, backing by a high status authority perhaps added credibility to coordinators' communications. As with coordinators' eventual endorsement, referral partners' adoption was perceived as linked to perceptions of relative advantage (Rogers, 1995), with coordinators describing the benefits of NERS as a means of alleviating concerns about removal of a previous service or adoption of a trial within the implementation of NERS. However, consistent with health professionals' reports in previous studies, which point to a lack of time as a barrier to promoting physical activity (Lawlor, Keen and Neal, 2000), policy representatives cited objections to a paper-based referral system, seen as taking too long to complete, as well as requests for financial incentives for which resources were not available. Consistent with a previous study in which absence of feedback was cited as a key factor limiting referral behaviour (Graham et al., 2005), and with notions that making positive outcomes observable to adopters reinforces adoption decisions (Rogers, 1995; Bartholomew et al., 2006), coordinators highlighted the value of feedback in increasing referrals.

Communicating the scheme to leisure centres was typically seen as relatively easy, perhaps because of the similar status of exercise coordinators and leisure service managers, and because of relationships developed in previous roles. In several areas, coordinators commented that centres had begun to see advantages of NERS in terms of income generation from classes and post-scheme memberships. However, in other areas, typically those with relatively small referral volumes, and where income generated by small class sizes was perhaps insufficient to incentivise or reinforce adoption (Rogers, 1995), coordinators cited challenges accessing space for classes, or described limiting use to quiet times in the centre. Some also commented on a reluctance of mainstream staff to accept patients after the scheme, perhaps indicating a lack of local acceptance that the scheme was an adequate introduction to exercise for those with clinical conditions. In most areas where challenges communicating with

centres were described, the scheme tended to be offered in a limited number of centres or relatively small numbers of classes were offered per centre. In conjunction with relatively poor completion levels in these areas, this may suggest that conflicts had implications for patient experience, through for example, limiting the number of classes, or restricting class times. In areas where reluctance to accept completers into mainstream classes after 16 weeks was reported, long term maintenance of behavioural changes may be made particularly difficult.

Key challenges moving from initiation to implementation were cited as arising from the limited support and training offered to coordinators, particularly amongst newly promoted coordinators. Policy representatives described initial ambiguity over whose role it was to provide management support and training, with this seen as the role of local authorities, but not being delivered by either party in the initial stages. In hindsight, not identifying the need for a national coordinator to support these roles was cited as a key weakness. Given that innovations perceived as complex are less likely to be successfully implemented (Rogers, 2003), and that innovations augmented with sufficient training and support are more likely to be implemented (Greenhalgh et al., 2004), identification of support and training needs prior to implementation must be key priorities in implementing complex interventions. In addition, in several areas, coordination difficulties arose from coordinators leaving their post at a critical phase in implementation. These local authorities perhaps had insufficient 'slack' (i.e. excess) resources to cover these periods (Greenhalgh et al., 2004). Alternatively, the scheme may not yet have become seen as sufficiently important to the organisation for cover to be provided, perhaps consistent with policy representatives' surprise at the limited provision of management support within local authorities.

In discussing routinising the innovation into practice, coordinators described a process of mutual adaptation, where practices of leisure services changed to support the new innovation, whilst the innovation changed to suit local needs (Ringwalt et al., 2004). Long term sustainability of outcomes was often seen as contingent on development of post-programme exercise options, perceived as leading to an increased culture of low intensity classes in leisure centres and community centres, increasing availability for patients and non-patients alike. In addition, coordinators commented on the value of reinvention, including rebranding the scheme in order to achieve local ownership and

acceptability to the local system, or integration of additional services into NERS, often seen as key to building links with other organisations and meeting the diverse needs of local patients. This is consistent with a view of innovations as not constructed by developers and passively replicated, but as reconstructed, given meaning and shaped to fit organisational ways of working and local needs (Boczkowski, 1999).

Reinvention is problematic where removing core elements, or introducing ideas which conflict with innovation functioning. Where no reinvention takes place however, innovations are less likely to achieve routinisation (O'Loughlin et al., 1998; Ringwalt et al., 2004), perhaps due to limited local fit, or limited commitment to an innovation over which implementers perceive no ownership (Rogers, 1995). Coordinators appeared to distinguish between passive reinvention emerging due to misunderstandings of the innovation or limited support to implement it, likely leading to limited fidelity delivery of some core elements, and more active reinvention, seen as building upon the foundation provided by protocols and increasing the chance of the innovation becoming routinised into its contexts (Goodman and Steckler, 1989; O'Loughlin et al., 1998).

The one perceived threat to sustainability however was the increasing volume of referrals, with demand beginning to outstrip supply. Some suggested targeting the intervention, or offering reduced intervention to those requiring less support, whilst others saw narrower criteria as a retrogressive step. Notably, the areas achieving the highest referral volumes also achieved the poorest rates of adherence, with coordinators in both areas focusing on need to ensure that the scheme was offered only to those who needed it. Whilst like NERS, most previous ERS have used relatively inclusive referral criteria (Stevens et al., 1998; Taylor et al., 1998; Harland et al., 1999; Harrison et al., 2005b; Isaacs et al., 2007) this model may be sustainable only in a climate of limited adoption by health professionals.

In summary, moves to improve consistency through national standardisation will likely engender challenges ensuring adoption amongst local implementers, due to associated restriction of autonomy and reduction in local ownership. Where offering clear advantages over previous practice, realisation of advantages over time may reduce resistance. However, the dynamics of communication channels likely play a

key role in determining the success of efforts to persuade implementers of the value of change, with civil servants likely not optimally positioned to achieve these ends. Given their heterophily with the professional and social status of local implementers and their orientation to the agency imposing the change, assigning the role of change agent to civil servants will likely produce a defensive interpersonal dynamic, heightening resistance. Assigning these roles to a respected peer may reduce power imbalances and increase implementers' perceptions of the credibility of communications and that their concerns are being understood. In innovations involving multiple layers of diffusion, the success of these communications in persuading implementers of the need for change will like have knock on effects for communications throughout the system.

In addition to careful attention to communication structures, attention is needed to ensuring provision of adequate training and infrastructure to support such moves. A more prolonged piloting phase may have assisted in understanding the feasibility of protocols and proposed delivery mechanisms prior to full trial (Craig et al. 2008a). In addition, efforts to maintain quality and standardise delivery of policy innovations should be aware of the value of reinvention. Reinvention is problematic where arising from misunderstandings, shortage of skills or introduction of conflicting activities, though is likely useful and adaptive where ensuring sustainability of new innovations, perceptions of the innovation as locally owned and fit with local contexts.

6.4 Consistency of implementation with programme theory

6.4.1 Aims

Thus far, this chapter has described the programme theory underpinning NERS, and explored its diffusion into local practice. Exploration of diffusion indicated some significant perceived weaknesses in communication structures as well as limited training and support for implementation at the national and local levels; challenges perceived by policy representatives and some coordinators as contributing to divergences in implementation. Whilst coordinators were experienced implementers of ERS, many aspects of NERS protocols were likely unfamiliar, such as motivational interviewing and enhanced emphasis on maintaining contact with patients outside of classes. Hence, it is perhaps likely that these activities were less well implemented as a

result of these difficulties. Guided by the programme logic model described in Section 6.2, this section provides quantitative data on the implementation of NERS, before considering implications of divergences for programme theory.

6.4.2 Data sources

The summary figures presented in this section draw upon routine monitoring data (consultation conduct, topics covered and programme attendance) structured interviews with 12 coordinators and 38 professionals (content of exercise programme and application of the £1 discount) and observations of consultations recorded by 23 professionals (motivational interviewing delivery and length of consultations). Summary statistics were calculated using SPSS version 16.

6.4.3 Findings

6.4.3.1 Fidelity

6.4.3.1.1 The exercise programme: group based exercise, direct supervision and discount.

All centres offered at least one form of supervised group based activity. In most areas (n=8), at least one professional offered one-to-one exercise sessions, though a minority (n=13; 22.8%) reported doing so. Coordinators in 11 areas reported that each patient was supervised by an exercise professional for all classes in the first 4 weeks before other opportunities were made available. In the remaining area, supervised classes were also available, but patients could attend other activities from scheme entry. According to coordinators, all programme activities supervised by a NERS professional were charged at £1, with the exception of one area who whilst charging £1 for classes such as circuits, charged £1.50 for supervised gym use.

6.4.3.1.2 Patient consultations and follow up contact.

Table 8 presents the percentage of consultations containing each component specified in protocols. At least one component of health check data (e.g. blood pressure, resting heart rate, waist circumference or weight) was recorded for 909 (98.6) attendees at first consultations. However, MI fidelity was poor in all recordings (mean=8.7%, σ =5.2%, range 0.0 to 23.9%). For the most part, consultations were dominated by form completion and measurements, involving limited discussion of behaviour change, with

73.9% of professionals according to coder 1, and 65.2% according to coder 2, speaking more than half the time (82.6% agreement between coders). Whilst goals were recorded for 901 (98.7%) patients attending a first consultation, only 33.1% of goal records contained both a measurable goal and a timeframe, though this varied from 0% to 77.4% between areas. MI coders also noted concerns about the processes followed to set goals, with professionals often leading the goal setting process and introducing a need for weight loss before the patient identified this as a goal.

Table 8. Fidelity of consultations, determined by the percentage of delivered consultations containing components prescribed by protocols.

Consultation	Component	Frequency and percentage of delivered consultations including listed component
First		913
	- Health check	904 (99.0)
	- Motivational interviewing	0 (0)*
	- Goal setting	901 (98.7)
	- Measurable, time bounded goals	298 (32.6)
4 week		682
	- Goal reviewing	645 (94.6)
	- Discussion of how finding classes	675 (99.0)
	- Discussion of how feeling after 4 weeks of exercise	654 (95.9)
16 week		473
	- Goal review	437 (92.4)
	- Health check	439 (92.8)
8 month		432
	- Goal review	384 (88.9)
	- Change in conditions	407 (94.2)
	- Discussion of how feeling	419 (97.0)
12 month		283
	- Goal review	253 (89.4)
	- Health check	219 (77.4)

* estimate based on observed sample of 25 consultations

All areas offered contact several weeks after programme entry and on scheme exit. As indicated in Table 8, in most cases, the exercise professional indicated that they had covered the topics prescribed by protocols. However, whilst protocols stated that non-

attendees should be contacted 4 weeks into the programme, in most cases, only patients still attending NERS were contacted at 4 weeks. In total, 682 (63.1%) patients were contacted at 4 weeks, with 621 (57.5) of these patients still attending the scheme. Follow up of non-attendees varied between areas, with 20.1% of non-attending scheme entrants contacted at 4 weeks overall, varying from 0% to 62.3% between areas. Approximately 1 in 3 non-attendees contacted by an exercise professional (n=20) subsequently returned to complete the programme. Whilst almost a quarter of patients did not receive a health check on scheme exit, conversations with coordinators indicate that this was often because patients chose to conduct these conversations by telephone.

After scheme exit, 77.8% of scheme completers were contacted at 8 months, whilst area-level percentages ranged from 13.3% to 96.7%. A small number of patients listed as non-completers (n=68) were also contacted. Of the patients contacted at 8 months, 100 reported not currently participating in any physical activity. Discussion of relapse prevention strategies was indicated for 55 (55.0%) of these patients, including agreement to maintain telephone contact or allow the patient to return to NERS classes, or advice on returning to activity. Overall, 56.0% of scheme completers were contacted at 12 months, whilst area level percentages ranged from 8.7 to 84.2%. A small number of patients listed as non-completers were also contacted (n=18). A composite score based on the percentage of non-attending scheme entrants contacted at 4 weeks, and percentages of scheme completers contacted at 8 and 12 months, indicated 51.3% overall fidelity to patient follow-up protocols, ranging from 8.9% to 78.4% at the area level.

6.4.3.1.3 Delivery of non-prescribed elements: the exercise programme and exit routes.

In all areas, the NERS programme comprised a minimum of group-based gym and circuits classes, with 55 (98.2%) sites offering at least one of these activities, and 32 (57.1%) offering both. A minority of centres, distributed across 7 areas, also included pool-based activities (n=14; 25.0%) and weekend classes (n=14; 25.0%). Most offered evening exercise classes (n=33; 58.9%), with all areas bar one offering evening classes in at least one centre. Whilst 47 (83.9%) centres in 11 areas provided patient-only group classes, supervised gym sessions were held during general opening times in all

but 7 (12.5%) centres. Two had a small separate gym specifically for the referral scheme, whereas in the remaining 5 (8.93%), all from the same area, the gym was closed to the public during NERS sessions. Five areas also offered non-centre based outdoor activities.

In terms of exit routes, all areas offered leisure-centre based activities (e.g. gym memberships, swimming and mainstream classes). Additional activities were offered in 10 areas including outdoor activities (e.g. walking groups), community classes (e.g. salsa or line dancing, yoga) and sports clubs (e.g. canoeing, badminton, bowling or sailing clubs). The most common combination of alternatives was leisure centre based activities and outdoor activities, offered in 5 areas (see Table 9). Only 1 offered all 4 types.

Table 9. Number of areas offering a range of types of exit route options to scheme completers

	Number of areas
Leisure centre based activities only	2
Leisure centre based activities and community classes	1
Leisure centre based activities and outdoor activities	5
Leisure centre based activities, outdoor activities and community classes	2
Leisure centre based activities, outdoor activities and sports clubs	1
Leisure centre based activities, community classes, outdoor activities and sports clubs	1

6.4.3.2 Dose delivered

The exercise programme and discount. The mean time elapsed between scheme entry and exit consultations was approximately 21 weeks (144.3 days; σ =46.9), though the median was somewhat lower at 131 days (approximately 19 weeks). Overall, 20 consultations (4.4%) were conducted less than 15 weeks after scheme entry, 96 (21.2%) within one-week of the 16 week target (i.e. between 15 and 17 weeks) and a majority (n=338; 74.5%) more than 17 weeks after scheme entry.

The average number of group classes in each centre was 4.1 (σ =3.9). However, 4 centres ran large numbers of classes (i.e. 10 to 23 classes), with the median being somewhat lower at 3 classes per centre. Overall 46 facilities (80.7%) spread across all

12 areas offered at least 2 group classes per week. In the remaining 10 (7 leisure centres, 2 community centres and 1 other facility) the recommended dose of 2 sessions per week could only be achieved by attending multiple sites. One however was a countryside service based site used only occasionally, whilst 4 others were in the same area as one another, where implementers reported that facilities were near one another. As described above, scheme completers were able to access NERS classes for a median of approximately 19 weeks. However, 24 (63.2%) professionals, divided across 10 areas, reported allowing continued access to NERS classes after the programme.

Whilst 9 coordinators reported imposing no limits on how often patients could use the discount, one limited use to twice per week, another reported having done so in the past but having changed this policy midway through the trial, whilst one reported that use was rationed to twice per week during times when space was limited. In addition, four areas did not restrict the discount to NERS activities, allowing it to be used for any independent centre use, with one further area applying the discount to independent pool use, but not gym use or classes. In two areas excluding independent centre use from the £1 rate, lower level discounts were available for these activities. All coordinators reported that the £1 rate was offered only for the duration of the exercise programme. However, it was replaced in 9 areas by discounts of between 22% and 58% for scheme completers, with one area offering one free month as a reward for completion prior to discounted membership. One additional area waived the initial joining fee for scheme completers, though offered no discount on subsequent membership.

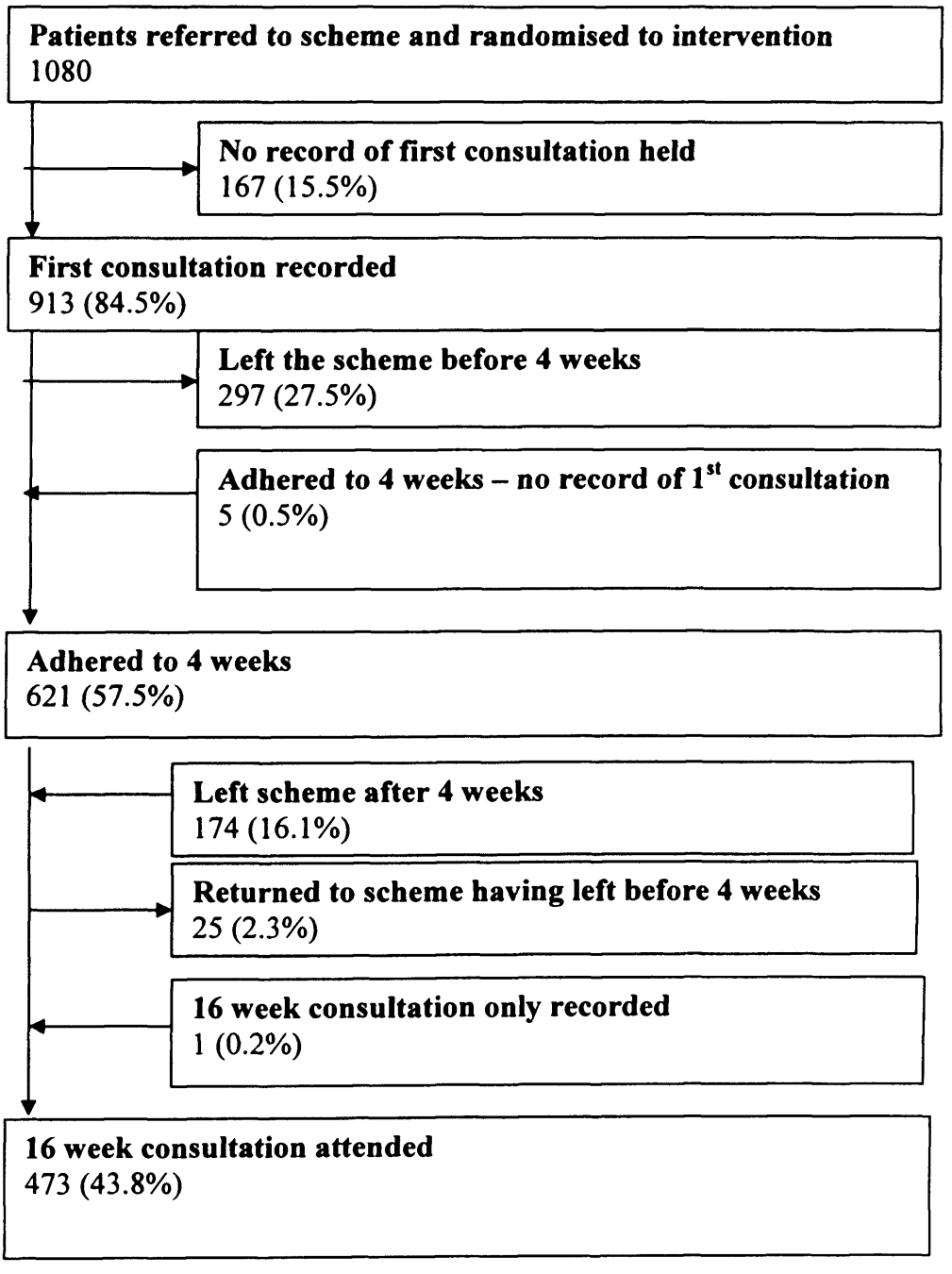
Consultations. Recordings of first consultations averaged 34.8 minutes, though were highly variable between professionals (σ =13.1, range 12.5 to 57.4). No data on subsequent consultation length was available.

6.4.3.3 Reach

The flow of participants through the intervention is displayed in Figure 5. Of 1080 referrals, there was no record of scheme entry for 167 patients (15.5%), although for 5, there was a record of 4 week contact and for one, a record only of 16 week consultation, presumably indicating that scheme entry data were missing. Hence, 161

(14.9%) patients received only health professional advice and no further intervention, whilst 919 (85.1%) entered the scheme. Among patients for whom scheme entry was recorded, records indicate that 297 (27.5%) patients did not adhere to the first four weeks. Of these 297 however, 25 subsequently returned to the programme, as indicated by records for 16 week consultations, whereas 272 (25.2%) are presumed to have left between the initial consultation and 4 week contact. Overall, 621 (57.5%) referred patients were still attending the scheme at 4 weeks, of whom 174 (16.1%) dropped out prior to scheme completion, with 473 (43.9%) patients attending a 16 week consultation. At the area level, recorded attendance at a first consultation ranged from 79.4% to 92.9%, adherence to 4 weeks from 30.9% to 70.9% and scheme completion from 10.7 to 61.5%.

Figure 5. Flow of participants through the 16 week NERS scheme



6.4.4 Summary and implications

Although some authors have highlighted the need for interventions to be delivered with adequate fidelity and dose if programme theory is to be tested (Steckler and Linnan, 2002), implementation checks often reveal substantial deviations from protocols (Carroll et al., 2007). However, as described in Chapter 3, implementation is rarely discussed in the context of a clear definition of programme theory, and hence the implications of divergent delivery for programme functioning are often unclear. In the present study, prior articulation of programme theory provided a framework for assessment of implementation, which enabled implications of implementation shortcomings for programme functioning to be considered.

Fidelity assessments indicated some common characteristics across areas. Protocols stipulated that patients were to be offered a programme of mostly group-based activity, and were to be under direct supervision of the professional for a minimum of four weeks. In all areas, direct professional supervision in group exercise classes was offered, though one area also allowed patients to access other exercise opportunities immediately. The £1 discount was applied to NERS supervised exercise in all areas, though areas were almost evenly divided as to whether non-NERS activities were also included within this rate. Hence, whilst some minor local reinvention took place, with some local areas going further than the NERS protocols, a common core was offered across areas.

In addition, all areas offered consultations on scheme entry, partway through the scheme and on scheme exit. However scrutiny of their content revealed that patient consultations typically diverged substantially from programme theory. As described in Section 6.3, coordinators often commented that implementation had been somewhat hampered by limited communications, training and support. Perhaps as a result of these shortcomings, whilst activities likely highly familiar to implementers were delivered, complex or unfamiliar new practices such as motivational interviewing, goal setting and extended patient follow-up were typically delivered poorly.

Motivational interviewing was not delivered in any area, and for the most part, consultations were dominated by data gathering activities, which appeared to

supersede intended motivational functions of consultations. Hence, MI is removed in a revised NERS model (see Figure 6). The use of theory-led motivational counselling approaches is recommended by the Department of Health (2001) in order to address the often poor levels of adherence to ERS. Non-delivery likely weakens the capacity of the scheme to elicit and strengthen patients' internal motivations, perhaps rendering the scheme more effective for patients who enter already somewhat internally motivated. Following the evaluation team highlighting the lack of use of MI, a further two-day training course was implemented towards the end of the trial to improve confidence and practice in the use of these techniques. As almost all participants had received initial consultations by the time training was received, it will not be possible to integrate analysis of the impacts of MI fidelity into outcomes analyses (Thyrian et al., 2007). A sub-study evaluating impacts of training on subsequent MI fidelity will be reported in Chapter 7.

Policy representatives stated that goal setting was to follow on from MI, with MI eliciting talk regarding motivations for attendance, which would be harnessed into specific, personally valued goals (Miller and Rollnick, 2002). Therefore, MI non-delivery may have had knock on effects for goal setting. Indeed, goal-setting within NERS typically involved non-specific goals, likely compromising the goal reviewing functions of subsequent consultations and making it difficult to ascertain whether goals were achieved. Notes by MI coders also indicated a tendency for professionals to direct the goal setting process, and to introduce weight-loss goals without the patient having identified weight loss as a goal, presumably because this was stated on referral forms as the health professional's reason for referring the patient.

This is however perhaps problematic firstly because it involves introducing potentially emotionally sensitive goals, rather than guiding patients towards identifying goals for themselves. Even if the health professional has identified weight loss as their reason for referring the patient, it should not be assumed that this is the patient's own principle motivation. Furthermore, to achieve the caloric expenditure necessary to achieve weight loss goals of typically 1lb per week simply through two hours of exercise would require greater exercise intensity than offered within NERS. Hence, perceived failure to achieve goals is perhaps likely to lead to demotivation (Locke and Latham, 2002) unless the patient also changes their diet.

As described in Section 6.3, in response to the volume of patients entering NERS for weight management, some coordinators had tailored the scheme to include additional nutritional coaching. According to self-efficacy theory however (Bandura, 1977), behaviour change is most likely where valued outcomes are directly attributed to the behaviour in question. Hence, with change attributable only in part to the exercise programme, weight losses may play a stronger role in motivating dietary change than adherence to exercise. Whilst goal setting and monitoring remain in the NERS model, links with intermediate outcomes are represented by broken arrows, to indicate that this aspect of programme theory is compromised by variable quality (see Figure 6). As with MI, concerns regarding the delivery of goal setting were fed back to policy representatives during the trial. Chapter 7 examines whether goal setting quality improved over time, and explores associations of goal setting quality and type with adherence.

Further divergences were observed in patient follow-up protocols. Whilst recommending that patients be contacted at 4 weeks regardless of attendance, 4 week contact was typically only made with patients still attending the scheme. Some areas did not contact a single non-attendee, whilst others contacted a majority, whilst one in 3 non-attendees with whom contact was made returned to the scheme. Though professionals may have contacted patients they felt were more likely to come back, closer adherence to these protocols may have led to higher adherence. Similarly, after leaving the scheme, whilst most completers were contacted at 8 and 12 months to review progress, this again varied by area. As with goal setting, links between follow-up contact and intermediate outcomes are represented by broken arrows in the revised model, to indicate that this aspect of programme theory is compromised by variable delivery (see Figure 6). Whilst non-delivery of activities intended for delivery in first consultations were identified early, given that follow-up activities were to be conducted in later phases of the scheme, it was not possible to communicate concerns regarding non-delivery until the trial was drawing to a close. Impacts for trial outcomes may be usefully explored at a later stage.

Assessments of dose delivered also revealed important variations. Large standard deviations suggest that programme duration was highly variable, with 16-week

consultations typically held after 19 weeks according to median values; approximately double the 10-week dose offered in many previous ERS trials (Isaacs et al. 2007; Stevens et al. 1998; Taylor et al. 1998). The amount of time taken over first consultations varied from 12 minutes to almost an hour. Previous studies demonstrating impacts of MI on physical activity have typically allowed 45 minutes or more for this activity alone (Harland et al., 1999; Brodie and Inoue, 2005; Bennett et al., 2007b; Carels et al., 2007; Hardcastle et al., 2008) and even the longest NERS consultations were likely too short to feasibly include all intended components. Thirdly, there was substantial variability between centres in terms of the number of classes available.

Furthermore, intervention dose was extended by most professionals through allowing patients to continue accessing NERS classes indefinitely, rather than withdrawing support at 16 weeks as suggested in protocols. Given that continued access may minimise loss of contact with the professional and other patients (Wormald et al., 2006), participants allowed continued access may be more likely to adhere beyond 16 weeks. However, as described in Section 6.3, as numbers increased after randomisation, the sustainability of the referral criteria were beginning to be questioned. Hence, space would likely become unavailable for scheme completers, with the tendency for indefinite access unsustainable. In addition, most areas extended intervention dose through negotiating discounts for completers with leisure service providers. Hence, important caveats in interpreting trial outcomes are that whether activity continued beyond expiry of longer-term discounts and after eventual removal of access to NERS classes cannot be established. A modified NERS model includes introduction of post-programme discounts as a key programme activity potentially moderating the impact of removing the £1 discount, as well as continued access to NERS classes as a moderator of the impact of the withdrawal of the scheme on physical activity adherence levels (see Figure 6).

Whilst stipulating that the programme should consist mostly of group-based exercise, protocols were non-prescriptive about what types of exercise would be offered, leaving these decisions to local providers. All areas offered group-based exercise, although some offered a more restricted programme consisting of only circuit classes and gym sessions, whilst others included outdoor and pool based activities. Most

professionals did not report offering one-to-one sessions. However, these were anticipated to be a minority activity for patients not wanting to enter a group model, and it is likely that many professionals did not have any such patients. In addition, some offered weekend or evening sessions whilst others did not. Furthermore, a variety of exit route opportunities were offered, ranging from leisure centre based activities, to community based classes, sports clubs and outdoor activities. As described in Section 6.3, variability in activities often reflected local tailoring in response to patient wants and needs. However, previous studies have highlighted a desire amongst exercise referral participants for a greater diversity of activities than sometimes offered within such schemes (Wormald and Ingle, 2004), also identifying a lack of flexibility in available times as a barrier to attendance for working patients (Wormald and Ingle, 2004; Taylor and Fox, 2005). Hence, future analyses may perhaps focus usefully on exploring variability in adherence by activity types and programme flexibility.

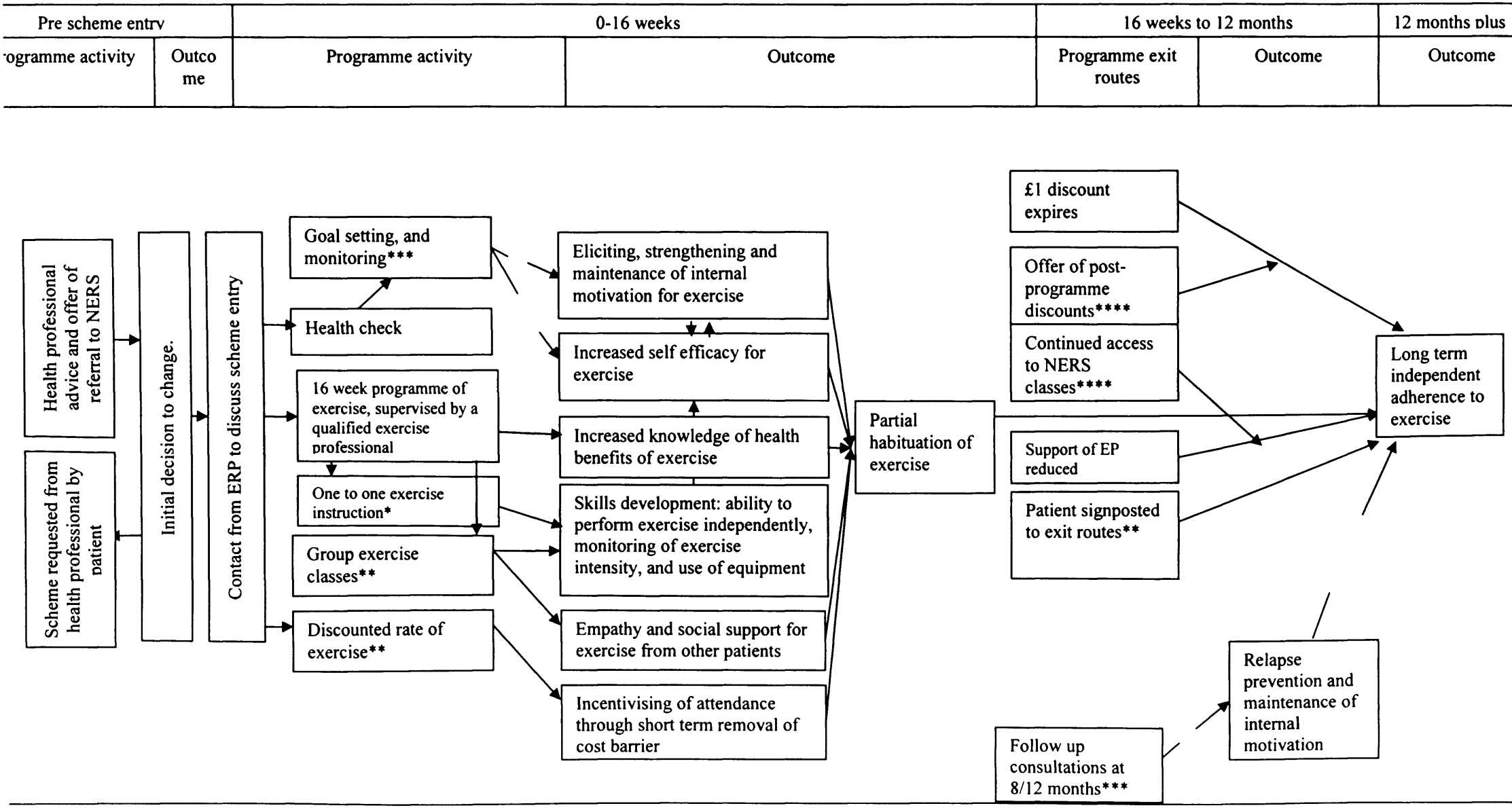
Finally, attendance figures indicated that 1 in 7 patients received no intervention after health professional advice. This is better than average according to Williams and colleagues (Williams et al., 2007), whose review concluded that almost a third of referred patients typically do not enter. Approximately 44% of patients completed the scheme, which is towards the top of the range reported for previous schemes of 12-52% (Williams et al., 2007; Lee et al., 2009). Hence, despite not all components being fully implemented, the scheme achieved one of the best adherence levels in the ERS literature, likely attributable to activities including a relatively high level of professional support and supervision, and group based classes. This however varied substantially between individual areas, who achieved adherence rates both lower than the bottom end (11%) and higher than the top end (62%) of the previously reported range. The largest drop off occurred during the first few weeks, with approximately 1 in 3 entrants not returning for a four week consultation, although some subsequently returned to the programme.

Given the time pressures typical to a pragmatic policy evaluation, the present study elicited programme theory as it was rolled out as part of the trial, rather than within a separate developmental pilot phase. Consequently, although a comprehensive assessment is offered of many key aspects of programme delivery, some components

were not adequately captured by the monitoring systems developed prior to elicitation of programme theory at the start of the programme. For example, whilst the range of exit routes offered at the area level is described, it would have been desirable to identify specific exit route opportunities to which patients were directed at the patient level. In addition, given that much data on fidelity is based on routine records, some of the variation reported in the conduct and content of patient reviewing protocols may reflect poor record keeping rather than low fidelity. Furthermore, the study is limited by the fact that the data components collected for the purpose of the process evaluation are largely cross sectional. Building such assessments into programme monitoring structures would provide a comprehensive longitudinal assessment of intervention implementation and useful quality control checks of fidelity, providing that measures are also in place to ensure the quality of monitoring data.

Nevertheless, the analyses presented in this section demonstrate the importance of understanding the consistency of programme implementation with programme theory prior to interpreting outcomes, and exploring how the delivered programme is experienced by patients. Given the changing nature of the intervention as it diffused into practice, a new model is presented which will allow interpretation of outcomes based on the intervention as delivered (Figure 6), rather than naively assuming the intended intervention and the delivered intervention to be exactly the same, as has been the case in all previous ERS evaluations.

Figure 6. A modified theory of change for the National Exercise Referral Scheme, taking into account scheme implementation



* Minority activity, for patients who don't wish to enter under the group model
** Offered in all areas, but with variable range of exercise types available
*** Offered in all areas, though with highly variable fidelity
**** Offered in most areas

6.5 Conclusions

This chapter has demonstrated the importance of making explicit programme theory, focusing upon how a complex intervention is diffused into practice and understanding its consistency with programme theory before drawing conclusions regarding its outcomes. NERS programme theory contained a range of theoretically plausible principles guided by behaviour change theory. However, exploration of programme diffusion highlighted key shortcomings in communication structures, training provision and support for implementation of the new protocols, as well as a range of issues which shaped the rate of local scheme adoption, including clashes with the financial priorities of leisure centres, perceptions of previous schemes as not needing to change and periods of time with no-one to fulfil the coordinator role. In many areas, the scheme took longer to be adopted into practice than in others, with one area selected to be in the trial failing to get underway until randomisation was complete.

Perhaps in part as a result of these perceived shortcomings in communication, training and support, some substantial uncoupling from programme theory emerged. Aspects of delivery representing significant changes in practice, such as motivational interviewing, goal setting and patient follow-up protocols, tended to be implemented poorly. Hence, these activities likely played a limited role in shaping the outcomes of NERS. In addition to these key omissions, additional local activities were often introduced to support long-term maintenance of changes, including extension of NERS discounts, and a widespread tendency to allow indefinite access to NERS classes. Hence, key caveats in interpreting outcomes will include that patients were likely still in receipt of a discount at long-term follow-up and that impacts were likely supported by a higher level of support than will be feasible in the longer term. In several areas, maintenance of impacts was seen as contingent on altering the local context through providing additional opportunities for post-scheme activity, and it is perhaps likely that as these were developed over time, impacts of NERS were enhanced. Furthermore, coordinators' reports questioned whether the open referral criteria used in NERS during the trial period will remain sustainable as numbers increase after the trial, and longer term applications of NERS may ultimately focus on somewhat more targeted audiences than did the trial. Chapter 7 will now provide an

in-depth examination of formative aspects of the NERS process evaluation, focusing on efforts to improve delivery of MI and goal setting as the trial progressed.

7 Formative aspects of the NERS process evaluation: improving fidelity of motivational communication strategies

7.1 Chapter aims

Chapter 6 reported that whilst adherence to NERS was relatively high by comparison to many previous ERS, most referred patients still did not receive a full programme, with adherence substantially higher in some areas than others. Furthermore, efforts to diffuse two key components of NERS protocols into local practice, whose intended function was to enhance adherence, were unsuccessful. Concerns regarding the non-implementation of MI and the poor implementation of goal setting were communicated to policy representatives during the trial. In the case of motivational interviewing, concerns were expressed before any data were collected, as training protocols revealed that only one hour of training had been offered. Though no recommendations were offered at this stage, these comments triggered administration of additional training midway through the trial. In the case of goal setting, a tendency for setting of vague goals was communicated to policy representatives, and via policy representatives, to implementers. This chapter attempts to inform improvement of these two key limitations, addressing the following research questions:

- For whom and under what circumstances does a two-day training course in motivational interviewing lead to increased consistency with motivational interviewing?
- Are patients for whom measurable and time-bound goals are agreed more likely to adhere to NERS?

The first half of the chapter presents data from a mixed-methods evaluation of MI training, whilst the second half chapter uses routine monitoring data to explore whether goal setting quality increased over time and to explore the links between goal setting quality and programme adherence.

7.2 *Integration of motivational interviewing into practice*

7.2.1 Context and aims

NERS professionals attended a two-day (approximately 13 hours) training course, combining brief presentations with live demonstrations, video presentations, discussion, self-assessment, supervised practice, coaching and 'real play' exercises. Practitioners were asked to think of a behaviour which they were ambivalent about and given exercises in which they alternated between acting as patient or practitioner in order to gain experience of MI strategies, whilst experiencing an insight from the patients' perspective about contrasts between MI consistent and non-consistent behaviours. This section explores responses to this course, examining perceptions of the usefulness and feasibility of MI for practice, as well as examining the circumstances in which practice begins to become more consistent with MI after training.

7.2.2 Data sources

This study component combines cross-sectional qualitative interviews shortly after attendance at training with quantitative monitoring of MI delivery in practice samples before and 6 months after training. Eligible participants for the qualitative aspect were a sub-sample of 31 professionals who attended one of 3 training courses between August 2008 and January 2009, of whom 27 took part in an interview after training. Ten area coordinators interviewed as part of the evaluation were based in areas whose professionals had at the time of interview attended this training. Whilst not explicitly exploring views on MI, many coordinators volunteered opinions on training and spoke of MI whilst describing how they saw the programme as bringing about change in patients' behaviour. Finally an interview with the MI trainer explored views on delivering training, perceptions of professionals' responses to it, expectations for its impact and perceived feasibility issues in integrating MI into NERS. Quantitative monitoring of consultation delivery included the subsample of professionals who attended the October course. Of the 23 professionals who provided baseline samples examined in Chapter 6, 14 attended this course. These 14 professionals were invited to provide further recordings, 6 months after training, with 11 doing so. Study design and flow of participants through the study are summarised in Figure 7.

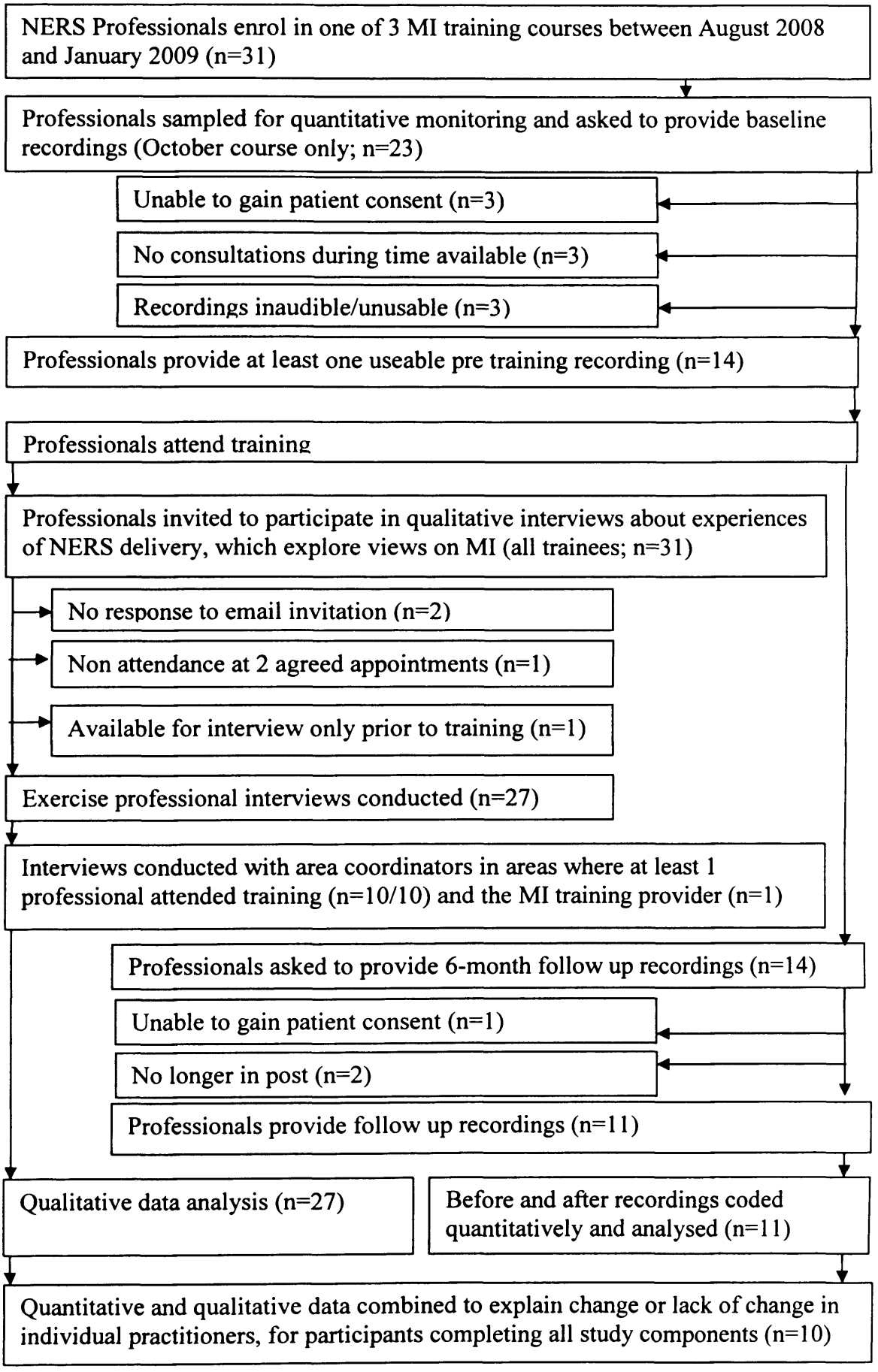


Figure 7. Overview of study design and flow of NERS exercise professionals through quantitative and qualitative aspects of the study.

7.2.2.1 Analysis

7.2.2.1.1 Qualitative analysis

Coordinator data are described under a single heading 'Coordinators' views on the role of MI within NERS', followed by an overview of the views of the MI training provider, described under the heading 'The MI training provider's views on the potential impacts of training on MI delivery in NERS'. Key themes from exercise professional interviews were: i) awareness of MI and the purpose of the first consultation before and after training, ii) perceived usefulness of training and impacts on practice, and iii) implementation issues. Data from exercise professionals are described under each heading, with verbatim quotes presented.

7.2.2.1.2 Quantitative analysis

The BECCI scale (Lane et al. 2005) demonstrated excellent inter-rater reliability at both time points in the subsample participating in this study element (r=0.94 pre training; r=0.96 post training). Means, standard deviations and medians for pre and post training MI fidelity scores were calculated. As 4 professionals at baseline and 2 at follow up provided only one recording, for the 11 professionals providing recordings, the first was used as the score for MI fidelity rather than an average of two or more scores. Spearman's rank correlation coefficients comparing first and second recordings for 7 professionals at baseline and 9 at follow up indicated that MI adherence was typically stable between recordings (r=0.71 and r=0.87). Significance of change in practices after training was assessed using a Wilcoxon Sign test. Mean and standard deviation consultation lengths and frequencies for remaining consultation characteristics were calculated.

7.2.2.1.3 Integrated qualitative and quantitative analysis

Ten professionals provided both quantitative before and after measures and qualitative data. Data sources were linked using a form a single-case design (Kazi, 2003), whereby quantitative change over time is tracked at the individual level and compared against qualitative responses to training, to attempt to develop theories regarding why change may or may not have occurred. Qualitative data for each individual participant were coded into substrata of the thematic categories described above and tabulated

against change in practice. Individual participants were grouped into clusters by area and responses to training compared against those of their area coordinator.

7.2.3 Findings

7.2.3.1 Coordinators' views on the role of MI within NERS

Whilst three coordinators whose professionals had received training in MI made no reference to its role when discussing how NERS affected behavioural change, and one discussed MI only when discussing views of training within the scheme, others commented that integration of MI offered substantial promise for the development of the scheme in their area. Some commented that professionals had gained significant confidence in dealing with patients with complex conditions through adopting MI principles. There was however some ambiguity over what MI was and the purposes it served. Some saw MI as a means of improving communication between patient and professional, being of real relevance only to professionals who were relatively poor communicators. MI was sometimes seen as largely what instructors did once they had been in the role for sufficient time, developing tacitly through experience, with new instructors simply lacking the experience to communicate effectively. By contrast, others saw adopting MI as involving a radical rethinking of how consultations were to be approached, introducing principles which fundamentally conflicted with prior practices. MI was sometimes seen as a tool to enable the professional to extract information from clients, with 3 coordinators commenting that it had assisted the information gathering process. In addition, coordinators sometimes conflated MI with goal setting, or saw MI simply as a strategy for eliciting meaningful goals.

7.2.3.2 The MI training provider's views on potential impacts of training on MI delivery in NERS

The trainer commented that the course was well received by professionals, and that most had appeared to see the value of change. However, the trainer also argued that the 2-day course 'gives you intellectual insight but doesn't necessarily give you emotional insight such that you can easily change your practice'. Real change was seen as requiring structures for self-reflection and further expert guidance, with the training provider arguing that 'like with any new skill, you need to rehearse it, you need to get guidance, you need to get coaching and your confidence goes up then'. In

addition to a need for ongoing training to support integration, the training provider identified barriers to implementation relating to structures of NERS consultations, and professionals' perceptions of their role within these consultations. The training provider remarked that when asking professionals what they hoped to gain from the course, professionals consistently commented that 'their job was collecting information. They were very clear that that was what their job was...they were very clear that what they hoped that we would teach them was how to be more efficient in doing that'. This perception of the primacy of data collection was linked to a 'genuine tension between the desire to collect data and what actually would help the person sat in front of you'. The tendency to prioritise data collection over MI was seen as needing to be reversed, with the training provider arguing that if MI is to be integrated, professionals 'have to think that this (MI) is the most important thing to do. You can collect the data at the end if the time is available'. The training provider commented that once pressures of data collection were adhered to, time allowed for consultations lay at the outer margins of the minimum required to incorporate MI and that it 'sets the wrong tempo for MI if you start off by saying, right I'm going to fill these forms in and ask you lots of questions'.

7.2.3.3 Professionals views on the introduction to motivational interviewing

7.2.3.3.1 Awareness of MI and the purpose of the first consultation before and after training

Most professionals described the 2-day course as their first in-depth experience of MI, often describing the initial 1-hour training as superficial, or not recalling it at all. In addition to increasing their understanding of MI, several professionals commented that training had made them aware of a substantial disjuncture between how they had seen the purposes of the initial consultation and how its purposes had been conceived in the design stage by national implementers. Whilst professionals often commented on the value of the first consultation in allowing them to build rapport and put patients at ease, its primary purpose had often been seen by professionals as one of form filling and data gathering. Several professionals expressed a view that introduction to an MI-based approach had helped them to move away from this view, towards seeing the consultation as an opportunity to elicit talk about change.

> [C5] We had training a few weeks ago, regarding motivational interviewing, and it kind of opened, probably, most of our eyes, to what it was actually for. I think before we just thought it was a paper exercise where you just get everybody's details and go from there really, and just get them started up.

Rejection of MI was in a minority of cases however linked to maintenance of the view of consultations as being primarily for gathering information, and a perception that MI did not assist this task.

> [B34] I found it boring and, personally and, I just didn't feel it helped as there's only so much you can listen to obviously as you're timescale and what you want from your client, you want how they're feeling, what medication, what can you do for them and then obviously you want to move on

7.2.3.3.2 Perceived usefulness of training and impacts on practice

Most professionals saw value in MI and spoke of a willingness to adopt MI principles in practice. Some described how they had changed practices since training, often offering substantial critical reflection on the perceived shortcomings of previous practice. In some instances for example, professionals identified prior tendencies to ask closed questions and to tell patients what they should do, which they felt had begun to alter after training. Through these changes, many commented that rapport and the amount of change talk had greatly improved.

> [J40] Since we've come back from that course the way that we sort of speak to our clients has changed quite dramatically...instead of just asking question after question we sort of get them to sit there and they talk. Initially I thought no they're not going to talk but they do. They tell you the answers, you don't even need to ask the questions, they're already telling you it...I think you engage more with the clients.

A minority however were unconvinced of the value of changing their own practices, sometimes commenting that their current approach worked.

> [L12] I did find the course difficult because like I say, I've obviously, not set in my ways, maybe I have got a way of working with my clients and it seems to work.

Others commented that as clients were already someway along the continuum of readiness to change, MI was unnecessary

> [C17] With our clients, mostly a large percentage of them are there anyway because they really want to be or they have been asked to. So there is ambivalence there but it's not restrictive kind of thing.

7.2.3.3.3 Implementation issues

Whilst some felt that the duration of training was adequate to facilitate practice, many commented that MI had yet to become normalised into daily practice, and that they would require time and practice to feel confident in using the newly taught skills. Hence, as with the MI training provider, the two day course was seen by some professionals as the start of the road to implementing MI, but by no means as the end of it.

> [J39] You do it on the course but it's putting it into practice we need to do now...you feel almost embarrassed first of all

In addition, several professionals discussed the tension between data collection and MI described by the training provider above. Some commented that structures of the initial consultation and the volume of data collection made integration of MI principles difficult or impossible. One for example expressed a desire for further clarification on how the tasks could be integrated.

> [G30] The only thing I found hard was that we have to have the information that we need to put on the database, and how without asking questions do we get that because it's not necessary that they will actually give that information, so I think I would really like to see the people who taught it. I'd like to see them do a consultation

Another commented that structures of the initial consultation would need to be radically rethought if MI were to be integrated, with the implicit value judgments and closed questions involved in data collection components clashing with the patient-centred spirit of MI.

> [B32] The new skills they've tried to teach us on this course you know you couldn't do it with the present system we've got.... the forms are a hindrance in a way to gaining what you want to gain from the client, which is change

7.2.3.4 Quantitative monitoring of MI delivery before and after training

7.2.3.4.1 MI fidelity before and after training

Mean MI fidelity for the 11 professionals providing baseline and follow up recordings rose from 10.0% prior to training (SD=6.52) to 16.7% after training (SD=13.49). Medians at both time points were lower than means, at 8.0% and 10.2% pre and post training respectively, indicating an undue influence of small numbers of relatively high scores. A Wilcoxon Sign Test indicated that overall improvement in MI fidelity was not significant (Z=.80, p=.42).

7.2.3.4.2 Other consultation characteristics

Overall, 7 professionals scored higher for MI fidelity after training, whilst 4 scored marginally lower (see Table 10). Whilst most changes were small, 3 professionals demonstrated clear improvements, with fidelity increasing by more than 20%, from below 10% at baseline to between 30 and 40% at follow up. These professionals and another who scored 4th highest at follow up, all opened consultations with a 10-25 minute unstructured discussion of behaviour change, before completing measures. Two others opened by asking the patient to describe a typical day, but did little to relate this to behaviour change, moving onto forms after about 5 minutes. In the remaining 5, as was typical in baseline consultations, talk about behaviour change typically occurred only during completion of structured questionnaires. Four professionals within two areas opened consultation with lengthy discussion of medical conditions and medications, before moving onto lifestyle questionnaires. Pre-training consultations averaged 35.2 minutes (SD=12.8). Post training mean length remained almost identical (Mean=34.7; SD=17.2). Whilst prior to training, there was a moderate correlation between consultation length and MI fidelity (r=.59, p=0.09), longer consultations were no more likely to be MI adherent at follow up (r=-.04, p=.91).

7.2.3.5 Linkage of qualitative implementer data to quantitative monitoring of practice.

Table 10 presents a matrix of sub-themes arising in interviews with each of the 10 professionals who provided both quantitative longitudinal data and took part in a semi-structured interview, with these linked to change in practice. Notably, only in areas

where the coordinator commented on the value of integrating MI into practice did professionals consistently offer critical reflection on their own previous practice. Although two professionals in these areas scored only slightly higher after training, and another regressed to a lower level than at baseline, one professional in each area demonstrated improvements of 20% or more, with the 5 highest post training fidelity scores coming from practitioners in these 3 regions. In two areas where the coordinator did not mention MI when discussing how NERS helped patients become more active, negative views on the usefulness of training were expressed by professionals, with concerns regarding the feasibility of integrating MI into current structures expressed in one. In both of these areas, no professional demonstrated substantial change in practice, with 3 of the 4 lowest post-training fidelity scores coming from these 2 regions. In the remaining area, where the coordinator commented that training had mostly reinforced his professionals' current practice, the professional providing recordings spoke positively of training, yet did not offer critical reflection on her own practice, with fidelity regressing to a lower level than at baseline. However, quantitative data were only obtained from 1 professional in this area, with all remaining professionals offering more substantial critical reflection than did this one professional.

Whilst overall stability of practices between first and second recordings was high ($r=0.87$), the only professionals whose practices demonstrated low stability were two of the three practitioners who performed highest at follow up (the third provided only a single recording). Both of these professionals, whilst maintaining higher than average MI fidelity in both consultations, scored more than 10% lower for second recordings than they did for first recordings. Hence, although demonstrating higher proficiency than at baseline, improvements appeared not to be consistent across consultations.

Table 10. Cross tabulation of themes emerging from exercise professional and coordinator interviews against change in MI fidelity according to coding of pre and post training consultations for individual professionals.

Area	Area coordinator views	Exercise professional views						Change in MI fidelity score *	Stability of post training practice**
		Role of 1st consultation	Perceived usefulness for practice			Implementation issues			
	Value of integrating MI into consultations raised	Discussion of changed perception of role of 1st consultation	Positive comments regarding usefulness for practice	Negative comments regarding usefulness for practice	Critical reflection on previous practice and discussion of how MI has been integrated since training	Concerns regarding need for practice or training	Feasibility clash with current structures		
C	X	X	X		X	X		+	High
		X	X		X	X		+++	Low
A			X			X		-	High
E				X				+	High
			X				X	-	High
D	X	X	X		X			+++	Low
			X		X	X		-	High
F				X				-	Moderate
J	X		X		X	X		+	n/a
			X		X			++++	n/a

* - no improvement, + 0-10% improvement, ++ 10-20% improvement, +++ 20-30% improvement, ++++ >30% improvement

** low stability >10% discrepancy between recording one and two, moderate 5-10% discrepancy between recording one and two, high <5% discrepancy between recording one and two, n/a only one recording provided

7.2.4 Summary and implications

Whilst a crucial first stage in diffusing a new innovation into practice is to raise potential adopters awareness of it (Rogers 2003), in Section 6.3, some local coordinators commented that the core intent of NERS protocols had not always been well communicated, leading to divergent local understandings. Similarly, in this sub-study, both the MI training provider and a number of professionals identified tendencies for professionals to see the first consultation primarily as a data collection exercise, rather than as an opportunity to serve the baseline motivational functions intended by programme developers. Hence, in relation to MI, the innovation-decision process (Rogers 2003) essentially broke down at the first stage, with communications failing to raise intended implementers' awareness of the proposed change in practice. However, following additional training, professionals commonly described recognition and acceptance of intended motivational functions of the consultation.

As described by Greenhalgh et al. (2004), once successfully communicated to intended implementers, perceived relative advantage over current practice is a prerequisite for innovation adoption. If seen as offering no benefit, a new innovation will typically be dismissed with no further consideration. Following training, though a minority of professionals described MI as offering little advantage over previous practice, most spoke positively of the training, often offering substantial critical reflections on previous practice and linking efforts to use MI to increased rapport with patients. Hence, perceived relative advantage and willingness to adopt MI appeared to be the norm, with observable impacts beginning to reinforce the adoption decision for some (Rogers 2003).

Even where seen as offering advantages over prior practice however, non-adoption or weak implementation of new innovations commonly arises from factors such as high complexity or limited compatibility with current structures (Rogers 2003). Following training, many professionals expressed reservations about confidence, and a perceived need for practice or further training, appearing to see the new innovation as too complex to be used well unless additional support was offered. These reservations, and similar concerns expressed by the MI trainer, are consistent with growing evidence that putting MI into practice requires skills development beyond a 2-day course

(Miller, 2004; Bennett et al., 2007a; Smith JL, 2007; Mitcheson et al., 2009). Although some commented that the 2-day course was adequate to inform proficient practice, such reports should be treated with caution given the limited change observed in practice samples, and previous studies which indicate that attending a 2-day workshop might convince clinicians they do not need further training, whilst practice samples show insufficient change to benefit patients (Miller and Mount, 2001). An accurate self-assessment of competence is likely contingent on a full understanding of MI, with perceptions that it is easy to use after only 2-days of training likely revealing a limited understanding of its complexity (Miller and Rollnick 2009).

Concerns were also raised regarding compatibility with other aspects of NERS consultations. As described in Chapter 6, these averaged 35 minutes, including a substantial structured data gathering element, seen as having a number of detrimental impacts on MI delivery. In addition to difficulties arising from the aforementioned tendency for data collecting functions to be seen as the key purpose of the consultation, the data gathering element reduced the time available for MI, with structured questioning and implicit value judgments within lifestyle questionnaires seen by some as causing an insurmountable clash between incompatible tasks. Rollnick et al. (2008) have recently described a tendency for large volumes of structured assessment as a barrier to delivery of MI in health care settings, given that this risks rendering the patient passive in the consultation process, which is antithetical to the patient-centred spirit of MI.

Hence in summary, the 2-day training course appeared to be successful in communicating the intended purposes of the first consultation, with most perceiving MI as offering relative advantages over prior practice. However, high perceived complexity and limited compatibility appeared to contribute either to non-adoption or to concerns regarding weak implementation. Six months after training, no significant change in MI fidelity was observed at the aggregate level, with all scores remaining in the bottom half of the possible range. Overall, scores at both time points were comparable to those in a recent study using the same scale to examine GPs' routine practice (Moran et al., 2008).

However, several professionals did demonstrate substantial movement towards an MI consistent approach. Where doubts were expressed regarding relative advantage or compatibility, no improvements in consistency with MI were observed. However, improvements were observed amongst some professionals who commented on a need for further training or practice to become competent users, with limited competence in these professionals perhaps leading to weak implementation rather than to complete non-adoption. Substantial improvements occurred only in areas headed by coordinators who had commented on the potential value of integrating MI into practice after training and where all professionals expressed positive views on training, consistent with a view that transferring training into practice is easier where colleagues are supportive of integration (Baldwin and Ford, 1988). Alternatively, enthusiasm amongst coordinators may have reflected views reported back by professionals having received training. Either way, enthusiasm or scepticism was clustered by area, and only in areas where there was agreement throughout the local system that it offered potential for practice, did practice begin to change. However, practice remained variable between professionals even within areas where support appeared unanimous and was somewhat unstable even amongst those in whom major improvements were observed.

A number of strengths and limitations of this study component merit consideration before considering implications for practice. Strengths include use of a validated procedure for quantifying MI fidelity (Lane et al., 2005). In addition, the study is among the first to integrate qualitative data into a before and after assessment of MI fidelity, offering potential explanations for the individual level changes observed in quantitative data. Weaknesses include the fact that MI proficiency at each time-point was typically based on a single recording, given that some professionals were only able to provide 1 recording. Multiple recordings at both time points would ideally have been obtained for all professionals, to ensure representativeness of routine practice. The amount of data that could be collected was limited by the fact that ethical approval was being sought as implementers developed plans to provide further training, with approval received weeks before the October course, leading to non-participation of several eligible professionals who reported willingness to provide recordings, but had no consultations in the window of time available. The relatively small sample size limited statistical power, although statistical significance was not of

primary importance to this study, which focused principally upon the circumstances under which individual practitioners' behaviours began to change.

Nevertheless, the study has important implications for programmes aiming to integrate MI into practice. A tendency is emerging in ERS trials to claim to have incorporated motivational counselling strategies such as MI into consultations (Hosper et al., 2008; Lawton et al., 2008; Sorensen et al., 2008). However, significant abandonment of learned professional practices and adoption of patient-centred counselling roles is almost always assumed to be straightforward. This study is the first study to examine attempted diffusion of MI into the practices of ERS implementers and demonstrates the naivety of such assumptions. Unless it can be clearly demonstrated that sufficient steps have been followed to ensure effective implementation, any claims of MI delivery should be met with scepticism.

Implementers attempting to use MI should seek consultation with experts in the field on baseline training requirements, ensuring that implementers have adequate skills and understanding prior to practice, and providing an opportunity to communicate the principle planned functions of consultations throughout the whole system. Facilitating delivery by professionals with no counselling background is an ambitious project, likely requiring a greater level of investment in support than where delivered by self-selected practitioners. Indeed, in one recent study, following training in MI and ongoing feedback, skills of smoking cessation practitioners continuously improved throughout a 2 and a half year period (Forsberg et al., 2010) and only at the end of this period did all practitioners achieve minimum competence standards.

This task is likely made difficult within NERS by the fact that professionals were experienced in delivery of ERS which did not use MI, and were given a year to develop ways of operating within NERS, before the expectation of MI delivery was reinforced through in-depth training. Asking professionals who have developed their own ways of approaching consultations to adopt MI will likely require them to accept that MI may be an improvement on current practice (Spillane et al., 2002); a position most though not all NERS professionals appeared happy to accept. It is perhaps unlikely that MI will be adopted with proficiency by the minority of professionals who see little value in it. Hence, as well as providing training at the earliest opportunity,

willingness to adopt MI should perhaps be ascertained during staff recruitment, with MI forming an integral part of professionals' job description.

Whilst most professionals appeared willing to adopt MI, additional support is needed to translate willingness into ability to practice (Madson et al., 2009). Progression towards an MI consistent approach will likely require systematic integration of mechanisms for critical self-reflection, monitoring and feedback into the structures of NERS (Miller, 2004; Bennett et al., 2007a; Smith JL, 2007; Mitcheson et al., 2009). Such structures may be useful in providing additional support to those recognising a need for it and in preventing practitioners from prematurely believing that they have become competent. Where coordinators are new to MI, as was the case in NERS, feedback structures should be contracted to external parties with proven expertise in MI.

Finally, whether consultation structures allow for integration of MI needs careful consideration. Consultations must allow sufficient time for MI. Furthermore, MI will often not be the sole purpose of the consultation, in which case, careful consideration needs to be given to whether implementers are being asked to combine MI with activities with which it is incompatible. Structured assessments pose a substantial challenge for integration of MI and if assessments cannot be removed or substantially reduced, the manner in which professionals are expected to integrate the two competing purposes deserve careful consideration(Rollnick et al., 2008).

7.3 Variability in goal setting quality: change over time and implications for adherence

7.3.1 Aims

As described in Chapter 4, MI if delivered well facilitates effective goal setting processes through eliciting talk regarding the latent motivations driving the patients' decision to attend, which may then be usefully harnessed into specific, achievable and personally relevant goals. Hence, non-delivery of MI likely contributed to the weak delivery of goal setting described in Chapter 6. As described in Chapter 4, goal setting theory stipulates that specific goals (i.e. goals which are measurable and time-bound) tend to be more motivating than 'do your best' goals, given that they reduce ambiguity

in what is to be achieved and how to achieve it. However, within NERS, goals tended to be non-specific. In addition, a tendency emerged for heavy emphasis on weight loss goals. Given the need for goals to be achievable, and the fact that significant weight loss is unlikely through a low intensity exercise programme such as NERS, weight loss goals will likely serve a limited function in facilitating change. The poor quality of goal setting was communicated to implementers when identified by the trial manager in early stages of the trial. This section explores whether goal setting quality improved over time and whether goal setting quality is linked with higher levels of adherence. It is hypothesised that patients for whom measurable and time-bound goals are set will be more likely to complete the programme than those for whom more unspecific goals are set, and that associations between goal setting quality and adherence will vary by goal type, with weight loss goals less strongly linked to adherence than other goals.

7.3.2 Data sources and analysis

Routine monitoring data and baseline trial data were used for a subsample of 901 patients taking part in the randomised controlled trial of NERS who: i) were randomised to the intervention group, ii) attended a first consultation and iii) for whom a goal was recorded in the database. Goal records were coded dichotomously as either: i) measurable (e.g 'to lose 2lb' rather than 'to lose weight') and time-bound (i.e. having an endpoint at which goal achievement may be assessed and new goals set, such as 'in 4 weeks time'), or ii) not measurable and time-bound. For patients for whom measurable and time-bound goals were set, these were broken down further into types of goals. Patients' age, sex, and WIMD scores were derived from baseline questionnaires. Reasons for referral and referral/scheme entry date were recorded in the database. The patient was considered to have completed the programme if the database indicated that they had attended a scheme exit consultation at the end of the 16-week programme.

Patterning of goal setting quality between patient groups was first explored by examining frequencies and percentages of patients for whom measurable and time bound goals were set, according to sex, age, deprivation and reasons for referral. Significance of associations was assessed using univariable binary logistic regression models. Frequencies and percentages of each goal type within the subgroup of patients

for whom measurable and time bound goals were set were then calculated. In order to explore links between goal setting quality and adherence, further logistic regression models were used, with scheme completion/non-completion the binary dependent variable, and a binary variable for goal setting quality the sole independent variable (i.e. measurable and time-bound versus not measurable and time-bound). Subsequently, multivariate regressions adjusted for age, sex, deprivation, reason for referral, time between referral and scheme entry and time since first referral to NERS. In order to explore variation in adherence by goal type, given the small numbers of patients for whom goals other than weight-loss or scheme attendance were set, measurable and time-bound goals were divided into: i) weight loss goals, ii) attendance related goals, iii) other goals. A binary variable was created to represent each goal type. Unadjusted associations of goal setting quality with scheme adherence were first calculated using binary logistic regression, before models were adjusted for age, sex, education, reason for referral, time to scheme entry and time since first referral. Odds ratios and 95% confidence intervals are presented. Analyses were adjusted for clustering at the exercise professional level using STATA version 11.

7.3.3 Findings

7.3.3.1 Response rates

Of the 1080 patients randomised to the NERS intervention, a record of a scheme entry consultation was held for 913 patients. Of these patients, goal records were held for 901 patients. Analyses focus on these 901 patients. Three however were not linked to a specific exercise professional, and were dropped from analyses adjusted for clustering, whilst age, referral date or date of scheme entry could not be determined for a further 59 patients, due to missing data, giving a sample of 839 patients for final models.

7.3.3.2 Sample description

As indicated in Table 11, participants were predominantly female. Mean (and standard deviation) age was 52.2 (14.3) years. Characteristics of patients for whom goal records were obtained were comparable to the whole sample of intervention patients in these 12 areas. Almost three-quarters were referred for weight management, with approximately 2 in 5 referred for non-weight related CHD risk factors and 1 in 4 for mental health reasons.

Table 11. Sample description for NERS patients for whom a goal record and adherence details were obtained compared against whole sample

		All intervention patients	Patients for whom goal record obtained
Females		708 (65.6)	603 (66.9)
WIMD	**Low**	361 (34.4)	313 (35.8)
	Medium	358 (34.1)	290 (33.2)
	High	330 (31.5)	271 (31.0)
Reason for referral	**Non-weight related CHD risk factors**	457 (42.3)	386 (42.8)
	Weight management	786 (72.8)	661 (73.4)
	Mental health	310 (28.7)	253 (28.1)

7.3.3.3 Patterning of goal setting quality by patient characteristics and stability of practice over time

As indicated in Table 12, measurable and time-bound goals were significantly more likely to be recorded for patients referred for weight management. Where weight management was not listed as a reason for referral, measurable and time-bound goals were set for only approximately 1 in 4 patients, compared to more than 1 in 3 patients referred for weight management. No associations between other reasons for referral, or demographic predictors were observed. However, a positive association was observed between the length of time since the scheme began and the odds of measurable time-bound goals being set, with measurable and time bound goals set less frequently in the early days of the programme.

Table 12.Percentages of patients for whom measurable and time bound goals were set, by demographic factors, reasons for referral and length of time since scheme began. Odds ratios are from univariable binary logistic regression.

		Percentage of patients for whom measurable, time bound goals set	Odds ratio (95% CI)
Sex	Male	106 (35.6)	1.18
	Female	192 (31.8)	(0.88 to 1.59)
Age	Low (16 to <40 years)	88 (33.3)	-
	Medium (41 to 60 years)	106 (35.0)	1.09 (0.79 to 1.50)
	High (>60 years)	92 (30.3)	0.87 (0.64 to 1.19)
WIMD	Low	108 (34.5)	-
	Medium	79 (27.2)	0.70 (0.45 to 1.09)
	High	103 (38.0)	1.16 (0.71 to 1.90)
Weight management	No	59 (24.6)	**1.74**
	Yes	239 (36.2)	**(1.25 to 2.41)**
Mental Health	No	222 (34.3)	0.82
	Yes	76 (30.0)	(0.59 to 1.15)
Non weight related CHD risk	No	160 (31.1)	1.24
	Yes	138 (35.8)	(0.93 to 1.66)
Time since 1st referral	-	-	**1.06 (1.00 to 1.13)**

7.3.3.4 Types of measurable and time bound goals set

In 177 (59.4%) of the 298 goal records which contained a measurable and time-bound element, the only such element was a weight-loss goal, such as 'to lose 4lb in the next 4 weeks'. A further 59 (19.8%) contained elements which related both to weight loss and to another form of goal, whereas the remaining 62 (20.8%) contained only elements relating to goals other than weight loss. Of the 121 patients for whom non-weight related goals were set, most related to scheme attendance (e.g. 'to attend 2 classes per week for 4 weeks' - n=99; 33.2% of measurable and time-bound goals). Twenty-four participants set goals relating to issues such as activity outside of the programme such as walking the dog or going swimming with the family, functional

fitness such as 'to be able to walk 100m without feeling breathless in 4 weeks', or goals relating to participation in an event, such as 'to complete the Cardiff 10k'.

7.3.3.5 Associations of goal setting quality with adherence

Among patients for whom goal records contained no measurable and time-bound element, 310 (51.4%) completed the programme. Amongst patients for whom goal records contained at least one measurable and time bound element, an increase in adherence rates of less than 2% was observed (n=158; 53.0%). This difference was not significant either in univariable analyses (OR=1.06; 95% CI=0.64 to 1.73) or multivariate analyses (OR=1.11; 95% CI=0.65 to 1.91).

As indicated in Table 13, where goals were disaggregated by type, patients with weight loss goals were very marginally less likely to adhere than those with non-specific goals, whereas patients with attendance related goals were 9% more likely to complete the programme than those with non-specific goals. However, in logistic regression analyses containing only binary variables for goal setting type and in multivariate analyses, these trends were not significant (see Table 13).

Table 13. Frequencies and percentage completion of NERS according to the type of goals set, and odds ratios from logistic regression models adjusted for clustering at the exercise professional level

	weight loss goals	attendance goals	other goals
Frequencies and percentages	118 (50.0)	60 (60.6)	12 (50.0)
Unadjusted odds ratios and 95%CI (n=898)	0.85 (0.58 to 1.24)	1.54 (0.63 to 3.75)	0.92 (0.37 to 2.29)
Adjusted odds ratio and 95% CI* (n=839)	0.86 (0.55 to 1.33)	1.72 (0.73 to 4.09)	1.19 (0.46 to 3.08)

*adjusted for age, sex, deprivation, reason for referral, time to scheme entry and time since 1st referral

7.3.4 Summary and implications

There was a tendency for more specific goals to be set for patients in the later stages of the NERS trial, likely arising from feedback from the evaluation team to national coordinators midway through the trial that preliminary data checks indicated that goals were typically non-specific, with this feedback then relayed to local implementers. However, further coding by goal type revealed a somewhat narrow range of goals,

with measurable and time-bound goals relating almost exclusively to weight loss or programme attendance, and set most frequently for patients referred for weight management, with weight loss goals by far the most common form of measurable and time-bound goal.

Linkage of goal setting quality with attendance records indicated that patients for whom goals contained a measurable and time-bound element were only marginally more likely to complete the programme than those for whom non-specific goals were set (Locke and Latham, 20022006). Adherence rates amongst patients with weight loss goals were almost identical to those for patients with non-specific goals. By contrast, amongst patients whose goals focused upon a specified level of attendance, a 9% higher adherence rate was observed than amongst patients with non-specific goals. Meaningful analysis of remaining goal types was not possible given that goals other than weight loss or attendance were set for only 24 patients.

As discussed in Chapter 6, the limited associations of weight loss goals with adherence likely stem from limited achievability of weight loss through a low intensity exercise programme, with perceived failure leading to demotivation (Jones et al. 2005). Higher adherence rates in patients with goals relating to scheme attendance may have been in part because these goals were simply more achievable. In addition, Bodenheimer and Handley (Bodenheimer and Handley, 2009) argue that in health behaviour settings, goals relating to concrete action plans tend to be more specific, and hence more motivating, than outcome-oriented goals such as weight loss. The patient for whom action-oriented goals have been agreed likely leaves the consultation not just with an idea of where they want to get to, but having considered how they will get there. However, whilst perhaps failing to reach significance largely due to the relative infrequency of these goals and hence limited statistical power, it is also likely that attendance goals were not sufficiently relevant to patients' internal motivations to trigger change, reflecting goals of the exercise professional rather than those of the patient.

Given that professionals received only one hour of combined training in goal setting and motivational interviewing, instruction to set measurable and time-bound goals in conjunction with limited proficiency likely led to satisfaction of this requirement

through focusing on the most easily measurable goals. Given the intrinsically measurable nature of weight loss, and the high volume of patients referred for weight management, professionals may have found it easier to negotiate specific goals for these patients. Indeed, as reported in Chapter 6, during coding of consultations for MI fidelity, coders noted that professionals were often directing the goal setting process towards weight loss goals before weight loss had been identified by the patient as a motivation for attendance, perhaps because this was listed as the health professionals' reason for referring the patient.

However, more in-depth elicitation and careful harnessing of patients' own internal motivations for change, linking change to personally relevant goals, may be necessary if goal setting is to contribute to motivating change. As will be described in Chapter 8, and as has been described in previous qualitative studies (Stathi et al., 2004), physiological changes are often a secondary priority for ERS patients, with maintenance of autonomy or central life goals such as occupational functioning often seen as more important. However, goals relating to functional fitness or everyday living were surprisingly rare. It may be that with additional training and structures to support integration of motivational interviewing into practice, professionals become able to explore patients' motivations more effectively, leading to negotiation of a wider range of achievable goals linking change to patients' values.

It is important to note that the observational nature of this study component means that cause and effect cannot be established. Goal setting quality likely did not vary at random, but probably emerged from a combination of patient and professional characteristics. Secondly, goal types and adherence levels were based upon written records completed by implementers. It may be that recording of more vague goals in some cases represents poor recordkeeping, with more specific goals agreed but not recorded. Finally, the study focuses upon programme adherence rather than longer term change.

Nevertheless, the study offers some important insights into use of goal setting within ERS. Whilst recommendations that goals be measurable and time-bound were followed to an increasing extent as the trial progressed, the types of goals set did not appear to be significantly more effective than 'do your best' goals in promoting

adherence. It appears unlikely therefore that goal-setting played a significant role in the causal chain within NERS during the trial period. It is recommended that in improving the programme, excessive focus upon weight loss goals should be discouraged. Where patients identify weight loss as a primary goal, they should where possible be guided towards identifying additional more achievable goals which may contribute towards development of an exercise habit, with weight loss and maintenance viewed as a more distal goal arising from long term maintenance of this habit. Furthermore, whilst there was a non-significant trend towards higher adherence amongst patients for whom attendance related goals were set, these goals are likely of limited relevance to the patients' internal motivations, and if set should perhaps be accompanied by goals linked more closely to patients' values.

7.4 Conclusions

Diffusing MI into the practice of implementers such as exercise professionals likely requires fundamental change in learned professional practices, and the complexity of this goal should therefore not be underestimated. Such efforts must ensure that implementers are persuaded of the value of change, should seek expert guidance on training needs, should incorporate monitoring and feedback structures and need to consider carefully whether it is feasible to implement MI alongside other consultation activities. Claims to have integrated MI into routine practice or into evaluated interventions should not be trusted unless accompanied by assessments of fidelity.

It appears likely that neither MI nor goal setting played a significant role in the causal chain for NERS outcomes. Where measurable and time-bound goals were set, these were typically no more effective than 'do your best' goals, likely in large part because these focused on a narrow range of goals which were not necessarily achievable or relevant to patients' internal motivations. Goal setting will likely be made significantly easier by implementation of MI. However, additional guidance may be of value in order to provide implementers with an understanding of how to set measurable and time bound goals which remain consistent with additional tenets of goal-setting theory, being both relevant to patients internal motivations and achievable.

8 Participant experience and social patterning in programme reach

8.1 Chapter aims

Having clearly defined the programme as delivered in Chapter 6, a brief recap of which is presented in Figure 8, this chapter aims to address the following research question

- How and for whom does the delivered intervention facilitate adherence and behavioural change?

As described in Chapter 4, qualitative ERS studies have tended to focus squarely on patients' views of the exercise programme itself, failing to understand the experience of attending ERS in the context of patients' backgrounds, pathways into ERS and motivations for attendance, or to explore concerns about long term maintenance of activity. In addition, although exercise professionals are also likely well positioned to offer insights into which types of patients respond most positively to which activities, and why, qualitative research has typically been limited to patient interviews. The role of patient interviews in exploring these issues is limited, given that these typically involve small numbers of patients, and over-represent the views of older female patients due to the high volume of females and older patients referred. This chapter therefore uses qualitative interviews with exercise professionals and patients in order to explore experiences of the programme, focusing on patients wants and needs on entering NERS, the programme itself and concerns about leaving NERS. Having generated an understanding of perceived active ingredients and processes of change, quantitative profiling of uptake and adherence tests emerging hypotheses.

Figure 8. NERS components as conceived and in practice

Planned key components of NERS	NERS components in practice
Health professional referral to NERS	Health professional referral to NERS
Baseline consultation including • Health check and lifestyle assessment • Motivational interviewing • Goal setting	Baseline consultation including • Health check and lifestyle assessment
16 week exercise programme including • £1 per class rate • Patient only group exercise classes as well as supervised gym use • Supervision by a level 3 qualified instructor • Four week contact to discuss goals and experiences of the programme • Contact of non-attendees at this stage to encourage back to programme	19 week (median) exercise programme including • £1 per class rate • Patient only group exercise classes as well as supervised gym use (indefinite access beyond programme offered in most areas) • Supervision by a level 3 qualified instructor • Contact to discuss goals and experiences of the programme. • Non-attendees typically not contacted.
Scheme exit consultation including • Repeat health check and lifestyle assessment • Discussion of goal progress • Signposting to exit routes	Scheme exit consultation including • Repeat health check and lifestyle assessment • Discussion of goal progress • Signposting to exit routes (including replacement of NERS discount with local discounts in most areas)
8 and 12 month follow up consultations to discuss progress since the scheme	8 and 12 month follow up consultations to discuss progress since the scheme – highly variable delivery

8.2 Participant experiences: views of professionals and patients

8.2.1 Data sources and analyses

Participants were 38 exercise professionals involved in the delivery of NERS in 12 LHB areas and 32 patients (see Chapter 5). Analysis of professional interviews was finalised first, with analysis of patient interviews building on emerging findings.

Key themes from professional interviews related to:

i) Individual differences in patient needs and responses to the programme: baseline motivation and adherence
   a. The referral process: motivating patients or identifying motivated patients?
   b. Patient diversity, confidence, motivation and adherence

ii) The exercise programme: roles of the professional and other patients in supporting motivation, confidence and adherence
   a. Educating, supporting and motivating patients
   b. Identifying and overcoming anxieties about the exercise environment
   c. Group-based activity and emerging social support
   d. Adequacy of training and perceived training needs

Key themes from patient interviews related to:

i) Entering the scheme: routes into NERS and motivations for attendance,

ii) Experiences of the exercise programme
   a. Perceived impacts
   b. Roles of the exercise professional in facilitating change
   c. Roles of other exercisers in facilitating change
   d. The exercise classes: accessibility, flexibility and types

iii) Leaving the programme: the transition to independent activity.

Data are described under each of these headings, with verbatim quotes used to convey a range of views, including consensual and conflicting views between and within interviews.

8.2.2 Findings: exercise professional interviews

8.2.2.1 Individual differences in patient needs and responses to the programme: baseline motivation and adherence

8.2.2.1.1 The referral process: motivating patients or identifying motivated patients?

Approximately half of professionals identified a distinction between patients who sought the programme for themselves, and those advised to take part by a health professional. In all cases where this contrast was identified, health professional advice

was seen as a weaker determinant of adherence than the patient's independent decision to seek help.

> [8] Even if the doctor has told some clients that they need to go and do some exercise, that's still not enough of a culture shock for them, but the ones that decide, or saw the leaflet in the doctors and had to ask the doctor about it, generally they stick around.

In some areas, the majority of patients were seen to have asked their GP for a referral rather than being advised to enter.

> [2] The majority of them are really keen to start exercising to be honest. Like most of them have gone to the GP and asked to be put in the scheme so they are highly motivated.

High levels of drop out in other areas were however attributed to failures of the referral process to identify appropriately motivated patients.

> [40] The drop out is I think a little bit too high at the moment. But the people that are really motivated coming in, they honestly their lives have changed so much, it just makes it worthwhile and it is worth doing. If the right people are being referred in.

Only one professional commented on a role for GPs as motivators of change, suggesting that the esteemed role of GP may have led some patients to act on advice.

> [5] If they go to their GP and their GP makes them aware of their behaviour, then they can 'oh god this is not my family nagging me now, this is somebody that's medically trained' and it makes them aware of their condition a little bit more.

Hence, rather than seeing the referral process as playing an active role in motivating patients, professionals sometimes focused upon a need for the scheme to provide a service to patients who were already sufficiently motivated.

> [29] We've had one or two that have come who don't want to be there at all. I think it's because the GP has sent them ... you can't force no one, and we shouldn't be anyway because it's those that want to come on the scheme that want the help, I think that's where we should put our strengths to.

In areas where problems with referral of unmotivated patients were seen as limiting uptake and adherence, most did not explicitly discuss how or at what stage motivation should be identified. One professional however commented on roles both for health professionals and implementers in ensuring that patients' fully understood the scheme prior to seeking their agreement to refer them through, ensuring that the scheme was offered only to patients interested in taking part.

> [37] Quite often I think physios or any other health professional refer us in because they've, for want of a better word, run out of ideas. So they come to us these people, who haven't had the whole thing explained to them, so they're not sure really what's going on. I'm not sure if that means we need to promote it more in the community with more advertising to let people know it's here or whether that's for health professionals to work through.

8.2.2.1.2 Patient diversity, confidence, motivation and adherence

Among patients referred to NERS, needs and responses to the programme were commonly seen as varying according to conditions and demographic factors. In particular, although some saw mental health patients as benefiting most from the programme where they engaged with it, these patients were seen as facing a range of intrapersonal barriers to adherence, with motivation commonly lowest as a result.

> [8] People with the mental health issues, if we get them along and keep them coming and keep them interested they see huge benefits but they tend to be the group that drop off.

Patients referred for mental health reasons were often seen as having additional anxieties about starting the scheme and as requiring additional interpersonal support to integrate into the exercise environment. However, some professionals commented that if these barriers were addressed, mental health patients tended to adhere more than other groups.

> [25] Just give them that little bit more support when they come in just chat to them a little bit more and make sure they work in pairs...I find that they are the ones most likely to stick at it more to be honest because they reap the benefits...the first couple of weeks are hard but they're the ones that stick at it

Whilst professionals commonly described the majority of patients as vulnerable or lacking confidence on entry to the scheme, mental health patients were seen as

benefiting most from improvements in confidence, through both exercise itself and social aspects of participation.

> [5] I think a lot of people suffer with anxiety, depression and things you see a lot more openness...They get more confidence obviously through exercise, maybe their body shape changes and that gives them that perception, also they are interacting with people when they are here, so their social skills improve in that respect, and also they may get a buzz out of exercising then which increases, more purpose in their lives, rather than what was their day to day life before.

In addition, some professionals highlighted perceived difficulties in motivating very overweight patients, often seeing this as arising from the view that substantial weight loss was unlikely through the low intensity exercise programme alone.

> [16] When you are looking at weight loss you know, they can starve themselves and they'll see quicker changes but you know that's not obviously what they, what I'm trying to get them to do.

Almost all professionals commented that patients as a whole were often anxious about the exercise environment. In addition to higher prevalence of anxieties amongst mental health patients, variation by demographic factors was cited, with anxieties often seen as particularly prevalent amongst older patients for whom the environment was more alien, and women or overweight patients, who were conscious of being looked at by other patients.

> [6] I think people with say like depression, they sort of, they kind of think people are talking about them and stuff like that – that's the sort of feeling I get – I've had one or two actually ask me if so and so was saying something – you know – just a little bit insecure that way. And I mean one or two of the older people, who are retired, they sort of feel that it's a young persons' sort of thing, going to the gym.

By contrast however, as the majority of referrals were older women, some commented that younger patients or men referred through to the scheme sometimes felt less comfortable in the group environment, benefitting less from social aspects of participation.

> [1] The youngsters that we get through are injuries or depression, do kind of drop off because they probably feel a little bit awkward because of the older people.

More than a third of professionals discussed perceived socioeconomic variation in responses to the scheme, with all bar one of these professionals commenting that engaging clients, in terms of both uptake and adherence, had been more difficult in deprived areas. Professionals attributed lower engagement to multiple factors such as a lower tendency for poorer patients to place value on maintaining health, limited financial resources or a lack of buy in among GPs in more deprived areas.

> [6] Its probably one of the hardest valleys to get the GPs to sort of buy into the scheme ... Its an ex-mining valley sort of thing, and its very negative, its like 50% unemployment. So they are kind of 'poor me' sort of thing, and they won't do anything to sort of progress themselves, if it doesn't involve say a pub or a restaurant, they're not interested.

8.2.2.2 The exercise programme: roles of the professional and other patients in facilitating change

8.2.2.2.1 Educating, supporting and motivating patients

Professionals typically described their roles as comprising overlapping tasks of educating and motivating patients and providing empathic support, with some commenting that fitness instruction became a secondary role given the vulnerabilities of the client group.

> [19] My role is a motivator and mentor almost and um a support and someone that people can relate to and talk to openly about their situation so, I would say that's the first thing and then you are almost a fitness instructor second. Because the clients are quite vulnerable.

Whilst as described, some saw failures to engage initially unmotivated patients as arising from referral of inappropriate patients, this was sometimes seen partly as a matter of not having found successful strategies for helping patients to appreciate the value of becoming more active.

> [37] You don't want to offend them by saying look you have to listen to what I'm saying, you don't want to preach, you want them to sort of understand what's going on and hopefully we can educate them into it. But quite often it

> falls on deaf ears and I'm sure that's what the health professionals experience as well

Once patients got underway in their programmes however, professionals typically saw their role as central to maintaining patients' motivation and confidence, through providing constant support and reassurance.

> [3] Keeping their motivation up, making sure that they are not left alone in, for some of them, an alien environment, so they see a friendly face, they may have seen before, and just reassuring them, and giving them the confidence and motivation to carry on.

Patients were often seen as lacking the knowledge of how to exercise safely given the limits of their conditions, with education seen as crucial in enabling them to become independent exercisers. However, comments regarding education were typically inseparable from talk of the need to provide interpersonal support to allow patients to become confident exercisers.

> [5] Knowledge obviously, because of the training we've been through. Knowledge about conditions and what exercise then would suit them. Like I say, all the professionals here are supportive of their client's needs and understanding towards different problems they may occur, maybe anxiety, or confidence, or other problems with illnesses.

Some spoke of becoming valued components of patients' support networks, with patients often seeking someone they could trust and with whom they could discuss issues affecting wellbeing which were not always directly related to physical activity.

> [27] I had a client this week who I only said 'are you ok' and she started crying to me, so it was clearly nothing to do with the gym, but she felt that she could just talk to us and I suppose let off her emotions so. So yeah, clients look for us to be able to help them, but also I suppose be like a friend you know, be their support network.

Whilst some appeared happy to provide this level of support, a minority saw themselves as almost becoming counsellors to patients, with patients' seeing them as someone they could talk to about other difficulties, distracting from the main aim of promoting activity.

> [34] What you want from your client, you want how they're feeling, what medication, what can you do for them and then obviously you want to move on. You know listening, it's a good idea to listen but, you know when you've got your clients who want go into too much depth over things, and a bit too personal... it's not what you're here for, I don't think.

Hence, some commented on the need to balance potential benefits of providing support to maintain patients' motivation and confidence, against risks of patients coming to depend on an unsustainable level of support.

> [14] Certainly with the older ones, there is a slight dependency trap with them, they do still like to come when you are there. It's quite hard sometimes to get them to exercise on their own.

8.2.2.2.2 Identifying and overcoming patients' anxieties

According to most professionals, a large part of developing skills and confidence involved helping patients to overcome anxieties about the exercise environment. Anxieties were typically seen as arising from feeling intimidated by fitter exercisers, fears about an unfamiliar environment, or fears about being exercised beyond comfortable limits.

> [1] They are just worried about what people will think of them, they think the people there, everyone there is going to be fit, in their lycra and looking really smart but so that's the main thing, they just don't know, it's the fear of the unknown, they don't know what we are going to do with them

For some, identifying and overcoming anxieties began in the time between referral and scheme entry, with implications for scheme uptake. One professional for example commented that advertising patient-only classes during initial telephone contact had led to good responses, whereas another talked of sending out information packs about what the scheme would entail prior to scheme entry, in response to feedback on the anxiety-provoking nature of scheme entry.

> [22] They say yes it's quite daunting coming into the leisure centre for the first time, they're not too sure what they are going to be doing ...so we are trying to design a leaflet now which we are going to put out with the card itself saying exactly what they are required to do.

Initial consultations were often cited as a key opportunity to overcome anxieties, through reassuring patients that professionals would serve as a familiar point of

contact, as well as offering assurance that patients would not be expected to do anything they were uncomfortable with. Highlighting at this stage that the patient would be surrounded by other patients in the same position was seen as playing a substantial role in assuaging anxieties.

> [39] They're very often afraid of the gym so we try and take away those barriers by being beside them in the gym for the first couple of weeks. We explain that we're going to be there and it's going to be a regular familiar face. They're quite reassured to know that whoever else is with us, are in the same position as they are.

Some described patients' anxieties as arising from preconceptions about the exercise environment, which would largely dissipate if the professional accompanied patients and allowed them to see that concerns were unfounded.

> [14] It's a perception that people have, that it's going to be full of very fit people and they are going to be um, you know, you know, sort of feel out of it or not be able to cope with it or anything. Once they have had a go and they realise that they can do it you know, it's not nearly as difficult or as hard as they initially feel it is going to be and then they are fine. It's just their own barriers that people put in place.

However, anxieties were commonly seen as lasting to some degree throughout the programme. Hence, professionals typically commented on a need to structure classes in a manner which made patients feel at ease, through methods such as the aforementioned patient only group classes. Whilst most areas limited group classes to patients, supervised gym sessions were typically held during public opening hours. In order to make gym use less daunting therefore, professionals commonly talked of booking sessions during quiet times, in order to reduce the number of fitter exercisers in the gym.

> [7] We try to look at trying to get them in during quieter times when the age range is more mixed or more for their age range, as opposed to coming in at peak time when there are a lot of students here working at a lot higher intensity than they would be.

Some areas however offered patient only gyms or closed the gym to other users during NERS sessions, allowing the entire programme to be exclusively conducted in the

presence of other patients. This model was seen as highly beneficial in promoting programme adherence.

> [3] It is an exclusive gym for that type of client, it tends to make it easier for the person to sort of integrate, and come in on board, whereas in the mainstream gym, they may have pre-conceived ideas about who's going to be in there, how many's going to be in there, what they are going to do being, so that tends to sort of put them off really.

This was however sometimes seen as a short-term solution, with patients commonly seen as anxious about progressing beyond this environment after the programme.

> [1] The negative side is they don't want to go into the main gym, so it's kind of, they're wrapped in cotton wool because they've got their own which is great for the first sixteen weeks, but then we do have to try and push them on slightly, to integrate them in the main gym.

8.2.2.2.3 Group processes and emerging social support networks

All professionals commented that other patients played a key role in supporting one another's adherence, often speaking of the empathy patients offered to one another, having been referred for similar reasons and suffering similar limitations, destigmatising the process of struggling to overcome illness whilst attending classes.

> [7] It's knowing that they are not on their own really the group tend to motivate themselves, and they will talk about what works for them, if they are having a bad day they will say they're having a bad day, they don't seem to be intimidated, they have got problems, it seems to be quite natural and they are not on their own.

Experiencing the scheme with other patients was seen as providing patients with realistic exemplars. New patients could see others at later stages of their programmes doing things they weren't yet capable of, and could be encouraged to believe that they too were capable of achieving those improvements in time, rather than having confidence undermined by comparison with healthier mainstream exercisers.

> [40] They realise they're not the only person that's put weight on for example, that feels uncomfortable coming in. That there's several people that come into the gym and they say 'I've had this, I've had that' whereas not everybody wants to say why they're there but they're sort of 'look at me now' and you know 'I'm doing this and I'm doing that' and it sort of helps them to think that they can achieve their goals.

Whilst some saw social support networks as emerging spontaneously through the programme's group structures, others saw explicit efforts to foster social interaction as key to avoiding the aforementioned 'dependency trap', with emerging social support allowing the professional to withdraw over time.

> [29] I find that as an instructor they tend to follow you around 'will you be teaching the session'. 'Oh I'll miss that one then and wait for you to come back. You tend to get people like that ... I try and partner them up with a stronger person that I know. It's like putting them with someone in a similar situation that will also give them support.

Several talked of organising regular social events for patients' past and present or of 'buddying' patients with others of similar ability. In one area, professionals had agreed a policy which ensured that patients finished the programme in clusters, and were guided into maintenance classes together.

> [31] I never finish one on their own, even if it means that I keep them after assessing them another couple of weeks until someone else is exiting the scheme. I always try and buddy them...because we're feeding into maintenance classes, we find that a lot easier because there's the one's who've come off the scheme maybe a month before them there and then you've got maybe four more going in, and so they're really just the same group but at a different timeslot.

Whilst many commented that patients often continued to exercise after the programme with friends they had made, some remarked that loss of social aspects of patient classes, and a need to reintegrate into a different exercise environment after the scheme, were key reasons why some struggled to maintain physical activity levels after the 16 week programme.

> [26] They love the people that they're with and they feel comfortable in that surroundings. And they're obviously feeling comfortable with myself, so you know what it's like, they don't like. A lot of people don't like change do they?

Concerns about patients' inability to progress without ongoing support of the professional or of the patient-only environment had led to widespread tendencies for allowing completers to access NERS classes indefinitely, rather than progressing them on at 16 weeks.

> [30] It's just a shame that we have to let them go at 16 weeks. But as I say, they've all got classes to go into and I wouldn't let anyone go who I thought was incapable of being by themselves, which is why I have kept a few on, and they probably won't ever leave. But that's fine. If I didn't do that I wouldn't be safe in the knowledge that they'd be out by themselves.

8.2.2.2.4 Adequacy of training and perceived training needs

Professionals typically spoke highly of the opportunities for professional development offered within NERS. Whilst often commenting on having struggled to keep on top of the workload associated with the volume of training in the schemes' early stages, the level of training provided was often seen as increasing patients' confidence in the programme.

> [9] People like to know how much are you trained and when you say, the training courses you have been on they are like really impressed. Do you see what I mean? You know private gyms don't have the things that, well they might do if you pay for a personal instructor but you know, just your bog standard me person.

Many however described a need for increased emphasis on dealing with the diverse clinical populations in their training. Approximately a quarter of professionals expressed a desire for additional in depth training in specific conditions, with some commenting that having a team of professionals specialising in specific conditions may be more effective than a system where all were expected to deal with all conditions.

> [19] Well I believe that there is going to be Stroke COPD, back care um, mental health, so these are done modules are going to be critical because we do feel that at the moment we lack in the knowledge in those areas.... we want to have like a champion if you know what I mean, of a particular condition so that they can be the mentor for that condition then.

Some areas had topped up professionals' training with additional nutritional qualifications to enable them to support patients' dietary change efforts, whereas in others, professionals expressed a desire for further nutrition training.

> [4] Um maybe we could have done with a bit more on nutrition I think we'd all like to see a bit more of that, and I think maybe the scope to go on some

> specific courses like maybe maybe falls prevention, or stroke, or mental health or that type of thing I'm hoping that maybe come in the future now.

In addition, a need for further training in supporting the specific needs of mental health patients was expressed by several professionals.

> [14] If they are depressed and you have the days you don't feel like coming, you are not going to come. You know again, the mood thing, their barriers as well are harder to break down. So a little bit more training in that area would be useful.

8.2.3 Findings: patient interviews

8.2.3.1 Entering the scheme: routes into NERS and motivations for attendance

Some patients commented that they had entered NERS largely because their health professional had advised them to attend.

> [Female aged 63, centre 1, 4 weeks] I had a catastrophic accident – my knee – about eighteen months ago, so I've been having surgery and physiotherapy following those. They [physiotherapists] wanted me to come.

However, many others described having made an independent decision to become more active and having initiated referral themselves. In some centres, a majority of patients were referral seekers, with patients commenting that health professionals had been unaware of the scheme until made aware of it by patients.

> [Female aged 56, Centre 4, 8 weeks] I was looking for something to do exercise-wise, I do swimming, but I thought this was great to do a bit more exercise and keep yourself more fitter you know. So I got in contact with the leisure centre, and then through the leisure centre through my doctor....he didn't really know much about it at first, because I think it was just starting up...but anyway he got the form and signed it.

Referral seekers often reported a history of activity interrupted by illness, injury or family circumstances. Some had attempted to join a gym independently but had been refused membership due to health conditions, and saw NERS as a means of overcoming this barrier to achieving a self-determined desire to become more active.

> [Female aged 48, Centre 2, 11 weeks] I explained that I had blood pressure and that I was diabetic. So they suggested that I go back to my doctors and ask for a referral which I did, and I haven't looked back since. But I intended to go to the gym anyhow.

Some, who described entering following health professional advice, focused primarily upon health and fitness related goals suggested by their GP.

> [Female aged 44, centre 5, 9 weeks] To keep as fit as I can, you know, because my GP today said, she thinks that I'm really, you know, I keep as active as I can with all the complaints that I've got, she's happy about that.

However, several patients advised to enter the scheme also reported a history of activity, often having made independent attempts to return to an active lifestyle. Such patients sometimes reported being advised to enter the scheme to treat injuries which prevented a return to activity, because efforts to exercise without supervision had not been effective.

> [Male aged 74, centre 5, first day] So when I started doing the exercises myself, I didn't know whether it was right or wrong... He says go up and get it done properly, because my whole body is starting to seize up.

Patients commonly cited multiple motivations for attendance, including physiological improvements, such as reduced blood pressure and weight loss. However, both referral seekers and those advised to attend commonly also linked behaviour change to personal values such as playing a pro-active role in eradicating illness.

> [Female aged 37, centre 4, first class] I already feel better about myself...we are actually doing something, rather than just taking a pill.

Patients of retirement age commonly emphasised maintenance of autonomy and ability to perform everyday activities as principle motivations for entering NERS

> [Female aged 82, centre 5, first class] My main aim is to stay fit, and active, and mobile. I don't mean run a marathon or anything like that, day to day mobility. In charge of everything I want to do – gardening, household things, all that is extremely hard for me now.

For younger patients, primary motivations commonly centred around maintaining occupational functioning or returning to work, with many having entered the scheme due to injuries which prevented or threatened to prevent them from working.

> [Female aged 43, centre 5, first class] With me, I'm the only wage earner, my husband had two MIs two years ago and kind of forgot to take out protection, so I have to keep five people, I cannot stop work.

8.2.3.2 Experiences of the exercise programme

8.2.3.2.1 Perceived impacts

Although some patients were at very early stages, many who had been attending for several weeks were beginning to perceive progress towards their goals. Some, highlighted improvements in mood and increased social contact as valued outcomes.

> [Female aged 51, centre 4, 8 weeks] It's good, as I say, it's getting me out of the house. Like this morning, I've come in not very well, but by the end of the class I was having a laugh and a joke with everybody, so it does work.

Others highlighted medical improvements, such as reduced breathlessness and blood pressure, increased mobility or reduced pain.

> [Female aged 48, centre 2, 11 weeks] I've got arthritis in one of my knees and I used to find it difficult in the morning, would be stiff and aching, but now no problems.

The one area of disappointment among several patients was however a lack of impact upon their weight. However, one such patient described how as she experienced the benefits of feeling fitter, weight loss became peripheral to other benefits in motivating adherence.

> [Female aged 61, centre 3, 13 weeks] I don't worry so much about that now, because being fitter is better. I think the weight loss will come, but it hasn't as yet...I'm hoping it will go, but I certainly feel fitter.

8.2.3.2.1.1 Roles of the exercise professional in facilitating change: skills development, confidence and motivation.

Most patients spoke highly of the expertise of professionals and the support offered. Patients in the early stages of their programme highlighted the value of support of the

professional in guiding them through use of machinery, building confidence in using unfamiliar equipment safely.

> [Female aged 82, centre 5, first class] They are intimidating, well, I speak as an older person, I know some people who think, no way am I going there to try those machines, seeing people do what they do, you think, I can't do that.

As well as fears regarding use of machinery, fear of harm through exercising beyond one's capabilities was cited by some, particularly amongst patients at the beginning of their programme. Having an exercise professional nearby for advice and support was reassuring to patients who feared that left to their own devices, they would overexert themselves.

> [Female aged 43, centre 5, first class] The thought that there would be somebody who could actually advise me on what to do, so I didn't knacker myself...I wouldn't dare try it by myself. Because every time I've tried it by myself I've ended up overdoing it.

Whilst some valued the role of the exercise professional in holding them back from overexertion whilst they developed the confidence to exercise without causing harm, others highlighted the value of motivational support which enabled them to push themselves harder than they would have by themselves.

> [Female aged 69, centre 3, 8 weeks] Well – when I first came, as I said, I wasn't walking very far – we were supposed to be walking around the hall, which wasn't walking very fast or very far, so [exercise professional] grabbed hold of my arm and said 'come on, come for a walk with me' and he was walking around the hall with me – that's the sort of support you get.....They always check to see that you're all right, that you are not going to over do it and cause yourself an injury or whatever, they are always checking on that.

Professionals were commonly seen as effectively monitoring progress and putting in steps in order to ensure further progression, although some patients at later stages of the programme commented on a need for more or different forms of exercise in order to progress, with catering for patients with diverse fitness levels in the same class sometimes making it difficult to tailor exercise to everyone's needs.

> [Female aged 58, centre 4, 8 weeks] I'm puffed out by the time I've finished it, so clearly, I don't need that much more, but I need something more. I think you're right or certainly I need a lot of women need something to be doing the tummy muscles

8.2.3.2.2 Roles of other exercisers in facilitating change: empathy and social comparison

Patients placed substantial emphasis upon the importance of shared experience with other patients in a similar position. Many highlighted the value of being surrounded by others who were going through a similar experience and were able to demonstrate empathic understanding of the challenges they were facing.

> [Female aged 4, centre 5, first class] That's another thing we've got in common, unless you've got it, you do not know what it's like and how it gets you down, and how it just never goes away.
> [Female aged 51, centre 5, first class] And I think talking to other people who actually know what it feels like helps.

Patients at early stages of the scheme often commented that interaction with patients at later stages provided positive exemplars of the benefits of participating, enhancing beliefs that they could achieve valued benefits through continued participation.

> [Male aged 51, centre 2, 2 weeks] And he's being doing it – I don't know, a month or so, and he said he felt really – the more he's doing the fitter he's feeling. There's two ladies there – they were the same, you know, they were saying a few months ago they couldn't walk to the shops and back. To see them yesterday and today there – you know – great – and I'm thinking, if those can do it, you know, there's hope for me, no disrespect to them.

Exclusion of more intimidating groups, alongside simultaneous emergence of support from other patients, was seen as beneficial in helping patients to overcome anxieties about the exercise environment.

> [Female aged 68, centre 3, 13 weeks] When you come into the gym, it's this group, you don't feel out of place. We all know one another and we are all more or less in the same sort of situation. Whereas when you go on, full of gleaming bodies, and everybody's superfit, you just, it doesn't help, you don't do it.

8.2.3.2.3 Opinions of the exercise classes: accessibility, flexibility and types of exercise

Whilst many patients appeared happy with the range of activities offered, in areas with more restricted programmes, some expressed a desire for greater diversity of exercise types.

> [Female aged 59, centre 3, 13 weeks] It would be lovely to have a swimming session – it would be bliss to come out of there and go swimming, but it's schools all day, isn't it?

Furthermore, in areas where the scheme was offered in a limited number of centres, or where large distances between centres were cited, many stated that accessing the programme involved substantial travelling time, contingent upon access to a car. Class times, whilst cited by some as flexible and convenient, in other centres were cited as too inflexible to allow working patients to access two classes per week.

> [Female aged 53, centre 4, 2 weeks] I need different times, you know, that's what I do need. And to just do one class, it's just not enough.

8.2.3.3 Leaving the programme: the transition to independent activity

Some patients identified clear plans for how they would maintain increases in physical activity, sometimes describing having already identified exit route classes which they planned to enter, or focusing upon a desire to return to former hobbies such as horse riding or walking. Such patients were most commonly those with a prior history of activity, some of whom had entered NERS as a means of overcoming the refusal of gyms to accept them as members.

> [Female aged 48, Centre 2, 11 weeks] I'm going straight into the other scheme then where you pay 20 pounds a month and try all different classes then and I'm hoping to try a few different things then.

Many others however, who were perhaps more dependent on the programme to provide ideas for long term maintenance of activity, seemed more unsure how they might maintain change. In some cases, formation of action plans appeared to have been hampered by a lack of information about available options after the programme. When asked how they planned to maintain activity, in almost half of interviews, patients responded by asking interviewers for information on exit routes.

> [Female aged 58, centre 4, 8 weeks] I think we need to know what's going to happen after the sixteen weeks. We don't know so we're feeling a bit in limbo really. What is going to happen?

Some expressed concerns that without a firm commitment to an agreed time and place, other aspects of daily life would crowd out time for physical activity, with NERS providing justification for taking time out from other commitments, but this time becoming harder to protect after programme completion.

> [Female aged 68, centre 3, 13 weeks] I'm not that confident actually because it is very easy to let it slide, especially with everyday life. When you come here, you have to leave home to come and do it, and everybody realises this. When you are at home, I find, people find things for me to do, and you know, you just, yourself is not, well I'm not saying not considered, but you are busy.

Amongst patients in the early stages of the programme, some expressed concerns regarding whether they would be able to maintain increases without ongoing support from the professional, in terms of continued guidance on how to exercise safely or continued motivational support.

> [Female aged 24, centre 5, second class] I need her quite a bit around. But maybe as I get through, I'll get used to it. And not really needing her as much. I don't know, it's just that – being able to say well 'do you think I could do this, or do that' whereas when that finishes, you know, you won't have anyone just to have there.

Amongst patients nearing the end of the programme perceived challenges commonly included loss of social support from other patients and enduring anxieties about the environment. Having other people to exercise with was described as reassuring to patients who felt that exercising alone was unsafe or that they would want another person there if they were to hurt themselves. Many described having others to exercise with as essential if they were to maintain motivation.

> [Female aged 58, Centre 4, 8 weeks] I think the motivation, for me, coming to a class is fine, and I can do that... if I had to get up and do it myself I wouldn't bother, I know I wouldn't because I tried it before.

In addition, some cited enduring anxieties about the exercise environment as a likely barrier to remaining active. One programme completer commented that initial anxieties about the environment had faded as the scheme progressed.

> [Female aged 55, Centre 2, 16 weeks] I said I'm not going up there [to the gym] I'd be shamed. But [my daughter] encouraged me to do it, and [exercise

> professional] I met, and she was so nice and made me feel so relaxed and looked after me, I joined, and then I got to know everyone you know....they are so nice, they make me feel part of it then, you know what I mean, I don't feel embarrassed coming in here at all, I actually enjoy it now.

Others however expressed reservations about moving into mainstream exercise settings. The following quote was from a participant entering the final quarter of her programme in a centre where all activities, including gym use, were held in areas closed off to users other than NERS patients.

> [Female aged 59, centre 3, 13 weeks] I read an article recently which said I am not going to go into a room where everybody's there because they don't want to look like me, and that's true, they were all there because they don't want to be overweight, they don't want to be unfit, and it's embarrassing, and it just makes you feel worse. So that's the huge benefit of this class, and I'm still not ready to go back, any time of day, into one of those gyms.

Additional concerns were raised by some about the cost of remaining active after the programme discount was withdrawn.

8.2.4 Summary and implications

Interviews with exercise professionals and patients offered a number of key insights into the perceived active ingredients of NERS and emerging social patterning in participants' experiences of them. This discussion draws together findings from these interviews, before Section 8.3 moves onto quantitatively assess emerging hypotheses regarding social pattering in uptake and adherence.

8.2.4.1 Entering the scheme: the referral process and individual differences in patient needs

In talking about the referral process, debate emerged amongst professionals as to whether this should seek to motivate patients, or should start from the assumption that patients will only change where already motivated to do so. Whilst Markland and Tobin (2010) comment that motivation is likely external on scheme entry given that it is based on recommendation from an authority figure, professionals commented that rather than acting on instruction, many patients had actively sought referral, and that self-referring patients were more likely to adhere. Hence, the process of internalising exercise motivations was seen as more likely where motivations were already

somewhat internalised at baseline, with health professional advice not typically seen as a strong active ingredient in promoting change.

Some previous studies have focused upon identifying patients' stage of change prior to offering primary care based physical activity promotion (Kirk et al., 2003; Bennett et al., 2008). This is however a controversial approach, in terms of the ethics of offering a scheme only to one group of patients based upon a subjective judgment of motivation (Mutrie and Woods, 2003). Rather than a fixed trait, motivation is a fluctuating product of social interaction (Miller and Rollnick, 2002), and the appearance of being more or less ambivalent may reflect the manner in which the offer of referral is presented to the patient. Hence, interpreting resistance as a sign that the patient is unwilling to change may lead to the scheme being withheld from patients who would benefit. As recommended by Department of Health Quality Assurance Frameworks (Department of Health, 2001), and discussed throughout Chapter 7, it may be that with future integration of effective motivational communication strategies into baseline consultations, the scheme may offer greater effectiveness for patients who are initially more ambivalent about change.

Consistent with reports of professionals, patients were divided as to whether they reported seeking the scheme for themselves or were advised to take part by a health professional. In some cases, consistent with one previous study (Wormald and Ingle, 2004), patients commented that health professionals had only become aware of the scheme after patients requested referral. Referral seekers often reported a history of activity, interrupted by health conditions or circumstances, with the opportunity to enter NERS sometimes simply removing barriers which had thwarted self-determined change efforts, such as the refusal of gyms to accept members with health conditions. Whilst some patients advised to enter by their GP identified motivations only at a medical level, or referred to goals that their health professional had set for them, many still linked change to more personally valued goals. Several, for example were referred to overcome injuries which prevented a return to a previously active lifestyle. Referral to NERS was in such instances however often seen as a means of holding patients back from causing themselves harm through attempting to return to activity without guidance, rather than a means of motivating them to take up exercise.

Consistent with previous research, many older patients focused upon maintaining autonomy and ability to perform everyday activities or participate in valued leisure pursuits (Stathi et al., 2004). By contrast, occupational functioning emerged as a principle motivation amongst many younger patients, with many wanting to return to or remain in work. Hence, for many, health conditions were beginning to impinge on central life goals (Bodenheimer and Handley, 2009), with increasing activity seen as a means of minimising this impingement. Given the sharp decline in attendance in the early weeks of the programme (see Chapter 6), it is likely that many patients for whom entry was less clearly linked to internal motivations did not attend beyond the programme's early stages, and that patients interviewed during participation represented a more internally motivated sub-sample of referred patients. However, a diverse range of motivations were expressed, contrasting with the narrow range captured within the goal setting processes described in Chapter 7.

In addition to describing patterning in responses to NERS by baseline motivation, professionals offered a range of insights into perceived patterning by medical and socio-demographic factors. Professionals commonly described mental health patients as needing support to maintain motivation and foster social interaction, describing higher levels of anxieties about entering the scheme and assimilating into the exercise environment amongst these patients. Given the association between social assimilation and internal motivation (Markland and Tobin, 2010), where anxieties act as a barrier to assimilation, this may undermine motivation. However, several professionals commented that where were supported and encouraged to interact with other patients, mental health patients appeared to benefit substantially from the programme. Hence, the degree of professional support and opportunity for social interaction within NERS may have gone some way towards overcoming the commonly cited tendency for lower adherence in mental health patients (Dugdill et al. 2005; Crone et al. 2008; James et al. 2008). However, given the perceived contingency of adherence amongst mental health patients on the interpersonal skills of the professional, it may be that additional specific training is needed by some professionals to support adherence in mental health patients, with some highlighting this as an area of training need.

Older patients and women were also seen by professionals as often exhibiting additional anxieties on scheme entry. However, these patients were also seen as

benefiting most from social aspects of classes, as most patients were also older female participants, with assimilation into this group network perhaps easier for patients who fitted this demographic. The explicitly group-based structures may have gone some way towards offsetting the tendency for lower adherence amongst female participants observed in some ERS (Dugdill et al., 2005; Gidlow et al., 2007; James et al., 2008), though may also have exacerbated the tendency for higher adherence in older patients, through providing an interpersonal context favouring these patients. A large proportion of professionals also commented that the scheme was better received and hence easier to implement in more affluent areas, with lower perceived engagement attributed to factors including lower motivation, limited financial resources and limited buy-in among GPs in areas of deprivation. Findings on socioeconomic patterning in uptake and adherence of ERS are at present equivocal (Gidlow et al., 2007; Sowden et al., 2008).

8.2.4.2 Active ingredients of the exercise programme

Patients and professionals both discussed professional support as an important aspect of the programme in facilitating behavioural change. Professionals saw their role as involving overlapping tasks of education and social support in order to foster increases in knowledge and confidence, whilst supporting motivation, with talk of knowledge, confidence and motivation interwoven and these factors seen as developing in parallel. Patients highlighted the importance of a knowledgeable instructor, who educated them on using unfamiliar equipment and exercising within the limits of their conditions, whilst providing reassurance, support and supervision to allow them to feel confident putting knowledge into practice. Consistent with findings reported by Riddoch and colleagues (Riddoch et al., 1997), professional support appeared particularly important for patients who were anxious about exercise, with many commenting that without the level of support offered, fears regarding exacerbation of illness or injury would have prevented participation.

Consistent with previous research demonstrating the importance of assimilation into the exercise environment (Wormald and Ingle, 2004; Wormald et al., 2006), professionals placed substantial emphasis upon their own role and that of other patients in helping patients feel at ease in the exercise environment. Professionals identified initial contact about joining the scheme as an opportunity to reassure

patients both that they would accompany patients into the exercise environment and that the patient would be exercising with others of a similar ability and background. Similarly, patients described one another as providing an empathic context for change efforts, assuaging anxieties and helping one another to assimilate into an otherwise intimidating environment. Given that patients were often referred for similar reasons, and faced similar challenges, they were likely able to offer an accurate empathic understanding of one another's positions (Rogers, 1959), as well as providing role models, demonstrating what could be achieved through sustained participation. This in turn likely contributed to a sense of competence which might have been thwarted by unrealistic comparison with fitter exercisers. Given the previously reported links between social assimilation and development of internal motivation (Markland and Tobin, 2010), the centrality of self-efficacy and competence to internal motivation (Deci and Ryan, 1985) and the value of empathic and non-threatening social contexts in facilitating behavioural change (Miller and Rose, 2009), patients likely played a significant role in increasing one another's internal motivation, self-efficacy and behavioural change.

Whilst most patients spoke positively of NERS, a number of centre specific limitations were highlighted. Some cited travelling substantial distances which could only be covered by car, likely limiting access amongst those without a car. Furthermore, consistent with previous qualitative research (Wormald and Ingle, 2004; Taylor and Fox, 2005), working patients cited difficulties with limited class timings. Whilst many appeared satisfied with the activities available, patients in areas offering a more restrictive range of class types expressed a desire for a wider range of alternatives, whilst some commented on needing more or different exercises to achieve their goals as they became fitter, with meeting the needs of all abilities in one class seen as challenging.

Many patients who had attended NERS for several weeks cited medical improvements, ability to do more or walk further. Consistent with professionals' reports, mood lifting benefits of the scheme were cited by some patients referred for mental health reasons, with additional social contact often valued by patients as an end in itself. Two recent studies demonstrate mental health benefits of group-based ERS programmes, although neither impacted physical activity, perhaps indicating that such

benefits are more strongly mediated by social processes than by exercise itself (Isaacs et al., 2007; Mutrie et al., 2007). However, one key area of disappointment for several patients was failure to lose weight. Given the intensity and volume of activity within NERS, substantial weight loss is perhaps unrealistic, and de-motivating impacts of unrealistic expectations ought to be carefully considered (Jones et al., 2005). One patient described weight loss as becoming less important over time, with motivation becoming strengthened by emerging valued changes. However as described in Chapter 7, it is perhaps likely that where weight loss is a sole motivation, non-achievement will lead to de-motivation. Focus upon weight as a distal goal, with principle focus on outcomes more achievable in the short term may more effectively promote participation.

8.2.4.3 Maintaining change beyond the programme

Whilst often describing changes during the programme, in discussing maintenance beyond the programme, professionals expressed concerns about a perceived dependence of some patients on professional support and the group environment, often commenting on a tendency to become valued components of patients' support networks. As reported by Wormald and Ingle (2004), patients commented that a structured programme provided commitment to exercise, leading to activity that patients would not have participated in at home, with some commenting that the scheme provided justification for setting aside time for oneself and taking time out from other commitments (Schmidt et al., 2008). However, for many, loss of the programme was expected to allow constraints such as family commitments, which had taken priority over physical activity before the scheme, to reassert control and to return activity to baseline levels, consistent with socio-ecological models, which suggest that interventions focusing upon change in individual determinants will often be thwarted by enduring social and environmental constraints which gave rise to sedentary behaviour prior to intervention (Stokols 1996).

Some patients commented that they would struggle to maintain changes without support, feeling that they may still be too afraid of aggravating illness or injury, or might struggle to maintain motivation. Hence, self-efficacy and motivation were in some instances contingent on professional supervision, consistent with a recent study in which patients expressed disappointment at being 'dropped' after the programme

(Isaacs et al., 2007). However, concerns about loss of professional support were typically limited to patients in the early stages of the scheme. Although patients exhibiting higher levels of initial dependence might have dropped out soon after being interviewed, dependence on professional support likely faded as confidence increased. The scheme trialled by Isaacs et al. (2007), whilst offering full professional supervision, offered only a 10 week programme, perhaps allowing insufficient time for confidence to develop before withdrawing the scheme. As described in Chapter 6, the NERS programme averaged almost double this duration, at 19 weeks. Furthermore, as described in Chapter 6, patients were typically not 'dropped' after the programme but allowed to continue attending NERS classes indefinitely, with only the discount withdrawn, and even this replaced by a lower level discount.

As with professional support, the creation of a patient-only environment was seen by some professionals as engendering dependence, with assimilation into this environment sometimes seen as a barrier to moving patients on towards alternative environments. Hence, the habituation of exercise into patients' daily lives anticipated by national implementers (see Chapter 6) was in many cases contingent upon continuation of the programme. Key emerging strategies to lessen dependence on the structures of the programme included active fostering of long term social support networks through filtering of completers into maintenance classes or exiting patients in clusters, with emerging social networks taking over the role of the programme in motivating adherence. As described above, a widespread tendency emerged for patients to be offered indefinite access to NERS classes (see Chapter 6).

In patient interviews, concerns relating to anxieties about moving into a mainstream environment, and loss of social networks were often expressed by patients coming to the end of their programme. Although some reported that anxieties faded over time, others felt that they would struggle to move into a general gym population after the scheme. Whilst most areas offered patient-only classes and supervised gym use, with gym use taking place during general opening hours, concerns appeared particularly prominent in one centre where all sessions were conducted in private. Higher anxieties amongst patients in this more exclusive system may indicate simply that such patients would not have attended under a less exclusive system. However, professionals operating within centres with fully exclusive systems commented that whilst

facilitating adherence, it often became difficult to move patients beyond this environment after the programme. It may be that such a system offers shelter from anxieties about mainstream environments, but does not help them to be overcome. Given the anxieties expressed regarding entry to a mainstream exercise environment, exit route options outside of leisure centres may be crucial. As described in Chapter 6, most though not all areas offered exit routes including walking groups and community based classes in addition to centre-based options.

Nevertheless, some patients described plans for post-scheme activity, having identified exit route classes and memberships some weeks before completing the scheme. Consistent with studies showing strong associations between past behaviour and subsequent physical activity (McAuley et al., 2007), confidence often appeared highest amongst those who reported a history of exercise. Others appeared more uncertain of how they might remain active, in part due to a lack of information about post 16 week options. This perhaps arose largely from the evolving nature of the scheme, with exit routes being negotiated at the time of interview. This lack of information may however have made sustained change more difficult through inhibiting formation of implementation intentions (Latimer, Ginis and Arbour, 2006). Some also cited difficulties with prospects of meeting increased costs of exercise after the scheme, although as reported in Chapter 6, most areas responded to such concerns by negotiating local post-scheme discounts.

In summary, interviews with professionals and patients suggested that health professional advice may play a limited role in promoting change, unless patients have already begun the process of linking change to their own internal motivations. As described in Chapter 7, delivering activities such as motivational interviewing and effective goal setting on scheme entry may lead to greater impacts among patients entering on advice of a health professional. Programmes dealing with patients referred for mental health reasons should ensure that professionals are well trained to provide the support necessary to help patients overcome additional intrapersonal barriers and assimilate into the exercise environment.

The high dose of professional support in NERS appeared crucial in supporting skills development, confidence and motivation. Similarly, patient-only classes appeared to

be a valued programme component, with patients providing an empathic context and realistic exemplars for social comparison, supporting one another's confidence and motivation and assuaging anxieties about the leisure centre environment. However, this social context was seen by some as favouring older females given their overrepresentation. Additional patterning in responses to the programme were seen as arising from challenges accessing the service amongst working patients and those without transport, with uptake often perceived to be lower in more deprived areas.

Furthermore, there is a need to focus upon understanding how patients' transition from the scheme ought to be best supported to avoid dependence on the structures of the programme which perhaps led to the widespread tendency for an unsustainable level of support to be offered to many NERS patients. Given the anxieties expressed about the prospect of entering mainstream gym environments, exit route options beyond leisure centres should be provided. The effectiveness of emerging strategies to support this transition, such as filtering patients into maintenance classes or exiting them in clusters deserves attention.

8.3 Social patterning in reach

The qualitative data described in Section 8.2 provides a range of insights into perceived active ingredients of NERS, as well as individual differences in responses to the programme. The aim of this section is to explore some of the hypotheses generated throughout Section 8.2 and to quantify variability in uptake and adherence by patient characteristics. A first emerging hypothesis was that scheme adherence would be higher among patients entering the scheme with change more internally motivated. Though no measure of baseline motivation was available, patterning by baseline physical activity level will be explored. Mental health patients were also seen as experiencing additional barriers to assimilating into the exercise environment, whilst the group environment was seen as favouring older patients and female patients, due to their overrepresentation. In addition, patients of working age described challenges accessing the service due to limited class timings. Hence, patterning by reason for referral, age and sex will be explored. Furthermore, accessing the service was often seen as contingent on access to a car, whilst professionals commonly cited lower engagement amongst patients in more deprived areas, and hence associations of car ownership and deprivation with scheme reach will now be explored.

8.3.1 Data sources and analysis

This section involves secondary analysis of routine monitoring data and baseline questionnaire data for the 1080 patients referred to NERS and randomised to the intervention group. Age, deprivation, sex and baseline physical activity measures were derived from baseline questionnaires, with reason for referral and scheme entry/completion variables obtained from routine monitoring data.

The frequency and percentage of patients who i) did not enter following referral, ii) entered but dropped out in advance of completion and iii) completed NERS are presented by each level of each patient characteristic. Binary logistic analyses are subsequently used to examine associations with uptake (i.e. attendance or non attendance at a 1st appointment) and with adherence (i.e. attendance or non attendance at a scheme exit appointment). Variables demonstrating significant associations with uptake and adherence in univariable analyses are analysed in multivariable models.

8.3.2 Findings

8.3.2.1 Characteristics of patients referred to NERS

As indicated in Table 14, more than half of patients reported no physical activity at baseline, whilst approximately 1 in 6 were moderately inactive and a further 1 in 6 moderately active. A small number were already active on referral to NERS. Patients were predominantly female, whilst three quarters were car owners. Mean (and standard deviation) age was 51.9 (14.5) years, ranging from 16 to 88. Weight management was the most common reason for referral, followed by non-weight related CHD risk factors (e.g. diabetes, high blood pressure) and mental health.

Table 14. Sample description for patients within the intervention arm of the NERS trial

		All intervention patients
Baseline activity	**Inactive**	623 (57.7)
	Moderately inactive	170 (15.7)
	Moderately active	181 (16.8)
	Active	79 (7.3)
Females		708 (65.6)
Car owner		772 (72.1)
WIMD tertile	**Low (2.3 to 13.1)**	361 (34.4)
	Medium (13.2 to 26.2)	358 (34.1)
	High (26.2 to 81.0)	330 (31.5)
	Non-weight related CHD risk factors	457 (42.3)
Reason for referral	**Weight management**	786 (72.8)
	Mental health	310 (28.7)

8.3.2.2 Patterning in uptake

As indicated in Table 15, patients who were already active at baseline were slightly more likely than all other patients to enter NERS. Women and older patients were slightly more likely to enter NERS. Higher levels of scheme entry were observed amongst patients living in the least deprived area. Almost twice as many non-car owning patients did not enter the scheme compared to car owners. Patients referred for mental health reasons were slightly less likely to enter NERS than patients referred only for physical health reasons.

As indicated in Table 16 however, the only significant correlates of scheme uptake were car ownership and area level deprivation. Those living in moderately deprived areas were significantly less likely to enter than those in the least deprived areas, whilst contrasts between the most and least deprived areas were significant at the 10% level. Car owners were significantly more likely to access the service than non-car owners. Whilst associations of deprivation with uptake are reduced where entered alongside car ownership, a marginally significant difference between low and medium deprivation areas remains.

Table 15 Frequencies and percentages of patients within each independent variable category achieving each level of adherence to NERS

		Did not enter	0-16 weeks	Completed
Baseline activity level	Inactive	97 (15.6)	274 (44.0)	252 (40.4)
	Moderately inactive	27 (15.9)	61 (35.9)	82 (48.2)
	Moderately active	24 (15.4)	64 (35.4)	93 (51.4)

	Active	8 (10.1)	36 (45.6)	35 (44.3)
Sex	Male	64 (17.2)	140 (37.6)	168 (45.2)
	Female	97 (13.7)	306 (43.2)	305 (43.1)
	Low (16 to <40 years)	55 (17.0)	160 (49.4)	109 (33.6)
Age	Medium (41 to <59years)	55 (15.1)	149 (40.8)	161 (44.1)
	High (60 years+)	46 (13.0)	121 (34.1)	199 (53.0)
WIMD tertile	Low (2.3 to 13.1)	42 (11.6)	160 (44.3)	159 (44.0)
	Medium (13.2 to 26.2)	63 (17.6)	141 (39.4)	154 (53.0)
	High (26.2 to 81.0)	53 (16.1)	128 (38.8)	149 (45.2)
Car ownership	No	63 (21.1)	117 (39.3)	118 (39.6)
	Yes	97 (12.6)	324 (42.0)	351 (45.5)
Weight management	No	49 (16.7)	122 (41.5)	123 (41.8)
	Yes	112 (14.3)	324 (41.2)	350 (44.5)
Mental Health	No	109 (14.2)	294 (38.2)	367 (47.7)
	Yes	52 (16.8)	152 (49.0)	106 (34.2)
Non weight related CHD risk	No	98 (15.7)	274 (44.0)	251 (40.3)
	Yes	63 (13.8)	172 (37.6)	222 (48.6)

8.3.2.3 Patterning of adherence

As indicated in Table 15, with the exception of patients who were already active at baseline, completion rates increased in a linear manner by baseline activity level, with 40% of inactive patients completing NERS, compared with 48% and 51% of moderately inactive and moderately active patients respectively. Women were marginally less likely to complete NERS, despite slightly higher entry levels. Older patients were most likely to complete NERS, with more than half of over 60s completing NERS by comparison to 1 in 3 patients aged 40 or younger. Although higher levels of scheme entry were observed amongst patients living in the least deprived area, completion rates were similar for all 3 groups. Non-car-owners were slightly less likely to complete NERS. Patients referred for non-weight related CHD-risk factors were most likely to complete NERS, followed by those referred for weight management and mental health patients.

Table 16. Odds ratios and 95% CIs for patterning in uptake and adherence

		Univariable analyses		Multivariate analyses	
		Uptake	**Adherence**	**Uptake**	**Adherence**
Baseline activity	Moderately inactive	0.98 (0.62 to1.55)	1.37 (0.98 to 1.93)		**1.49 (1.05 to 2.12)**
	Moderately active	1.21 (0.75 to 1.95)	**1.56 (1.12 to 2.17)**		**1.60 (1.13 to 2.25)**
	Active	1.64 (0.77 to 3.51)	1.17 (0.73 to 1.88)		1.22 (0.74 to 1.99)
Sex (Male)		0.76 (0.54 to 1.08)	1.08 (0.85 to 1.40)		
Age	Medium (41 to <60 years)	1.15 (0.77 to 1.73)	**1.56 (1.14 to 2.12)**		**1.47 (1.06 to 2.02)**
	High (61 years+)	1.37 (0.90 to 2.10)	**2.22 (1.63 to 3.03)**		**2.09 (1.48 to 2.95)**
WIMD	Medium	**0.62 (0.40 to 0.94)**	0.96 (0.71 to 1.29)	**0.64 (0.42 to 0.99)**	
	High	0.69 (0.45 to 1.06)	1.05 (0.77 to 1.41)	0.82 (0.52 to 1.28)	
Car ownership		**1.87 (1.31to 2.65)**	1.27 (0.97 to 1.67)	**1.92 (1.33 to 2.76)**	
Weight management		1.20 (0.83 to 1.74)	1.12 (0.85 to 1.46)		
Mental Health		0.82 (0.57 to 1.17)	**0.57 (0.43 to 0.75)**		**0.72 (0.53 to 0.98)**
CHD risk		1.17 (0.83 to 1.64)	**1.40 (1.10 to 1.79)**		1.06 (0.81 to 1.40)

As indicated in Table 16, age, reason for referral and baseline activity were significant predictors of adherence in univariable analysis. Older patients, patients referred for CHD risk factors, non-mental health patients and patients who were already moderately active at baseline were most likely to adhere. Despite higher uptake levels, adherence was not significantly patterned by car ownership or deprivation. In multivariate analyses, associations of CHD risk factors with adherence become non-significant, likely because association with adherence is confounded by age. Associations of age, mental health status and baseline activity however remain significant, though associations of mental health status are reduced to a borderline level of significance. After adjustment for age and reason for referral, the association of baseline physical activity level with adherence increases in strength, with the contrast between 'inactive' and 'moderately inactive' patients becoming significant.

8.3.3 Summary and implications

Though most patients reported no structured exercise at baseline, a small number of referred patients were classed as already active. Given the increasing demands on the service following randomisation described in Chapter 6, mechanisms to identify and remove active patients from the service prior to entry (Johnston et al. 2005) may be of value. Whilst slightly more likely to enter NERS, 'active' referred patients were also most likely to drop out. Though this contrasts with findings in some previous ERS (Taylor et al. 1998), where completion occurred disproportionately in highly active patients, the low intensity group-based focus of NERS may have been less appealing to active patients, leading many to withdraw.

The most active group aside, adherence increased in a linear manner with baseline activity, with patients who did no activity at baseline least likely to adhere, whilst more than half of moderately active patients completed NERS. It is likely that factors which contributed to inactivity, including intrapersonal factors such as baseline motivation, as well as higher level contextual impediments (Stokols 1996), also contributed to constraining programme adherence, with inactive patients having the largest mountain to climb and perhaps requiring the greatest support to overcome barriers to engaging in the scheme.

Although many professionals commented that mental health patients often adhered to a similar extent where supported into the exercise environment, quantitative profiling indicated that, consistent with several recent studies (Dugdill et al., 2005; Crone et al., 2008; James et al., 2008), mental health patients were significantly less likely to complete NERS. This trend lay on the border of statistical significance after adjusting for other factors and lower adherence in mental health patients is likely at least in part explained by factors such as the lower age of these patients compared to patients referred for CHD risk factors. Given the perceived contingency of adherence among mental health patients on the ability of practitioners to support patients' assimilation into the social environment, ERS practitioners may require specific training in supporting the needs of mental health patients, among whom additional barriers to adherence may limit the usefulness of exercise as a treatment (Seime and Vickers 2006).

Older patients and females were overrepresented within NERS, with this trend likely due to factors such as increased contact with health professionals amongst women and older patients (Green and Pope, 1999). All previous leisure centre-based ERS trials have attracted a similar 2:1 ratio of female to male patients, whilst many have operated inclusion criteria limited to older patients. However, the previously cited tendency for significantly higher completion by males was not replicated, with women achieving only slightly lower adherence than men (Dugdill et al., 2005; Gidlow et al., 2007; James et al., 2008), though older referred patients were substantially more likely to complete NERS than were younger patients. Hence, professional reports that the group based structure of NERS made it particularly effective in facilitating adherence in females and older patients are perhaps supported. The benefits of group exercise are highlighted in two recent qualitative studies limited to female patients (Schmidt et al. 2008; Emslie et al. 2007). Whilst not limited to female patients, given that most patients in NERS classes were older women, social assimilation was likely easier than in a more male dominated environment. Patterning by age is perhaps also related to patient reports that class times often made it difficult for working patients to access the scheme.

Consistent with reports of professionals, significantly higher uptake was observed in the most affluent tertile. However, these patients were also more likely to drop out,

with completion rates approximately equal across tertiles. This contrasts with data in Chapter 6 which indicates higher completion in the 4 most affluent counties. It may be that adherence is higher in more affluent counties, though within these counties, patients in poorer sub-areas are as likely or more likely to adhere. Previous findings on patterning in uptake and adherence by SES are somewhat equivocal (Gidlow et al., 2007; Sowden et al., 2008) and multi-level influences of socioeconomic status deserve attention. Consistent with patient interviews, where accessing NERS was sometimes described as contingent upon access to a car, non car-owners were almost twice as likely not to enter NERS. This is perhaps also consistent with previously reported tendencies for lower uptake of ERS in more rural areas (Gidlow et al. 2007).

Limitations of these analyses include that adherence is based on data collected by practitioners, the quality of which likely varied between areas, as well as reliance on self-report measures of baseline activity. Furthermore, analyses of uptake are less well powered than analyses of adherence, given that only 15% of patients did not enter NERS, compared to a more even split of adherers or non-adherers. Nevertheless, these analyses offer significant insights into for whom the NERS programme may have been more or less effective.

8.4 Conclusions

Whilst to date, studies examining processes of change within ERS have typically involved either quantitative profiling of adherence or qualitative scrutiny of patient experience, this chapter demonstrates the value of combining both approaches to allow patterning in responses to be explained rather than simply described. Exercise professional and patient interviews brought into question whether health professional advice played an active role in motivating patients to become more active, or whether the scheme was likely only effective for patients who were already motivated to change. Whilst no measure of baseline motivation was available, adherence was significantly higher amongst patients who were already moderately active compared to inactive patients. Though lower baseline activity is likely determined by a multitude of factors, of which motivation is perhaps only one, impacts for patients amongst whom change is less internally motivated at baseline may be enhanced through delivery of activities intended to enhance internal motivation, including motivational interviewing and goal setting.

Key active ingredients identified by patients and professionals were the high dose of professional support and patient-only group exercise, with professionals supporting development of skills and confidence, whilst patients provided an empathic social context and realistic role models. However, responses to these ingredients varied between patient groups, with assimilation into the social environment perceived by professionals as being easier for older patients and women given the overrepresentation of these patients, whilst mental health patients were seen as needing additional interpersonal support to assimilate into classes. Quantitative profiling confirmed that NERS was entered disproportionately by women and older patients. By contrast to previous studies, women were almost as likely as men to complete NERS, though consistent with previous studies, the scheme was less effective in meeting the needs of younger patients and mental health patients. Hence, attention is needed to better meeting the needs of these patient groups.

In addition, access issues deserve attention, given that consistent with reports of patients that the scheme was often difficult to access without a car, car-owners were substantially more likely to access the service. ERS should aim to offer services in as many locations as possible and consider how patients living in more remote areas might be supported to attend the scheme. Somewhat more equivocal findings emerged in relation to influences of socioeconomic status on participation, with lower uptake in more deprived regions consistent with reports of implementers, but also higher drop out in these regions, resulting in equal completion rates across WIMD tertiles.

Future research should also focus upon understanding how patients' transition from the scheme ought to be best supported to avoid behaviour change becoming contingent on the structures of the programme. The effectiveness of emerging strategies to support this transition, such as filtering patients into maintenance classes or exiting them in clusters deserves close attention. Given the anxieties expressed regarding entry the gym environment, exit route options beyond leisure centres should be provided. Adherence is no guarantee of long term change, and future analyses may focus usefully on whether patterning in adherence translated into patterning in long term impacts.

9 Discussion and conclusions

9.1 Chapter aims

This chapter reflects on the findings presented throughout Chapters 6 to 8, beginning by discussing advances over previous ERS literature offered by the evaluation approach adopted, and implications for the implementation of ERS and interpretation of NERS trial outcomes. Reflections will then be offered on applying the process evaluation framework developed during this thesis and lessons for future evaluations.

9.2 Advances over previous ERS evaluations: implications for implementation and interpretation of outcomes

This thesis offers a range of insights into the implementation and functioning of an ERS above and beyond those offered within the current evidence base. It is amongst the first ERS evaluations to present a theoretical model underpinning the intervention, allowing for internal and external scrutiny of the plausibility of the intended programme model. Furthermore, no previous ERS studies have examined implementation processes or the extent to which the intervention was delivered as planned, implicitly treating implementation as an unproblematic process. Adoption of Diffusion of Innovations theory as a guiding framework (Rogers 2003) and combination of methods allowed identification of strengths and weaknesses in implementation as well as tangible insights into how these emerged. These findings demonstrated the naivety of the assumption implicit to almost all previous ERS research that intervention as conceived and as delivered are one and the same, and hence the need to define the programme as delivered, not just as conceived before outcomes can be interpreted (Schultz et al. 2010). This also facilitated identification of contextual circumstances impacting intervention functioning and local efforts to change both the innovation and the local context in order to potentiate the causal mechanisms of the intervention.

Whilst implementation and diffusion have been largely ignored within ERS literature, an increasing number of qualitative studies have explored patient experiences of ERS (e.g. Wormald and Ingle 2004; Schmidt et al. 2008), whilst quantitative analyses of emergent social patterning in scheme reach are also increasingly being presented (e.g.

Sowden et al. 2008). However, analyses of these issues have not previously been conducted in the same study or in relation to a clearly defined intervention package. In this thesis, qualitative exploration of processes of change and perceived patterning in responses to programme structures preceded quantitative profiling of scheme reach, with combination of methods allowing trends not simply to be described, but also to be explained with reference to how interactions with specific programme structures varied between patient groups.

Furthermore, whilst presented independently from outcomes, the study is also the first to examine many of these issues within a wider evaluation package including effectiveness assessment as recommended within MRC guidelines, and offers substantial insights to guide the interpretation of outcomes. Reflections on the process evaluation model developed and applied to NERS will be presented in Section 9.3. First, this section focuses on implications for implementation of ERS and for interpretation of trial outcomes.

9.2.1 Implications for implementation of ERS

MRC guidelines recognise that in attempting to maintain consistency of implementation with programme theory, the extent and difficulty of behavioural change required amongst implementers represent key dimensions of intervention complexity (Craig et al. 2008a). Efforts to change local practices in order to incorporate evidence-based practices will likely be met with resistance, due to investments in local schemes and belief in their effectiveness (Sowden and Raine. 2008). Given that such moves will likely involve restricting autonomy and reducing local ownership, and might be interpreted as criticism of cherished practices, attention needs to be paid to developing communication structures which might persuade local implementers of the benefits of change. As described in Chapter 6, appointment of respected peers and opinion leaders to mediate between the change agency and implementers is likely a useful means of pacifying power imbalances inherent to authority-innovation decisions.

MRC guidelines also recognise the range of stakeholders in whom behavioural change is required as another key dimension of implementation complexity (Craig et al. 2008a). ERS involve multiple layers of diffusion, and hence effective

communications with referral partners and leisure service providers are crucial in ensuring that the scheme diffuses to its target audiences. In NERS, the esteem of a national scheme was commonly seen as aiding communications with local health professionals. Given that doubts regarding effectiveness have previously been cited as barriers to using ERS (Graham et al. 2005), referral may become more fully routinised into health professionals' practice should trial data indicate effectiveness. However, given the emerging concerns regarding sustainability of referral criteria, the degree of targeting needed to ensure sustainability of large-scale ERS deserves consideration.

Whilst communications with leisure centres were seen as largely positive, compromise is perhaps to be expected in the times that facilities can be made available, in order to avoid conflict with centre's financial priorities. This may have the negative effect of limiting the suitability of class times for working patients who are only available during peak hours. However, within NERS, such compromises were typically described in areas with small referral volumes. As scale increases in these areas, offering up space for classes may become more financially beneficial, reinforcing adoption. A further concern which deserves attention is reported refusals of providers in some areas to accept patients into mainstream services after completion due to their conditions. Where ERS deal with relatively high risk patients whose conditions are unlikely to be fully resolved by the end of their programme, increased likelihood of such refusals is perhaps to be anticipated, presenting a key barrier to long-term maintenance of outcomes. Hence, communicating with mainstream providers to increase their confidence that NERS offers an effective introduction to exercise for clinical populations is perhaps crucial.

In moving from initiation to implementation, limited consideration of the practicalities of implementation will likely cause challenges (Craig et al. 2008a). Development of training and infrastructure for a Wales-wide scheme were identified as key goals when intent was signalled for the development of a national scheme in 2005 (Welsh Assembly Government 2005a). However, infrastructure and training provision was developing, often in response to perceived shortcomings as they emerged during the definitive trial, rather than having been identified during a developmental pilot phase. Chapter 6 described significant perceived shortages in management training for coordinators in NERS, as well as a perceived lack of guidance arising from the initial

lack of a full time national coordinator. Furthermore, several areas operated without a local coordinator for lengthy periods, with limited structures available to cover these periods, with these areas typically achieving poor patient recruitment or scheme completion rates. The limited management training arose in part from ambiguity over whether it was the responsibility of the scheme or the local authority to provide this. These roles and responsibilities should ideally be clarified during development. As described in Chapter 7, whilst professionals spoke highly of many aspects of training, they had also not been well trained to deliver some intended aspects of the service.

As recommended within Department of Health (2001) guidelines, this thesis demonstrates the value of analysing routinely collected monitoring data in order to identify variations in implementation quality. This should not be a short-term evaluation activity, but an ongoing part of any scheme, with data regularly reviewed to maintain quality. However, whilst able to identify variability in components such as goal setting and patient follow-up protocols, some core components were not well captured by monitoring structures, and hence required development of temporary monitoring procedures purely for this thesis. Implementation plans should include monitoring structures which cover all core components of the scheme, particularly those which are likely to prove difficult to implement.

Close monitoring is increasingly crucial as schemes begin to include activities such as motivational interviewing, whose implementation requires behaviour change processes among practitioners as complex as those anticipated in patients (Miller and Rose 2009). In such scenarios, MRC guidelines recommend involvement of behavioural scientists in implementation (Craig et al. 2008a) with congruence needed between the skills and experiences of implementers and those they are tasked with supervising. In NERS however, coordinators responsible for overseeing local implementation, whilst experienced in delivering ERS, had little experience of MI. Hence, the scheme attempted to incorporate activities in which no-one employed by the service was sufficiently experienced. This is not uncommon to ERS, in which movements towards expecting instructors to adopt patient-centred counselling approaches are typically treated as entirely unproblematic. Combining monitoring structures with expert feedback and ongoing training and advice may help to ensure

that claims to deliver approaches such as MI move beyond the protocol document and into practice.

Chapter 6 also describes the processes of mutual adaptation needed for a new ERS to become an integrated part of local systems. The intervention was seen as needing some reinvention, such as rebranding or integration with other local services in order to retain local ownership and achieve fit with local systems. Activities, locations and communications were described as needing to be tailored to address the perceived needs of local areas which varied in terms of factors such as education level and urbanisation. Whilst maintaining a common core, an understanding of the needs of local populations is crucial in ensuring effective local delivery of national policy.

The relationship of NERS with its context was reciprocal, and where limited fit was perceived, this in some cases triggered adaptations to the local context. These included emergence of new low intensity exercise opportunities in areas where realisation emerged that a 16-week programme would have no impact if the local context did not provide opportunities for maintenance of changes. Hence, NERS likely had impacts beyond the programme, influencing the culture of leisure services and increasing availability of low intensity exercise options for patients and the general public. Notably, this need for additional post-scheme opportunities was identified in areas where schemes had run for several years, with the shortage of such opportunities apparently overlooked by previous schemes. Indeed, apparent ineffectiveness of many previously evaluated ERS may in part arise from the often small and short-term nature of such schemes preventing them from becoming sufficiently routinised into their local contexts to trigger processes of mutual adaptation which are likely necessary to achieve fit between intervention and context and potentiate mechanisms of change (Rogers 2003).

Outcomes of the NERS trial will be reported elsewhere, and the analyses and interpretations presented in this thesis were conducted in advance of knowledge of outcomes (Oakley et al., 2006). However, as described throughout, a key function of process evaluation is to provide information to guide the subsequent interpretation of outcome effects. Attention will therefore now turn to insights into causal processes within NERS.

9.2.2 Implications for interpretation of outcomes

Whilst containing a number of evidence-based principles, NERS came uncoupled from many of these as it diffused into practice. Neither MI nor goal setting appeared to play a meaningful role in producing outcomes. As described in Chapter 4, failure to deliver MI due to underestimation of implementation complexity is common (Miller and Rollnick, 2002; Miller and Rollnick, 2009) and where process evaluation is not conducted, false conclusions are likely drawn. Without these data, the NERS trial would likely have appeared in systematic reviews of MI based interventions, as has at least one previous ERS which provided no quality assurance measures (Harland et al. 1999).

Perhaps unsurprisingly given the non-delivery of activities to elicit and strengthen internal motivations in baseline consultations (Department of Health 2001), qualitative data from professionals and patients suggest that the scheme was most effective amongst patients for whom behavioural change was already somewhat internally motivated, with entry on the basis of health professional advice typically seen as a weak determinant of change. Indeed, whilst no measure of baseline internal motivation or of whether the patient had sought referral or been advised to enter was available, patients inactive at baseline were least likely to complete NERS. It is an oversimplification to conflate baseline activity with motivation, which likely represents one of a multitude of factors contributing to lower levels of activity. However, it remains plausible that impacts may be strengthened for some patients by future integration of MI and goal setting.

Nevertheless, NERS achieved a relatively high completion rate compared to previous ERS at 44%. However, individual areas achieved adherence rates more diverse than the 12-52% range reported in previous schemes, with one area achieving only 11% compared to 62% in another. Various factors likely contributed to this variability, including absence of a local coordinator for much of the trial in the area achieving the weakest adherence level, with lower adherence where population-adjusted referral rates were highest perhaps suggesting that health professionals in some areas were better at identifying patients most likely to adhere. In addition, implementation checks indicated substantial local variability in the range of class times, types and locations,

fidelity to patient follow-up protocols and programme dose, whose links with adherence have yet to be explored.

The key ingredients perhaps contributing to the relatively high overall adherence rate include professional supervision and opportunity to enter a supportive social network. Professional supervision was seen by professionals and patients as important in putting patients at ease in an unfamiliar environment, supporting development of skills to exercise without making conditions worse, and supporting confidence and motivation. The patient-only group setting was seen as eliminating anxieties about exercising in front of fitter exercisers, providing an empathic social context and realistic behavioural models.

However, responses varied across patient groups and settings. The social context of the scheme perhaps favoured older female patients, due to their overrepresentation, whilst mental health patients perhaps also struggled to assimilate into the group environment. Attention is likely needed to training professionals to support the needs of mental health patients and to attracting more younger and male patients. Patterning also emerged from challenges accessing the service without transport, perhaps limiting uptake in rural areas, and from time constraints on working age patients. Whether this patterning is reflected in long-term outcomes deserves attention.

Key caveats in interpreting intervention effects include the fact that most local areas extended discounts, due to concerns that patients would not continue to stay active if costs returned to normal levels. Given that half of patients in one recent qualitative study (Schmidt et al. 2008) stated that they would not remain active when attendance returned to full price, whether extending these discounts eliminates or simply delays regression to baseline activity until expiry of the extended discount cannot be established. Secondly, most professionals offered indefinite access to NERS classes during the trial, and outcomes will therefore often have been produced by a level of support which would not continue after the trial. The intervention was also offered on average for a significantly longer duration than stipulated in protocols, with a median time from entry to exit of 19 weeks, 3 weeks longer than intended duration and almost double the 10-weeks in many previous ERS (Taylor et al. 1998; Stevens et al. 1998; Isaacs et al. 2007).

Qualitative data revealed a variety of local approaches to the transition to independent activity, with professionals commonly emphasising maintenance of social support networks beyond the scheme, through strategies such as exiting patients in clusters or filtering them directly into maintenance classes, as vital to maintaining scheme outcomes. No previous trials describe post-scheme activities. However, it is likely that local emergence of post-scheme activities contributed significantly to maintaining intervention effects for some patients.

9.3 The process evaluation framework: strengths and areas for development

The comprehensive understanding of implementation and insights into functioning of the NERS programme described above were achieved through development and application of a comprehensive hierarchical framework for process evaluation. As demonstrated in Chapter 3, process evaluations have tended to focus on one or two aspects of implementation or participant experience, with little justification for the aspects chosen for study and little clear linkage between process components or between process evaluation and interpretation of outcomes. The framework in this thesis offers significant advances over previous guidance in that it not only recommends aims, but also offers methodological guidance and proposes clear linkage between components, allowing process evaluation to fulfil the functions described within MRC guidance to: 'assess fidelity and quality of implementation, clarify causal mechanisms and identify contextual factors associated with variation in outcomes' (Craig et al. 2008a). Reflections on each component and methodological challenges emerging during their design and conduct will now be offered, with linkage between components in contributing to broader study aims described before discussing linkage of process evaluation with outcomes. Implications for the conduct of process evaluation within similar policy trials and within researcher-initiated studies adhering more closely to the development and evaluation process described by the MRC are discussed throughout.

9.3.1 Process evaluation components

9.3.1.1 Eliciting programme theory

MRC guidelines recommend an iterative development and evaluation process beginning with identification of programme theory prior to feasibility testing and evaluation (Craig et al., 2008a; Craig et al., 2008b). In researcher-initiated studies therefore, theory development will likely have taken place during formative phases, with process evaluation within the definitive evaluation beginning by simply describing programme theory. However in NERS, as perhaps common to many policy trials, development was conducted prior to commissioning evaluation, with no explicit programme theory presented.

Although outcomes of the development process were represented in protocol documents, these did not transparently reflect the priority given to particular programme components, or their perceived roles in bringing about change. Whilst Michie and colleagues (2009) argue that publication of detailed intervention protocols is crucial to understanding what was delivered, these may be of limited usefulness in identifying the functional components of the intervention. For example, whilst several pages were dedicated to health checks, only a single reference was made to goal setting and MI. Hence, whilst this thesis planned initially to use protocol documents as a basis for implementation checks, it rapidly became clear that this was not possible. Only through negotiation of a parsimonious model with policy representatives, describing key programme inputs and hypothesised causal processes as recommended by Armstrong et al. (2008), was a clear definition of the intended intervention achieved. Ideally, this should have been conducted at the outset of the evaluation, allowing identification of targets for implementation checks at the earliest opportunity.

In retrospect, a useful extension to this phase would have been to use discussions with policy representatives to elicit a model of the implementation process, in terms of the steps followed to diffuse protocols into local practice, and their anticipated outcomes. Early elicitation of a model both of the programme and its implementation would perhaps have allowed for earlier engagement with relevant literature bases, and anticipation of potential shortcomings, allowing early provision of feedback to guide

implementation. As will be described in Section 9.3.1.3 for example, whilst concerns were expressed at the outset regarding the shortage of training in motivational interviewing, earlier exploration of implementation activities beyond training, and engagement with the MI implementation literature would likely have highlighted additional shortcomings which in practice only emerged in evaluating subsequent training.

9.3.1.2 Understanding programme diffusion

As described in Chapter 3, many process evaluations move straight from description of the intended programme towards measurement of implementation. However, measuring implementation without exploring the processes through which the intervention takes shape across contexts may reveal problems, but will likely offer little insight into how divergences might be corrected. Frameworks such as that of Steckler and Linnan (2002) highlight the need to understand impacts of context on programme implementation. However, relationships between context and implementation are treated as somewhat unidirectional, with identification of factors which impede or facilitate implementation recommended, rather than development of an understanding the dynamic processes through which context and intervention interact and adapt to one another. In this thesis, Diffusion of Innovations theory (Rogers 2003), applied with a critical realist lens focusing on how implementation arises from interaction of implementers with the innovation in context (Greenhalgh et al. 2004), offered a useful theoretical framework for exploring these issues, highlighting the value of enhanced synthesis between the diffusion and process evaluation literatures. This provided insights not only into how contextual factors impeded or facilitated the implementation and functioning of the intervention, but also the human activities involved in the diffusion process and the actions of local implementers in modifying the intervention and the local context.

Challenges in diffusion are perhaps particularly pertinent in policy trials or other large multi-site trials, with implementation scale enhancing the diversity of settings and complicating the communication of the innovation throughout the whole system. Challenges are perhaps greater still in policy trials, which move to definitive evaluation in advance of extensive piloting of implementation structures (Craig et al. 2008a; Craig et al. 2008b). Nevertheless, even where tested in one or two locations

prior to multi-site definitive evaluation, it cannot be assumed that the intervention will translate unproblematically across contexts, and hence exploration of challenges ensuring consistency of implementation across new contexts should be included within definitive evaluation.

Given the complexity and unpredictability of these issues, this aspect of process evaluation will almost certainly emphasise qualitative methods. However, it is crucial in collecting large volumes of qualitative data to be mindful of avoiding artificial impacts on the intervention (i.e. Hawthorne effects). Hawe and colleagues (2004b) for example describe using written implementation diaries in order to track evolution of a complex intervention and implementers' perceptions and experiences over time. However, the authors question whether such methods provide additional opportunity for reflective practice, changing how the intervention is delivered.

In this study, a single retrospective interview was conducted with coordinators and policy representatives on the experience of delivering the scheme almost a year after its inception. Data collection likely therefore had minimal impact on implementation behaviour. However, perhaps the disadvantage of collecting data only at one time point is that this minimises the extent to which the dynamic nature of the diffusion process can be captured. Data are situated in the temporal context in which they are collected, with perceptions of the intervention changing significantly over time. For example, whilst most coordinators spoke positively about national standardisation, policy representatives' accounts indicated that the move had been met with substantial initial resistance. Interviewing coordinators during the transition may therefore have led to quite different data. However, interviewing coordinators at this stage would allow limited reflection on implementing NERS and its routinisation into practice. Whilst an alternative would have been to interview fewer coordinators at multiple time points, time pressures during a chaotic period such as transition to a new scheme may have meant that collecting data at this point was a burden, damaging rapport, whilst this degree of discourse between implementers and evaluators may have introduced Hawthorne effects. Collecting retrospective data from multiple perspectives was helpful in overcoming the tendency for implementers to focus on more positive experiences later in the implementation process. However,

understanding temporal aspects of implementation processes without undermining the trial remains a challenge for process evaluation.

9.3.1.3 Measuring implementation

As described in Chapter 3, process evaluations often move to implementation assessment without describing programme theory, making it difficult to understand implications for programme functioning. However, wedding implementation checks to programme theory, explicitly presenting these as measures of the consistency of implementation with programme theory, and reflecting on theoretical implications of variations in implementation allowed clear specification of the theory ultimately tested within the evaluation. This not only facilitated greater understanding of outcomes and identification of areas where improved implementation was needed, but will also facilitate future comparisons with other interventions on the basis of functional components and theoretical similarity (Shepherd et al. 2009; Armstrong et al. 2008; Michie et al. 2009).

In researcher-initiated evaluations, monitoring structures should be developed alongside the programme and framed to capture key components forming the basis of programme theory. Use of data collected as part of routine practice wherever possible reduces the aforementioned likelihood of Hawthorne effects arising from activities integrated purely for evaluative purposes (Hawe et al., 2004b; Audrey et al., 2006b). This also allows monitoring of practice over time throughout the scheme, with checks within NERS for example showing improvement in goal setting practice over time. By contrast, it is likely only feasible to collect researcher implemented measures at one or two time points. In policy evaluation settings however, evaluators may have limited influence on monitoring structures. Hence, a key stage of designing implementation checks becomes identifying what data are available via routine data sources, before developing additional structures specifically for process evaluation where key components were not captured.

The conduct of implementation checks identified a number of key challenges likely common to many similar evaluations. Firstly, reliance upon implementer self-reports or data collected by implementers is perhaps inevitable, given that implementation activities will not be easily observable other than by those responsible for programme

delivery or receipt. However, this may lead to excessively positive appraisal (Audrey et al., 2006b), as well as concerns surrounding the quality of data collected.

In this thesis, exercise professionals acted as data collectors for two study components. The first was recordings for the MI sub-study. Instructions were provided for professionals on how to conduct recordings, though in some instances, data were returned without accompanying consent forms. Whilst in all such cases, professionals reported having obtained consent but having misplaced forms, data were destroyed. Given that ethical approval was received weeks before introduction of additional training courses, the rush to collect pre-training data meant that it had not been possible to pilot these methods in order to ensure that instructions were easy to follow. For follow-up data collections, difficulties were anticipated, and instructions clarified. Arrangements were made to collect completed data in person from area coordinators, at which point folders were checked for completeness.

The second component was the aforementioned routine monitoring database, completed throughout the trial. A key use of this data was calculation of attendance rates, indicated by dates that consultations took place. However, consistency checks indicated substantial discrepancies between the number of dates entered and the number of patients for whom objective data were present, triggering doubts regarding validity. Email enquires to coordinators revealed that some entered patients' date of attendance when the appointment was booked, whereas others entered data only after the consultation took place. Hence, dates were disregarded as evidence of attendance, and a complicated command file developed between the author and the trial manager in order to search the database for evidence that the consultation took place.

These problems could have been ironed out through developmental piloting in which monitoring structures were tested (Craig et al., 2008a; Craig et al., 2008b), which for example would have revealed that using appointment date as evidence of attendance was prone to error. Within researcher-initiated evaluations, piloting of monitoring methods in order to ensure high quality monitoring data is crucial alongside pilot tests of the intervention itself. However, in policy trials, the rapid movement to implementation which likely compromises some aspects of programme implementation also perhaps risks compromising the quality of implementation

assessment, through allowing insufficient time for piloting of methods. Where using implementers as data collectors, systems should be as simple as possible, whilst structures to ensure that completion of monitoring data is itself monitored and standardised are necessary.

An additional issue related to formative functions of implementation checks and structures for ongoing feedback of emerging weaknesses. Where limited formative work has taken place, process evaluation will likely serve extensive formative functions within the definitive evaluation. The intervention-evaluation distinction can become somewhat blurred at this stage and process evaluation may find itself performing tasks which should ideally be part of the programme, with gaps in monitoring structures perhaps pointing to areas where implementation complexity has been underestimated. As described above for example, the lack of structures to monitor MI fidelity was identified as a key weakness in implementation, with temporary cross-sectional structures imposed for the purposes of fidelity checks.

In addition, whilst formative functions of process evaluation within definitive trials are increasingly recognised (Wilson et al., 2009), evaluators should perhaps consider whether triggering immediate action by feeding back concerns as soon as they emerge is desirable, or whether concerns should be communicated only once robust recommendations can accompany them. Concerns regarding the shortage of MI training and the limited specificity of goal setting were both fed back, leading to rapid action. However, in neither case did this address the identified problem. For example, MI training took place in contexts where there had still been little consideration of the need for expert monitoring and feedback structures, or the compatibility of MI with consultation structures, and hence the proposed solution was in place before the causes of the problem had been understood. Further training therefore resulted in increased willingness to implement, though no clinically significant behavioural change. As described in Section 9.3.1.1, using discussions with policy representatives to describe implementation activities at the outset of the process evaluation may have allowed the process evaluation to serve these formative functions more efficiently, facilitating early recommendations which in practice largely emerged only through studying impacts of mid-trial training.

9.3.1.4 Participant experiences

In relation to understanding participant experiences, much discussion in process evaluation literature has focused on whether measures should focus on engagement (Steckler and Linnan 2002) or exposure (Baranowski and Stables 2000), with measures such as patient exposure or satisfaction emerging as the primary or sole focus of many process evaluations. In addition, many focus on these issues without providing data on implementation. However, acting on feedback and improving the intervention, will require an understanding of what was delivered. Hence, conducting and presenting analyses of these data in the context of a clear prior definition of the intervention within NERS allowed for more concrete suggestions for improving intervention delivery. Where adopting a perspective which views outcomes as a product of the interaction of patients with an intervention (Pawson and Tilley 1997), qualitative exploration of patients' interaction with an intervention which has been clearly defined in advance, is useful in illuminating causal processes.

Key challenges however include the representativeness of patients sampled to provide qualitative data. Whilst implementer interviews included almost all implementers, schemes such as NERS will involve thousands of patients, with a very small percentage of these approached for qualitative data. In the present study, only programme attendees were interviewed, with perspectives of patients who did not take up the programme not included. In addition, due to the relative homogeneity of patients who attended NERS, females and older patients were overrepresented, with views of patients whom the scheme was less effective in engaging less well represented. Including perspectives of exercise professionals on patterning in patient responses to the scheme provided one means of gaining a broader overview of variability in responses to the intervention and triangulating findings from patients on the processes through which programme activities supported change. Where feasible however, obtaining views from non-attendees or groups amongst whom engagement is lowest may provide valuable insights into what works, and for whom.

In addition, data were collected from patients and professionals at only one time point. As described in relation to programme diffusion, this has implications for data collected. For example, reports such as those that health professionals had not been

previously aware of NERS may have been less prevalent had these data been collected at a later stage when the scheme was more fully routinised into practice. Ideally, given additional resources, data would perhaps have been collected at several stages of the scheme's implementation in order to examine patients' interactions with a dynamic and evolving service. Nevertheless, analyses provided significant insights into processes through which components of the intervention acted to produce change, as well as perceived processes through which social patterning in scheme reach emerged.

9.3.1.5 Social patterning in reach

The final stage reported in this thesis related to social patterning in reach. Whilst a growing number of studies report quantitative profiling of adherence (Gidlow et al. 2007; Sowden et al. 2008; Dugdill et al.. 2005), this analysis was strengthened by the fact that it followed qualitative analyses, which in turn had been preceded by implementation checks. Hence, rather than just describing patterning in relation to a non-specified intervention, the process evaluation taken as a whole allowed for explanation of the emergence of patterning and linkage to patients' interactions with specific programme structures and mechanisms.

It should be acknowledged that the processes associated with programme reach may not be the same as those associated with independently maintaining behavioural change. Hence, patterning in outcomes may differ from patterning in reach. Nevertheless, examining patterning in scheme reach offers a useful stage in mapping the process from referral to long-term behavioural change. When combined with analyses of intermediate change processes and long term behavioural change, this may help to identify for whom greater emphasis is needed on promoting adherence, and for whom the transition to independent activity proves most challenging. The ability to profile adherence within previous trials has often been limited by small sample sizes, whilst few studies profiling adherence have also provided long term effectiveness measures.

9.3.2 Linkage between study components in data collection, analysis and presentation of findings: combining methods

As described in Chapter 3, a key shortcoming of much process evaluation to date is the lack of clear linkage between components and elucidation of how these combine to contribute to wider aims of understanding implementation and outcomes. Hence, as described above, this thesis provides a hierarchical model in which each phase builds on the previous and feeds into the next, incrementally forming an understanding of the implementation of the intervention and how it produces change in context. Central to this framework is the combination of qualitative and quantitative methods; an approach common in process evaluation, but often conducted with little justification and with qualitative methods subservient to quantitative analyses. Indeed, more widely within health research, mixing methods has often been conducted poorly, with little justification for approaches selected and limited clarity of linkage between findings from each method (O'Cathain et al., 2007; O'Cathain, 2009).

In the hierarchical framework developed in this thesis, qualitative and quantitative methods served distinct yet complementary functions. Diffusion was explored qualitatively and conceived as the process through which implementation emerged, with quantitative implementation checks conceived as measuring outcomes of these processes. Hence, data on diffusion provided significant insights into how weaknesses in implementation emerged, allowing for identified problems to be accompanied by recommended solutions. Indeed, this approach of presenting solutions rather than just identifying problems played a key role in alleviating the potential challenges described in Chapter 2, in relation to communicating critical findings regarding a policy to representatives who were highly invested in that policy. Similarly, participant's interactions with NERS were conceived as the processes through which patterning in scheme reach emerged. Implementation and social patterning in uptake and adherence were conceived as intermediate outcomes in the causal chain linking the planned intervention to outcomes. Hence, qualitative data explored processes, whilst quantitative data measured intermediate outcomes.

However, whilst analyses are presented sequentially in order to tell the unfolding story of the intervention in a logical order, this does not reflect the outcomes of a

linear research process. For example, whilst implementation is presented as an outcome of diffusion, implementation checks needed to be in place from scheme initiation and collected throughout the trial, whilst qualitative data on programme diffusion were collected towards the end of the trial to allow reflection on a year of practice. Hence, whilst analyses from one phase provided hypotheses for the next, it was often only possible to test hypotheses for which data had been collected, rather than to collect new data. For example, whilst professionals highlighted the importance of baseline internal motivation in determining patient adherence, and a perceived tendency for higher adherence where patients requested the scheme rather than being referred, no measure of baseline internal motivation or referral seeking behaviour was available. In addition, as implementation checks revealed shortcomings which in turn triggered further action, examining the actions of these new diffusion activities became necessary and hence a degree of iteration between study components was required rather than a linear movement from one phase to the next.

In analysis and presentation of findings, efforts have however been made to draw out the implications of each phase of analysis for the next and to reflect back on what each phase adds to the previous in incrementally building an understanding of the implementation and functioning of the intervention. Separating presentation of process evaluation findings into distinct components, and reflecting on linkage between components, rather than presenting all data together before attempting to make sense of them, enabled fuller insights into the implications of the study as a whole for the implementation and functioning of the programme. This produced far more data than can realistically be reported in a single article without methods and findings becoming opaque or reduced to the superficial. Hence, dissemination efforts will maintain separation in order to aid clarity, whilst ensuring that clear links are made between study components (Stange et al., 2006).

9.3.3 The role of process evaluation within the wider evaluation: linkage to trial outcomes

Many of the lessons learned from this thesis will likely be of use in informing the implementation of similar complex interventions regardless of the effectiveness of NERS. For example, lessons learned regarding the integration of motivational interviewing into routine practice will be of significant interest to the growing number

of organisations attempting to use MI. Furthermore, findings surrounding the need for effective communication and support structures in implementing policy initiatives will likely be generalisable to other efforts to achieve high quality implementation of complex interventions. However, understanding how to implement an intervention which ultimately has no benefit is perhaps of limited usefulness (Bond et al., 2010) and the value of process data in understanding the specific intervention it describes is greatly enhanced when understood in the context of good quality information on outcomes. Hence, MRC guidelines recommend process evaluation as a vital accompaniment to, but not a substitute for robust effectiveness evaluation (Craig et al. 2008a).

As discussed throughout, this study is nested within a wider evaluation; a randomised policy trial. In recent years, several authors have argued against the use of randomised controlled trials in evaluating complex interventions, through focusing upon the variable nature of implementation in real world settings and the complex and contextually situated nature of causality (Pawson and Tilley 1997; Berwick 2008; Mackenzie et al. 2010). As described in Chapter 2, these difficulties are perhaps exacerbated in policy evaluation settings, where limited formative work has taken place, likely leading to weak or variable implementation in some trial areas. However, as argued by Oakley and colleagues (2006) and reflected within MRC guidelines, process evaluation offers a potential means of retaining the strengths of the RCT, whilst overcoming traditional shortcomings in terms of understanding implementation, causal processes and the interaction of the intervention with its context (Craig et al. 2008a). There ought not to be an either/or approach to evaluating these issues, and process and outcomes evaluation conducted together will be of significantly greater value than the sum of their parts.

Combining RCTs with process evaluation has not always achieved the aims set out within the MRC framework particularly well. Where process evaluation has been included, this has typically involved exploring patient responses to the intervention, often at the neglect of implementation or contextual issues (see Chapter 3). However, this study demonstrates that through including systematic and comprehensive process evaluation within multi-site trials, it is possible to supplement estimates of the amount

of change in key outcomes following introduction of a policy with rich insights into 'what works, for whom and under what circumstances' (Pawson and Tilley 1997).

Whilst a number of the original theoretical underpinnings of NERS were compromised by poor delivery, key active ingredients in practice (i.e. 'what works') were identified as professional supervision, which acted through facilitating increased knowledge, confidence and motivation, as well as providing opportunity to enter a supportive social network which provided an empathic context for behavioural change and realistic role models. In terms of 'for whom' these activities appeared to facilitate change, adherence was greatest amongst older patients, patients who were already moderately active at baseline and patients referred for reasons other than mental health, whilst the scheme predominantly attracted females. Implementation and functioning were affected by a range of contextual circumstances, including local perceptions of the need for national standardisation, cooperation of centres in providing space for classes, ease of patients' access to leisure centres and compatibility with patients' work/family commitments, whilst long term maintenance of outcomes was often perceived as contingent on creation of opportunities for continued low intensity exercise.

Process evaluation has therefore provided rich data which will contribute substantially to ongoing improvement of the programme and interpretation of trial outcomes (Oakley et al., 2006). Subsequent analyses may involve direct integration of data on variable aspects of the programme and its contexts in order to evaluate hypotheses generated by this study. For example, whether the emerging patterning in programme reach is reflected in outcomes may be usefully explored. Whether the local variation in fidelity to patient follow-up protocols, or variability in the flexibility of the programme offered in each site were linked to impacts on patients' physical activity may usefully be explored in order to further develop an understanding of the active ingredients of NERS and the mechanisms through which they support change.

9.4 Summary and conclusions

In conclusion, this thesis demonstrates the feasibility and usefulness of a comprehensive mixed-methods approach to process evaluation within a randomised policy trial, focusing on eliciting programme theory, diffusion, implementation,

participant experiences and social patterning in programme reach. Through applying this framework in a manner which was systematic yet flexible, process evaluation served formative functions, offering recommendations for improving implementation of ERS, as well as providing data to allow outcomes to be fully understood in relation to a fully specified model of delivery, the causal processes through which these components produced change and the roles of contextual factors and human agency in shaping the intervention and its impacts. The study, particularly when understood alongside trial outcomes, will add significantly more to an understanding of what works, for whom and under what circumstances than has a procession of trials and observational studies which focus solely on aggregate effectiveness. It is hoped that the framework developed and tested in this thesis will be transferable to other public health interventions, assisting evaluators and programme developers in their efforts to effectively implement ambitious complex interventions and understand the outcomes produced by them.

10 Appendices

10.1 Appendix 1 – literature search strategy

The literature review covers a broad base of literature, and aims to give an overview of relevant methodological and empirical literature, though does not purport to be a systematic review. Chapter 2 focuses on political and methodological issues surrounding the evaluation of complex interventions, and aims to highlight current debates regarding how such evaluations ought to be conducted. Reading for the chapter began with key texts on evidence-based policy, before focusing on key policy documents, grey literature and journal articles discussing the need for evidence-based policy and debates surrounding the evaluation of complex interventions. Literature were identified via a range of sources including conference presentations, reading of key books, discussions with researchers in the field, articles suggested by project supervisors as well as internet searches using terms including 'evidence-based policy', 'evaluation', 'randomised controlled trials', 'complex interventions', 'realistic evaluation' 'process evaluation', 'policy trials' and 'policy evaluation'.

The following chapter reviews process evaluations conducted since 2003, identifying articles via the search terms 'process evaluation' and 'health promotion or public health or health behaviour or physical activity or diet or smoking or alcohol'. Empirical reports of studies focusing on the delivery and receipt of interventions were included in the review of process evaluations. A number of non-empirical articles discussing process evaluation conduct and theory were also identified via these searches and discussed later in the chapter, whilst further articles were obtained via searches of databases including web of science and Google scholar and through inspecting reference lists of published articles.

The final literature review chapter focuses on exercise referral schemes and their constituent parts. A comprehensive review of ERS studies was conducted in 2007, as this thesis commenced. Hence a starting point was to retrieve all articles cited by this review, along with a number of other review articles. Regular web searches were made for articles published since this review, using terms 'exercise referral' and 'exercise on prescription'. Furthermore, checks were made for articles citing this review to identify emerging studies. Included studies focused on effectiveness or processes of change within a programme of exercise, accessed via health professional referral. The chapter also discusses a range of core components of ERS, including health professional advice, motivational interviewing and goal setting. Key theoretical articles and books were sought to understand the theoretical background of these components, before searches were made for review articles examining their use in physical activity settings.

10.2 Appendix 2 – Participant interview schedules

The following is a list of potential topics to be covered during interviews. The list is not prescriptive or exclusive. Bullet points represent potential prompts. These may not be relevant in all interviews.

Opinion of Scheme
How was your experience of the referral process?
How have you found the support and guidance offered by your instructor?
What do you think about the activities / classes available?
What have your experiences of the leisure centre environment been like?

Reasons for attendance
What would you say are your main reasons for attending the program?

Perceived changes and processes of change
Do you think that the scheme has had any affect on your: -physical activity levels
-mood/well-being
-energy
-social network
-health

Before beginning the scheme, what do you feel prevented you from being physically active? Has the scheme affected any of these issues? Are there any issues which you feel the scheme does not address?
What were your thoughts and feelings about physical activity before taking part in the scheme? To what extent have these thoughts and feelings changed or stayed the same?

Self efficacy and intentions for future activity
Do you feel confident that you will be able to maintain increases in physical activity once the programme is over? (If so - how has the scheme helped you to develop this confidence?- if not, why not?)
How will you attempt to do this?

Possible alterations to the scheme/Barriers and facilitators to scheme participation
What aspects do you feel have been beneficial in helping you to participate in the scheme?
Are there any things that have made it difficult to participate?
Do you think that the scheme could be improved in any way?

Other comments
Do you have any other thoughts or comments about the scheme?

10.3 Appendix 3 – Exercise professional interview schedule

Section 1 – structured interview

The first part of this interview is fairly structured, mostly involving closed questions which mainly require either yes/no answers or simply a number, although I may ask for a little elaboration to some responses. I will be completing a form based on your responses as I talk to you. The second part of the interview will be less structured and I will ask a number of open questions in order for you to provide me with an in depth overview of your experiences of delivering the scheme.

Initially, I'm going to ask a few questions relating to the 16 week programme of activity that participants receive in each of the facilities where you deliver the scheme:

1. Which facilities you deliver the scheme at?
   Facility 1.....
   Facility 2.....
   Facility 3.....
   Facility 4.....

2a. At these facilities, do patients attend group exercise sessions as part of their programme?
   Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

2b. What types of group exercise sessions are available at each of these facilities (please list)?
   Facility 1.....
   Facility 2.....
   Facility 3.....
   Facility 4.....

2c. Do you organise/supervise any exercise sessions outside of the exercise facility for clients during their 16 week programme?
   Facility 1 - Yes (please describe)/No, Facility 2 – Yes (please describe)/No,
   Facility 3 - Yes (please describe)/No, Facility 4 – Yes (please describe)/No

.....
.....
.....
.....
...

2d. If yes to question 2a, are group exercise sessions exclusively for NERS patients?
   Facility 1 - Yes all/Yes some/No, Facility 2 - Yes all/Yes some/No,
   Facility 3 - Yes all/Yes some/No, Facility 4 - Yes all/Yes some/No

2e. If no to question 2d, are there restrictions on who can attend these classes?
   Facility 1 - Yes (please describe)/No, Facility 2 – Yes (please describe)/No,
   Facility 3 - Yes (please describe)/No, Facility 4 – Yes (please describe)/No

.....
.....
.....
.....

3. How many group exercise sessions per week do you deliver on average, in each facility?
   Facility 1.....Facility 2.....Facility 3.....Facility 4.....

4. Do you deliver classes alone, or jointly with another instructor?
   Facility 1.....Facility 2.....Facility 3.....Facility 4.....

5. What is the average class size at each facility, approximately?
   Facility 1.....Facility 2.....Facility 3.....Facility 4.....

6. How many one-to-one exercise sessions per week do you deliver on average, in each facility?
   Facility 1.....Facility 2.....Facility 3.....Facility 4.....

7. Are weekend NERS classes available to clients at any of these facilities?
   Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

8. Are evening NERS classes available to clients at any of these facilities?
   Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

9. Do clients use the gym as part of their 16 week programme (if yes, go to q11, if no to all go to q12?)
   Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

9a. Is this just for a one off induction, or regular exercise sessions throughout?

Facility 1 - one-off/regular, Facility 2 - one-off/regular, Facility 3 - one-off/regular, Facility 4 - one-off/regular

10. Do clients receive individually tailored gym programmes during their 16 weeks on the scheme?
Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

11. Is the gym open to the general public whilst being used by patients?
Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

12. Is the pool at the facility used for NERS classes?
Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

12a. Is the pool open to the general public during NERS classes?
Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

13. Are changes made to clients' programmes after their 4 week assessment?
Yes (please describe)/No

.....
.....
.....
.....
.....
.....

Now moving on to think about after the 16 week programme is completed:

14. What happens to participants after their 16 week programme is completed?

.....
.....1

5. Are participants allowed to continue attending classes run by NERS exercise professionals after the initial 16- week programme at any of these facilities?
Facility 1 - Yes/No, Facility 2 - Yes/No, Facility 3 - Yes/No, Facility 4 - Yes/No

16. At any of these facilities, are there any classes or activities specifically for scheme participants beyond the 16 week programme?
Facility 1 - Yes (please describe below)/No (go to question 18)
Facility 2 - Yes (please describe below)/No (go to question 18)
Facility 3 - Yes (please describe below)/No (go to question 18)
Facility 4 - Yes (please describe below)/No (go to question 18)

.....
.....
.....
.....

...17. Who pays for these activities?

.....

18. What is the cost to participants?

.....

19. At any of these facilities, do participants receive discounted services after their 16 week programme?
Facility 1 - Yes (please describe below)/No
Facility 2 - Yes (please describe below)/No
Facility 3 - Yes (please describe below)/No
Facility 4 - Yes (please describe below)/No

.....
.....
.....
.....
.....
.....

Section 2 - Semi structured interview

The following is a list of potential topics to be covered during interviews. The list is not prescriptive or exclusive. Bullet points represent potential prompts. These may not be relevant in all interviews.

1. Role of the exercise professional

- What does your role as an exercise professional on the scheme involve?
- What kinds of support do you feel that patients look for from you as an exercise professional
- How diverse is the type of support required by different client groups?

2. Opinion of scheme and integration into context

- What are your thoughts about:
  - the referral process?
  - the exercise facilities?
  - the supportiveness of LC staff/ECs/WAG?
  - the activities/classes available?
  - training provided?
  - the scheme as a whole?
- Is there anything about the area in which you work that has made it easier or more difficult to implement the scheme?

3. One-to-one consultations and motivational interviewing

- What do you see as the main purposes of your one-to-one consultations with clients:
  - on entry to the scheme?
  - after 4 weeks?
  - after 16 weeks?
  - after the programme?
- To what extent do you base your consultations on motivational interviewing principles (and in what ways)?
- Do you feel that the training received at the beginning of the scheme enabled you to do this with confidence (and how about after the refresher training course you recently received)?

4. Perceived changes and processes of change

- Do you think that the scheme has been effective in causing changes in participants':
  - -physical activity
  - -mood/psychological well-being/ health outcomes
- Do you think the scheme is more effective in increasing physical activity for some client groups than others (who)?
- Do you think the intervention should be targeted towards specific population subgroups?
- What is it about the scheme that you think helps some clients to become more active?
- Leisure centre environment
- Social aspects of participation

5. Motivation

- How motivated, or ready to change, do you feel that clients are when they enter the programme?
  - What does it mean to be 'motivated'?
  - Are any client groups generally more or less motivated on entry to the programme (who, why)?
  - What do clients tend to be motivated by?
  - What kinds of issues tend to limit clients' motivation?

6. The randomised controlled trial

- Do you have any thoughts about the methods that are being used to evaluate the scheme?
- Has running NERS as part of a randomised controlled trial has had any impacts on the way the scheme has been implemented?

7. Other comments

Do you have any other thoughts or comments about the scheme? How it has been set-up? How it is funded? Or how it is being evaluated?

10.4 Appendix 4 – Exercise coordinator interview schedule

Section 1- structured interview

1. What rate are patients charged for group exercise during their NERS programme?.....
2. Are there any exceptions to this rate? (please describe).....
3. What services are available at the £1 rate?
   i. Supervised classes Yes/No
   ii. Supervised gym use Yes/No
   iii. Independent gym use Yes/No
   iv. Independent pool use Yes/No
4. For how long does this discount last?.....
5. Is there any limit on the number of times per week clients can use this discount during this period?
   .....
   ...
6. Are there any activities which remain at this rate after 16 weeks (if yes, please describe and how funded)?
   .....
   ...
7. Is each patient under the supervision of one single professional for all activities throughout their initial 4 weeks, or is supervision shared between professionals? One/Shared
8. Are patients limited to classes run by the professional for the first 4 weeks of their programme or allowed to access other activities from the start? Limited/not limited
9. Is each patient linked to a single professional for all consultations, or are these shared between professionals? One/Shared
10. Have you introduced any new activities since the end of randomisation to the trial?
    a. If yes, what are these?
    .....
11. Do you run any maintenance classes/gym sessions specifically for post 16-week NERS patients? (please describe)
    .....
    a. If yes, what are these?.....
12. Are any discounted services available after the programme?
    .....
    a. If yes, what are these? .....

Section 2 - Semi structured interview

The following is a list of potential topics to be covered during interviews. The list is not prescriptive or exclusive. Bullet points represent potential prompts. These may not be relevant in all interviews.

1. **Role of the exercise co-ordinator and professional**

- What does your role as an exercise coordinator on the scheme involve?

2. **National protocols and local contexts**

- What are your views in principle, on the movement from local areas running their own schemes, to the implementation of a single set of national protocols for exercise referral in Wales?
    - Benefits and disadvantages?
    - How effectively has NERS achieved National standardisation?
- What do you think of the NERS protocols?
    - Any aspects you like or dislike?
- How well do these protocols work *in your area*?
    - Any barriers or facilitators to implementation?
    - Have they been adapted to your local context in any way?
- In your opinion, what are the most important aspects of the scheme for bringing about change in patients behaviour?
- After the programme is complete, what exit routes are clients directed to?
    - How are patients introduced to exit routes? (4 weeks or 16 weeks)
- Can you think of any significant changes to the delivery of the scheme in your area between the start of the trial and now?

3. **Opinion of scheme**

- What are your thoughts about:
  - the facilities available in your area?
  - the supportiveness of leisure centres and integration of the scheme?
  - the activities/classes available (range, flexibility of times, change over time)?
  - training provided (to yourself and to EPs – any additional requirements)?
  - communications with the National coordinator/WAG?
  - the referral process? (during the trial and now)

4. **Opinion of the trial**

- Do you have any views on the acceptability of using random allocation and RCT methods to evaluate exercise referral?
- What effects, if any, did the trial format have on the implementation of the scheme?
- How effectively were the evaluation methods and their justifications communicated to you?
- How did you find the process of explaining the methods to GPs?
    - Acceptability to GPs?
    - Transition from previous scheme?
- How important do you feel it is to develop evidence for the effectiveness of practices such as exercise referral?
- Has anything changed about the scheme since randomisation ended in October?

5. **Uptake and reach**

- Do you feel that uptake of the scheme is adequate at the moment?
- Has uptake increased or decreased since the start of the trial?
- Is there anything about your area that you think has affected uptake positively or negatively?
- Are there any particular client groups for whom uptake appears particularly low or high?
- What kind of things have you done throughout the scheme to maximise uptake?
- How effectively does the scheme's reach currently extend throughout the area you are responsible for?

6. **Recruitment**

- How have you marketed the scheme to potential referrers?
- How receptive to *the scheme* have you found referrers to be? (any difficulties or objections, how dealt with, change over time - why)
- Have referral behaviour and your relationship with referrers changed since the end of randomisation?

7. **Sustainability**

- Do you feel that the scheme is sustainable in its present form? (why or why not? Any threats to sustainability?)

8. **Suggestions for improvement and other comments**

- What could be changed about the delivery of the scheme to make it effective for more people?

11 References

Abrams D B, Orleans C T, Niaura R S, Goldstein M G, Prochaska J O, and Velicer W (1996) Integrating individual and public health perspectives for treatment of tobacco dependence under managed health care: a combined stepped-care and matching model. *Annals of Behavioral Medicine* 18: 290-304.

Adams J, and White M (2003) Are activity promotion interventions based on the transtheoretical model effective? A critical review. *British Journal of Sports Medicine* 37: 106-114.

Adams J, and White M (2005) Why don't stage-based activity promotion interventions work? *Health Education Research* 20: 237-243.

Altman D G (1996) Better reporting of randomised controlled trials: the CONSORT statement. *British Medical Journal* 313: 570-571.

Amoroso C, Harris M F, Ampt A, Laws R A, McKenzie S, Williams A M, Jayasinghe U W et al. (2009) The 45 year old health check: Feasibility and impact on practices and patient behaviour. *Australian Family Physician* 38: 358-362.

Amrhein P C, Miller W R, Yahne C E, Palmer M, and Fulcher L (2003) Client commitment language during motivational interviewing predicts drug use outcomes. . *Journal of Consulting and Clinical Psychology* 71: 862-878.

Armstrong R, Waters E, Moore L, Riggs E, Cuervo L G, Lumbiganon P, and Hawe P (2008) Improving the reporting of public health intervention research: advancing TREND and CONSORT. *Journal of Public Health*: fdm082.

Ashenden R, Silagy C, and Weller D (1997) A systematic review of the effectiveness of promoting lifestyle change in general practice. *Family Practice* 14: 160-176.

Audrey S, Holliday J, and Campbell R (2006a) It's good to talk: Adolescent perspectives of an informal, peer-led intervention to reduce smoking. *Social Science & Medicine* 63: 320-334.

Audrey S, Holliday J, Parry-Langdon N, and Campbell R (2006b) Meeting the challenges of implementing process evaluation within randomized controlled trials: the example of ASSIST (A Stop Smoking in Schools Trial). *Health Education Research* 21: 366-377.

Baer J S, Beadnell B, Garrett S B, Hartzler B, Wells E A, and Peterson P L (2008) Adolescent change language within a brief motivational intervention and substance use outcomes. *Psychology of Addictive Behaviors* 22: 570-575.

Baert P (2005) *Philosophy of the social sciences: towards pragmatism.* Cambridge: Wiley.

Baldwin T, and Ford J (1988) Transfer of training: A review and directions for research. *Personnel Psychology* 41: 63–104

Bandura A (1977) Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review* 84: 191-215.

Baranowski T, and Stables G (2000) Process evaluations of the 5-a-Day projects. *Health Education and Behaviour* 27: 157-166.

Bartholomew K L, Czyzewski D I, Swank P R, McCormick L, and Parcel G S (2000) Maximizing the impact of the Cystic Fibrosis Family Education Program: Factors related to program diffusion. *Family & Community Health* 22: 27-47.

Bartholomew L K, Parcel G S, and Kok G (1998) Intervention mapping: A process for developing theory- and evidence-based health education programs. *Health Education & Behavior* 25: 545-563.

Bartholomew L K, Parcel G S, Kok G, and Gottlieb N H (2006) *Planning health promotion programs: an intervention mapping approach.* 2nd ed. San Francisco: Jossey-Bass.

Basch C E, Sliepcevich E M, Gold R S, Duncan D F, and Kolbe L J (1985) Avoiding Type III Errors in Health Education Program Evaluations: A Case Study. *Health Education & Behavior* 12: 315-331.

Baum F (1995) Researching public health: Behind the qualitative-quantitative methodological debate. *Social Science & Medicine* 40: 459-468.

Becker S, and Bryman A (2004) *Understanding research for social policy and practice: themes, methods and approaches.* The Policy Press.

Befort C A, Nollen N, Ellerbeck E F, Sullivan D K, Thomas J L, and Ahluwalia J S (2008) Motivational interviewing fails to improve outcomes of a behavioral weight loss program for obese African American women: a pilot randomized trial. *Journal of Behavioral Medicine* 31: 367-377.

Bennett D (2005) Replacing positivism in medical geography. *Social Science and Medicine* 60: 2685-2695

Bennett G A, Moore J, Vaughan T, Rouse L, Gibbins J A, Thomas P, James K et al. (2007a) Strengthening Motivational Interviewing skills following initial training: A randomised trial of workplace-based reflective practice. *Addictive Behaviors* 32: 2963-2975.

Bennett J A, Lyons K S, Winters-Stone K, Nail L M, and Scherer J (2007b) Motivational interviewing to increase physical activity in long-term cancer survivors - A randomized controlled trial. *Nursing Research* 56: 18-27.

Bennett J A, Young H M, Nail L M, Winters-Stone K, and Hanson G (2008) A telephone-only motivational intervention to increase physical activity in rural adults - A randomized controlled trial. *Nursing Research* 57: 24-32.

Bere E, Veierod M B, Bjelland M, and Klepp K I (2006) Outcome and process evaluation of a Norwegian school-randomized fruit and vegetable intervention: Fruits and Vegetables Make the Marks (FVMM). *Health Education Research* 21: 258-267.

Berg B (2004) *Qualitative Research Methods for the Social Sciences (5th ed.)* Boston: Pearson.

Bertens M G B C, Eiling E M, van den Borne B, and Schaalma H P (2009) Uma Tori! Evaluation of an STI/HIV-prevention intervention for Afro-Caribbean women in the Netherlands. *Patient Education and Counseling* 75: 77-83.

Berwick D M (2008) The Science of Improvement. *Journal of the American Medical Association* 299: 1182-1184.

Bleijlevens M, Hendriks M, van Haastregt J, van Rossum E, Kempen G, Diederiks J, Crebolder H et al. (2008) Process factors explaining the ineffectiveness of a multidisciplinary fall prevention programme: A process evaluation. *BMC Public Health* 8: 332.

Block G, Sternfeld B, Block C, Block T, Norris J, Hopkins D, Quesenberry C et al. (2009) Development of Alive! (A Lifestyle Intervention Via Email), and Its Effect on Health-related Quality of Life, Presenteeism, and Other Behavioral Outcomes: Randomized Controlled Trial. *Journal of Medical Internet Research* 19.

Bloor M, Frankland J, Thomas M, and Robson K (2000) *Focus Groups in Social Research.* London: Sage.

Boczkowski P J (1999) Mutual shaping of users and technologies in a national virtual community. *Journal of Communication* 49: 86-111.

Bodenheimer T, and Handley M A (2009) Goal-setting for behavior change in primary care: An exploration and status report. *Patient Education and Counseling* 76: 174-180.

Bolam B, McLean C, Pennington A, and Gillies P (2006) Using new media to build social capital for health: A qualitative process evaluation study of participation in the CityNet project. *J Health Psychol* 11: 297-308.

Bond L, Craig P, Egan M, Skivington K, and Thomson H (2010) Health improvement programmes: really too complex to evaluate? (letter). *British Medical Journal* 340: c1332-.

Bourdieu P, and Passeron J (1977) *Reproduction in education, society and culture.* London: Sage.

Bovend'Eerdt T J H, Botell R E, and Wade D T (2009) Writing SMART rehabilitation goals and achieving goal attainment scaling: a practical guide. *Clinical Rehabilitation* 23: 352-361.

Braun V, and Clarke V (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology* 3: 77-101.

Brehm J W (1966) *A Theory of Psychological Reactance.* New York: Academic Press.

Britt E, Hudson S M, and Blampied N M (2004) Motivational interviewing in health settings: a review. *Patient Education and Counseling* 53: 147-155.

Britten N (1995) Qualitative Research: Qualitative interviews in medical research. *British Medical Journal* 311: 251-253.

Britton P C, Williams G C, and Conner K R (2008) Self-Determination Theory, Motivational Interviewing, and the treatment of clients with acute suicidal ideation. *Journal of Clinical Psychology* 64: 52-66.

Brodie D A, and Inoue A (2005) Motivational interviewing to promote physical activity for people with chronic heart failure. *Journal of Advanced Nursing* 50: 518-527.

Broers S, Smets E, Bindels P, Evertsz F B, Calff M, and de Haes H (2005) Training general practitioners in behavior change counseling to improve asthma medication adherence. *Patient Education and Counseling* 58: 279-287.

Brown C H, Wang W, Kellam S G, Muthén B O, Petras H, Toyinbo P, Poduska J et al. (2008) Methods for testing theory and evaluating impact in randomized field trials: Intent-to-treat analyses for integrating the perspectives of person, place, and time. *Drug and alcohol dependence* 95: S74-S104.

Brug J, Spikmans F, Aartsen C, Breedveld B, Bes R, and Fereira I (2007) Training dietitians in basic motivational interviewing skills results in changes in their counseling style and in lower saturated fat intakes in their patients. *Journal of Nutrition Education and Behavior* 39: 8-12.

Brustad M, Skeie G, Braaten T, Slimani N, and Lund E (2003) Comparison of telephone vs face-to-face interviews in the assessment of dietary intake by the 24h recall EPIC SOFT program the Norwegian calibration study. *European Journal of Clinical Nutrition* 57: 107-113.

Bryman A (2006) Integrating quantitative and qualitative research: how is it done? *Qualitative Research* 6: 97-113.

Bryman A (2007) Barriers to integrating quantitative and qualitative Research. *Journal of Mixed Methods Research* 1: 8-22.

Burke B, Arkowitz H, and Menchola M (2003) The efficacy of motivational interviewing: a meta-analysis of controlled clinical trials. *Journal of Consulting and Clinical Psychology* 71: 843-861.

Butler K, Rayens M, Zhang M, Maggio L, Riker C, and EJ. H (2009) Tobacco dependence treatment education for baccalaureate nursing students. *Journal of Nursing Education.* 48: 249-254.

Campbell A, Mutrie N, White F, McGuire F, and Kearney N (2005) A pilot study of a supervised group exercise programme as a rehabilitation treatment for women with breast cancer receiving adjuvant treatment. *European Journal of Oncology Nursing* 9: 56-63.

Campbell J P, Maxey V A, and Watson W A (1995) Hawthorne effect: Implications for prehospital research. *Annals of Emergency Medicine* 26: 590-594.

Campbell M, Fitzpatrick R, Haines A, Kinmouth A L, Sandercock P, Spiegelhalter D, and Tyrer P (2000) Framework for design and evaluation of complex interventions to improve health. *British Medical Journal* 321: 694-696.

Campbell M K, Resnicow K, Carr C, Wang T, and Williams A (2007) Process Evaluation of an Effective Church-Based Diet Intervention: Body & Soul. *Health Education & Behaviour* 34: 864-880.

Carels R A, Darby L, Cacciapaglia H M, Konrad K, Coit C, Harper J, Kaplar M E et al. (2007) Using motivational interviewing as a supplement to obesity treatment: A stepped-care approach. *Health Psychology* 26: 369-374.

Carlin J B, Taylor P, and Nolan T (1998) School based bicycle safety education and bicycle injuries in children: a case-control study. *Injury Prevention* 4: 22-27.

Carol H W (1997) Theory-based evaluation: Past, present, and future. *New Directions for Program Evaluation* 1997: 41-55.

Carroll C, Patterson M, Wood S, Booth A, Rick J, and Balain S (2007) A conceptual framework for implementation fidelity. *Implementation Science* 2: 40.

Chatzisarantis N L D, and Hagger M S (2009) Effects of an intervention based on self-determination theory on self-reported leisure-time physical activity participation. *Psychology & Health* 24: 29-48.

Chinn D J, White M, Howel D, Harland J O E, and Drinkwater C K (2006) Factors associated with non-participation in a physical activity promotion trial. *Public Health* 120: 309-319.

Claassen D (2007) Financial incentives for antipsychotic depot medication: ethical issues. *Journal of Medical Ethics* 33: 189-193.

Clark A M (1998) The qualitative-quantitative debate: moving from positivism and confrontation to post-positivism and reconciliation. *Journal of Advanced Nursing* 27: 1242-1249.

Clark A M, MacIntyre P D, and Cruickshank J (2007) A critical realist approach to understanding and evaluating heart health programmes. *Health (London)* 11: 513-539.

Cochrane T, and Davey R (1998) Evaluation of exercise prescription for 25 general practices and a large leisure complex in Sheffield. *Journal of Sports Sciences* 16: 17-18.

Coffey A, and Atkinson P (1996) *Making sense of qualitative data: complementary research strategies*. London: Sage.

Connelly J (2002) The Behavior Change Consortium studies: missed opportunities--individual focus with an inadequate engagement with personhood and socio-economic realities. *Health Education Research.* 17: 691-695.

Coote A, Allen J, and Woodhead D (2004) *Finding out what works: building knowledge about complex community-based interventions*. London: Kings Fund.

Corbett K, Thompson B, White N, and Taylor M (1990) Process evaluation in the community intervention trial for smoking cessation (COMMIT). *International Quarterly of Community Health Education* 11: 291-309.

Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, and Petticrew M (2008a) *Developing and evaluating complex interventions: new guidance*. Medical Research Council.

Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, and Petticrew M (2008b) Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 337: a1655.

Creegan C, and Hedges A (2007) *Towards a policy evaluation service: Developing infrastructure to support the use of experimental and quasi-experimental methods* London: Ministry of Justice.

Creswell J W, and Clark P L K (2007) *Designing and conducting mixed methods research*. London: Sage.

Creswell J W, Fetters M D, and Ivankova N V (2004) Designing A Mixed Methods Study In Primary Care. *Ann Fam Med* 2: 7-12.

Crone D, Johnston L, and Grant T (2006) Maintaining quality in exercise referral schemes: a case study of professional practice. *Primary Health Care Research and Development* 5: 96-103.

Crone D, Johnston L H, Gidlow C, Henley C, and James D V B (2008) Uptake and participation in physical activity referral schemes in the UK: An investigation of

patients referred with mental health problems. *Issues in Mental Health Nursing* 29: 1088-1097.

Crone D, Smith A, and Gough B (2005) 'I feel totally at one, totally alive and totally happy': a psycho-social explanation of the physical activity and mental health relationship. *Health Education Research* 20: 600-611.

Curran S, Gittelsohn J, Anliker J, Ethelbah B, Blake K, Sharma S, and Caballero B (2005) Process evaluation of a store-based environmental obesity intervention on two American Indian Reservations. *Health Education Research* 20: 719-729.

Curry L A, Nembhard I M, and Bradley E H (2009) Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation* 119: 1442-1452.

Damush T M, Stump T E, Saporito A, and Clark D O (2001) Predictors of older primary care patients' participation in a submaximal exercise test and a supervised, low-impact exercise class. *Preventive Medicine* 33: 485-494.

Danielzik S, Pust, S, Müller, MJ. (2007) School-based interventions to prevent overweight and obesity in prepubertal children: process and 4-years outcome evaluation of the Kiel Obesity Prevention Study (KOPS). *Acta Paediatrica Supplement. 96(454):19-25, 2007 Apr.*

Davidson R A (1986) Source of Funding and Outcome of Clinical-Trials. *Journal of General Internal Medicine* 1: 155-158.

Davies H T O, Nutley S M, and Smith P (2000) *What works? Evidence-based policy and practice in public services.* Bristol: Policy press

Day F, and Nettleton B (2001) The Scottish Borders general practitioners exercise referral scheme (GPERS). *Health Bulletin* 59: 343.

Deci E L, Koestner R, and Ryan R M (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 125: 627-668.

Deci E L, and Ryan R M (1985) *Intrinisc motivation and self determination in human behaviour.* New York: Plenum.

Denzin N K, and Lincoln Y S (2000) *Handbook of qualitative research.* 2nd ed. Thousand Oaks, CA: Sage.

Department of Health (2001) *National quality assurance framework for exercise referral schemes.* London: The Stationery Office.

Department of Health (2007) *DH statement on Exercise Referral (March 2007)* [Online]. Available at: http://www.erpho.org.uk/Download/Public/16041/1/DH%20Statement%20on%20exercise%20referral%20Mar07.pdf [Accessed: April 2010]

Dinan S, Lenihan P, Tenn T, and Iliffe S (2006) Is the promotion of physical activity in vulnerable older people feasible and effective in general practice? *British Journal of General Practice* 56: 791-793.

Dishman R, DeJoy D, Wilson M, and Vandenberg M (2009a) Move to Improve: A Randomized Workplace Trial to Increase Physical Activity. *American Journal of Preventive Medicine* 36: 133-141.

Dishman R K, Vandenberg R J, Motl R W, Wilson M G, and DeJoy D M (2009b) Dose relations between goal setting, theory-based correlates of goal setting and increases in physical activity during a workplace trial. *Health Educ. Res.*: cyp042.

Donnelly J E, Smith B K, Dunn L, Mayo M M, Jacobsen D J, Stewart E E, Gibson C et al. (2007) Comparison of a phone vs clinic approach to achieve 10% weight loss. *Int J Obes* 31: 1270-1276.

Dugdill L, Graham R C, and McNair F (2005) Exercise referral: the public health panacea for physical activity promotion? A critical perspective of exercise referral schemes; their development and evaluation. *Ergonomics* 48: 1390-1410.

Dunbar A (1992) National fitness survey: a report on activity patterns and fitness levels. *London, UK: Sports Council and Health Education Authority.*

Dunn C, DeRoo L, and Rivara F (2001) The use of brief interventions adapted from motivational interviewing across behavioral domains: A systematic review. . *Addiction* 1725-42.

Dusenbury L, Brannigan R, Falco M, and Hansen W B (2003) A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Educ. Res.* 18: 237-256.

Eakin E, Glasgow R E, and Riley K M (2000) Review of primary care-based physical activity intervention studies. *Journal of Family Practice* 49: 158-168.

Eakin E G, Lawler S P, Vandelanotte C, and Owen N (2007) Telephone Interventions for Physical Activity and Dietary Behavior Change: A Systematic Review. *American Journal of Preventive Medicine* 32: 419-434.

Eaton C B, and Menard L M (1998) A systematic review of physical activity promotion in primary care office settings. *British Journal of Sports Medicine* 32: 11-16.

Edmunds J, Ntoumanis N, and Duda J L (2007) Adherence and well-being in overweight and obese patients referred to an exercise on prescription scheme: A self-determination theory perspective. *Psychology of Sport and Exercise* 8: 722-740.

Edmunds J, Ntoumanis N, and Duda J L (2008) Testing a self-determination theory-based teaching style intervention in the exercise domain. *European Journal of Social Psychology* 38: 375-388.

Emslie C, Whyte F, Campbell A, Mutrie N, Lee L, Ritchie D, and Kearney N (2007) 'I wouldn't have been interested in just sitting round a table talking about cancer'; exploring the experiences of women with breast cancer in a group exercise trial. *Health Education Research.* 22: 827-38

Escoffery C, Glanz K, and Elliott T (2008) Process evaluation of the Pool Cool Diffusion Trial for skin cancer prevention across 2 years. *Health Education Research.* 23: 732-743.

Fairburn C, and Beglin S (1994) Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders* 16: 363-370.

Feathers J T, Kieffer E C, Palmisano G, Anderson M, Janz N, Spencer M S, Guzman R et al. (2007) The development, implementation, and process evaluation of the REACH Detroit Partnership's diabetes lifestyle intervention. *The Diabetes Educator* 33: 509-520.

Festinger L (1957) *A theory of cognitive dissonance.* CA: Stanford University Press.

Finnegan J R J R, Murray D M, Kurth C, and McCarthy P (1989) Measuring and tracking education program implementation: The Minnesota Heart Health Program experience. *Health Educ Behav* 16: 77-90.

Fitzgerald L, Ferlie E, Wood M, and Hawkins C (2002) Interlocking interactions, the Diffusion of Innovations in Health Care. *Human Relations* 55: 1429-1449.

Fogarty J S (1997) Reactance theory and patient noncompliance. *Social Science & Medicine* 45: 1277-1288.

Foresight (2007) *Tackling obesities: Future choices.* London: Government Office for Science.

Forsberg L, Forsberg L, Lindqvist H, and Helgason A (2010) Clinician acquisition and retention of Motivational Interviewing skills: a two-and-a-half-year exploratory study. *Substance Abuse Treatment, Prevention, and Policy* 5: 8.

Forsetlund L, Talseth K O, Bradley P, Nordheim L, and Bjorndal A (2003) Many a slip between cup and lip: Process evaluation of a program to promote and support evidence-based public health practice. *Evaluation Review* 27: 179-209.

Fox K, Biddle S, Edmunds L, Bowler I, and Killoran A (1997) Physical activity promotion through primary health care in England. *British Journal of General Practice* 47: 367-369.

Freeman R (2009) Health promotion and the randomised controlled trial: a square peg in a round hole? *BMC Oral Health* 9: 1.

Frey J H, and Fontana A (1991) The group interview in social research. *Social Science Journal* 28: 175-187.

Friedberg M, Saffran B, Stinson T J, Nelson W, and Bennett C L (1999) Evaluation of Conflict of Interest in Economic Analyses of New Drugs Used in Oncology. *Journal of the American Medical Association* 282: 1453-1457.

Ganley B, Sheets I, Buccheri R, Thomas S A, Doerr-Kashani P, Bolla C, Stoker D et al. (2004) Collaboration Versus Competition: Results of an Academic Practice Alliance. *Journal of Community Health Nursing* 21: 153 - 165.

Gaume J, Gmel G, and Daeppen J-B (2008) Brief alcohol interventions: Do counsellors' and patients' communication characteristics predict change? *Alcohol and Alcoholism* 43 62–69.

Gerald L B, Bruce F, Brooks C M, Brook N, Kimerling M E, Windsor R A, and Bailey W C (2003) Standardizing contact investigation protocols. *The International Journal of Tuberculosis and Lung Disease* 7: S369-S374.

Gidlow C, Johnston L H, Crone D, and James D (2005) Attendance of exercise referral schemes in the UK: A systematic review. *Health Education Journal* 64: 168-186.

Gidlow C, Johnston L H, Crone D, and James D V B (2008) State of the Art Reviews: Methods of Evaluation: Issues and Implications for Physical Activity Referral Schemes. *American Journal of Lifestyle Medicine* 2: 46-50.

Gidlow C, Johnston L H, Crone D, Morris C, Smith A, Foster C, and James D V (2007) Socio-demographic patterning of referral, uptake and attendance in physical activity referral schemes. *Journal of Public Health* 29: 107-113.

Gilbert T (2006) Mixed methods and mixed methodologies: The practical, the technical and the political. *Journal of Research in Nursing* 11: 205-217.

Gilligan C (1982) *In a different voice: psychological theory and women's development.* Cambridge MA: Harvard University Press.

Glaser B G, and Strauss A L (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Chicago:

Glasgow R E, McKay H G, Piette J D, and Reynolds K D (2001) The RE-AIM framework for evaluating interventions: what can it tell us about approaches to chronic illness management? *Patient Education and Counseling* 44: 119-127.

Glasgow R E, Vogt T M, and Boles S M (1999) Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 89: 1322-1327.

Gnich W, Sheehy C, Amos A, Bitel M, and Platt S (2008) A Scotland-wide pilot programme of smoking cessation services for young people: process and outcome evaluation. *Addiction* 103: 1866-1874.

Godwin M, Ruhland L, Casson I, MacDonald S, Delva D, Birtwhistle R, Lam M et al. (2003) Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. *BMC Medical Research Methodology* 3: 28.

Goodman R M, and Steckler A (1989) A model for the institutionalization of health promotion programs. *Family & Community Health* 11: 63-78.

Gorard S (2002) Political control: A way forward for educational research. *British Journal of Educational Studies* 50: 378-389.

Gorard S (2004) Sceptical or clerical? Theory as a barrier to the combination of research methods. *Journal of Educational Enquiry* 5: 1-21.

Gorard S, and Taylor C (2004) *Combining methods in educational and social research.* London Open University Press.

Graham R C, Dugdill L, and Cable N T (2005) Health professionals' perspectives in exercise referral: implications for the referral process. *Ergonomics* 48: 1411 - 1422.

Greater Glasgow Health Board (2004) *An evaluation report of the Glasgow Exercise Referral Scheme.* Glasgow: Greater Glasgow Health Board.

Greaves C J, Middlebrooke A, O'Loughlin L, Holland S, Piper J, Steele A, Gale T et al. (2008) Motivational interviewing for modifying diabetes risk: a randomised controlled trial. *British Journal of General Practice* 58: 535-540.

Green C A, and Pope C R (1999) Gender, psychosocial factors and the use of medical services: a longitudinal analysis. *Social Science & Medicine* 48: 1363-1372.

Green J, and Tones K (1999) For debate. Towards a secure evidence base for health promotion. *Journal of Public Health* 21: 133-139.

Greenhalgh T, Robert G, and Bate P (2003) How to spread good ideas. A Systematic Review of the Literature on Diffusion, Dissemination and Sustainability of Innovations in Health Service Delivery and Organisation.

Greenhalgh T, Robert G, Macfarlane F, Bate P, and Kyriakidou O (2004) Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations. *The Milbank Quarterly* 82: 581-629.

Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, and Peacock R (2005) Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. *Social Science & Medicine* 61: 417-430.

Griffin S F, Wilcox S, Ory M G, Lattimore D, Leviton L, Castro C, Carpenter R A et al. (2010) Results from the Active for Life process evaluation: program delivery fidelity and adaptations. *Health Education Research* 25: 325-342.

Hanson W E, Creswell J W, Clark V L P, Petska K S, and Creswell J D (2005) Mixed methods research designs in counseling psychology. *Journal of Counseling Psychology* 52: 224-235.

Hardcastle S, Taylor A, Bailey M, and Castle R (2008) A randomised controlled trial on the effectiveness of a primary health care based counselling intervention on physical activity, diet and CHD risk factors. *Patient Education and Counseling* 70: 31-39.

Hardcastle S, and Taylor A H (2001) Looking for more than weight loss and fitness gain: Psychosocial dimensions among elder women in a primary-care exercise-referral program. *Journal of Aging and Physical Activity* 9: 313-328.

Hardcastle S, and Taylor A H (2005) Finding an exercise identity in an older body: "It's redefining yourself and working out who you are". *Psychology of Sport and Exercise* 6: 173-188.

Harland J, White M, Drinkwater C, Chinn D, Farr L, and Howel D (1999) The Newcastle exercise project: a randomised controlled trial of methods, to promote physical activity in primary care. *British Medical Journal* 319: 828-832B.

Harrison R A, McNair F, and Dugdill L (2005a) Access to exercise referral schemes - a population based analysis. *Journal of Public Health* 27: 326-330.

Harrison R A, Roberts C, and Elton P J (2005b) Does primary care referral to an exercise programme increase physical activity 1 year later? A randomized controlled trial. *Journal of Public Health* 27: 25-32.

Hawe P, Shiell A, and Riley T (2004a) Complex interventions: how "out of control" can a randomised controlled trial be? *British Medical Journal* 328: 1561-1563.

Hawe P, Shiell A, and Riley T (2009) Theorising Interventions as Events in Systems. *American Journal of Community Psychology* 43: 267-276.

Hawe P, Shiell A, Riley T, and Gold L (2004b) Methods for exploring implementation variation and local context within a cluster randomised community intervention trial. *Journal of Epidemiology and Community Health* 58: 788-793.

Heim S, Stang J, and Ireland M (2009) A Garden Pilot Project Enhances Fruit and Vegetable Consumption among Children. *Journal of the American Dietetic Association* 109: 1220-1226.

Heisler M, Bouknight R R, Hayward R A, Smith D M, and Kerr E A (2002) The relative importance of physician communication, participatory decision making, and patient understanding in diabetes self-management. *Journal of General Internal Medicine* 17: 243-252.

Hettema J, Steele J, and Mller W R (2005) Motivational interviewing. *Annual Reviews of Clinical Psychology* 1: 91-111.

Hillsdon M, Foster C, Naidoo B, and Crombie H (2004) *The effectiveness of public health interventions for increasing physical activity among adults: a review of reviews*. NHS: Health Development Agency.

Hillsdon M, Thorogood M, White I, and Foster C (2002) Advising people to take more exercise is ineffective: a randomized controlled trial of physical activity promotion in primary care. *International Journal of Epidemiology* 31: 808-815.

Hodgins D C, Ching, L. E., & McEwen, J. (2009) Strength of commitment language in motivational interviewing and gambling outcomes. *Psychology of Addictive Behaviors* 23: 122-130.

Holliday J, Moore G, and Moore L (2009) Changes in child exposure to secondhand smoke after implementation of smoke-free legislation in Wales: a repeated cross-sectional study. *BMC Public Health* 9: 430.

Holstein J, and Gubrium J (2004) The active interview. In: Silverman D [ed.] *Qualitative research: Theory, method and practice*. (2nd edn.) London: Sage, pp. 140-161.

Hosper K, Deutekom M, and Stronks P (2008) The effectiveness of "Exercise on Prescription" in stimulating physical activity among women in ethnic minority groups in the Netherlands: protocol for a randomized controlled trial. *Bmc Public Health* 8: 406.

House of Commons Health Committee (2009) *Health inequalities: third report of session 2008-9*. London: Stationery Office.

Humphries K H, and Gill S (2003) Risks and benefits of hormone replacement therapy: The evidence speaks. *CMAJ* 168: 1001-1010.

Hunt M K, Barbeau E M, Lederman R, Stoddard A M, Chetkovich C, Goldman R, Wallace L et al. (2007) Process evaluation results from the healthy directions-small business study. *Health Education & Behavior* 34: 90-107.

Hunt M K, Lederman R, Stoddard A M, LaMontagne A D, McLellan D, Combe C, Barbeau E et al. (2005) Process Evaluation of an Integrated Health Promotion/Occupational Health Model in WellWorks-2. *Health Education & Behaviour* 32: 10-26.

Huston P, Hogg W, Newbury A, and Martin C (2006) A process evaluation of an intervention to improve respiratory infection control practices in family physician offices. *Canadian Journal of Public Health. Revue Canadienne de Sante Publique.* 97: 475-479.

Hyler S E, Skodol A E, Kellman H D, Oldham J M, and Rosnick L (1990) Validity of the Personality Diagnostic Questionnaire--revised: comparison with two structured interviews. *American Journal of Psychiatry* 147: 1043-1048.

Iliffe S, Masud T, Skelton D, and Kendrick D (2008) Promotion of exercise in primary care. *British Medical Journal* 337: a2430-.

Iliffe S, Tai S S, Gould M, Thorogood M, and Hillsdon M (1994) Prescribing Exercise in General-Practice. *British Medical Journal* 309: 494-495.

Inchley J, Muldoon J, and Currie C (2007) Becoming a health promoting school: evaluating the process of effective implementation in Scotland. *Health Promotion International* 22: 65-71.

Isaacs A J, Critchley J A, Tai S S, Buckingham K, Westley D, Harridge S D R, Smith C et al. (2007) Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. *Health Technology Assessment* 11: 1-+.

Jackson C, Bell F, Smith R, and Dixey R (1998) Do adherers and non-adherers to a GP exercise referral scheme differ in their long-term activity levels? *Sports Science* 16: 1.

Jalleh G D, RJ. James, R. Ambridge, J. (2007) Process evaluation of the Act-Belong-Commit Mentally Healthy WA campaign: first 12 months data. *Health Promotion Journal of Australia.* 18: 217-220.

James D V B, Johnston L H, Crone D, Sidford A H, Gidlow C, Morris C, and Foster C (2008) Factors associated with physical activity referral uptake and participation. *Journal of Sports Sciences* 26: 217-224.

Jancey J M, Clarke A, Howat P A, Lee A H, Shilton T, and Fisher J (2008) A Physical Activity Program to Mobilize Older People: A Practical and Sustainable Approach. *Gerontologist* 48: 251-257.

Johnson C C, Lai Y L, Rice J, Rose D, and Webber L S (2010) ACTION Live: Using process evaluation to describe implementation of a worksite wellness program. *Journal of Occupational and Environmental Medicine* 52: S14.

Johnson R B, and Onwuegbuzie A J (2004) Mixed methods research: A research paradigm whose time has come. *Educational researcher* 33: 14-26.

Johnston L H, Warwick J, De Ste Croix M, Crone D, and Sldford A (2005) The nature of all 'inappropriate referrals' made to a countywide physical activity referral scheme: Implications for practice. *Health Education Journal* 64: 58.

Jolly K, Duda J, Daley A, Eves F, Mutrie N, Ntoumanis N, Rouse P et al. (2009) Evaluation of a standard provision versus an autonomy promotive exercise referral programme: rationale and study design. *BMC Public Health* 9: 176.

Jonas S, and Phillips E (2009) *ACSM's Exercise is Medicine: A Clinician's Guide to Exercise Prescription.* New York: Lippincott Williams & Wilkins.

Jones F, Harris P, Waller H, and Coggins A (2005) Adherence to an exercise prescription scheme: The role of expectations, self-efficacy, stage of change and psychological well-being. *British Journal of Health Psychology* 10: 359-378.

Jowell R (2003) *Trying it Out: The role of pilots in policy-making*. London: The Cabinet Office.

Jurg M E, Kremers S P J, Candel M J J M, Van der Wal M F, and Meij J S B D (2006) A controlled trial of a school-based environmental intervention to improve physical activity in Dutch children: JUMP-in, kids in motion. *Health Promot. Int.* 21: 320-330.

Kallings L V, Leijon M, Hellenius M L, and Stahle A (2008) Physical activity on prescription in primary health care: a follow-up of physical activity level and quality of life. *Scandinavian Journal of Medicine & Science in Sports* 18: 154.

Kazi M (2003) Realist Evaluation for practice. *British Journal of Social Work* 33: 803-818.

Keller C, Fleury J, Gregor-Holt N, and Thompson T (1999) Predictive ability of social cognitive theory in exercise research: An integrated literature review. *Online Journal of Knowledge Synthesis for Nursing* 6: art. no.-2.

Kellogg Foundation W K (2004) *Logic model development guide*. Battle Creek, MI: W.K Kellogg Foundation.

Khun S, and Manderson L (2007) Community and school-based health education for dengue control in rural Cambodia: a process evaluation. *PLoS Neglected Tropical Diseases* 1.

King A C, Ahn D K, Oliveira B M, Atienza A A, Castro C M, and Gardner C D (2008) Promoting Physical Activity Through Hand-Held Computer Technology. *American Journal of Preventive Medicine* 34: 138-142.

Kirk A, Mutrie N, MacIntyre P, and Fisher M (2003) Increasing Physical Activity in People With Type 2 Diabetes. *Diabetes Care* 26: 1186-1192.

Kloek G, van Lenthe F, Meertens Y, Koelen M, and Mackenbach J (2006) Process Evaluation of a Dutch Community Intervention to improve Health Related Behaviour in deprived neighbourhoods. *Sozial- und Präventivmedizin/Social and Preventive Medicine* 51: 259-272.

Kolt G S, Oliver M, Schofield G M, Kerse N, Garrett N, and Latham N K (2006) An overview and process evaluation of TeleWalk: a telephone-based counseling intervention to encourage walking in older adults. *Health Promotion International* 21: 201-208.

Komro K, Perry C L, Veblen-Mortenson S, Farbakhsh K, Toomey T L, Stigler M H, Jones-Webb R et al. (2008) Outcomes from a randomized controlled trial of a multi-

component alcohol use preventive intervention for urban youth: Project Northland Chicago. *Addiction* 103: 606-618.

Kratz R, Ponce N, and Yancey A (2008) Process evaluation of the Los Angeles Unified School District Nutrition Network. *Preventing Chronic Disease* 5(2): A42.

Kuchipudi V, Hobein K, Fleckinger A, and Iber F L (1990) Failure of a 2-hour motivational intervention to alter recurrent drinking behavior in alcoholics with gastrointestinal disease. *Journal of Studies on Alcohol* 51: 356-360.

Kuper A, Lingard L, and Levinson W (2008) Critically appraising qualitative research. *British Medical Journal* 337: 687-689.

Lamb S E, Bartlett H P, Ashley A, and Bird W (2002) Can lay-led walking programmes increase physical activity in middle aged adults? A randomised controlled trial. *Journal of Epidemiology and Community Health* 56: 246-252.

Lane C (2005) *The Behaviour Change Counselling Index (BECCI). Manual for Coding Behaviour Change Counselling* [Online]. Available at: http://www.cardiff.ac.uk/sonms/resources/20020930BECCIManual.doc [Accessed:

Lane C, Hood K, and Rollnick S (2008) Teaching motivational interviewing: using role play is as effective as using simulated patients. *Medical Education* 42: 637-644.

Lane C, Huws-Thomas M, Hood K, Rollnick S, Edwards K, and Robling M (2005) Measuring adaptations of motivational interviewing: the development and validation of the behavior change counseling index (BECCI). *Patient Education and Counseling* 56: 166-173.

Lane C, Johnson S, Rollnick S, Edwards K, and Lyons M (2003) Consulting about lifestyle change: Evaluation of a training course for specialist diabetes nurses. *Practical Diabetes International* 20: 204-208.

Latimer A E, Ginis K A M, and Arbour K P (2006) The efficacy of an implementation intention intervention for promoting physical activity among individuals with spinal cord injury: A randomized controlled trial. *Rehabilitation Psychology* 51: 273-280.

Lawlor D A, and Hanratty B (2001) The effect of physical activity advice given in routine primary care consultations: a systematic review. *Journal of Public Health Medicine* 23: 219-226.

Lawlor D A, Keen S, and Neal R D (2000) Can general practitioners influence the nation's health through a population approach to provision of lifestyle advice? *British Journal of General Practice* 50: 455-459.

Lawton B A, Rose S B, Elley C R, Dowell A C, Fenton A, and Moyes S A (2008) Exercise on prescription for women aged 40-74 recruited through primary care: two year randomised controlled trial. *British Medical Journal* 337: a2509.

Lee A S W, Griffin S J, and Simmons R K (2009) An evaluation of the effectiveness of 'Active for Life': An exercise referral scheme in West Suffolk. *Public Health* 123: 670-672.

Lee T W, Locke E A, and Lantham G (1989) Goal setting theory and job performance. *Goal concepts in personality and social psychology.* New Jersey: Lawrence Erlbaum.

Leicester G (1999) The seven enemies of evidence-based policy. *Public Money & Management* 19: 5-7.

Leijon M E, Bendtsen P, Nilsen P, Festin K, and Stahle A (2009) Does a physical activity referral scheme improve the physical activity among routine primary health care patients? *Scandinavian Journal of Medicine & Science in Sports* 19: 627.

Lillian Southwick B, and Rui W (1991) The Role of Psychological Reactance in Drinking Following Alcohol Prevention Messages¹. *Journal of Applied Social Psychology* 21: 1111-1124.

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

Liu X, Li L, Cui H, and Jackson V W (2009) Evaluation of an emergency department-based injury surveillance project in China using WHO guidelines. *Injury Prevention* 15: 105.

Lobb R, Gonzalez Suarez E, Fay M E, Gutheil C M, Hunt M K, Fletcher R H, and Emmons K M (2004) Implementation of a cancer prevention program for working class, multiethnic populations. *Preventive Medicine* 38: 766-776.

Locke E A, and Latham G P (2002) Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist* 57.

Locke E A, and Latham G P (2006) New Directions in Goal-Setting Theory. *Current Directions in Psychological Science* 15: 265-268.

Lord J C, and Green F (1995) Exercise on prescription: does it work? *Health Education Journal* 54: 453-464.

Lowe J B, Ball J, Lynch B M, Baldwin L, Janda M, Stanton W R, and Aitken J F (2004) Acceptability and feasibility of a community-based screening programme for melanoma in Australia. *Health Promotion International* 19: 437-444.

Lubans D R, Foster C, and Biddle S J H (2008) A review of mediators of behavior in interventions to promote physical activity among children and adolescents. *Preventive Medicine* 47: 463-470.

Lynch R, Maio G, Moore G F, Moore L, Orford S, Robinson A, Taylor C et al. (2007) *ESRC/HEFCW Scoping Study into Quantitative Methods Capacity Building in Wales.* Economic and Social Research Council.

Macintyre S (2003) Evidence based policy making. *British Medical Journal* 326: 5-6.

Mackenzie M, O'Donnell C, Halliday E, Sridharan S, and Platt S (2010) Do health improvement programmes fit with MRC guidance on evaluating complex interventions? *British Medical Journal* 340: c185-.

Macniven R O, B. Shilton, T. Bauman, A. (2008) The initial development and process evaluation of AusPAnet: the Australian Physical Activity Network. . *Health Promotion Journal of Australia.* 19: 36-39.

Madson M B, and Campbell T C (2006) Measure of fidelity in motivational enhancement: A systematic review. *Journal of Substance Abuse Treatment* 31: 67-73.

Madson M B, Loignon A C, and Lane C (2009) Training in motivational interviewing: A systematic review. *Journal of Substance Abuse Treatment* 36: 101-109.

Maiorana A, Kegeles S, Fernandez P, Salazar X, Cáceres C, Sandoval C, Rosasco A M et al. (2007) Implementation and evaluation of an HIV/STD intervention in Peru. *Evaluation and Program Planning* 30: 82-93.

Marino R W, C. Minichiello, V. Schofield, M. Calache, H. (2005) A qualitative process evaluation of an oral health promotion program for older migrant adults.. *Health Promotion Journal of Australia.* 16: 225-228.

Markland D, Ryan R M, Tobin V J, and Rollnick S (2005) Motivational interviewing and self-determination theory. *Journal of Social and Clinical Psychology* 24: 811-831.

Markland D, and Tobin V J (2010) Need support and behavioural regulations for exercise among exercise referral scheme clients: The mediating role of psychological need satisfaction. *Psychology of Sport and Exercise* 11: 91-99.

Marshall S J, and Biddle S J H (2001) The transtheoretical model of behavior change: A meta-analysis of applications to physical activity and exercise. *Annals of Behavioral Medicine* 23: 229-246.

Marteau T M, Ashcroft R E, and Oliver A (2009) Using financial incentives to achieve healthy behaviour. *BMJ* 338: b1415-.

Martens M, van Assema P, Paulussen T, Schaalma H, and Brug J (2006) Krachtvoer: process evaluation of a Dutch programme for lower vocational schools to promote healthful diet. *Health Educ. Res.* 21: 695-704.

Martin, and Woolf M (1999) The retrospective evaluation of a general practitioner exercise prescription programme. *Journal of Human Nutrition and Dietetics* 12: 32-42.

Martins R K, and McNeil D W (2009) Review of Motivational Interviewing in promoting health behaviors. *Clinical Psychology Review* 29: 283-293.

Matland R E (1995) Synthesizing the implementation literature: The Ambiguity-Conflict Model of policy implementation. *Journal of Public Administration Research and Theory* 5: 145-174.

Mavimbe J, Braa J, and Bjune G (2005) Assessing immunization data quality from routine reports in Mozambique. *BMC Public Health* 5: 108.

May T (2001) *Social research, issues, methods and process*. 2nd ed. California: Open University.

McAuley E, Morris K, Motl R, Hu L, Konopack J, and Elavsky S (2007) Long-term follow-up of physical activity behavior in older adults. *Health Psychology* 26: 375-380.

McConnon Á, Kirk S F L, and Ransley J K (2009) Process evaluation of an internet-based resource for Weight Control: Use and views of an obese sample. *Journal of Nutrition Education and Behavior* 41: 261-267.

McElroy M (2002) *Resistance to exercise: a social analysis of inactivity*. Leeds: Human Kinetics.

McEvoy P, and Richards D (2006) A critical realist rationale for using a combination of quantitative and qualitative methods. *Journal of Research in Nursing* 11: 66-78.

McGraw S A, McKinlay S M, McClements L, Lasater T M, Assaf A, and Carleton R A (1989) Methods in program evaluation: The process evaluation system of the Pawtucket Heart Health Program. *Evaluation Review* 13: 459-483.

McGraw S A, Stone E J, Osganian S K, Elder J P, Perry C L, Johnson C C, Parcel G S et al. (1994) Design of process evaluation within the Child and Adolescent Trial for Cardiovascular Health(CATCH). *Health Education Quarterly*: 5-26.

McKee M, Britton A, Black N, McPherson K, Sanderson C, and Bain C (1999) Methods in health services research: Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ* 319: 312-315.

McKenna J, Naylor P J, and McDowell N (1998) Barriers to physical activity promotion by general practitioners and practice nurses. *British Journal of Sports Medicine* 32: 242-247.

McLeroy K R, Bibeau D, Steckler A, and Glanz K (1988) An Ecological Perspective on Health Promotion Programs. *Health Education & Behaviour* 15: 351-377.

McMunn A, Bartley M, Hardy R, and Kuh D (2006) Life course social roles and women's health in mid-life: causation or selection? *Journal of Epidemiology and Community Health* 60: 484.

McQueen D V (1987) *Research in Health Behaviour, Health Promotion and Public Health.* Edinburgh: Research unit in health and behavioural change.

Meier S, Stock C, and Kramer A (2007) The contribution of health discussion groups with students to campus health promotion. *Health Promotion International.* 22: 28-36.

Mendlinger S, and Cwikel J (2008) Spiraling Between Qualitative and Quantitative Data on Women's Health Behaviors: A Double Helix Model for Mixed Methods. *Qual Health Res* 18: 280-293.

Mensh E, and Mensh H (1991) *The IQ mythology: Class, race, gender, and inequality.* Southern Illinois Univ Pr.

Merom D R, C. Mahmic, A. Bauman, A. (2005) Process evaluation of the New South Wales Walk Safely to School Day. *Health Promotion Journal of Australia.* 16: 100-106.

Michie S, Fixsen D, Grimshaw J, and Eccles M (2009) Specifying and reporting complex behaviour change interventions: the need for a scientific method. *Implementation Science* 4: 40.

Miller J, and Glassner B (2002) The 'Inside' and the 'Outside': Finding Realities in Interviews. In: Silverman D [ed.] *Qualitative Research: Theory, Method and Practice.* London: Sage.

Miller S T, and Beech B M (2009) Rural healthcare providers question the practicality of motivational interviewing and report varied physical activity counseling experience. *Patient Education and Counseling* 76: 279-282.

Miller W R (1983) Motivational Interviewing with Problem Drinkers. *Behavioural Psychotherapy* 11: 147-172.

Miller W R (1996) Motivational Interviewing: Research, Practice, and Puzzles. *Addictive Behaviors* 21: 835-842.

Miller W R (2001) When is it motivational interviewing? Commentary. *Addiction* 96: 1770-1772.

Miller W R, Benefield R G, and Tonigan J S (1993) Enhancing motivation for change in problem drinking: A controlled comparison of two therapist styles. *Journal of Consulting and Clinical Psychology* 61,.

Miller W R, and Mount K A (2001) A small study of training in motivational interviewing: Does one workshop change clinician and client behavior? *Behavioural and Cognitive Psychotherapy* 29: 457-471.

Miller W R, and Rollnick S (2002) *Motivational interviewing: Preparing people for change.* New York: Guilford Press.

Miller W R, and Rollnick S (2009) Ten things that motivational interviewing is not. *Behavioural and Cognitive Psychotherapy* 37: 129-140.

Miller W R, and Rose G R (2009) Toward a theory of motivational interviewing. *American Psychologist* 64: 527-537.

Miller W R, Yahne C E, and Tonigan J S (2003) Motivational interviewing in drug abuse services: A randomized trial. *Journal of Consulting and Clinical Psychology* 71: 754-763.

Miller W R Y, Carolina E.; Moyers, T. B, Martinez, J, Pirritano, M (2004) A Randomized Trial of Methods to Help Clinicians Learn Motivational Interviewing *Journal of Consulting and Clinical Psychology* 72: 1052-1062.

Mitcheson L, Bhavsar K, and McCambridge J (2009) Randomized trial of training and supervision in motivational interviewing with adolescent drug treatment practitioners. *Journal of Substance Abuse Treatment* 37: 73-78.

Moher D, Schulz K F, and Altman D (2001) The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials (Reprinted from JAMA 2001;285:1987-1991). *Journal of the American College of Surgeons* 193: A71-A76.

Moore L, Moore G, Tapper K, Lynch R, Desousa C, Hale J, Roberts C et al. (2007) Free breakfasts in schools: design and conduct of a cluster randomised controlled trial of the Primary School Free Breakfast Initiative in Wales [ISRCTN18336527]. *BMC Public Health* 7: 258.

Moran J, Bekker H, and Latchford G (2008) Everyday use of patient-centred, motivational techniques in routine consultations between doctors and patients with diabetes. *Patient Education and Counseling* 73: 224-231.

Morgan D L (2007) Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research* 1: 48-76.

Morgan O (2005) Approaches to increase physical activity: reviewing the evidence for exercise-referral schemes. *Public Health* 119: 361-370.

Morton K L, Biddle S J H, and Beauchamp M R (2008) Changes in self-determination during an exercise referral scheme. *Public Health* 122: 1257-1260.

Moyers T B, and Martin T (2006) Therapist influence on client language during motivational interviewing sessions: Support for a potential causal mechanism. *Journal of Substance Abuse Treatment* 30: 245-251.

Moyers T B, Martin T, Christopher P J, Houck J M, Tonigan J S, and Amrhein P C (2007) Client language as a mediator of motivational interviewing efficacy: Where is the evidence? . *Alcoholism: Clinical and Experimental Research* 31(Suppl. 3).

Moyers T B, Martin T, Manuel J K, Hendrickson S M L, and Miller W R (2005) Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment* 28: 19-26.

Munro J F, Nicholl J P, Brazier J E, Davey R, and Cochrane T (2004) Cost effectiveness of a community based exercise programme in over 65 year olds: cluster randomised trial. *Journal of Epidemiology and Community Health* 58: 1004-1010.

Murphy S, Raisanen L, Moore G, Edwards R, Linck P, Williams N, Ud Din N et al. (2010) A pragmatic randomised controlled trial of the Welsh National Exercise Referral Scheme: protocol for trial and integrated economic and process evaluation. *BMC Public Health* 10: 352.

Mutrie N, Campbell A M, Whyte F, McConnachie A, Emslie C, Lee L, Kearney N et al. (2007) Benefits of supervised group exercise programme for women being treated for early stage breast cancer: pragmatic randomised controlled trial. *British Medical Journal* 334: 517.

Mutrie N, and Woods C (2003) How can we get people to become more active? A problem waiting to be solved. In: McKenna J, and Riddoch C [eds.] *Perspectives on Health and Exercise*. Basingstoke: Palgrave MacMillan.

Neumark-Sztainer D, Haines J, Robinson-O'Brien R, Hannan P J, Robins M, Morris B, and Petrich C A (2009) 'Ready. Set. ACTION!' A theater-based obesity prevention program for children: a feasibility study. *Health Educ. Res.* 24: 407-420.

Neumark-Sztainer D, Story M, Hannan P J, and Rex J (2003) New moves: a school-based obesity prevention program for adolescent girls. *Preventive Medicine* 37: 41-51.

NICE (2006a) *Four commonly used methods to increase physical activity: Brief interventions in primary care, exercise referral schemes, pedometers and community-based exercise programmes for walking and cycling*. London: National Institute for Health and Clinical Excellence.

NICE (2006b) Rapid review of the economic evidence of physical activity interventions. London: National Institute for Health and Clinical Excellence.

NICE (2006c) *A rapid review of the effectiveness of exercise referral schemes to promote physical activity in adults*: National Institute for Health and Clinical Excellence Public Health Collaborating Centre.

Nicolaides-Bouman A, van Rossum E, Habets H, Kempen G, and Knipschild P (2007) Home visiting programme for older people with health problems: process evaluation. *Journal of Advanced Nursing* 58: 425-435.

Nishikido N, Matsuda K, Fukuda E, and Motoki C (2007) Development and process evaluation of the participatory and action-oriented empowerment model facilitated by occupational health nurses for workplace health promotion in small and medium-sized enterprises. *Industrial Health.* 45:62-73

Nutbeam D, Smith C, Murphy S, and Catford J (1993) Maintaining evaluation designs in long-term community-based health promotion programs - Heartbeat Wales case-study. *Journal of Epidemiology and Community Health* 47: 127-133.

O'Cathain A (2009) Editorial: Mixed methods research in the health sciences: A quiet revolution. *Journal of Mixed Methods Research* 3: 3-6.

O'Cathain A, Murphy E, and Nicholl J (2007) Why, and how, mixed methods research is undertaken in health services research in England: a mixed methods study. *Bmc Health Services Research* 7.

O'Cathain A, Murphy E, and Nicholl J (2008) The quality of mixed methods studies in health services research. *J Health Serv Res Policy* 13: 92-98.

O'Loughlin J, Renaud L, Richard L, Gomez L S, and Paradis G (1998) Correlates of the sustainability of community-based heart health promotion interventions. *Preventive Medicine* 27: 702-712.

Oakley A (2000) *Experiments in Knowing: Gender and method in the social sciences.* Cambridge: Polity Press.

Oakley A, Strange V, Bonell C, Allen E, and Stephenson J (2006) Health services research - Process evaluation in randomised controlled trials of complex interventions. *British Medical Journal* 332: 413-416.

Onwuegbuzie A, and Leech N (2005) On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology* 8: 375-387.

Orton S U, KE. Rosen, B. McIver, J. Menkens, AJ. (2006) The Management Academy for Public Health: program design and critical success factors. *Journal of Public Health Management & Practice.* 12(5):409-18, 2006 Sep-Oct.

Owen N, Glanz K, Sallis J F, and Kelder S H (2006) Evidence-based approaches to dissemination and diffusion of physical activity interventions. *American Journal of Preventive Medicine* 31: 35-44.

Papadaki A, and Scott J A (2006) Process evaluation of an innovative healthy eating website promoting the Mediterranean diet. *Health Education Research.* 21: 206-218.

Pawson R, and Tilley N (1997) *Realistic evaluation.* London: Sage.

Pearlman D N, Dowling E, Bayuk C, Cullinen K, and Thacher A (2005) From Concept to Practice: Using the School Health Index to Create Healthy School Environments in Rhode Island Elementary Schools. *Preventing Chronic Disease* 2(Special Issue): A09. .

Pearson N, Atkin A J, Biddle S J H, and Gorely T (2010) A family-based intervention to increase fruit and vegetable consumption in adolescents: a pilot study. *Public Health Nutrition* 13: 876-885.

Perry C K, Rosenfeld A G, Bennett J A, and Potempa K (2007) Heart-to-Heart - Promoting walking in rural women through motivational interviewing and group support. *Journal of Cardiovascular Nursing* 22: 304-312.

Perry C L, Sellers D E, Johnson C, Pedersen S, Bachman K J, Parcel G S, Stone E J et al. (1997) The Child and Adolescent Trial for Cardiovascular Health(CATCH): Intervention, implementation, and feasibility for elementary schools in the United States. *Health Education & Behavior* 24: 716-735.

Perry C L, Williams C L, Komro K A, Veblen-Mortenson S, Stigler M H, Munson K A, Farbakhsh K et al. (2002) Project Northland: Long-term outcomes of community action to reduce adolescent alcohol use. *Health Education Research* 17: 117.

Petrunoff N L, Beverley. Watson, Natalie. Morrisey, David. (2009) Suitability of a structured Fundamental Movement Skills program for long day care centres: a process evaluation. *Health Promotion Journal of Australia* 20: 65-68.

Petticrew M, Cummins S, Ferrell C, Findlay A, Higgins C, Hoy C, Kearns A et al. (2005) Natural experiments: an underused tool for public health? *Public Health* 119: 751-757.

Phillips D C, and Burbules N C (2001) *Postpositivism and educational research.* Oxford: Bowman and Littlefield.

Phongsavan P M, D. Wagner, R. Chey, T. von Hofe, B. Silove, D. Bauman, A. (2008) Process evaluation in an intervention designed to promote physical activity among adults with anxiety disorders: evidence of acceptability and adherence. *Health Promotion Journal of Australia.* 19: 137-143.

Pierson H M, Hayes S C, Gifford E V, Roget N, Padilla M, Bissett R, Berry K et al. (2007) An examination of the Motivational Interviewing Treatment Integrity code. *Journal of Substance Abuse Treatment* 32: 11-17.

Pirie P L, Stone E J, Assaf A R, Flora J A, and Maschewskyschneider U (1994) Program-Evaluation Strategies for Community-Based Health Promotion Programs - Perspectives from the Cardiovascular-Disease Community Research and Demonstration Studies. *Health Education Research* 9: 23-36.

Pope C, and Mays N (1995) Qualitative Research: Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *British Medical Journal* 311: 42-45.

Pope C, Ziebland S, and Mays N (2000) Qualitative research in health care: Analysing qualitative data. *British Medical Journal* 320: 114-116.

Power R, Langhaug L F, Nyamurera T, Wilson D, Bassett M T, and Cowan F M (2004) Developing complex interventions for rigorous evaluation--a case study from rural Zimbabwe. *Health Education Research* 19: 570-575.

Rabiei K, Kelishadi R, Sarrafzadegan N, Abedi H, Alavi M, Heidari K, Bahonar A et al. (2009) Process evaluation of a community-based program for prevention and control of non-communicable disease in a developing country: The Isfahan Healthy Heart Program, Iran. *BMC Public Health* 9: 57.

Rankin D, Truman J, Backett-Milburn K, Platt S, and Petticrew M (2006) The contextual development of healthy living centres services: An examination of food-related initiatives. *Health & Place* 12: 644-655.

Redfern J, Briffa T, Ellis E, and Freedman S B (2008) Patient-Centered Modular Secondary Prevention Following Acute Coronary Syndrome: A RANDOMIZED CONTROLLED TRIAL. *Journal of Cardiopulmonary Rehabilitation and Prevention* 28: 107-115.

Reed J, and Payton V (1997) Focus groups: issues of analysis and interpretation. *Journal of Advanced Nursing* 26: 765-771.

Resnicow K, Davis R, and Rollnick S (2006) Motivational interviewing for pediatric obesity: Conceptual issues and evidence review. *Journal of the American Dietetic Association* 106: 2024-2033.

Riddoch C, Puig-Ribera A, and Cooper A (1997) *Effectiveness of physical activity promotion schemes in primary care: a review.* Bristol: Health Education Authority.

Ringwalt C L, Vincus A, Ennett S, Johnson R, and Rohrbach L A (2004) Reasons for teachers' adaptation of substance use prevention curricula in schools with non-white student populations. *Prevention Science* 5: 61-67.

Ritchie D, Parry O, Gnich W, and Platt S (2004) Issues of participation, ownership and empowerment in a community development programme: tackling smoking in a low-income area in Scotland. *Health Promot. Int.* 19: 51-59.

Rogers C (1959) A theory of therapy, personality and interpersonal relationships as developed in the client-centered framework. In: Koch S [ed.] *Psychology: A study of a science. Vol. 3: Formulations of the person and the social context.* New York: : McGraw Hill.

Rogers E M (1995) *Diffusion of Innovations.* 4th ed. New York: Free Press.

Rogers E M (2003) *Diffusion of Innovations.* 5th ed. New York: Free Press.

Rohde P, Lewinsohn P M, and Seeley J R (1997) Comparability of telephone and face-to-face interviews in assessing Axis I and II disorders. *American Journal of Psychiatry* 154: 1593-1598.

Roland M, and Torgerson D J (1998) Understanding controlled trials - What are pragmatic trials? *British Medical Journal* 316: 285-285.

Rollnick S, Butler C C, McCambridge J, Kinnersley P, Elwyn G, and Resnicow K (2005) Consultations about changing behaviour. *British Medical Journal* 331: 961-963.

Rollnick S, and Miller W R (1995) What is motivational interviewing? *Behavioural and Cognitive Psychotherapy* 23: 325-334.

Rollnick S, Miller W R, and Butler C C (2008) *Motivational Interviewing in Health Care*. New York: guildford Press.

Romé Å, Persson U, Ekdahl C, and Gard G (2009) Physical activity on prescription (PAP): Costs and consequences of a randomized, controlled trial in primary healthcare. *Scandinavian Journal of Primary Health Care* 27: 216-222.

Ronda G, Van Assema P, Ruland E, Steenbakkers M, and Brug J (2004) The Dutch Heart Health Community Intervention `Hartslag Limburg': design and results of a process study. *Health Education Research* 19: 596-607.

Rosecrans A M, Gittelsohn J, Ho L S, Harris S B, Naqshbandi M, and Sharma S (2007) Process evaluation of a multi-institutional community-based program for diabetes prevention among First Nations. *Health Education Research*: cym031.

Rothwell P M (2005) External validity of randomised controlled trials: "To whom do the results of this trial apply?" *The Lancet* 365: 82-93.

Rubak S, Sandbaek A, Lauritzen T, Borch-Johnsen K, and Christensen B (2009) General practitioners trained in motivational interviewing can positively affect the attitude to behaviour change in people with type 2 diabetes. *Scandinavian Journal of Primary Health Care* 27: 172-179.

Rubak S, Sandboek A, Lauritzen T, and Christensen B (2005) Motivational interviewing: a systematic review and meta-analysis. *British Journal of General Practice* 55: 305-312.

Rubin R S (2002) Will the real SMART goals please stand up? *The Industrial-Organizational Psychologist* 39: 26-27.

Rutter M (2006) Is Sure Start an effective preventive intervention? *Child and Adolescent Mental Health* 11: 135.

Rychetnik L, Frommer M, Hawe P, and Shiell A (2002) Criteria for evaluating evidence on public health interventions. *Journal of Epidemiology and Community Health* 56: 119-127.

Saini P, and Lacroix J (2009) Self-setting of physical activity goals and effects on perceived difficulty, importance and competence. In: *Proceedings of the 4th International Conference on Persuasive Technology*. Claremont, California, ACM.

Salmon J O, Ball K, Crawford D, Booth M, Telford A, Hume C, Jolley D et al. (2005) Reducing sedentary behaviour and increasing physical activity among 10-year-old

children: overview and process evaluation of the 'Switch-Play' intervention. *Health Promot. Int.* 20: 7-17.

Sannibale C, Fucito L, O'Connor D, and Curry K (2005) Process evaluation of an out-patient detoxification service. *Drug and Alcohol Review* 24: 475-481.

Sanson-Fisher R W, Bonevski B, Green L W, and D'Este C (2007) Limitations of the randomized controlled trial in evaluating population-based health interventions. *American Journal of Preventive Medicine* 33: 155-161.

Saunders R P, Evans M H, and Joshi P (2005) Developing a Process-Evaluation Plan for Assessing Health Promotion Program Implementation: A How-To Guide. *Health Promotion Practice* 6: 134-147.

Saunders R P, Ward D, Felton G M, Dowda M, and Pate R R (2006) Examining the link between program implementation and behavior outcomes in the lifestyle education for activity program (LEAP). *Evaluation and Program Planning* 29: 352-364.

Sayer A (2000) *Realism and Social Science*. London: Sage.

Schmidt M, Absalah S, Nierkens V, and Stronks K (2008) Which factors engage women in deprived neighbourhoods to participate in exercise referral schemes? *Bmc Public Health* 8: 371.

Schulz K F, Altman D G, Moher D, and for the C G CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *British Medical Journal* 340: c332-.

Seime R J, and Vickers K S (2006) The challenges of treating depression with exercise: From evidence to practice. *Clinical Psychology-Science and Practice* 13: 194-197.

Shek D, Lee T, and Sun R (2007) Process evaluation of the implementation of the Secondary 2 Program of Project P.A.T.H.S. in the experimental implementation phase. *The Scientific World Journal* 8: 83-94.

Shek D T L M, Hing Keung ; Sun, Rachel C.F.; Lung, Daniel W.M. (2008) Process evaluation of the Tier 1 Program (Secondary 1 Curriculum) of the Project P.A.T.H.S.: findings based on the full implementation phase. *The Scientific World Journal.* 14: 35-46

Shepich J, Slowiak J M, and Keniston A (2007) Do subsidization and monitoring enhance adherence to prescribed exercise? *American Journal of Health Promotion* 22: 2-5.

Shepperd S, Lewin S, Straus S, Clarke M, Eccles M P, Fitzpatrick R, Wong G et al. (2009) Can we systematically review studies that evaluate complex interventions? *PLoS Medicine* 6: e1000086.

Shilts M, Horowitz M, and Townsend M S (2004) Goal setting as a strategy for dietary and physical activity behavior change: a review of the literature. *American Journal of Health Promotion* 19: 81-93.

Sibbald B, and Roland M (1998) Understanding controlled trials: Why are randomised controlled trials important? *BMJ* 316: 201-.

Silva M N, Vieira P N, Coutinho S R, Minderico C S, Matos M G, Sardinha L B, and Teixeira P J Using self-determination theory to promote physical activity and weight control: a randomized controlled trial in women. *Journal of Behavioral Medicine* 33: 110-122.

Silverman D (2005) *Doing qualitative research: second edition.* London: Sage.

Simmons D, and Voyle J A (2003) Reaching hard-to-reach, high-risk populations: piloting a health promotion and diabetes disease prevention programme on an urban marae in New Zealand. *Health Promot. Int.* 18: 41-50.

Simmons R, Ogilvie D, Griffin S, and Sargeant L (2009) Applied public health research - falling through the cracks? *Bmc Public Health* 9: 362.

Singh A S, Chinapaw M J M, Brug J, and van Mechelen W (2009) Process evaluation of a school-based weight gain prevention program: the Dutch Obesity Intervention in Teenagers (DOiT). *Health Education Research.* 24: 772-777.

Singh S (1997) Why are GP exercise schemes so successful (for those who attend)? *Journal of Management and Medicine* 11: 233-237.

Smith B J (2004) Promotion of physical activity in primary health care: update of the evidence on interventions. *Journal of Science and Medicine in Sport* 7: 67-73.

Smith JL A P, Brooks AC, Carpenter KM, Levin D, Schreiber EA, et al. (2007) Providing live supervision via teleconferencing improves acquisition of motivational interviewing skills after workshop attendance. *The American Journal of Drug and Alcohol Abuse.* 33: 163-168.

Sorensen J B, Kragstrup J, Skovgaard T, and Puggaard L (2008) Exercise on prescription: a randomized study on the effect of counseling vs counseling and supervised exercise. *Scandinavian Journal of Medicine & Science in Sports* 18: 288.

Sowden S L, Breeze E, Barber J, and Raine R (2008) Do general practices provide equitable access to physical activity interventions? *British Journal of General Practice* 58: e1-e8.

Spillane J P, Reiser B J, and Reimer T (2002) Policy Implementation and Cognition: Reframing and Refocusing Implementation Research. *Review of Educational Research* 72: 387-431.

Spradley J (1980) *Participant Observation.* New York: Holt, Rinehart, and Winston.

Spruijt-Metz D, Nguyen-Michel S T, Goran M I, Chou C P, and Huang T T K (2008) Reducing sedentary behavior in minority girls via a theory-based, tailored classroom media intervention. *International Journal of Pediatric Obesity* 3: 240-248.

St. Louis R M, Parow J E, Eby D W, Bingham C R, Hockanson H M, and Greenspan A I (2008) Evaluation of community-based programs to increase booster seat use. *Accident Analysis & Prevention* 40: 295-302.

Stafstrom M, and Larsson S (2007) The Trelleborg Project: A process evaluation of a multi-sector community intervention to reduce alcohol consumption and related harm. *Substance Use & Misuse* 42: 2041 - 2051.

Standage M, Duda J L, and Ntoumanis N (2006) Students' motivational processes and their relationship to teacher ratings in school physical education: A self-determination theory approach. *Research Quarterly for Exercise and Sport* 77: 100-110.

Stange K C, Crabtree B F, and Miller W L (2006) Publishing multimethod research. *Annals of Family Medicine* 4: 292-294.

Stathi A, Fox K R, and McKenna J (2002) Physical activity and dimensions of subjective well-being in older adults. *Journal of Aging and Physical Activity* 10: 76-92.

Stathi A, McKenna J, and Fox K R (2003) The contribution of physical activity to the subjective well-being of older adults enrolled in exercise referral schemes. *Journal of Sports Sciences* 21: 359-360.

Stathi A, McKenna J, and Fox K R (2004) The experiences of older people participating in exercise referral schemes. *Journal of The Royal Society for the Promotion of Health* 124: 18-23.

Steckler A, Ethelbah B, Martin C J, Stewart D, Pardilla M, Gittelsohn J, Stone E et al. (2003) Pathways process evaluation results: a school-based prevention trial to promote healthful diet and physical activity in American Indian third, fourth, and fifth grade students. *Preventive Medicine* 37: S80-S90.

Steckler A, and Linnan L [eds.] (2002) *Process evaluation for public health interventions and research.* San Francisco: Jossey-Bass.

Steele R, Mummery K W, and Dwyer T (2007) Development and process evaluation of an Internet-based physical activity behaviour change program. *Patient Education and Counseling* 67: 127-136.

Stein J, Lewin S, Fairall L, Mayers P, English R, Bheekie A, Bateman E et al. (2008) Building capacity for antiretroviral delivery in South Africa: A qualitative evaluation of the PALSA PLUS nurse training programme. *BMC Health Services Research* 8: 240.

Stephenson J, and Imrie J (1998) Why do we need randomised controlled trials to assess behavioural interventions? *BMJ* 316: 611-613.

Stevens W, Hillsdon M, Thorogood M, and McArdle D (1998) Cost-effectiveness of a primary care based physical activity intervention in 45-74 year old men and women: a randomised controlled trial. *British Journal of Sports Medicine* 32: 236-241.

Stokols D (1996) Translating social ecological theory into guidelines for community health promotion. *American Journal of Health Promotion* 10: 282-298.

Strange V, Allen E, Oakley A, Bonell C, Johnson A, Stephenson J, and The Ripple Study T (2006) Integrating process with outcome data in a randomized controlled trial of sex education. *Evaluation* 12: 330-352.

Strecher V J, and Rosenstock I M (1997) The Health Belief Model. In: Glanz K, Lewis F M, and Rimer B K [eds.] *Health Behaviour and Health Education.* (2nd edn.) San-Fransisco: Jossey-Bass, pp. 41-59.

Strecher V J, Seijts G H, Kok G J, Latham G P, Glasgow R, DeVellis B, Meertens R M et al. (1995) Goal setting as a strategy for health behavior change. *Health Education & Behaviour* 22: 190-200.

Sturges J E, and Hanrahan K J (2004) Comparing Telephone and Face-to-Face Qualitative Interviewing: a Research Note. *Qualitative Research* 4: 107-118.

Sullivan G (2002) Reflexivity and subjectivity in qualitative research: The utility of a Wittgensteinian framework. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-line Journal], 3(3). Available at: http://www.qualitative-research.net/fqs-texte/3-02/3-02sullivan-e.htm [Date of Access: Jan 02, 2007].*

Sy A, and Glanz K (2008) Factors Influencing Teachers' Implementation of an Innovative Tobacco Prevention Curriculum for Multiethnic Youth: Project SPLASH. *Journal of School Health* 78: 264-273.

Taylor A (1997) Changes in perceived barriers towards exercising: Findings from a 9 month randomized controlled study of a general practitioner exercise referral scheme. *Journal of Sports Sciences* 15: 107-108.

Taylor A H, Doust J, and Webborn N (1998) Randomised controlled trial to examine the effects of a GP exercise referral programme in Hailsham, East Sussex, on modifiable coronary heart disease risk factors. *Journal of Epidemiology and Community Health* 52: 595-601.

Taylor A H, and Fox K R (2005) Effectiveness of a primary care exercise referral intervention for changing physical self-perceptions over 9 months. *Health Psychology* 24: 11-21.

The Cabinet Office (1998) Modernising government. The Stationery Office.

Thompson J L, Allen P, Helitzer D L, Qualls C, Whyte A N, Wolfe V K, and Herman C J (2008) Reducing Diabetes Risk in American Indian Women. *American Journal of Preventive Medicine* 34: 192-201.

Thomson H, Hoskins R, Petticrew M, Ogilvie D, Craig N, Quinn T, and Lindsey G (2004) Evaluating the health effects of social interventions. *British Medical Journal* 328: 282 - 285.

Thurston M, and Green K (2004) Adherence to exercise in later life: how can exercise on prescription programmeg be made more effective? *Health Promotion International* 19: 379-387.

Thyrian J R, Freyer-Adam J, Hannover W, Roske K, Mentzel F, Kufeld C, Bischof G et al. (2007) Adherence to the principles of Motivational Interviewing, clients' characteristics and behavior outcome in a smoking cessation and relapse prevention trial in women postpartum. *Addictive Behaviors* 32: 2297-2303.

Tilley N (2004) Applying theory-driven evaluation to the British Crime Reduction Programme: The theories of the programme and of its evaluations. *Criminal Justice* 4: 255-276.

Tones K (1997) Beyond the randomized controlled trial: a case for 'judicial review'. *Health Education Research* 12: 161-.

Torgerson D J, and Roberts C (1999) Understanding controlled trials: Randomisation methods: concealment. *British Medical Journal* 319: 375-376.

Tulloch H, Fortier M, and Hogg W (2006) Physical activity counseling in primary care: Who has and who should be counseling? *Patient Education and Counseling* 64: 6-20.

Van Dorsten B (2007) The use of motivational interviewing in weight loss. *Current Diabetes Reports* 7: 386-390.

van Harreveld F, Rutjens B T, Rotteveel M, Nordgren L F, and van der Pligt J (2009) Ambivalence and decisional conflict as a cause of psychological discomfort: Feeling tense before jumping off the fence. *Journal of Experimental Social Psychology* 45: 167-173.

van Sluijs E, van Poppel M, Stalman W, and van Mechelen w (2004) Feasibility and acceptability of a physical activity promotion programme in general practice. *Family Practice* 21: 429-436.

VanWormer J J, and Boucher J L (2004) Motivational Interviewing and Diet Modification: A Review of the Evidence. *The Diabetes Educator* 30: 404-419.

Wahab S, Menon U, and Szalacha L (2008) Motivational interviewing and colorectal cancer screening: A peek from the inside out. *Patient Education and Counseling* 72: 210-217.

Wanless D, Appleby J, Harrison A, and Patel D (2007) Our future health secured. *A Review of NHS Funding and Performance.*

Wegner L, Flisher A J, Caldwell L L, Vergnani T, and Smith E A (2008) Healthwise South Africa: cultural adaptation of a school-based risk prevention programme. *Health Education Research* 23: 1085-1096.

Weiss C (1997) Theory-based evaluation: Past, present, and future. *New Directions for Program Evaluation* 1997: 41-55.

Welsh Assembly Government (2003) *Healthy and Active Lifestyles in Wales: a Framework for Action.* . Cardiff: Welsh Assembly Government.

Welsh Assembly Government (2005a) *Climbing Higher: the Welsh Assembly Strategy for Sport and Active Recreation in Wales.* . Cardiff: Welsh Assembly Government.

Welsh Assembly Government (2005b) *Heads – We Win, a Strategic Framework for the Heads of the Valleys.* . Cardiff: : Welsh Assembly Government.

Welsh Assembly Government (2006) *Exercise referral – a guide to developing high quality schemes.* Cardiff: Welsh Assembly Government.

Welsh Assembly Government (2008) *Welsh Index of Multiple Deprivation 2008 (WIMD) - Local Authority Analysis* [Online]. Available at: http://wales.gov.uk/topics/statistics/publications/wimd08la/?lang=en [Accessed: January].

Welsh Assembly Government (2009) *Wales's Population: A Demographic Overview.* Cardiff: Statistics for Wales.

Wiles R, Demain S, Robison J, Kileff J, Ellis-Hill C, and McPherson K (2008) Exercise on prescription schemes for stroke patients post-discharge from physiotherapy. *Disability and Rehabilitation* 30: 1966-1975.

Williams C L, Carter B J, Kibbe D L, and Dennison D (2009) Increasing Physical Activity in Preschool: A Pilot Study to Evaluate Animal Trackers. *Journal of Nutrition Education and Behavior* 41: 47-52.

Williams N H, Hendry M, France B, Lewis R, and Wilkinson C (2007) Effectiveness of exercise-referral schemes to promote physical activity in adults: systematic review. *British Journal of General Practice* 57: 979-986.

Wilson D K, Griffin S, Saunders R P, Kitzman-Ulrich H, Meyers D C, and Mansard L (2009) Using process evaluation for program improvement in dose, fidelity and reach: the ACT trial experience. *International Journal of Behavioral Nutrition and Physical Activity* 6: 79.

Wing R, Jeffery R, Pronk N, and Hellerstedt W (1996) Effects of a personal trainer and financial incentives on exercise adherence in overweight women in a behavioral weight loss program. *Obesity Research* 4: 457-462.

Wolff N (2001) Randomised trials of socially complex interventions: promise or peril? *Journal of Health Services Research and Policy* 6: 123-126.

Wormald H, and Ingle L (2004) GP exercise referral schemes: Improving the patient's experience. *Health Education Journal* 63: 362-373.

Wormald H, Waters H, Sleap M, and Ingle L (2006) Participants' perceptions of a lifestyle approach to promoting physical activity: targeting deprived communities in Kingston-Upon-Hull. *Bmc Public Health* 6.

Xue F, Yao W, and Lewin R (2008) A randomised trial of a 5 week, manual based, self-management programme for hypertension delivered in a cardiac patient club in Shanghai. *BMC Cardiovascular Disorders* 8: 10.

Young D R, Steckler A, Cohen S, Pratt C, Felton G, Moe S G, Pickrel J et al. (2008) Process evaluation results from a school- and community-linked intervention: the Trial of Activity for Adolescent Girls (TAAG). *Health Educ. Res.* 23: 976-986.

Yu C H (2006) Misconceived relationships between logical positivism and quantitative research. *Philosophical foundations of quantitative research methodology.* Laham: MD: University Press of America.