Evaluation and evidence-based public health practices: their use, usability and usefulness in physical activity interventions.

Judith Fynn

Doctor of Philosophy (PhD)

University of East Anglia, UK

Norwich Medical School

10th February 2021

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution.

Abstract

Evidence-based public health ensures that actions to safeguard and improve the health of populations are based on sound evidence. This requires three processes: evaluation to generate evidence, dissemination, and use of evidence. This thesis aimed to improve understanding of these processes within multi-agency public health interventions; the research focuses on interventions promoting physical activity, which is a public health priority. Such interventions are challenging to implement and to evaluate, yet to achieve sustainable change to address the health of the population, evaluation is needed to understand their complexity and effectiveness. By exploring current practices, the thesis applied the insights gained to develop recommendations to improve practice and contribute to the underlying aim of closing the research-practice gap.

A scoping review was conducted to identify evaluation frameworks that could be used in evaluating physical activity interventions, and to appraise their applicability to different evaluation objectives and contexts. Secondly, a systematic review appraised the use and reporting of evaluation frameworks in physical activity evaluation studies. A collective case study approach was then applied to explore the use of strategies to support evidence-based practices within an applied context. This was based on a national physical activity programme, Sport England's Get Healthy Get Active programme. Multiple sources of evidence were analysed to explore influences on evaluation practice, knowledge exchange and the capacity to conduct and use evaluation.

This research highlighted the complex interconnections and context-specific nature of influences on evidence-based practices. Where systematic approaches, such as evaluation frameworks, are applied appropriately, these can improve evaluation and reporting. Yet, there are gaps in guidance, limitations in use and reporting of frameworks, and limited use made of evidence generated. Research-practice partnerships and networks can improve practice, but organisational structures and systems are needed to facilitate their implementation. The thesis considers implications of these findings for researchers and decision makers, who play a pivotal role in shaping the future of evidence-based public health.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate ‘take down’ action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Table of contents

Abstract 2

Table of contents 3

List of tables 7

List of figures 8

List of supplementary material 9

Acknowledgements 10

List of abbreviations 11

Publications and statement of authorship 13

Publications arising from this thesis 13

Statement of jointly authored publications 13

Chapter 1. Introduction to the thesis 15

1.1. Introduction 15

1.2. Research questions 15

1.3. Defining the research area 16

1.3.1. Evidence-based public health 17

1.3.2. Interventions 18

1.3.2.1. Complex interventions 18

1.3.2.2. Physical activity interventions 19

1.3.2.3. Intervention setting or context 20

1.3.2.4. Real-world interventions 20

1.3.3. Evaluation 21

1.3.3.1. Evaluation frameworks - what guidance is available, how applicable and usable are they, and how well are they used? 22

1.3.3.2. Partnerships, collaborations and networks – how effective are they in facilitating and improving evidence-based practice? 22

1.3.4. Dissemination and evaluation use 23

1.4. A conceptual framework for the study 24

1.4.1. Methodology 24

1.5. Overview of the chapters 27

Chapter 2. Evaluation frameworks, their availability, applicability and usability 30

Introduction 30

Background 30

Objectives 32

Method 32

Search strategy 32
Inclusion and exclusion criteria 33
Data extraction and synthesis 34
Results 36
Study selection 36
Identification of the evaluation frameworks available 38
Scope of the evaluation frameworks and development of a typology 40
Programme type 40
Evaluation objective 43
Mapping frameworks against evaluation components 43
Process evaluation components 47
Outcome evaluation components 50
Study design components 50
Discussion 51
Strengths and limitations 53
Conclusion 53
Chapter 3. A systematic review of the use and reporting of evaluation frameworks in physical activity evaluation studies 54
Introduction 54
Background 54
Methods 57
Protocol and registration 57
Search strategy 57
Study selection 58
Data extraction 59
Results 62
Study characteristics 64
Appraisal of the quality of reporting on the use of evaluation frameworks 65
Discussion 71
The extent to which evaluation frameworks have been used and reported 71
The quality of reporting with regards to how frameworks were used 72

Strengths and limitations 74
Conclusion 74
Chapter 4. Use of evaluation guidance in practice - exploring influences on evaluation practice within a national physical activity programme 76
Introduction 76
Background 77
Objectives 78
Method 78
The GHGA programme 78
Study design 79
Sampling and data collection for the documentary analysis 79
Sampling and data collection for the semi-structured interviews 79
Analysis of documents and interview data 80
Results 82
The Case study sample 82
Objective 1: To identify the logic of the programme and explore the relationships between programme and project aims 82
Objective 2: Influences on evaluation practices 90
1. Programme and project design 90
2. Evaluation design 95
2.1 Use of standardised tools 95
2.2 Use and reporting of the Standard Evaluation Framework 96
2.3 Use and reporting of optional evaluation components 101
3. Resources 101
4. Partnerships 101
5. Organisational structures, systems and processes 102
Objective 3: Appraisal of whether the programme was effective in generating high quality generalisable evidence that enabled it to meet its aims 103
Discussion 107
Strengths and limitations 110
Conclusion 110
Chapter 5. Partnerships, collaborations and networks for evaluation: their use in facilitating evaluation, dissemination and evaluation use 112

Introduction 112
Background 112
Objectives 115
Method 115
Study sample 115
Data collection 116
Data analysis 116
Results 118
1. The partners involved in the project and/or programme evaluation 118
2. Partnership characteristics and their influence on evaluation 122
2.1 Roles and responsibilities 124
2.2 Relationships 125
2.3 Collaboration 125
2.4 Communication 126
3. Evaluation use 127
4. A conceptual model of the relationships between partnerships, processes and partnership characteristics that facilitate evaluation, dissemination and evaluation use 129
Discussion 131
Strengths and limitations 133
Conclusion 133
Chapter 6. Discussion and conclusions - the use, usability and usefulness of strategies to support evaluation and evidence-based public health 134
6.1 Introduction 134
6.2 Summary of findings 134
6.2 Discussion and implications for research and practice 135
6.2.1. The use, usability and usefulness of evaluation frameworks and guidance 136
6.2.2 The role of partnerships and networks in supporting and improving evidence-based practices 138
6.3 Implications for research and practice 140
6.4 Reflections on the research approach and its strengths and limitations 141
6.5 Personal reflection of the candidate 143
6.5 Concluding comments 145
Appendices 147
Appendix 1. Glossary of key terms 147

Appendix 2 Research ethics approval 150
Appendix 3 Gatekeepers consent to conduct the research from Sport England 151
Appendix 4 Participant Information Sheet for the case study 152
Appendix 5 Consent Form for participants in the case study 155
Appendix 6 Topic guide for semi-structured interviews 156
References 158

List of tables

Table 2-1 Inclusion and exclusion criteria for the scoping review 34
Table 2-2 Evaluation components agreed for data extraction and mapping of frameworks 36
Table 2-3 Included frameworks grouped by decade of publication and source 39
Table 2-4 Frameworks intended for use in physical activity, dietary change or behaviour change programmes mapped against process and outcome evaluation components 44
Table 2-5 Frameworks intended for use in health promotion or public health programmes mapped against process and outcome evaluation components 45
Table 2-6 Frameworks intended for use in generic programmes mapped against process and outcome evaluation components 46
Table 2-7 Frameworks intended for use in evaluating physical activity, dietary change or behaviour change programmes mapped against study design, evaluation approach and reporting components 47
Table 2-8 Frameworks intended for use in evaluating health promotion or public health programmes mapped against study design, evaluation approach and reporting components 48
Table 2-9 Frameworks intended for use in evaluating generic programmes mapped against study design, evaluation approach and reporting components 49
Table 3-1 Search strategy applied in CINAHL database 58
Table 3-2 Inclusion and exclusion criteria for the systematic review 59
Table 3-3 Categories and indicators for assessing the quality of reporting of the use of evaluation frameworks 61
Table 3-4 Evaluation frameworks reported within the 69 studies 63
Table 3-5 Appraisal of use and reporting of an evaluation framework in studies using a single evaluation framework 66
Table 3-6 Appraisal of use and reporting of the use of evaluation frameworks in studies using multiple frameworks 70

Table 4-1 Summary of criteria included in the Standard Evaluation Framework for Physical Activity Interventions (SEF) 81
Table 4-2 Summary of the reported programme and project characteristics, aims and objectives 86
Table 4-3 Summary of influences on evaluation practice 92
Table 4-4 Study design and data collection methods included in project evaluation 98
Table 4-5 Summary of project reporting on SEF criteria related to programme details 99
Table 4-6 Summary of project reporting on SEF criteria related to participant demographics 100
Table 5-1 Sample of interview participants according to their role in the programme or projects 116
Table 5-2 Coding framework for the thematic analysis of interview data 117
Table 5-3 Partnership characteristics, and facilitators and barriers to successful partnerships and evaluation as perceived by stakeholders 122

List of figures

Figure 1-1 The evidence-based practice cycle 17
Figure 1-2 Conceptual framework for the research 26
Figure 2-1 PRISMA diagram of the screening process for the scoping review 37
Figure 2-2 Typology of evaluation frameworks intended for use in physical activity, dietary change or behaviour change programmes 41
Figure 2-3 Typology of evaluation frameworks intended for use in health promotion or public health programmes 42
Figure 2-4 Typology of evaluation frameworks intended for use in generic programmes 43
Figure 3-1 PRISMA diagram of screening process for the systematic review 62
Figure 4-1 Logic Model for the Get Healthy Get Active (GHGA) programme 84
Figure 4-2 Pathway diagram of the Get Healthy Get Active (GHGA) programme 85
Figure 4-3 Evidence generated from the Get Healthy Get Active programme mapped against the intended outcomes 106
Figure 5-1 The network of reported partners 120
Figure 5-2 The network, showing additional partners and relationships identified from the interviews 121
Figure 5-3 Conceptual model of the flow of information between partners and networks, and the relationships between processes, partnership characteristics and practices to support evaluation, dissemination and evaluation use 130

List of supplementary material

Additional File 2.1 Descriptions of evaluation frameworks included in the scoping review

Additional File 2.2 Characteristics of evaluation frameworks included in the scoping review

Additional File 3.1 Study characteristics of the evaluation studies included in the systematic review

Acknowledgements

I would like to thank all those who made the successful completion of this thesis possible. Firstly, I would like to thank Norwich Medical School and the Centre for Diet and Activity Research (CEDAR) for funding my studentship. I am very grateful for the opportunity this has given me.

A special thanks to my supervisory team, Professor Andy Jones, Professor Wendy Hardeman, and Dr Karen Milton, who have guided and supported me through the ups and downs of the research and writing of the thesis. They have provided the perfect balance of constructive critique and attention to detail with pragmatism and encouragement. I would also like to thank Dr Charlotte Salter for her support, specifically her advice on qualitative research and framework analysis, but also for being a great sounding board for my ideas. For me this has been the dream team, and I am extremely grateful to you all for the invaluable advice and for believing in me, thank you.

Many people contributed to this work. I would like to thank Sport England for their permission to undertake the case study of the Get Healthy Get Active programme. In particular, Darcy Hare and Suzanne Gardner for all their support, providing access to documents, answering my many questions, and commenting on thesis chapters. I am also grateful to everyone involved in the GHGA projects who shared documents and took time to participate in interviews and share their insights. I had the opportunity to meet many great people - thank you. A particular thanks to Nick Cavill, Darcy Hare, and Harry Rutter for providing expert opinion for the scoping review, and Joseph Murphy for working with me on the systematic review.

Thanks also to various people who I have met whilst at UEA whose support and conversations were always appreciated. Dr Gill Price for her friendly encouragement and her work in facilitating the training and development programme, for me these were a welcome opportunity to meet people. Florence, Vicky and Kimberley who I shared offices with at various times; and Lee, Joanna, and Andy for opportunities to teach and do additional work. A special thanks to Laura Haag, for her friendship and willingness to work with me on various projects, including the Research Student Forum and Business Boost consultation.

Outside of UEA, particularly in this last year when everyone's life has changed so much and we have all been conducting our work through the lens of a webcam, I would like to thank my swimming buddies. Harriet, Alex, Jacqui and Karen, the trips to the river, sea and lido were 'essential', and I am truly thankful for your friendship through this time.

Lastly, I would like to thank my family. Paul and Kyriel, who have had to put up with me day in, day out, and have read various sections as I have battled to refine and edit the chapters; and my other boys, Tarn, Tamar and Linton. Thank you for your patience, interest and encouragement; I seriously could not have done this without your support.

List of abbreviations

ACTIVE	Active Children through Incentive Vouchers – Evaluation
ALED	Active Living Every Day
APAN	Albany Physical Activity and Nutrition
5-As	Assess, Advise, Agree, Assist, Arrange
BCT	Behaviour Change Technique
BGDP	Bristol Girls Dance Project
BMI	Body Mass Index
CCG	Clinical Commissioning Group
CDC	Center for Disease Control and Prevention
Center TRT	Center of Excellence for Training and Research Translation
CBHEPA	Dutch Community-Based Health-Enhancing Activity Programs
CHAM JAM	Children's Hospital at Montefiore Joining Academics and Movement
CoM	Communities on the Move
COMMUNICATE	COMMUNIty-wide CAmpaign To promote Exercise
DoH	Department of Health
FAN	Faith, Activity, and Nutrition
FitEx	Fit Extension
FLEX	Fuelling Learning through Exercise
GPAT	Good Practice Appraisal Tool
GP	General Practitioner
GHGA	Get Healthy Get Active
GOTR	Girls On The Run
HEBS	Health Education Board Scotland
HP	Health Promotion
Healingo Fit	Health Integrated Gaming Online
HKOS	Healthy Kids Out of School
HT	Healthy Together
IDEFICS	Identification & prevention of Dietary & lifestyle-induced health EFfects In Children and infantS
IMIL	I am Moving, I am Learning
IPAQ	International Physical Activity Questionnaire
LEAP	Learning, Evaluation and Planning
LPAW	Lifestyle Physical Activity for Women
MAGNET	Mothers And dauGhters daNcing togEther Trial
MRC	Medical Research Council
MMIPP	Model for Management of Intervention Programme Planning

NECaSP	Newham's Every Child a Sports Person
NICE	National Institute for Health and Care Excellence
OPEN	Obesity Prevention through the European Network
PA	Physical Activity
PAC	Physical Activity Consultation
PACES	Partnerships for Active Children in Elementary Schools
PACE-UP	Primary care pedometer based walking trial
PAFES	Physical Activity, Sports, and Health Plan
PIP	Programme Impact Pathway
PIPA	Participatory Impact Pathway Analysis
PH	Public Health
PHE	Public Health England
RCT	Randomised Controlled Trial
RCP & ACP	Recreovia program & Academia da Cidade program
RE-AIM	Reach, Effectiveness, Adoption, Implementation, Maintenance
SAGE	Sustainability via Active Garden Education
SEF	Standard Evaluation Framework (for physical activity interventions)
SHAPES	Study of Health and Activity in Preschool Environments
SPAQ	Scottish Physical Activity Questionnaire
SLIMMER (SLIM)	iMplementation Experience Region Noord- en Oost-Gelderland
SPACE	Supporting Physical Activity in the Childcare Environment
STAND	Sedentary Time ANd Diabetes
STEPs	Strategies To Enhance Practice in YMCA-operated After School Programmes
STROBE	Statement for Reporting Observational studies in Epidemiology
TAME health	Testing Activity Monitors' Effect on health
TIDieR	Template for Intervention Description and Replication
WAVE	WAVE~Ripples for Change
WWPP	Walking with Poles Program
WAVES	West Midlands ActiVe lifestyle and healthy Eating in School children
WHO	World Health Organization

Publications and statement of authorship

Publications arising from this thesis

Fynn, J.F., Hardeman, W., Milton, K., Jones A.P. A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. *BMC Public Health. 2020;* 20(1): 1-16

https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-020-09062-0

Fynn, J.F., Hardeman, W., Milton, K., Murphy, J., Jones, A.P. A systematic review of the use and reporting of evaluation frameworks within evaluations of physical activity interventions. *Int J Behav Nutr Phys Act.* 2020; 17(1): 1-17

https://ijbnpa.biomedcentral.com/articles/10.1186/s12966-020-01013-7

Fynn, J.F., Hardeman, W., Milton, K., Jones A.P. Exploring influences on evaluation practice: A case study of a national physical activity programme. *Int J Behav Nutr Phys Act.* 2021;18(1):1-24

Fynn, J.F., Milton, K., Hardeman, W., Jones A.P. A model for effective partnerships and evidence-based practices. Unser review

Statement of jointly authored publications

The research reported is my own original work, which was carried out in collaboration with others as follows:

Chapter 1: Written by Judith Fynn

Chapter 2: Judith Fynn was the lead author of a paper published as:

Fynn, J.F., Hardeman, W., Milton, K., Jones A.P. A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. *BMC Public Health*. 2020; 20(1):1-16

Judith Fynn was the lead author. The research questions, search strategy, data extraction and synthesis were conceptualised and developed by Judith Fynn, Andy Jones and Wendy Hardeman. Judith Fynn conducted the screening and data extraction. Karen Milton checked and validated the screening of papers. Andy Jones and Wendy Hardeman checked and validated the data extraction. All authors contributed to the manuscript, critically reviewed and approved the final manuscript.

Chapter 3: Judith Fynn was the lead author of a paper published as:

Fynn, J.F., Hardeman, W., Milton, K., Murphy J., Jones, A.P. A systematic review of the use and reporting of evaluation frameworks within evaluations of physical activity interventions. *Int J Behav Nutr Phys Act.* 2020; 17(1): 1-17

Judith Fynn, Andy Jones, Wendy Hardeman and Karen Milton conceptualised the research questions, developed and agreed the search strategy, data extraction and synthesis. Judith Fynn conducted the search, screening, data extraction and analysis. Joey Murphy checked and validated the screening and data extraction. Judith Fynn drafted the manuscript. All authors critically reviewed and approved the final manuscript.

Chapter 4: Judith Fynn was the lead author of a paper published as:

Fynn, J.F., Hardeman, W., Milton, K., Jones A.P. Exploring influences on evaluation practice: A case study of a national physical activity programme. *Int J Behav Nutr Phys Act. 2021; 18(1): 1-24*

The research questions, data collection methods and analysis were conceptualised and developed by Judith Fynn. Judith Fynn conducted the research and drafted the manuscript, which was critically reviewed by Andy Jones, Wendy Hardeman and Karen Milton.

Chapter 5: Judith Fynn was the lead author of a paper which will be published as:

Fynn, J.F., Milton, K., Hardeman, W., Jones A.P. A model for effective partnerships and evidence-based practices. Under review

The research questions, data collection methods and analysis were conceptualised and developed by Judith Fynn. Judith Fynn conducted the analysis, and drafted the manuscript, which was critically reviewed by Andy Jones, Karen Milton and Wendy Hardeman.

Chapter 6: Written by Judith Fynn

Chapter 1. Introduction to the thesis

1.1. Introduction

Interventions to bring about behaviour change, such as increases in physical activity, are a public health priority (1, 2). The importance of physical activity for the physical, mental, and social health of individuals, communities and populations is well recognised (3-5). Yet, there remains high levels of inactivity amongst the population (1, 6, 7). If we are to meet targets to reduce physical inactivity to improve the health of the population, such as the World Health Organization's Global Action Plan target for a 15% reduction in physical inactivity by 2030 (1), it is essential to generate evidence about the complexity and effectiveness of interventions, and to use that evidence to inform practice and policy (8). Evidence-based public health seeks to ensure that decisions and actions are based on sound evidence; understanding how appropriate evidence can be generated and used is critical (9, 10).

Despite the value placed on evidence-based practices by researchers, policy makers, and practitioners, there continues to be calls for better evaluation and reporting to improve the evidence base (11-13). In particular, researchers and practitioners continue to debate the development and use of appropriate evaluation methods, what counts as evidence, and how to improve the reporting and use of practice-relevant evidence (10, 11, 14-16). Yet there are considerable gaps in our understanding of how best to facilitate and improve evaluation practices, and adoption and implementation of evidence. The research presented in this thesis aimed to address these gaps, by exploring strategies and recommendations that are intended to improve evaluation and the use of evidence, and to appraise their applicability and effectiveness. The use of evaluation frameworks to facilitate a systematic evaluation approach and research-practice partnerships to bring researchers and practitioners together to support evaluation are just two examples that this thesis seeks to explore. Whilst the focus is on physical activity interventions, the research sought to generate insights that would be applicable to other health behaviours and public health fields, where similar evidence-based approaches are required.

1.2. Research questions

To address the research aim, the following research questions were formulated:

1. What frameworks have been published that can be used for evaluation of physical activity and/or dietary change interventions?
2. What is the applicability of evaluation frameworks to different evaluation objectives, programmes, and contexts?
3. To what extent have evaluation frameworks been used within reported physical activity evaluation studies?

4. Which frameworks have been used within reported physical activity evaluation studies?
5. What is the quality of reporting with regards to how evaluation frameworks have been used within physical activity evaluation studies?
6. To what extent are strategies intended to facilitate real-world project evaluation effective?
7. What are the influences on evaluation practices in a real-world context?
8. How effective are programme level evaluation strategies at generating high quality, generalisable evidence?
9. What are the implications of influences on evaluation practice for the effective commissioning and evaluation of public health interventions?
10. How do partnerships and networks influence evaluation, dissemination, and evaluation use?
11. Who are the essential partners involved in evaluation of multi-agency interventions?
12. What are the implications of understanding influences on evaluation and partnership working for knowledge exchange and the capacity to do and use evaluation?

This introductory chapter provides the background to the thesis and highlights the gaps in understanding that informed the research questions and the development of the thesis. The latter part of the introductory chapter provides an overview of the chapters that follow. Briefly though, chapter two addresses questions one and two, and chapter three addresses questions three to five. Chapters four and five are based on a case study of a national physical activity programme, Sport England's Get Healthy Get Active (GHGA) programme. Questions six to nine are addressed in chapter four, and questions ten to twelve in chapter five.

1.3. Defining the research area

Increasing demands for evidence-based practice within public health have stimulated interest and expansion of evaluation and implementation research, and these two fields of research have evolved considerably over the last twenty years (12, 17, 18). Whilst evaluation is carried out for various purposes, a central purpose is to determine an intervention's effectiveness, thus identifying successful techniques that can inform future policy and practice. The World Health Organization defines evaluation as:

> *"the systematic examination and assessment of the features of an initiative and its effects, in order to produce information that can be used by those who have an interest in its improvement or effectiveness."* (19) (p.3)

Implementation research is:

> *"the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and, hence, to improve the quality and effectiveness of health services and care."* (17) (p.1)

The focus of this thesis is on practices and processes associated with these fields of research. One of the challenges in evaluation and implementation research is a lack of consistency and clarity in how concepts and terms are defined and applied (20). A common language that is accessible to practitioners and policy makers, as well as researchers, is needed to develop a shared understanding. The following section of this chapter, defines key terms and concepts, clarifies the relationships between them, and explains how they have been applied within this thesis. A glossary of key terms is provided in Appendix 1.

1.3.1. Evidence-based public health

Evidence-based practice is a central tenet of public health strategy development and implementation (21). Evidence-based public health involves: use of evidence to support decision making, intervention evaluation, learning from evaluation, reporting of findings, and use of evidence to inform and improve future decision making (22). This can be depicted as a cycle of evidence-based practice (Figure 1-1).

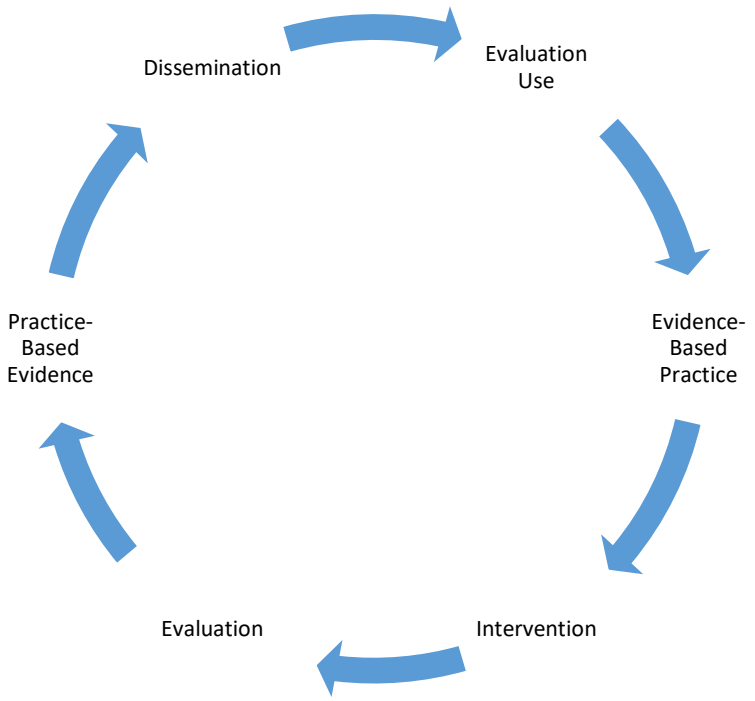

Figure 1-1 The evidence-based practice cycle

Figure 1-1 shows the relationship between the use of evidence to inform intervention development and implementation, and the use of evaluation to generate practice-based evidence (14, 23, 24). Practice-based evidence refers to the knowledge and insights generated from evaluation of 'real-world' interventions, and can be particularly important in generating practice-relevant evidence (25, 26). Integration of scientific evidence, with a consideration of evidence about an intervention's context, resources and stakeholder requirements, is central to evidence-based public health (24). Individual and organisational capacities, structures and systems that enable the generation of evidence through effective evaluation practices, and the flow of evidence and information, are vital to the functioning of the cycle (24).

Figure 1-1 is a simplified view of the sequence of activities, and previous studies have provided variations on this cycle (27-29). It serves here, though, to highlight the three fundamental practices, or processes, that underpin both evidence-based practice and this thesis - evaluation, dissemination, and evaluation use. Starting with the intervention, the importance of each of the elements depicted in Figure 1-1, the relationships between them, and how gaps in our understanding of the elements has informed the research questions, are explained below.

1.3.2. Interventions

The term 'intervention' is a general term that encompasses a broad spectrum of components, and includes interventions developed within a research context as well as those developed within a practice-based context. The terms 'public health' and 'health promotion' have both been used within this study to describe the interventions of interest. This is informed by a definition of public health interventions based on the socio-ecological model of public health (30, 31), and of health-promotion as interventions that adopt methods to enable people to improve their health or well-being (32). Following this approach, the focus is on interventions that seek to modify socio-ecological determinants of health, for example to bring about behaviour change, to address non-communicable health outcomes. This is a broad categorisation of interventions; in order to understand if and how an intervention works, for whom and in what contexts, and for evidence about their effectiveness to be used, it is essential that information regarding their contexts and components is defined and described clearly and consistently.

1.3.2.1. Complex interventions

Public health interventions are frequently described as 'complex'. Understanding what is meant by this term, and what makes an intervention complex, is important if we are to understand their effectiveness (33). Several authors have highlighted the importance of differentiating between complicated interventions as those that have multiple components, and complex interventions as those with multiple emergent outcomes (16, 33, 34).

Public health evaluations typically focus on the complexity of the intervention components, stakeholders and outcomes (16). A common approach to defining complex interventions has been to use the criteria provided in the MRC guidance on developing and evaluating complex interventions (35). This identifies several dimensions of complexity: the number and interactions between components; the number of behaviours required by those delivering or receiving an intervention; variability in the target population or outcomes; and the degree of flexibility in intervention implementation. More recently, there has been a growing appreciation of the need to consider the complexity and dynamic nature of the wider contextual system in which an intervention is implemented and evaluated (16, 36, 37).

Within this thesis the terms multi-agency, multi-sectoral and multi-component have been used to differentiate between complexity arising from an intervention's delivery context or content and modes of delivery. Multi-sectoral refers to the bringing together of different sectors, for example partnerships between the health and sports sectors to design, deliver or evaluate an intervention. Multi-agency, or inter-agency, is used to describe the bringing together of stakeholders from different groups or organisations, and may include those working within the same sector as well as those from different sectors, for example health charities, primary care, public health teams, or health researchers collaborating to address a common goal. Multi-component refers to interventions that have several elements, such as where physical activity interventions apply different modes of delivery or intervention functions (38), such as providing education or restructuring the environment.

As understanding and appreciation of the diverse factors that influence health behaviours has grown, interest in multi-component, multi-sectoral, and whole systems approaches to address public health priorities has expanded (13, 39, 40). Addressing physical inactivity amongst the population requires multi-sectoral and multi-dimensional actions for sustainable change to take place (1). Alongside this, the growth in appreciation and understanding of the wider health benefits of physical activity, such as social, mental and emotional health, highlights a requirement to capture evidence relating to multiple outcomes of interventions. The increasing number of calls for a systems approach to developing and delivering interventions, have implications for the evaluation methods needed (16).

1.3.2.2. Physical activity interventions

Interventions are often described by the specific behaviour which they aim to address, such as physical activity. Nevertheless, defining interventions by the target behaviour can mask their complexity. For example, in studies reporting on physical activity interventions there is often a lack of distinction between physical activity as a component of the delivery or as an outcome. The Behaviour Change Wheel identifies nine intervention functions; these are broad categories by

which an intervention might change behaviour. Their use is intended to facilitate clearer descriptions of intervention components to guide intervention development and reporting (38, 41). The intervention functions (education, persuasion, incentivisation, coercion, training, enablement, modelling, environmental restructuring and restrictions) have been applied in places within this thesis to provide a systematic and transparent approach to defining and reporting types of intervention. For example, they have been used in Chapter 3 to inform the inclusion and exclusion criteria to identify evaluation studies for the review, and to facilitate consistent reporting of included interventions.

1.3.2.3. Intervention setting or context

Defining intervention by setting or context is also difficult. For example, many interventions are described as 'community interventions', and frame this in relation to their location and/or target groups. However, there is a lack of consensus on how community is defined (42). Historically community has been considered as relating to location or place, whilst from a sociological viewpoint community is concerned with relationships between people. Understandings of community have evolved, and the term is increasingly defined by activity, purpose and commonalities of interest (42). Public Health England (PHE) (13) and the National Institute for Health and Care Excellence (NICE) (43) define 'community' as both place-based and where people share goals or affinity, and suggest community-centred approaches that promote relationships, mobilise local assets, and strengthen community capacities are more than simply community-based. Within this thesis a broad definition of community interventions has been adopted as being inclusive of both community-based and community-centred, for example in the inclusion criteria applied in the scoping review presented in Chapter 1. In the chapters that follow the initial scoping review, the term 'community' has been avoided to reduce the risk of ambiguity, and terms that describe the setting or context of delivery more precisely have been used. For example, the term 'setting' has been used to refer to the physical, geographical, or organisational space in which an intervention is implemented, and 'context' has been used as a general term that includes diverse internal and external factors that may influence an intervention, its implementation and/or its evaluation.

1.3.2.4. Real-world interventions

The terms 'real-world' and 'practice-based' have been used interchangeably to describe interventions that are part of normal service delivery or delivered in a practice setting, rather than within a research setting. The term 'programme' has been used to describe real-world interventions that represent a group of related 'projects'; the implication being that programmes are coordinated in such a way as to generate benefits beyond those available from individual projects (44). Generating evidence about the effectiveness of real-world interventions and

programmes is essential to evidence-based public health, and researchers and practitioners play a key role in this. 'Practitioner' is used to refer to those involved in decisions and actions related to intervention development, delivery, and evaluation from a practical standpoint, whilst 'researcher' is used to refer to those primarily engaged in research and evaluation from an academic standpoint.

1.3.3. Evaluation

The definition of evaluation provided by the World Health Organization (19), given at the start of section 1.3. above, highlights the importance of evaluation as a process, but also the importance of understanding its purpose and users. A more succinct, and often cited, definition is the one provided by Weiss (45) which stated:

> *"the overall aim of evaluation is to assist people and organizations to improve their plans, policies and practices on behalf of citizens."* (45) (p.469).

As strategy and programme development has become more evidence-based, and demands for accountability and quality assurance have grown, robust evaluation has become increasingly important to inform, and justify, decision making (8, 46-48). As the demands for practitioners to evaluate and use evidence have increased, the distinctions between research and practice, and between research and evaluation, have evolved. Pragmatic evaluation has emerged as an approach that seeks to balance the need for pragmatism within service delivery with demands for evaluation rigour. Researchers and practitioners are increasingly interested in pragmatic evaluation as an approach to facilitate evaluation of real-world interventions, and to close the gap between research and practice.

Complex interventions are difficult to evaluate, and much of the discourse within public health evaluation research has focused on barriers to evaluation, and recommendations for good practice e.g. the Medical Research Guidance on Developing and Evaluating Complex Interventions (35). Examples of barriers include: limited capacity and resources to conduct evaluation within applied contexts (11, 48); differing organisational structures and cultures (49-51); differing stakeholders' priorities and objectives for evaluation (52); differing values placed on forms of evidence (53); and a lack of awareness of appropriate tools (48, 54).

As understanding of the challenges to evaluation has developed, so too has the guidance available. Contemporary thinking in public health evaluation draws on many fields, including education and social sciences, and combines these with the scientific methods traditional within the fields of public health (55). Various recommendations, guidance and frameworks have been developed to support and improve evaluation practice. These reflect the growing understanding of the complexity of intervention design, implementation and evaluation. They include: guidance

on methodological approaches, such as theory-based or realist evaluation to identify and evaluate causal mechanisms, and link theory, process and outcomes (56, 57); recommendations for multiple methods to capture wider, longer-term and emergent outcomes (8, 11, 14, 58); and guidance for process evaluation to provide a more detailed understanding of an intervention's implementation, mechanisms of impact, and contextual factors (59). The use of evaluation frameworks to facilitate evaluation (46, 60), and research-practice partnerships (51, 61, 62) to improve evaluation practices and build capacity for evaluation are two strategies that this thesis has explored in more detail, and these are discussed below.

1.3.3.1. Evaluation frameworks - what guidance is available, how applicable and usable are they, and how well are they used?

A wide range of evaluation frameworks have been developed and published, from generic guidelines intended for use in a range of contexts, settings and sectors to checklists for use in interventions targeting specific health behaviours, health conditions, or populations. Within this thesis the term 'evaluation framework' is used to include any structured guidance that facilitates a systematic evaluation of the implementation or outcomes of an intervention. The ongoing calls for improvement in evaluation and reporting, despite the apparent plethora of frameworks, raises questions about their usability and the extent to which they are used. A recent review of evaluation frameworks for public health programmes (63) suggested that the wealth and breadth of frameworks may limit the ability of practitioners to access and use appropriate guidance. Questions remain regarding the applicability of evaluation frameworks to different evaluation objectives, intervention types, and contexts. These questions informed research questions one to five, and these are addressed in Chapters 2 and 3. It also raises questions as to their usability by different users; for example, there are questions about the extent to which frameworks developed by researchers are intended for practitioners and real-world interventions. These questions informed research questions six to nine, and are addressed in Chapter 4, which explored the use of a standardised framework, alongside other requirements for evaluation within a real-world physical activity programme.

1.3.3.2. Partnerships, collaborations and networks – how effective are they in facilitating and improving evidence-based practice?

Partnerships, collaborations and networks are advocated as a strategy to improve evaluation practices, knowledge exchange and use of evaluation. Research-practice partnerships, which bring practitioners and researchers together, can improve the quality of evaluation, help to build capacity for it (8, 11, 48, 50, 51), and also improve the use of evidence to inform programme development (61). For example, engagement of practitioners and policy makers in evaluation can improve understanding amongst researchers of what evidence is relevant and valued for decision

making in a real-world context, whilst engagement of research partners provides access to knowledge and expertise to help identify and implement appropriate and innovative evaluation methods (50, 64).

However, our understanding of the effectiveness of strategies in practice remains limited. Gaps remain in our understanding of influences on the capacity of practitioners to apply evaluation methods and to conduct evaluation (9, 49, 50, 54). Similarly, there are limitations in the evidence and gaps in our understanding of influences on partnership working and how partnerships may impact evaluation practices (51, 61, 62, 64, 65). The case study presented in Chapter 4 highlights the complex interconnections between influences on evaluation practices within multi-agency intervention implementation and evaluation, and in particular, it raises questions about how partnerships and networks may be developed and implemented to improve practice. These gaps in our understanding informed research questions ten to twelve which are addressed in Chapter 5.

1.3.4. Dissemination and evaluation use

Whilst robust evaluation and reporting are essential to build an evidence-base on which decision makers can draw, dissemination and evaluation use are critical to complete the evidence-based practice cycle. If good practice is not shared, then the translation from one setting to another, and wider scale up of effective interventions will remain limited. Dissemination is the process of communicating findings in ways that will facilitate their use in practice (66). Knowledge exchange is central to this, and these terms have been used interchangeably in the literature (67). Translational research and the use of evaluation are aspects of evaluation and implementation research that have afforded greater attention in recent years. Translational research explores which evidence and knowledge-transfer strategies are used within specific policies and programmes; it offers an approach to understanding the relationship between evidence building and review on one side of the evidence-based practice cycle, and evidence-based policy and practice on the other (68-70).

Evaluation use refers to the use of evidence generated from evaluation, and the effects of being involved in evaluation (55, 71). For example, effective use of evaluation by stakeholders may not only generate useful evidence of whether, how and why interventions work, but the evaluative process itself may improve understanding and the capacity to conduct and use evaluation. Alkin and King (55, 71, 72) and Cousins et al. (73, 74) have discussed the terminologies associated with ‘evaluation use’ and its evolution at length. The typology developed and used within their models has guided the application of the term ‘evaluation use’ within this thesis. For example, ‘process use’ is defined as the effects of being involved in evaluation and ‘findings use’ as the use of evidence generated. ‘Instrumental use’, defined as direct use of either process or findings, is

differentiated from 'conceptual use' (changing attitudes or improving knowledge) and 'symbolic use' (justifying decisions or actions). Evaluation use is an evolving and debatable concept; for the purposes of this study it is best conceptualised as a broad definition that encompasses the wider influences, and 'usefulness' to different users, as well as the types of use defined above. Although the term 'utility' is discussed in respect of the historical development of the field (55, 72), within this thesis the term 'usefulness' has been applied. The term 'utility' is defined as the state of being useful and can be used to describe both actual and potential uses, whereas the term 'usefulness' is defined as the quality or degree of being useful, and provides an indication of value to the user. Understanding the use made of an evaluation, and the usefulness to different users, may shed light on the effectiveness of the evaluation strategies adopted. This is explored within Chapter 4 and Chapter 5.

1.4. A conceptual framework for the study

This thesis aimed to improve understanding of evaluation practices, dissemination and evaluation use within multi-agency, multi-component public health interventions, using physical activity as an example. By exploring current practices, it also sought to apply the insights gained to develop recommendations to improve evidence-based public health, and to contribute to the underlying aim of closing the research-practice gap.

Figure 1-2 shows the conceptual framework which was developed to guide the research. Following Berman and Smyth's (75) proposed use of a conceptual framework, and their definitions, the framework provides an overview of the relationships between: the ontology, defined as the key concepts, language and context; the research questions; methods; and the epistemology, defined as the identification of new knowledge, its generalisability, and implications. This aligns with the concept of ontology applied in behaviour change research as "*a system specifying entities, definitions and inter-relationships for a given domain, with the potential to advance knowledge" (76)* (p.1). How this has been addressed in each of the chapters is explained in the latter part of this introductory chapter.

1.4.1. Methodology

Methodology refers to a system of methods used in a particular area of study or activity, normally associated with a particular paradigm. There is increasing appreciation of, and value placed on, multi-methodologies that combine aspects from different paradigms or fields of study. For example, within the fields of public health evaluation and implementation research there is an increasing interest in the use of multiple methods, qualitative methods, and systems thinking which has its origins in social sciences (10, 16). Systems approaches provide a framework to understand the relationships between actions, processes and influences within a wider whole-

system perspective, and a set of tools that are typically inter-disciplinary in nature (77). Guided by this general view, this study has drawn on principles within evaluation research to adopt an inter-disciplinary approach, to apply multiple methods, and to critically analyse multiple data sources. In essence, a pragmatic approach has been used to identify and apply methods that will produce the evidence needed to answer the research questions. The research design has sought to understand the contexts and underlying influences on evaluation and evidence-based practices, which are the outcomes of interest. Further, it has sought to ask evaluative questions and apply an evaluative logic to understand the complex influences on evaluation practices in an emergent research approach, with parallels to principles of developmental evaluation (78). By drawing on multiple methods, and recognising the partial nature of the knowledge generated at each stage, the thesis presents a set of linked studies to form a narrative in which the focus of inquiry in each chapter builds on, and is informed by, the findings of the previous chapter in an iterative approach to address the overall aim.

In each of the chapters qualitative methods have been used to provide an in-depth analysis, and to gain insights about a specific set of observations or phenomenon. A rigorous approach has been applied throughout to ensure that data collection, analysis and reporting has been systematic and transparent. The scoping review, presented in Chapter 2, followed the stages set out for a scoping review (79, 80): to identify the research question, apply a systematic search, and include consultation with experts. The scoping and systematic reviews used the PRISMA statements for the reporting of scoping reviews (81) and systematic reviews (82) respectively. Within the case study, logic models have been used in Chapter 4 and network analysis in Chapter 5.

The purpose of a case study is to provide an in-depth, multi-faceted understanding of a phenomenon in a real-world context (83). Case studies are often criticised for their limitations in terms of generalisability. A collective case study approach, as applied here is based on multiple projects, seeks to generate a broader understanding of a phenomenon (83). The use of multiple data sources sought to increase the internal validity. Framework analysis sought to allow consideration of the details of each case, and comparison between cases and data sources. Directed content analysis (84) allowed a deductive and inductive approach to be combined to facilitate a systematic in-depth and critical analysis. Details of the specific methods used in each of the studies are provided in the chapters.

Ontology	Research Questions	Methods	Knowledge and Implications
Evidence-based public health Three fundamental processes: evaluation, dissemination, use of evidence. **Physical activity interventions** • An example of public health interventions • Complexity: multi-agency, multi-sectoral, multi-component, and multiple outcomes • More precise defining of intervention components is needed e.g. behaviour types, intervention functions, delivery modes, target populations, settings and contexts • Real-world interventions **Evaluation** Challenges e.g. limitations in capacity to conduct evaluation in real-world interventions, differing priorities for evaluation and value placed on evidence. Recommendations: • Guidance on process and outcome evaluation • Evaluation of wider outcomes • Evaluation frameworks • Contextual influences • Multiple-methods • Pragmatic evaluation • Research-practice partnerships **Dissemination and evaluation use** Essential processes to complete the cycle, sharing good practice and use of evidence are critical for translation between settings and scale up of effective intervention components.	1.What frameworks have been published that can be used for evaluation of physical activity and/or dietary change interventions? 2.What is the applicability of evaluation frameworks to different evaluation objectives, programmes, and contexts?	Scoping review: systematic search and data extraction, consultation with experts, directed content analysis, categorisation and mapping of frameworks.	Improved understanding and signposting to relevant evaluation guidance. Identification of limitations in guidance and where further development is needed.
	3.To what extent have evaluation frameworks been used within reported physical activity evaluation studies? 4.Which frameworks have been used within reported physical activity evaluation studies? 5.What is the quality of reporting with regards to how they have been used?	Systematic review: development of a checklist of indicators to systematically and critically appraise the use and reporting of different evaluation frameworks.	Improved understanding of limitations in use and reporting of frameworks. Development of a checklist to improve future reporting and to review the quality of an evaluation report.
	6.To what extent are strategies intended to facilitate real-world project evaluation effective? 7.What are the influences on evaluation practices in a real-world context? 8.How effective are programme level evaluation strategies at generating high quality, generalisable evidence? 9.What are the implications of influences on practice for the effective commissioning and evaluation of public health interventions?	Collective case study: documentary analysis and semi-structured interviews, in-depth critical analysis of multiple sources of evidence from 23 projects funded through a national physical activity programme.	Improved understanding of: - influences on evaluation practices. - actions needed to improve evaluation and evidence-based practice within real-world public health interventions.
	10.How do partnerships and networks influence evaluation, dissemination and evaluation use? 11.Who are the essential partners involved in evaluation of multi-agency interventions? 12.What are the implications of understanding influences on evaluation and partnership working for knowledge exchange and the capacity to do and use evaluation?	Building on the findings from the case study: thematic analysis and network analysis combined to describe the network, and explore the relationships between processes and partnership characteristics to develop a conceptual model.	Improved understanding of: - influences of partnerships and networks on evaluation, dissemination, & evaluation use. - implications for implementation of partnerships to capitalise on the potential benefits.

Figure 1-2 Conceptual framework for the research

1.5. Overview of the chapters

The thesis chapters are presented as a linked narrative. Each chapter builds on the preceding one, and a brief introduction is provided at the start of each chapter to highlight the connections between them and how each contributes to the overall aim to improve understanding and implementation of evaluation, dissemination and evaluation use.

Chapter Two: Evaluation frameworks, their availability, applicability and usability

Paper: "A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation." (85)

Chapter two reports on a scoping review conducted to identify and appraise evaluation frameworks that could be used to support evaluation of physical activity and dietary change interventions. This focus was chosen, as real-world behaviour change programmes often aim to address both dietary change and physical activity behaviours, particularly where the goal is to address populations that are overweight or obese, or have associated co-morbidities. The purpose of this review was to develop a better understanding of the frameworks available and their applicability, or usability, in evaluating different interventions and meeting evaluation objectives. Seventy-one frameworks were identified and included in the appraisal. A typology of the frameworks according to evaluation objectives, programme type and framework format was developed, and each framework was mapped against a range of evaluation components. These can be used to signpost and support those engaged in evaluation to identify which frameworks may be most appropriate to their needs. The findings also highlighted where there is overlap and gaps in the guidance provided by the frameworks. To understand more fully potential limitations in using the frameworks Chapter three explored the use and reporting of evaluation frameworks within published evaluation studies. The scoping review identified 73 different frameworks; this included just four that were specific to evaluation of physical activity and three specific to dietary interventions, with the remaining being intended for use in public health or more generic interventions. The breadth and extent of the frameworks, and also the large number of evaluation studies identified in the process of conducting the search for the scoping review, informed a pragmatic decision to narrow the focus specifically to physical activity interventions in the subsequent chapters.

Chapter Three: A systematic review of the use and reporting of evaluation frameworks in physical activity evaluation studies

Paper: "A systematic review of the use of evaluation frameworks within published evaluations of community-centred physical activity programmes." (86)

Chapter three reports a systematic review that aimed to understand if and how the available evaluation frameworks were used and reported within published evaluation studies of physical activity interventions. Robust reporting of evaluation studies is vital to building an evidence base that researchers, policy makers and practitioners can draw on. Where an evaluation framework is applied appropriately, its use can facilitate a systematic evaluation and improve the reporting of an evaluation study. However, frameworks are often underused or underreported. A checklist of indicators was developed to appraise the evaluation studies; this can be used to facilitate the reporting of an evaluation study and to review the quality of an evaluation report.

Chapter Four: Use of evaluation guidance in practice - exploring influences on evaluation practice within a national physical activity programme

Paper: "Exploring influences on evaluation practice: A case study of a national physical activity programme." (87)

This chapter reports a collective case study conducted to explore evaluation practices and influences on practice within a real-world physical activity programme, Sport England's Get Healthy Get Active programme. The programme was developed to build an evidence base for the role sport plays in improving health through engaging inactive people in physical activity. Evaluation was central to the programme design and funding requirements; each project was required to engage an independent evaluation partner, and to use a standard evaluation framework and standard data collection methods. The programme also exemplified multi-sectoral and multi-component approaches to public health. Using the programme as a case study provided an opportunity to explore influences on evaluation practices. The study highlighted the context specific nature of influences and the complex interconnections between them. The discussion of this chapter reflects on the implications for the commissioning and evaluation of health promotion interventions. The findings suggested that the nature of relationships was an important influence on the effectiveness of partnership working and the use of the evaluation to individuals and organisations. Chapter 5 therefore looked at these two themes in more detail to better understand the relationships between partnerships, evaluation practices, and evaluation use.

Chapter Five: Partnerships, collaborations and networks for evaluation: their use in facilitating evaluation, dissemination and evaluation use

Paper: "A model for effective partnerships and evidence-based practices." In submission

Chapter 5 is based on the case study of the national physical activity programme reported in Chapter 4 and aimed to advance understanding of how partnership working can best be implemented to improve evidence-based practice. This chapter used the data gathered from the interviews conducted for the case study, and combined thematic analysis with network analysis to

describe the network of partners, and to explore their experiences and perceptions of partnerships, evaluation, and the use made of the evaluation by themselves, their organisations, or partners. Findings were used to develop a conceptual model of the relationships between partnerships, processes and partnership characteristics that facilitate evaluation, dissemination, and evaluation use.

Chapter Six: Discussion and conclusions - the use, usability and usefulness of strategies to support evaluation and evidence-based public health

Chapter 6 provides a summary of the findings and discusses how these contribute to addressing the overall aim of the thesis. Limitations, strengths and the generalisability of the thesis are discussed. A central tenet of the thesis, and each of the chapters, has been to apply the insights gained to consider the implications for practitioners, decision makers, and researchers. This is explored within this final chapter. Lastly, a personal reflection on the process and experiences of conducting the research and writing the thesis is provided.

Chapter 2. Evaluation frameworks, their availability, applicability and usability

Introduction

The overall aim of the thesis has been driven by a desire to understand and improve evidence-based practices, specifically evaluation, dissemination and evaluation use. An appreciation of the increasing demands on stakeholders to conduct evaluation, and for the value of evaluation to generate evidence that could be used to inform policy and practice and to build capacity to conduct and use evidence, stimulated and guided the focus of interest. Having scoped the evaluation research literature it was apparent that there was a plethora of guidance and recommendations on how best to improve evaluation (63). Yet, much of the evaluation literature focused on the barriers (11, 49, 50, 52), the missed opportunities to report robust evaluations (48), and calls for better reporting of evaluation studies (11-13, 88). This raised questions regarding the availability, usability, and applicability of evaluation frameworks and guidance for those wishing to evaluate public health interventions, such as those aiming to improve physical activity and dietary behaviours. The need to understand the guidance and frameworks available, and their applicability and usability underpins this chapter.

Background

Programmes that aim to increase physical activity and improve dietary behaviours in individuals, groups and populations play a central role in addressing local, national and global public health priorities (1, 2). Recent strategies have advocated approaches that are multi-sectorial, community-centred and evidence-based (2, 39, 89, 90). Understanding if, when, and how these programmes are effective is important to justify policy, programme and funding decisions, and to inform and improve future decisions and practice. In order to achieve this, there is a need for appropriate and comprehensive programme evaluation (11, 13).

Practice-based evidence is generated from formal evaluation of programmes in real-world settings and is a fundamental part of evidence-based public health (14, 26, 91). Those involved in the design, delivery and commissioning of physical activity and dietary change programmes are expected to evaluate programmes and contribute to the evidence base. However, real-world behaviour change programmes are complex and difficult to evaluate (15, 68). The challenges of programme evaluation may relate to contextual factors that influence the complexity of the programme itself, e.g. its setting, target population, intervention function(s), or intended outcome(s) (15), or to factors that influence the evaluation priorities and objectives, e.g. differing stakeholder evaluation needs and organisational, political or resourcing factors (52). Some of the practical challenges in conducting evaluation include the use of appropriate evaluation methods

and tools, understanding what counts as evidence and how that is applied, and the roles of practitioners and researchers in evaluating real-world programmes (11, 14, 46, 68, 92).

Evaluation frameworks facilitate a systematic approach to evaluation and can help mitigate against some of the above challenges. Frameworks can enable multiple stakeholders to gain a shared understanding of the programme and evaluation process, and help to identify and agree upon appropriate objectives and methods. In this way, they can facilitate a more comprehensive evaluation, and may improve the fit between researcher-led and practitioner-led evaluation approaches (46). A range of evaluation frameworks have been published. These include those developed specifically for use in programmes targeting specific health behaviours, conditions or populations (e.g. physical activity programmes (93-95)), those developed for health promotion and public health programmes more broadly (e.g. RE-AIM (47)), and generic frameworks intended to be applicable across a range of contexts, settings and sectors (e.g. Realist Evaluation (56)).

It is noteworthy that there is wide variation in the use of terminology used to describe frameworks, in the format of different frameworks, and in the context and ways in which they are intended to be used. Differentiating between frameworks, guidance, models or tools can be a challenge (96). In this review the term 'evaluation framework' is used to include any structured guidance which facilitates a systematic evaluation of the implementation or outcomes of a programme. A 'generic' framework is used to refer to one that is intended for use across a range of contexts, settings and sectors, as opposed to one that has been developed for use in a specific context or field. Several frameworks have been developed for evaluation of programme implementation (process evaluation), whilst others focus on programme effectiveness (outcome evaluation) or are intended to facilitate an overall or comprehensive evaluation. In order to understand the content and focus of the frameworks and the contexts in which they may be applied, we have referred to the individual elements encompassed within evaluation as an 'evaluation component'.

Many frameworks and developments in evaluation come from the research community, yet their intended audience and purpose is often unclear. For example, questions remain about the extent to which these frameworks are intended for use in practitioner-led or researcher-led evaluation, and their applicability to different evaluation objectives, programmes, and contexts.

Previous reviews of evaluation frameworks have been limited to frameworks which evaluate specific aspects of a programme, for example health inequalities (97), or methods used in health programme evaluations (98, 99). Within the field of implementation science, reviews have focused on frameworks for translation of research to practice (69, 100). The review by Denford et al. (63) made a valuable contribution by providing an overview of guidance available to support evaluation of public health programmes. However, it was limited to a subset of 48 documents

created or sourced by national and international organisations and published since 2000. As a result some key evaluation frameworks published before 2000 or within the academic literature were not included, such as RE-AIM (47) and Realist Evaluation (56). Denford et al. included various guidance documents intended for use in evaluating programmes targeting a broad range of health behaviours and health problems (e.g., smoking, asthma), as well as generic ones. Whilst they suggested that the wealth and breadth of available evaluation guidance may be a limiting factor in the ability of practitioners to access and apply appropriate guidance, the resulting review (63) and associated online catalogue (101) may still overwhelm practitioners seeking guidance on how to evaluate their specific programme.

To resolve some of this complexity we sought to develop a typology of frameworks, to help guide decision making by those involved in programme evaluation. The purpose was to appraise the frameworks that may be applicable for the evaluation of physical activity or dietary change programmes. By mapping the frameworks against a range of evaluation components (such as elements of process or outcome evaluation), we aimed to develop an overview of guidance included in each framework, enabling practitioners, commissioners and evaluators to identify and agree which frameworks may best meet their needs.

Objectives

1. To identify published frameworks that can be used for evaluation of physical activity and/or dietary change programmes.

2. To identify each framework's stated scope in order to assess their applicability to different evaluation objectives, programmes and contexts.

3. To identify and map which evaluation components are encompassed within each framework.

4. To use the findings to develop a typology of frameworks.

Method

A scoping review approach was used, as this allowed the extent and nature of the literature on evaluation guidance to be identified and an overview of the available frameworks to be developed (79, 80, 102). In line with the stages of a scoping review (79, 80), the process involved identification of the research question, a systematic search, consultation with experts, and mapping of the frameworks against different components of evaluation. We followed the PRISMA–ScR statement for the reporting of scoping reviews (81).

Search strategy

To identify any frameworks that could be applied to physical activity and/or dietary change programmes, we used a broad search strategy to find those intended for use in public health,

health promotion and generic programmes as well as those developed specifically for use in evaluating physical activity and dietary change programmes. Firstly, a search was conducted in Scopus. As a meta-database, including records from MEDLINE and EMBASE as well as other sources, Scopus is the world's largest abstract and citation database of peer-reviewed literature. It contains sources across a range of fields including medicine, sciences, humanities and social sciences. The following search strategy was used: (TITLE ((framework OR model OR guid* OR tool)) AND TITLE-ABS-KEY (("physical activity" OR exercise OR diet OR obes* OR overweight OR "public health" OR "health promotion")) AND TITLE-ABS-KEY (communit*) AND TITLE-ABS-KEY (evaluat*)). No date restriction was applied. The search was undertaken in March 2018. All sources identified from the search were downloaded into the Endnote reference manager, and any duplicates were removed.

Secondly, between April and September 2018, we searched for grey literature on the websites of key organisations interested in evaluation of physical activity and/or dietary change programmes, using 'evaluation framework' as a search term. This included the World Health Organization (WHO), Public Health England (PHE), Sport England, and the Centers for Disease Control and Prevention (CDC). Additional sources were identified from the authors' existing files. We consulted evaluation experts and stakeholders including academics, those involved in public health policy development and evaluation, and evaluation consultants within the domains of physical activity or dietary change, to augment the search results. These experts were contacted and asked to provide feedback on the list of frameworks we had identified by the search strategy and to identify any omissions. Reference lists were examined for additional relevant sources.

Sources were screened by title and abstract, and then by full text (JF). Full text screening was independently validated (KM) and disagreements resolved through discussion. Consensus could not be reached for six sources, which were checked by a third reviewer (AJ) and agreed through further discussion.

Inclusion and exclusion criteria

Inclusion and exclusion criteria were defined *a priori* and applied to all sources (JF). Table 2-1 provides details of the full inclusion and exclusion criteria. Sources were included from both the academic and grey literature that described a framework to support systematic evaluation of a physical activity and/or dietary change programme, including generic, public health or health promotion frameworks applicable to physical activity or dietary change programmes. Academic literature included journal articles and books. Grey literature was defined as all other printed and electronic documents published by organisations and agencies. Web-based sources were included

if they provided systematic guidance on how to conduct an evaluation but excluded if they were an organisation's general website without guidance. Only sources in English were included.

Table 2-1 Inclusion and exclusion criteria for the scoping review

Inclusion criteria	Exclusion criteria
Sources describing a framework or guidance to support evaluation of a programme e.g. process and/or outcome evaluation.	Sources describing a specific measurement tool.
Sources describing a framework or guidance to facilitate evaluation of physical activity, dietary change, public health or health promotion programmes.	Frameworks designed to support evaluation of programmes targeting other health behaviours (e.g. smoking, alcohol, substance abuse) or conditions not specifically linked to physical activity or dietary behaviours (e.g. HIV, mental health).
Sources describing a framework or guidance to support evaluation of a specific evaluation component that aligns with the underlying principles of real-world, community-based or health promotion programmes, e.g. community development, participation, wider health and non-health outcomes.	Sources describing frameworks or guidelines intended to support evaluation of technology-based programmes or cost-effectiveness, as these are related to distinct specialised areas of evaluation or health promotion approach.
Empirical and/or methodological studies reporting the development and/or validation of an evaluation framework, as well as conceptual or discussion papers describing a framework or guidance on evaluation.	Theoretical or conceptual models of conditions or interventions. Guidance on policy or action for management of disease, policy or clinical practices. Evaluation studies reporting the use of an evaluation framework.

Data extraction and synthesis

To address the first and second objective, a data extraction template was used to collate information about each framework. The name of each framework was identified. Where no framework name was provided in the source, a short name was given based on the authors' description in the title or abstract. To assess each framework's scope and applicability to the evaluation of physical activity and/or dietary change programmes, data extraction fields included the stated evaluation objective, the types of programme it was intended for, and additional data related to general characteristics of each framework, e.g. its intended audience, format and development process.

To address the third objective we developed a set of data extraction fields to enable us to appraise whether each framework provided any guidance on a range of evaluation components, and what that guidance comprised. We have used the term 'evaluation component' to refer to individual elements encompassed within evaluation; for example elements of process or outcome evaluation. The list of evaluation components included in the data extraction template was identified *a priori*, and developed through a process of consensus building. We initially identified a

list of evaluation components that were informed by recommendations for good practice in the evaluation literature, for example implementation, reach and unanticipated outcomes (15, 29, 59, 103). This was further developed through consultation with evaluation experts, who were contacted and asked to comment on the appropriateness of the evaluation components we had identified and to identify any gaps or additional components based on their personal experience and knowledge of programme evaluation. Table 2-2 shows the full list of evaluation components grouped into those related to: (1) process evaluation, (2) outcome evaluation and (3) study design. Grouping programme context, theory of change and logic models within process evaluation components aligns with its inclusion in the UK Medical Research Council (MRC) Process Evaluation guidance (59), and recognises the crucial role of logic models in the early stages of developing an evaluation plan, in reporting causal assumptions about how a programme works, and informing process and outcome questions and methods. Where possible, pre-defined categorical responses were developed to facilitate the data extraction, coding and synthesis.

Where authors had described the scope of a framework variably, and where terms were not mutually exclusive, multiple terms were noted in the data extraction table. For example, terms such as community or practice based were used interchangeably to describe a study, intervention, setting or population. Where frameworks gave more detailed guidance on specific evaluation components, we also extracted a summary of what the guidance comprised. For each evaluation component we assessed whether the framework simply mentioned or provided more detailed guidance on how to evaluate or break down the relevant component.

Data extraction was completed by JF. To verify the data extraction, a random sample of twenty sources was checked independently by AJ and WH. Differences were resolved through discussion and used to establish agreed definitions that were then applied to further data extraction.

Framework format, programme type and evaluation objectives are typically used to describe frameworks. We therefore used these aspects to develop our typology for the frameworks. For the purposes of categorising the frameworks within the typology we used the dominant term presented in the description and content of the source as the basis for identifying each framework's most defining characteristic. The extracted data was also used to map each framework against the evaluation components in order to provide an overview of the guidance encompassed within the frameworks. A narrative synthesis of the findings is presented.

Table 2-2 Evaluation components agreed for data extraction and mapping of frameworks

Groups of evaluation components	Evaluation components for data extraction
(1) Process Evaluation	Describing programme context
	Using theory of change or logic models
	Reach
	Implementation
	Maintenance
	Any other process measures stated
(2) Outcome Evaluation	Behavioural outcomes
	Health outcomes
	Non-health outcomes
	Unanticipated outcomes
(3) Study Design	Stakeholder involvement
	Participatory evaluation
	Evaluation linked to stages of programme
	Evaluation at different time points
	Study design/method
	Data collection
	Data analysis
	Dissemination and reporting of findings

Results

Study selection

The initial search in Scopus yielded 1604 sources once duplicates were removed. An additional 24 sources were identified from the grey literature search and consultation process, and a further 60 sources were identified from reference lists. Many articles were identified as ineligible from their title alone, mostly because they related to conceptual models, treatment models, or conditions not relevant to physical activity or diet. If there was any uncertainty regarding the potential eligibility of a paper, it was included in the next stage of the screening process. After screening of titles and abstracts 168 full-text sources were assessed for eligibility (PRISMA diagram, Figure 2-1).

At full-text screening 83 sources were included and 85 were excluded. Of those excluded, 37 were reported evaluation studies that used one or more framework(s) and three were sources that critically appraised framework(s) (104-106). The reference lists of these sources were searched to identify the index papers that described the frameworks mentioned.

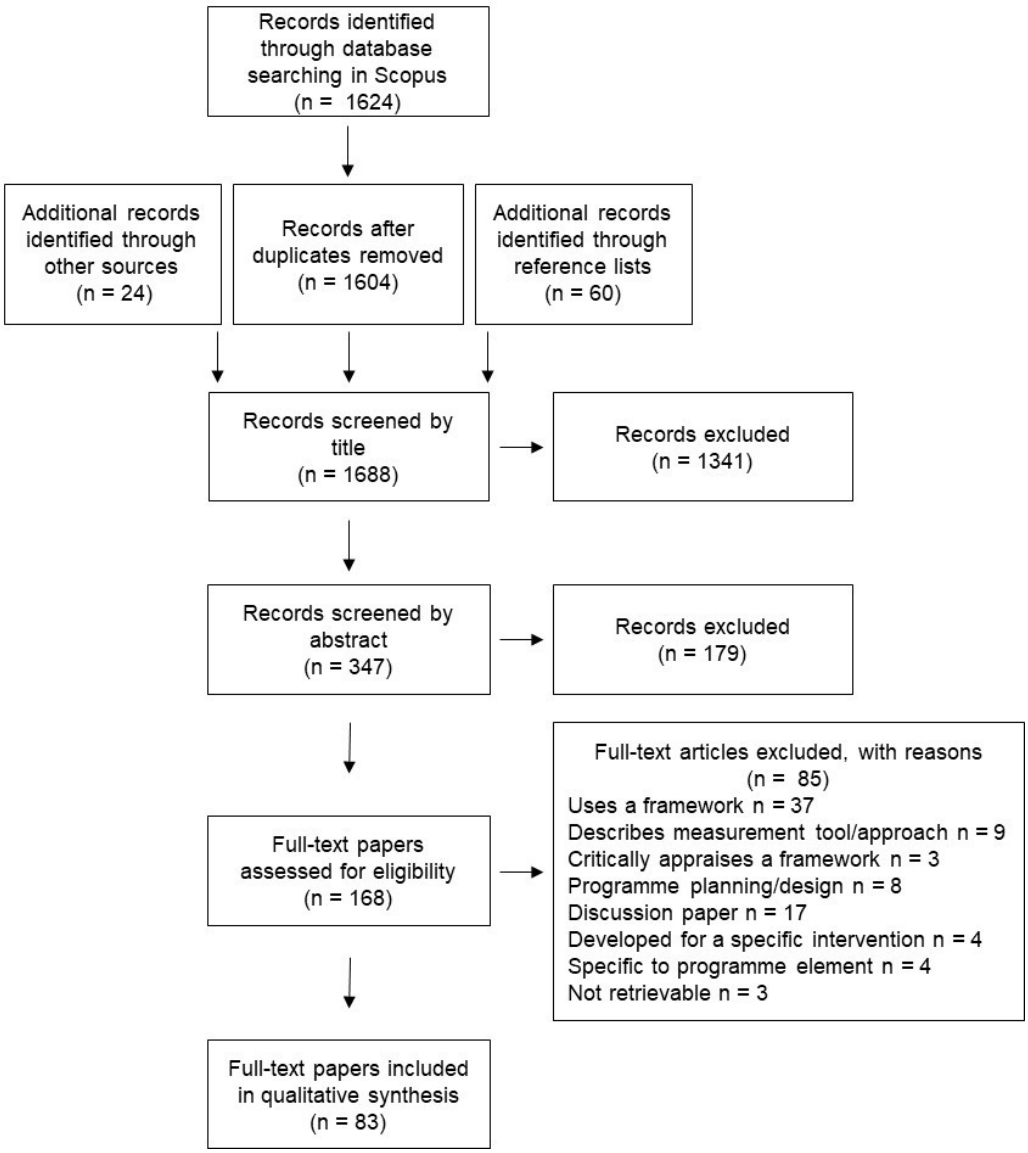

Figure 2-1 PRISMA diagram of the screening process for the scoping review

Sources which described programme and evaluation practices in general terms, e.g. (107), and those which described a specific measurement tool, e.g. photovoice (108) and memorable messages (109) were excluded. Other sources were also excluded if they reported a framework linked to a specific intervention and in such a way that it was not generalisable (e.g. Framework for Washington State's Healthy Communities Projects (110)). Planning frameworks that were solely for guidance on the design and development of an intervention were also excluded (e.g. (111-113), but a number were retained where they included guidance related to evaluation (114-118).

For frameworks which were described in more than one publication, for example in full and summary articles, we included both sources to facilitate data extraction and analysis, e.g. PRECEED-PROCEDE (114, 119), the CDC Framework (92, 120), UK MRC Guidance (15, 35, 121,

122), and Impact Pathway Analysis (123, 124). Data were extracted from 83 sources, describing 71 evaluation frameworks.

Identification of the evaluation frameworks available

A brief description of each framework is provided in Additional File 2.1 and an overview of their general characteristics is provided in Additional File 2.2. Table 2-3 lists the frameworks included in the review, grouped by decade of publication and source (academic/grey literature). All included frameworks were published during the last three decades (1990 onwards). Forty-two were described in academic publications and twenty-three in the grey literature. Six frameworks were reported in both the grey and academic literature (35, 59, 92, 120-128).

Table 2-3 also indicates the format of each framework. This ranged from highly structured to more flexible guidance. Thirty of the frameworks were presented as a set of steps; typically, these steps align with the stages of programme development and implementation. Twenty-four frameworks were presented as a set of indicators or questions, ranging from those that included a small number of key indicators (47, 129-131) to those that encompassed a longer checklist of evaluation criteria or questions (93-95, 125, 132). The remaining seventeen provided flexible evaluation guidance.

Table 2-3 Included frameworks grouped by decade of publication and source

	1990-1999	2000-2009	2010-2018
Academic Literature[1]	*Evaluation of Health Education (133)* *Evaluation of Healthy Community Initiatives (134)* *Health Workers Guide (135)* *Realistic Evaluation (56)* *Utilization-Focused Evaluation (136)* Framework for Outcome Assessment (137) Intervention Mapping (115, 138) MMIPP (139) PRECEDE-PROCEED (114, 119) **Stages of Evaluation Model (29, 140)** **Principles for Evaluating Community HP (141)** **RE-AIM (47)**	*California Healthy Cities Framework (142)* *Setting Standards (103)* Community Initiative Evaluation Model (143, 144) Evaluation in Health Promotion (145) Formative Model of Service Evaluation (146) Planning, Implementation and Evaluation Model (147) Process Evaluation for Public Health (148) Six Step Guide to Process Evaluation (149) **Concepts in process evaluation (150)** **Evaluating Legacy of community health initiatives (151)** **Getting to Outcomes (152)** **HEBS Framework (46)** **Levels of Coalition Evaluation (153)** **Participative Framework Health Inequalities (129)** **Participation, Partnerships & Equity (154)** **Settings for Health Promotion (117)** **Well Connected (131)**	*Cross-site Evaluation Tool (155)* *Empowerment Framework in Nutrition (156)* *Evaluating Complex Community-Based HP (157)* Generic Evaluation Toolkit (158) Systematic Evaluation Multiple Components (60) **Contextual Factors Framework (118)** **Coordinated Action Checklist (159)** **Multilevel Framework (160)** **OPEN tool (161)** **Process Evaluation in Group Settings (162)** **Process Evaluation Cluster Randomised Trials (163)** **Supportive Social Environments for Health (164)** **Three Dimensional Health Cube (130)**
Grey Literature[1]	*WHO Recommendations (19)*	*Kellogg Foundation Evaluation Handbook (165)* *Logic Model Development Guide (116)* *NICE Guidance: Behaviour Change (166)* Evaluating Community Projects (167) Framework for Community Health (168) Evaluating Sport and Physical Activity (169) Health Planners Toolkit (170) LEAP (171) Physical Activity Evaluation Handbook (172) Sport England Evaluation Framework (173) **SEF for Weight Management (95)**	Better Evaluation (174) Centre TRT's Framework (175) Community Toolbox (176) Evaluation Works: a toolkit (177) Public Health England (PHE) Guide (178) Magenta Book (179) Ontario Evaluation Workbook (180) Victoria Govt DoH Framework (181) **GPAT (132)** **SEF for Dietary Interventions (94)** **SEF for Physical Activity (93)**
Both	CDC Framework (92, 120)	*MRC Complex Intervention Guidance (35, 121)* Participatory Impact Pathway Analysis (PIPA) (123, 124, 182)	*MRC Process Evaluation Guidance (15, 59, 122)* *MRC Natural Experiments (127, 128)* **GENIE (125, 126)**

Italics = flexible guidance, Normal text = frameworks formatted as steps, **Bold = frameworks formatted as a set of indicators**

[1] Academic literature included journal articles and books. Grey literature was defined as all other printed and electronic documents published by organisations and agencies.

Sources generally described the framework development as being based on (i) some combination of literature review, consultation and testing, (ii) experiences of conducting evaluation(s), or (iii) prior frameworks or theory. Many of the more recently published frameworks referred to earlier ones as informing their development, such as realist evaluation (56), utilization-focused evaluation (136), PRECEDE-PROCEED (114) and intervention mapping (138). Several frameworks formatted as a set of steps mentioned the CDC framework (120) and other step-based frameworks (148, 150) as informing their development. Several frameworks formatted as a checklist referred to RE-AIM (47) as informing the indicators.

Seventeen frameworks provided guidance or links to sources for additional support or training in using the framework. Those that gave more detailed guidance of training and support, including links to additional resources, were predominantly published within the grey literature and had an online presence (165, 171, 173, 174, 177).

Scope of the evaluation frameworks and development of a typology

There was considerable heterogeneity in the terminology used to describe the scope of the frameworks. Authors described them variously in terms of purpose, content, or applicability to different programme and/or evaluation contexts. Additional File 2.2 shows the range of the descriptors used by authors. For example, thirty-one sources mentioned the frameworks were intended for use in real-world or practice-based settings, and twenty-two were intended for use in community-based programmes, with these terms often used interchangeably. Others were described as applicable to specific intervention functions (e.g. health education (125) or policy (19, 152, 175)), or specific intervention or study types (e.g. complex interventions (35, 112, 157), natural experiments (127) or cluster randomised trials (163)). These terms were not mutually exclusive so were not used to categorise the frameworks and develop the typology but are indicated within Additional File 2.2.

Programme type

Despite this variability in descriptors used by authors, we used the intended programme type as the primary categorisation to develop the typology, followed by the evaluation objective and the framework format. These characteristics enabled us to group the frameworks by applying the dominant description provided by the authors as an indication of a framework's most defining characteristics. Figures 2-2 to 2-4 show the typology which signposts to each framework within the categories.

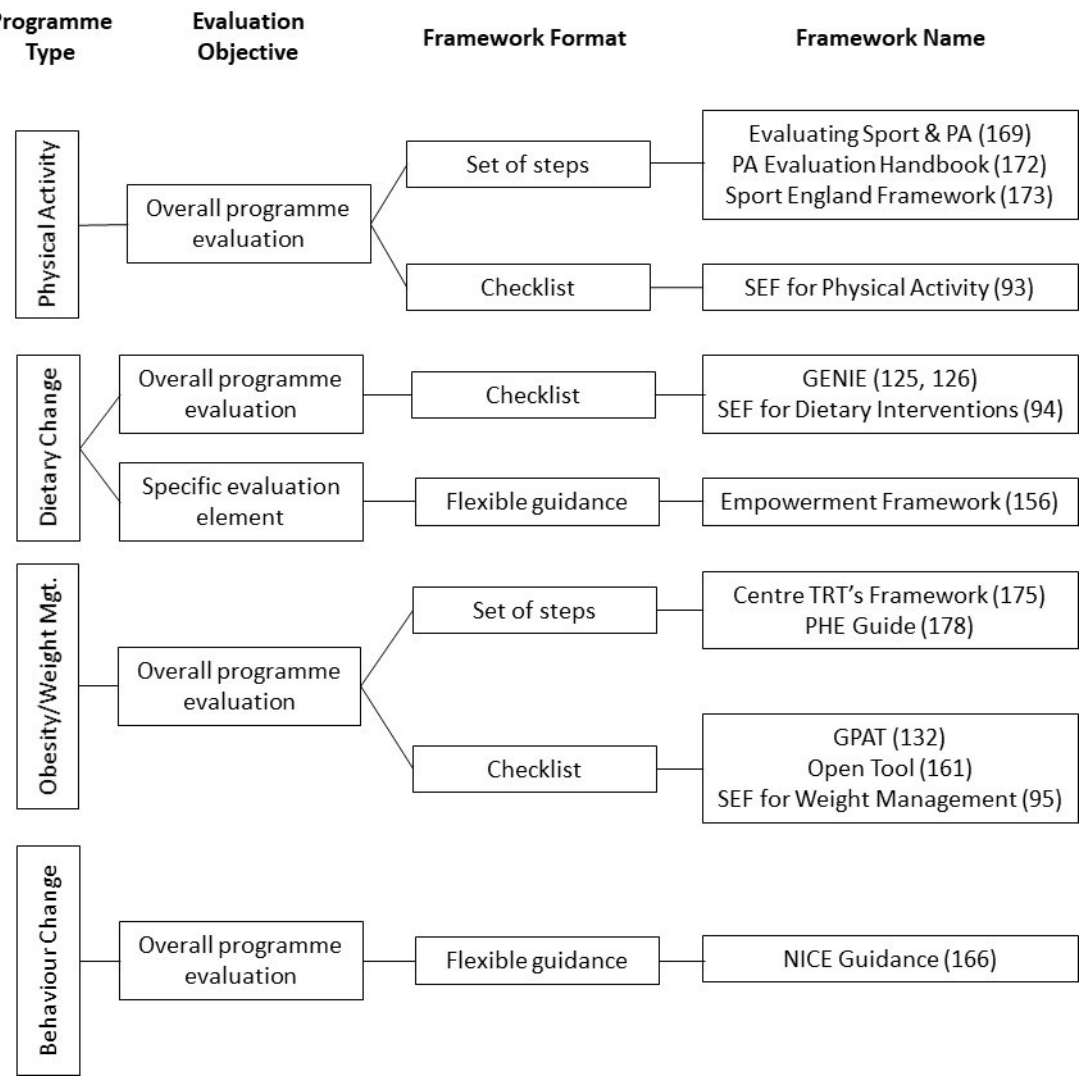

Figure 2-2 Typology of evaluation frameworks intended for use in physical activity, dietary change or behaviour change programmes

Twelve frameworks were stated as intended for use in physical activity and/or dietary change programme evaluation, and one as for use in behaviour change interventions (166) (Figure 2-2). Forty-eight were described as for use in public health or health promotion programmes. Some of these clearly stated how their components related to health promotion principles. However, several used the terms health promotion and public health interchangeably, and these were therefore grouped together (Figure 2-3). A further ten frameworks were described as applicable to a range of programme types and we have grouped these as intended for generic programme evaluation (Figure 2-4).

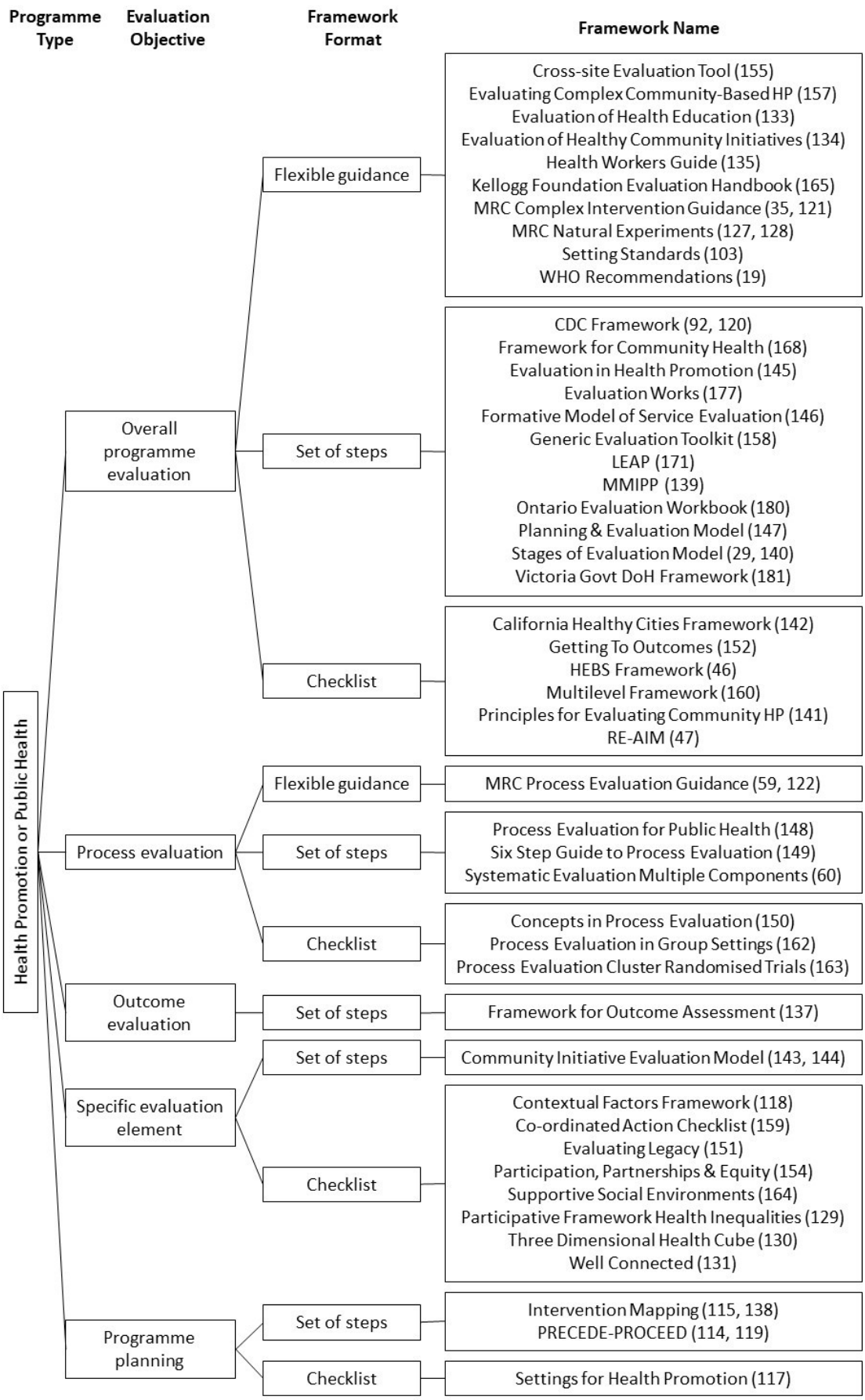

Figure 2-3 Typology of evaluation frameworks intended for use in health promotion or public health programmes

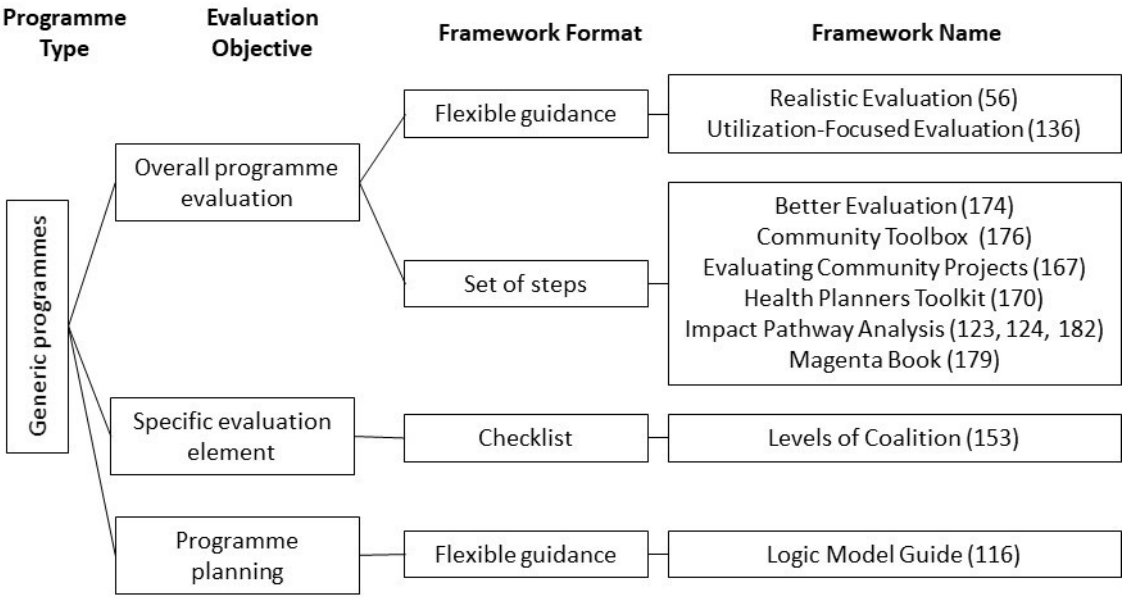

Figure 2-4 Typology of evaluation frameworks intended for use in generic programmes

Evaluation objective

Frameworks were also described variously in terms of their evaluation focus or objective, and we used this to further develop the typology shown in Figures 2-2 to 2-4. Fifty-two were stated as providing guidance on overall programme evaluation, nine as specific to process evaluation and one as specific to outcome evaluation. Several of the frameworks provided guidance on evaluating specific programme elements such as empowerment (156), partnerships and participation (131, 143, 153, 154, 159, 164), contextual factors (118), or legacy (151). Four frameworks were described as 'planning frameworks' but incorporated guidance on evaluation (114-117); these are grouped separately within the typology (Figures 2-2 to 2-4). Other frameworks that included guidance to facilitate both evaluation and planning, but were not specifically described as 'planning frameworks', e.g. (118) are not grouped separately.

Mapping frameworks against evaluation components

Frameworks were mapped against seven process and four outcome evaluation components (i.e. describing programme context, using theory of change, logic models, reach, implementation, maintenance, any other process measures, behaviour, health, non-health and unanticipated outcomes), as well as against the eight components of study design and reporting (see Table 2-2). Tables 2-4 to 2-9 provide an overview of the mapping. Describing programme context, theory of change, and logic models are crucial to informing process and outcome evaluation, we therefore included these alongside process evaluation components in Table 2-4 to 2-6. The mapping enabled us to develop an overview of the guidance included in each of the frameworks and

appraise their applicability to different evaluation objectives and to physical activity and/or dietary change programmes.

Many frameworks mentioned components without any further details (shaded in light grey in the tables), whilst others provided detailed descriptions of how the components may be broken down or evaluated (shaded in dark grey in the tables). For ease of navigation, the frameworks in Tables 2-4 to 2-9 are grouped and listed in the same order as in the typology (Figures 2-2 to 2-4). Most frameworks included guidance on a range of both process and outcome evaluation components. Eleven frameworks did not provide any guidance on outcome evaluation and were specific to process evaluation e.g. (131, 148-150). Frameworks intended to facilitate evaluation of specific programme elements focused on a narrower range of components that aligned with their stated purpose (118, 130, 151, 153, 154).

Table 2-4 Frameworks intended for use in physical activity, dietary change or behaviour change programmes mapped against process and outcome evaluation components

	Process evaluation							Outcome evaluation			
Framework Short Name	Describing Context	Theory of change	Logic model	Reach	Implementation	Sustainability	Other Process Measures	Behavioural Outcomes	Health Outcomes	Non-Health Outcomes	Unanticipated Outcomes
Evaluating Sport and Physical Activity (169)											
Physical Activity Evaluation Handbook (172)											
Sport England Evaluation Framework (173)											
SEF for Physical Activity (93)											
GENIE (125, 126)											
SEF for Dietary Interventions (94)											
Empowerment Framework in Nutrition (156)											
Centre TRT's Framework (175)											
PHE Guide (178)											
GPAT (132)											
OPEN Tool (161)											
SEF for Weight Management (95)											
NICE Guidance: Behaviour Change (166)											

Notes Tables 2-4 to 2-9: Light grey shading indicates the component is mentioned, dark grey shading indicates more detailed guidance on how to break down or evaluate the component

Table 2-5 Frameworks intended for use in health promotion or public health programmes mapped against process and outcome evaluation components

	Process evaluation							Outcome evaluation			
Framework Short Name	Describing Context	Theory of change	Logic model	Reach	Implementation	Sustainability	Other Process Measures	Behavioural Outcomes	Health Outcomes	Non-Health Outcomes	Unanticipated Outcomes
Cross-site Evaluation Tool (155)											
Evaluating Complex Community-Based HP (157)											
Evaluation of Health Education (133)											
Evaluation of Healthy Community Initiatives (134)											
Health Workers Guide (135)											
Kellogg Foundation Evaluation Handbook (165)											
MRC Complex Intervention Guidance (35, 121)											
MRC Natural Experiments (127, 128)											
Setting Standards (103)											
WHO Recommendations (19)											
CDC Framework (92,120)											
Framework for Community Health (168)											
Evaluation in Health Promotion (145)											
Evaluation Works (177)											
Formative Model of Service Evaluation (146)											
Generic Evaluation Toolkit (158)											
LEAP (171)											
MMIPP (139)											
Ontario Evaluation Workbook (180)											
Planning and Evaluation Model (147)											
Stages of Evaluation Model (29, 140)											
Victoria Govt DoH Framework (181)											
California Healthy Cities Framework (142)											
Getting To Outcomes [GTO] 152)											
HEBS Framework (46)											
Multilevel Framework (160)											
Principles for Evaluating Community HP (141)											
RE-AIM (47)											
MRC Process Evaluation Guidance (59, 122)											
Process Evaluation for Public Health (148)											
Six Step Guide to Process Evaluation (149)											
Systematic Evaluation Multiple Components (60)											
Concepts in Process Evaluation (150)											
Process Evaluation in Groups Settings (162)											
Process Evaluation Cluster-Randomised Trials (163)											
Framework for Outcome Assessment (137)											
Community Initiative Evaluation Model (143, 144)											

Table 2-5 Frameworks intended for use in health promotion or public health programmes mapped against process and outcome evaluation components

	Process evaluation							Outcome evaluation			
Framework Short Name	Describing Context	Theory of change	Logic model	Reach	Implementation	Sustainability	Other Process Measures	Behavioural Outcomes	Health Outcomes	Non-Health Outcomes	Unanticipated Outcomes
Contextual Factors Framework (118)	■	■			▒	▒	▒				
Co-ordinated Action Checklist (159)							▒				
Evaluating Legacy (151)	▒	▒	▒		▒	■	▒	▒	▒	■	
Participation, Partnerships and Equity (154)		▒		■	■		▒		■	■	
Supportive Social Environments (164)	■	▒	▒	▒	■	▒	▒			▒	
Participative Framework Health Inequalities (129)	■								■	▒	
Three Dimensional Health Cube (130)				▒	▒	■			▒		
Well Connected (131)	■				■						
Intervention Mapping (115, 138)	■	■	■	▒	■	■	▒	▒	▒	▒	
PRECEDE-PROCEED (114, 119)	■	■	■		■	▒		▒	■	▒	
Settings for Health Promotion (117)	■	■	■							▒	▒

Table 2-6 Frameworks intended for use in generic programmes mapped against process and outcome evaluation components

	Process evaluation							Outcome evaluation			
Framework Short Name	Describing Context	Theory of change	Logic model	Reach	Implementation	Sustainability	Other Process Measures	Behavioural Outcomes	Health Outcomes	Non-Health Outcomes	Unanticipated Outcomes
Realistic Evaluation (56)	■	■									
Utilization-Focused Evaluation (136)		■			■						▒
Better Evaluation (174)	■	■	■								▒
Community Toolbox (176)	■	■	■	▒	■		▒	▒	▒	▒	▒
Evaluating Community Projects (167)		▒		▒		▒					▒
Health Planners Toolkit (170)	■	■	■	▒	▒			▒	▒		▒
Impact Pathway Analysis (123, 124, 182)	■	■	■	■	▒		▒	▒	▒	▒	▒
Magenta Book (179)	■	■	■	■	■			■	■	■	▒
Levels of Coalition Evaluation (153)				■	■	▒			▒	▒	
Logic Model Development Guide (116)	■	■	■	■	■			■	▒		

Process evaluation components

Guidance on the key components of process evaluation were included in most frameworks, e.g. describing contextual factors of programmes, identifying and describing causal mechanisms or theories of change, reach and implementation. The frameworks providing the most comprehensive and detailed guidance on these components include the MRC guidance on process evaluation of complex interventions (15), Center of Excellence for Training and Research Translation (Center TRT) Framework (175), Victoria Government Department of Health (DoH) Evaluation Framework (181), the Physical Activity Evaluation Handbook (172) and the Standard Evaluation Frameworks (SEFs) (93-95). Other process evaluation components were included within fewer frameworks. For example, guidance on evaluation of sustainability was limited, with only thirteen frameworks providing more details of how to evaluate it, e.g. (130, 151). A small number of frameworks mentioned other process components such as adaptation, exposure, capacities, training, partnerships, satisfaction, and community changes; however, details of how to evaluate these components were limited. Over half the frameworks identified logic models as a useful tool in programme planning and evaluation. Several of these provide more detailed information, examples and/or templates to support the development of logic models (15, 116, 165, 178).

Table 2-7 Frameworks intended for use in evaluating physical activity, dietary change or behaviour change programmes mapped against study design, evaluation approach and reporting components

	Evaluation Approach/study design							
Framework Short Name	Stakeholder involvement	Participatory evaluation	Linked to stage of programme	At different time points	Study Design	Data Collection Methods	Data Analysis	Dissemination and reporting
Evaluating Sport and Physical Activity (169)	■		■	■	■	■	■	■
Physical Activity Evaluation Handbook (172)	■		■	■	■	■	■	■
Sport England Evaluation Framework (173)	■			■	■	■	■	■
SEF for Physical Activity (93)	■		■	■	■	■	■	■
GENIE (125, 126)				■			■	
SEF for Dietary Interventions (94)	■		■	■	■	■	■	■
Empowerment Framework in Nutrition (156)	■	■	■	■	■	■		
Centre TRT's Framework (175)	■		■	■				■
PHE Guide (178)	■		■	■	■	■	■	■
GPAT (132)	■			■		■		
OPEN Tool (161)	■					■		■
SEF for Weight Management (95)	■		■	■	■	■	■	■
NICE Guidance: Behaviour Change (166)	■		■		■	■	■	

Table 2-8 Frameworks intended for use in evaluating health promotion or public health programmes mapped against study design, evaluation approach and reporting components

	Evaluation Approach/study design							
Framework Short Name	Stakeholder involvement	Participatory evaluation	Linked to stage of programme	At different time points	Study Design	Data Collection Methods	Data Analysis	Dissemination and reporting
Cross-site Evaluation Tool (155)								
Evaluating Complex Community-Based HP (157)								
Evaluation of Health Education (133)								
Evaluation of Healthy Community Initiatives (134)								
Health Workers Guide (135)								
Kellogg Foundation Evaluation Handbook (165)								
MRC Complex Intervention Guidance (35, 121)								
MRC Natural Experiments (127, 128)								
Setting Standards (103)								
WHO Recommendations (19)								
CDC Framework (92,120)								
Framework for Community Health (168)								
Evaluation in Health Promotion (145)								
Evaluation Works (177)								
Formative Model of Service Evaluation (146)								
Generic Evaluation Toolkit (158)								
LEAP (171)								
MMIPP (139)								
Ontario Evaluation Workbook (180)								
Planning and Evaluation Model (147)								
Stages of Evaluation Model (29, 140)								
Victoria Govt DoH Framework (181)								
California Healthy Cities Framework (142)								
Getting To Outcomes [GTO] (152)								
HEBS Framework (46)								
Multilevel Framework (160)								
Principles for Evaluating Community HP (141)								
RE-AIM (47)								
MRC Process Evaluation Guidance (59, 122)								
Process Evaluation for Public Health (148)								
Six Step Guide to Process Evaluation (149)								
Systematic Evaluation Multiple Components (60)								
Concepts in Process Evaluation (150)								
Process Evaluation in Groups Settings (162)								
Process Evaluation Cluster-Randomised Trials (163)								
Framework for Outcome Assessment (137)								

Table 2-8 Frameworks intended for use in evaluating health promotion or public health programmes mapped against study design, evaluation approach and reporting components

	Evaluation Approach/study design							
Framework Short Name	Stakeholder involvement	Participatory evaluation	Linked to stage of programme	At different time points	Study Design	Data Collection Methods	Data Analysis	Dissemination and reporting
Community Initiative Evaluation Model (143, 144)								
Contextual Factors Framework (118)								
Co-ordinated Action Checklist (159)								
Evaluating Legacy (151)								
Participation, Partnerships and Equity (154)								
Supportive Social Environments (164)								
Participative Framework Health Inequalities (129)								
Three Dimensional Health Cube (130)								
Well Connected (131)								
Intervention Mapping (115, 138)								
PRECEDE-PROCEED (114, 119)								
Settings for Health Promotion (117)								

Table 2-9 Frameworks intended for use in evaluating generic programmes mapped against study design, evaluation approach and reporting components

	Evaluation Approach/study design							
Framework Short Name	Stakeholder involvement	Participatory evaluation	Linked to stage of programme	At different time points	Study Design	Data Collection Methods	Data Analysis	Dissemination and reporting
Realistic Evaluation (56)								
Utilization-Focused Evaluation (136)								
Better Evaluation (174)								
Community Toolbox (176)								
Evaluating Community Projects (167)								
Health Planners Toolkit (170)								
Impact Pathway Analysis (123, 124, 182)								
Magenta Book (179)								
Levels of Coalition Evaluation (153)								
Logic Model Development Guide (116)								

Outcome evaluation components

Guidance on outcome evaluation components was more variable than for process evaluation components. Frameworks designed for use in physical activity and/or dietary change related programmes provided more detailed information on evaluation of behavioural and health outcomes than the more generic evaluation frameworks. Evaluation of non-health outcomes was typically only mentioned briefly in the frameworks, with only seven providing any level of detail (142, 143, 151, 154, 172, 179, 181). Only about one third of the frameworks mentioned evaluation of unanticipated outcomes, and none provided further information on how to evaluate them.

Study design components

Tables 2-7 to 2-9 show the frameworks mapped against components related to study design, including evaluation at different time points, stakeholder involvement, participatory approaches, data collection and analysis, and reporting of findings. Most frameworks identified the importance of stakeholder involvement and/or participatory evaluation approaches. Few provided information on how to incorporate this, with a few exceptions that did provide detailed guidance on participatory evaluation methods (123, 124, 129, 143, 144).

Most frameworks mentioned the importance of conducting evaluation that is appropriate to a programme's stage of development, and many were presented as a set of steps aligned to stages of programme development and implementation. Most also mentioned evaluation at different time points (i.e. baseline and follow-up), mainly in relation to outcome measures only. Several frameworks used the terms formative and summative evaluation but gave limited information on how they were defining them, or how to do these types of evaluation. Exceptions to this were frameworks that gave a more detailed explanation of the role of formative and pilot studies in developing an intervention (29, 121).

Guidance on data collection and data analysis was highly variable. Several frameworks provided explanations of appropriate use of experimental designs and quantitative and qualitative methods (35, 56, 114, 150). Others provided more detailed guidance on specific data collection methods and measures (29, 60, 93-95, 147, 170, 180). Only thirteen frameworks provided information to guide data analysis. There was more consistency in the inclusion of guidance on data collection and analysis within the frameworks described as specific to physical activity and/or dietary change programmes than in the other categories of frameworks.

Finally, guidance on dissemination and reporting also varied. Many frameworks mentioned the importance of this aspect within the cycle of evidence-based practice, but few provided information about where and how to report findings to different target audiences.

Discussion

Our scoping review identified 71 evaluation frameworks, considerably more than previous reviews of evaluation frameworks within the field of public health (63, 69, 100). The broad search strategy we applied enabled us to identify frameworks developed within a range of domains that we could add to those included in these earlier reviews. The focused set of inclusion and exclusion criteria we then applied meant that we only included frameworks specific to or generalisable to physical activity and/or dietary change programmes. In addition to the 12 frameworks specifically intended for physical activity and/or dietary change programme evaluation, we identified a further 59 intended for public health, health promotion, behaviour change or generic programmes that were applicable to physical activity and/or dietary change programmes.

Our review has highlighted the plethora of frameworks available; previous reviews (63) reported this as a potential challenge to practitioners and evaluators navigating and making use of the available guidance. Our review also highlighted the variability in terms used by authors to describe the purpose and scope of the frameworks. Although we identified a growing number of frameworks developed by and for practitioners, e.g. (172, 173, 176, 177, 181), in many frameworks the intended audience was unclear. Terms used to describe programme types were poorly defined and were often used interchangeably. Some phrases such as 'natural experiment' and 'real-world' were used to refer to the evaluation approach and the intervention itself, whilst others (e.g. behaviour change and sustainability) were used to refer to both intervention processes and outcomes. Several frameworks which stated they were intended to support both programme planning and evaluation provided insufficient details about how these facilitated evaluation. The lack of clarity in the extent to which frameworks are intended to be used by researcher-led or practitioner-led evaluation, and in their applicability to different programmes and evaluation objectives, has implications for those using the available guidance. There needs to be a greater consensus of how terms are defined within public health evaluation. An agreed common language would enable those involved in programme evaluation to understand more clearly the applicability of the different frameworks and would help this research area to move forward.

Our typology and mapping resolves some of that complexity in purpose and scope of frameworks by signposting to relevant frameworks and by developing an overview of what guidance is encompassed within each. Our appraisal of frameworks has highlighted areas of overlap, strengths and limitations in the guidance available to support programme evaluation. For example, the inclusion of key process evaluation components (e.g. describing programme contexts and causal mechanisms, reach, and use of logic models) in most frameworks reflects the growing understanding of the importance of these aspects of evaluation to facilitate a more

detailed understanding of whether and how a programme works (11, 12, 29, 59, 103). These components represent strengths within the existing guidance, and areas where there is already an abundance of guidance.

The mapping process and appraisal also identified components where more guidance would be beneficial. We found limited guidance on participatory approaches, non-health and unanticipated outcomes, and wider programme components (e.g. resources, training, delivery, adaptation, partnerships, organisational structures), and sustainability. These components represent aspects of evaluation that require further development of guidance. Stakeholder involvement or participatory evaluation was mentioned in all but nine of the frameworks, reflecting the growing recognition of the importance of stakeholder engagement in evaluation decisions and processes (103, 157). However, detailed guidance on how to incorporate participatory evaluation methods was only provided by seven frameworks (103, 123, 131, 139, 143, 148, 154), and represents another area where further development of guidance would be beneficial. Compared to other categories within the typology, frameworks specific to physical activity programmes more consistently provided guidance on evaluation of health and behavioural outcomes, including the use of appropriate data collection and analysis methods. By their nature these components are specific and therefore may be difficult to define within more generic frameworks. Frameworks developed to facilitate evaluation of specific programme elements, such as sustainability (130, 151), and those intended to facilitate evaluation of partnerships (153, 154, 164) or community (143, 144, 154) also addressed some of the gaps within the more generic frameworks.

Our mapping and typology signpost to frameworks where guidance on specific components can be found. Although availability does not necessarily equate to accessibility or usability of information, the mapping of frameworks can be used to help understand some of the strengths and limitations within the guidance provided. Further investigation of whether and how frameworks have been used may provide insight into how fit for purpose they are, and the benefits and challenges of applying them within physical activity or dietary change programme evaluation. Furthermore, the typology and mapping can be used by practitioners, commissioners and evaluators of physical activity and/or dietary change programmes to identify frameworks relevant to their evaluation needs. They can also be used by researchers and those interested in developing evaluation guidance to identify evaluation components where it would be most useful to focus their efforts, rather than developing more guidance for components where there is already an abundance of guidance. Our categorisation could also be used by researchers publishing frameworks to more clearly report how these are intended to be used, and for those reporting evaluation studies to more clearly state how they have been used.

Strengths and limitations

Our broad search strategy enabled a comprehensive review which identified 71 frameworks within the academic and grey literature. By drawing on frameworks developed within different domains, we have added to previous reviews (63, 69) to map a wide range of evaluation frameworks applicable to physical activity and/or dietary change programmes.

Our scoping review methods, which included consultation with experts, helped to maximise the chances of identifying relevant frameworks, and of applying relevant components which were based on consensus to appraise the frameworks. It was not our intention to apply a formal consensus building method, however we recognise that the use of a more formalised process would be an alternative approach. By consulting both practice and research-based experts we are confident that the results will be of interest and value to both practitioners and researchers concerned with evaluation of physical activity and/or dietary change programmes.

There are limitations of the review. The review only included sources published in the English language. The heterogeneity and ambiguity in use of terminology was a methodological challenge during screening, data extraction and synthesis. Frameworks intended to support specialist evaluation aspects such as health economic evaluation and evaluation of programmes using digital technologies (e.g., mobile health) are critical to practice and policy decisions, however we excluded these frameworks due to their specificity and also due to the large number available. A separate review of the available guidance to support these specialist evaluation aspects would be beneficial.

Conclusion

We have added to previous reviews of evaluation frameworks and identified 71 frameworks applicable to physical activity and/or dietary change programme evaluation. There is an abundance of frameworks available to support programme evaluation. Our typology and mapping signpost to frameworks where guidance on specific components can be found, where there is overlap in their scope and content, and where there are gaps in the guidance. Practitioners and evaluators can use the typology and mapping to identify, agree upon and apply appropriate frameworks. Researchers who develop evaluation guidance can use them to identify evaluation components for which there are gaps in available guidance. This should help focus research efforts where it is most needed and promote uptake and use of appropriate evaluation frameworks in practice to improve the quality of evaluation and reporting. To gain a better understanding of the usability and applicability of frameworks to physical activity intervention evaluation Chapter 3 explores their use and reporting within published physical activity evaluation studies.

Chapter 3. A systematic review of the use and reporting of evaluation frameworks in physical activity evaluation studies

Introduction

Chapter 2 highlighted an abundance of available frameworks that could be used to support a systematic and robust evaluation of physical activity and dietary change interventions. The broad search strategy applied in Chapter 2 enabled an extensive number of frameworks that had been developed in a range of domains to be identified; this included just three that were specific to evaluation of dietary change and four to physical activity interventions. The remainder were intended for use in public health or generic interventions. The search also revealed a plethora of evaluation studies. For pragmatic reasons this informed the narrowing of the focus to physical activity interventions in the systematic review, and in the rest of the thesis. The search strategy for the systematic review reported in Chapter 3 was therefore designed to reflect this. Firstly, search terms were applied to identify evaluation studies of physical activity interventions only. Secondly, the names of frameworks identified in the scoping review as relevant to physical activity (Figure 2-2), health promotion or public health (Figure 2-3), and generic programmes (Figure 2-4), but not those specific to dietary change interventions (as shown in Figure 2-2), were applied as search terms.

Chapter 2 highlighted variability in terms used by authors to describe the purpose and scope of the frameworks. For example, in many frameworks it was not clear if the framework was intended to be used in researcher-led or practitioner-led evaluation. Despite the availability of guidance, this does not necessarily translate into uptake and use. Questions remain as to whether criticisms regarding limitations in the quality of evaluation studies result from ongoing limitations in uptake and use of evaluation frameworks, or if the expansion of the guidance developed has prompted greater use of frameworks and an improved quality of evaluation in more recent studies. In effect, do the criticisms and calls for better reporting of evaluation studies pre-date the development and expansion of available guidance, or is there still limited uptake and use made of evaluation guidance? There are also questions relating to how effective the frameworks are in improving the quality of evaluation studies, and whether any limitations in use of the frameworks relate to limitations in how fit for purpose the guidance is. Chapter 3 addresses these questions and presents a systematic review of the use and the quality of reporting of evaluation frameworks within evaluation studies of physical activity interventions.

Background

Increasing physical activity levels among the population is a public health priority (1, 2, 40). Yet the diversity of individual, environmental and societal influences on physical activity requires

interventions that reflect that diversity (40). This has led to various interventions targeting physical activity behaviour that are delivered to different populations and across many settings by a range of public, private and voluntary providers, many of which are multi-sectoral and multi-component. The complexity and heterogeneity in interventions poses challenges to understanding their effectiveness, and to generalising from one intervention to another (183, 184). Given the high rates of inactivity (4, 7) and the importance of physical activity for health (3), it is vital that we learn from the interventions delivered about what works, for whom, and in what contexts (8).

Over the past 20 years, there has been a growing appreciation of the importance of evaluation to inform evidence-based interventions to support population-wide changes in physical activity and to justify policy and practice (8, 47, 48). Evaluation can be defined as the "*systematic examination and assessment of the features of an initiative and its effects, in order to produce information that can be used by those who have an interest in its improvement or effectiveness*" (19), p3. Translation from one setting to another, and wider scale adoption of effective interventions, requires both rigorous evaluation and robust reporting of evaluations to build the evidence base (48, 70).

Several frameworks and guidance documents have been developed to facilitate the evaluation and reporting of intervention studies in public health. In this review the term 'evaluation framework' is used to include any structured guidance which facilitates a systematic evaluation of the implementation or outcomes of an intervention. A recent scoping review that we conducted identified 68 evaluation frameworks that could be used to guide evaluation of physical activity interventions (85). This included frameworks intended to support evaluation of physical activity interventions specifically (e.g. The Standard Evaluation Framework (SEF) for Physical Activity Interventions (93)), as well as frameworks intended to guide development and evaluation of various public health interventions, such as RE-AIM (47), and the Medical Research Council (MRC) guidance on the development and evaluation of complex interventions (35). We have included more general guidance, such as Logic Models (116), where these provide information or a structure to facilitate a systematic approach to identifying and reporting intervention objectives, activities and outcomes. Several checklists have also been developed to improve the completeness of reporting and quality of intervention descriptions; for example the STROBE Statement for Reporting Observational studies in Epidemiology (185, 186) and the Template for Intervention Description and Replication (TIDieR) (187). Further, the Behaviour Change Wheel (38) and the Behaviour Change Technique (BCT) Taxonomy V1 (188) provide a framework to facilitate intervention development, that can also be applied to help standardise how the content of behaviour change interventions are specified. Despite the publication of these frameworks and

guidance, there is a lack of evidence about whether frameworks are being used to guide evaluation.

There has been continued calls for better evaluation and reporting within public health (12, 88). In particular, the need for more detailed descriptions of intervention components and contextual factors to help evaluate how, why and in what contexts interventions may be effective, and to allow implementation of good practice (38, 41). Many of the frameworks and guidance have sought to address this and provide guidance on process evaluation and contextual factors. However, questions remain regarding if and how these frameworks are used within evaluation studies.

Two previous reviews have focused specifically on the use of RE-AIM (189) and the SEF for physical activity interventions (88). These reviews concluded that the reporting of framework components was inconsistent, and that details related to participants, recruitment and broader effects were particularly poorly reported, despite these being components of the frameworks used. Both reviews also highlighted a need for greater clarity in the reporting of how frameworks have been used. Heterogeneity in the format and guidance provided by frameworks may lead to heterogeneity in the way they are applied. This creates difficulties for those interested in further development of evaluation guidance, and those interested in understanding and comparing the effectiveness of interventions including reviewers of evaluation studies and practitioners or researchers wishing to implement or further develop interventions. This limits the contribution evaluation studies make to the evidence base. Given the extensive number of evaluation frameworks, a better understanding of current practices in the use and reporting of them is needed so that future recommendations related to the use of frameworks and evaluation can be developed appropriately.

The aim of this review was therefore to assess the use of evaluation frameworks and the quality of reporting of how they were used within evaluations of physical activity interventions. The primary objective was to explore whether evaluation frameworks are reported to have been used within evaluation studies of physical activity interventions, and which frameworks have been used. The second objective was to appraise the quality of reporting with regards to how evaluation framework use has been reported. Previous reviews (88, 189) have assessed use of a single evaluation framework against the criteria specified in that framework. To our knowledge, no previous review has developed a set of generic indicators to facilitate the appraisal of the use of multiple evaluation frameworks in reported studies. We therefore developed and applied a set of indicators that would enable a critical appraisal of the use and reporting of different evaluation frameworks in evaluation studies.

Methods

Protocol and registration

Search methods and inclusion criteria were specified in advance and registered on PROSPERO (CRD42018089472). We applied the PRISMA statement for reporting items for systematic reviews (82).

Search strategy

We searched Scopus, CINAHL, and EMBASE for published evaluation studies of physical activity interventions. We used free search terms and MeSH terms relating to evaluation, e.g. program* evaluation, programme effectiveness, process evaluation and outcome evaluation. We also included names of specific evaluation frameworks that we had identified in our scoping review of evaluation frameworks (85), to minimize the risk of missing frameworks that do not include the term evaluation in their title (e.g. RE-AIM). These terms were then combined with terms relating to physical activity behaviours (e.g. physical activity, sport, exercise, sedentary). Table 3-1 provides the full electronic search strategy for CINAHL. The context of this review was to understand current practice and use of frameworks in evaluation studies of physical activity programmes. Therefore, the search was limited to studies published between 2015 and the date of the search (25th March 2019). Only studies published in the English language were included.

All studies identified from the searches were downloaded into the Endnote reference manager and duplicates were removed. Screening of all studies was completed by the lead author. At each stage of the screening process (title, abstract and full paper) a sample of twenty percent of studies were checked and validated independently by a second author (JM). Disagreements were resolved through discussion.

Table 3-1 Search strategy applied in CINAHL database

	Search applied in CINAHL
1	*TITLE-ABS-SUBJECT ("program* evaluation") Published Date: 20150101-20191231*
2	*TITLE-ABS-SUBJECT ("service evaluation") Published Date: 20150101-20191231*
3	*TITLE-ABS-SUBJECT ("process evaluation") Published Date: 20150101-20191231*
4	*TITLE-ABS-SUBJECT ("implementation evaluation") Published Date: 20150101-20191231*
5	*TITLE-ABS-SUBJECT ("program* effectiveness") Published Date: 20150101-20191231*
6	*TITLE-ABS-SUBJECT ("outcome evaluation") Published Date: 20150101-20191231*
7	*TITLE-ABS-SUBJECT ("re-aim") Published Date: 20150101-20191231*
8	*TITLE-ABS-SUBJECT ("standard evaluation framework") Published Date: 20150101-20191231*
9	*TITLE-ABS-SUBJECT ("intervention mapping") Published Date: 20150101-20191231*
10	*TITLE-ABS-SUBJECT ("program impact pathway") Published Date: 20150101-20191231*
11	*TITLE-ABS-SUBJECT ("process evaluation of complex interventions") Published Date: 20150101-20191231*
12	*TITLE-ABS-SUBJECT ("developing and evaluating complex interventions") Published Date: 20150101-20191231*
13	*TITLE-ABS-SUBJECT ("framework for program evaluation in public health") Published Date: 20150101-20191231*
14	*TITLE-ABS-SUBJECT ("logic model") Published Date: 20150101-20191231*
15	*1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14*
16	*TITLE-ABS-SUBJECT ("physical activity") Published Date: 20150101-20191231*
17	*TITLE (exercise) Published Date: 20150101-20191231*
18	*TITLE (MH "exercise") Published Date: 20150101-20191231*
19	*TITLE-ABS-SUBJECT (sedentary) Published Date: 20150101-20191231*
20	*TITLE-ABS-SUBJECT (sport*) Published Date: 20150101-20191231*
21	*TITLE-ABS-SUBJECT (inactiv*) Published Date: 20150101-20191231*
22	*TITLE-ABS-SUBJECT (fitness) Published Date: 20150101-20191231*
23	*16 OR 17 OR 18 OR 19 OR 20 OR 21 OR 22*
24	*15 AND 23*

Study selection

Inclusion and exclusion criteria were defined *a priori* and applied to all papers (see Table 3-2 for full details). Our interest was in evaluation studies, therefore other articles including conceptual papers, reviews, and research protocols were excluded. To assess the use, and any limitations in the use, of evaluation frameworks across the full range of physical activity interventions we screened the papers to identify studies where increasing physical activity was the stated primary goal, irrespective of whether they reported the use of specified frameworks. We included evaluation studies of any physical activity intervention delivered in any individual, group or population setting (e.g. health care, schools, and geographical areas). We included studies of interventions delivered to the general population as well as to participants diagnosed with a disease (e.g. heart disease, diabetes) or as having one or more disease risk factors (e.g. inactive, obese). We then screened these to identify those studies that had referred to an evaluation framework, and to exclude those that had not mentioned one. We screened the reference lists of the included studies to identify any companion papers, for example, where process and outcome evaluations were reported separately.

Table 3-2 Inclusion and exclusion criteria for the systematic review

Included	Excluded
Published evaluation studies including real-world or service evaluations, randomised control trials, observational and natural experiments, feasibility and pilot studies, outcome and process evaluations, quasi-experimental, pre-post designs, effectiveness and impact studies. All types of evaluations using quantitative and/or qualitative methods will be included, whether they have used specified frameworks or not.	Commentaries or discussion papers, conceptual papers, published extracts, books, editorials, systematic reviews, clinical case-reports, research protocols and reported programme designs.
Reported evaluation studies of programmes that have increasing physical activity as the primary stated goal of the programme, including reduced sitting time or sedentary behaviour.	Reported evaluation studies of programmes that have other health behaviours as the primary stated goal of the programme, e.g. smoking, alcohol, substance abuse, eating disorder behaviours. Reported evaluation studies that state other behavioural outcomes or clinical measures as the primary goal of the programme, e.g. programmes aimed at weight loss, maintaining a healthy weight, prevention or management of diabetes, prevention of stroke or heart attack, improvement of aerobic or cognitive function, reduction of fall, improvement of physical performance/function through physical activity or exercise.
Evaluations of programmes that align with approaches to behaviour change, i.e. programmes that correspond to any of the nine intervention functions on the Behaviour Change Wheel (education, persuasion, incentivisation, coercion, training, enablement, modelling, environmental restructuring and restrictions)(38).	Evaluations of programmes that do not correspond to any of the nine intervention functions on the Behaviour Change Wheel (education, persuasion, incentivisation, coercion, training, enablement, modelling, environmental restructuring and restrictions).
Studies that referred to one or more evaluation frameworks.	Studies that did not refer to any evaluation framework.

Data extraction

To address the first objective, we extracted the names of any evaluation frameworks that had been reported as being used in any of the studies. For reporting purposes, we also noted the number of physical activity evaluation studies in which no framework was mentioned. To address the second objective, we extracted data from studies that reported the use of one or more evaluation frameworks. Criteria for data extraction were identified and agreed by all authors *a priori.* Data extraction was completed using a data extraction table.

To assess the context and circumstances in which evaluation frameworks had been used, we extracted data related to study characteristics. So that this review met PRISMA recommendations

for the reporting of systematic reviews (82) we used PRISMA guidelines to inform the data we extracted from the studies. In addition we used STROBE for the reporting of observational studies and natural experiments (186), and the TIDieR checklist (187) to guide our data extraction. We extracted data related to study population, intervention setting and components, study design, and process and outcome measures. To help us to characterise the intervention types we extracted data related to the nine intervention functions of the Behaviour Change Wheel, and the activities delivered, where these were explicitly reported. Intervention functions are broad categories to define the general means by which an intervention might change behaviour (e.g. education, enablement, and incentivisation) (38, 190). Their use in intervention development and reporting is intended to facilitate clearer descriptions of intervention components (38). This is essential for evaluation and implementation (41). We applied the nine intervention functions to guide a systematic approach to identify and report study characteristics.

To assess the quality of reporting of the use of the frameworks, we developed a set of data extraction criteria related to how the studies had described a framework and its application. To ensure that we identified a set of indicators that could be applied across any evaluation framework, rather than a specific framework, we used a similar approach to that described by Michie and Prestwich in their coding scheme for assessing the use and reporting of theory in intervention studies (191). We developed a set of indicators that would allow a systematic examination of how the use of a framework had been reported within each study. Each indicator required a yes/no/not sure response and supporting evidence. We adapted their categories and indicators which aligned closely to our own objectives. For example, Category 1 “Reference to underpinning theory” aligned to our objective to identify any “Reference to an evaluation framework”. Within this category we included four indicators that together assessed the extent to which the framework had been referred to and described to enable us to appraise whether or not the evaluation study was explicitly based on or informed by one or more frameworks. For other items, our indicators were more loosely based on those of Michie and Prestwich. Category 2 and 3 included three indicators to assess the extent to which the methods, data collection and outcomes reported were linked to the specified framework’s components. Category 4 included two indicators to assess the extent to which additional information on how the framework had been used is reported. This last category is important, as there may be good justification for reporting on some rather than all of the components in a framework, or adapting how a framework is applied within a specific evaluation study, but without that information it is difficult to appraise its use and reporting. Any one indicator taken in isolation might seem deficient, so the indicators are best considered together within each category and across the full checklist to provide an overall assessment of how use of a framework has been reported. The criteria were

discussed and agreed by all authors. The checklist of categories and indicators is shown in Table 3-3.

Table 3-3 Categories and indicators for assessing the quality of reporting of the use of evaluation frameworks

Category	Data Extraction Indicators (options for responses)
1. Reference to Framework.	1. Is the framework mentioned even if the study is not explicitly based on it? Yes/No/Not sure 2. Does the study refer to 1 or more frameworks? State number 3. Is the framework mentioned in the introduction? Yes/No/Not sure (Plus evidence) 4. Is a description of the framework components provided? Yes/No/Not sure (Plus evidence)
2. How the framework has been used to develop the evaluation methods and data collection. Are relevant components applied?	5. Is the evaluation stated as explicitly based on the framework components? Yes/No/Not sure (Plus evidence from the method of how the framework components have been applied to inform evaluation methods and data sources)
3. How the framework has been applied to the reporting of outcomes.	6. Are the outcome measures discussed in the result/discussion sections linked to the relevant framework components? Yes/No/Not sure (Plus evidence) 7. How many of the framework components are linked to data sources/measures? All the main framework components / At least one, but not all /None of the components are linked to data (Plus evidence)
4. Reporting use of framework fully.	8. Are any details of adaptations in how the framework has been applied provided? Yes/No/Not sure (Plus evidence) 9. Are any details of limitations and strengths in how the framework has been applied or suggestions for how it could be optimised provided? Yes/No/Not sure (Plus evidence)

Data extraction was completed by JF and validated by JM. For the data related to study characteristics, a sample of 20% of studies were checked and validated and any disagreements were resolved through discussion. For our checklist of indicators used to appraise the quality of reporting of framework use, we first tested the indicators by independently extracting data for a small sample of papers and discussed any differences to refine the process and reach a consensus

in how to apply the indicators to extract data. We then independently validated a sample of 20% of studies and calculated the level of agreement as a percentage in order to validate the data extraction process. Any further disagreements were resolved through discussion. We used narrative synthesis to summarise the use and reporting of frameworks within the included studies.

Results

The search identified 1524 studies once duplicates had been removed. The PRISMA diagram for the screening is shown in Figure 3-1. We identified a total of 292 evaluation studies of physical activity interventions. Only 69 (23%) of these mentioned using an evaluation framework. From the reference list of these 69 studies we identified an additional eight companion studies, however none mentioned using an evaluation framework so were not included. Three interventions were reported in more than one of the included studies; therefore the 69 included studies represent 64 different physical activity interventions.

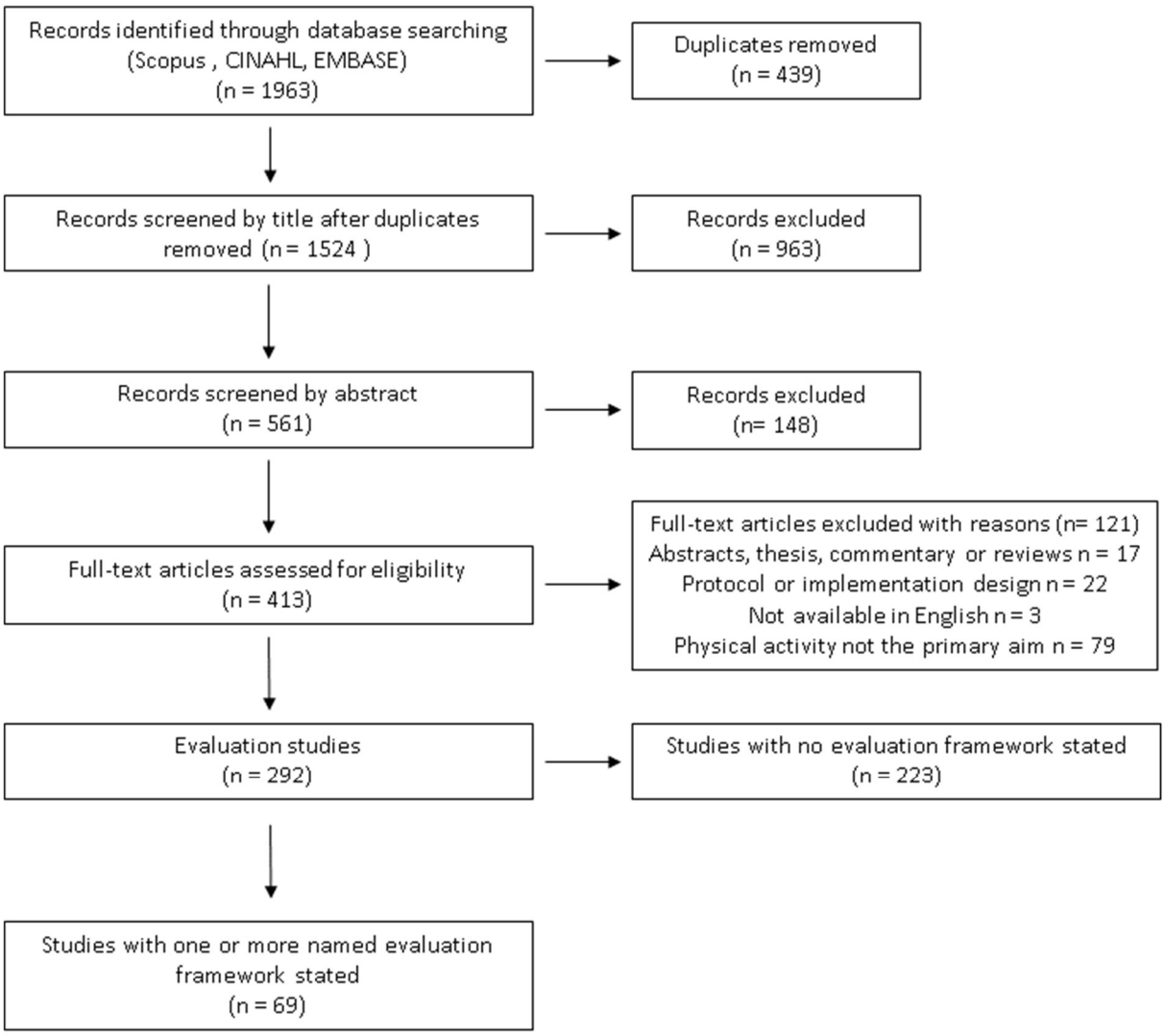

Figure 3-1 PRISMA diagram of screening process for the systematic review

Table 3-4 summarises the evaluation frameworks which were reported as being used and the number of studies using each framework. A total of 16 different evaluation frameworks were identified. These include frameworks that provide guidance on evaluation specifically, such as the Process Evaluation Plan (149), and frameworks that provide guidance on intervention planning and development but that facilitate evaluation and reporting, such as Precede-Proceed (114), Intervention Mapping (115) and Logic Models (116). The frameworks most frequently reported were RE-AIM (47), Saunders and Joshi's process evaluation plan (149) and Steckler and Linnans' process evaluation guidance for public health (148). RE-AIM (47) and the MRC guidance for development and evaluation of complex interventions (35) were the frameworks most frequently reported as being used as a single framework to inform the evaluation study. Realist evaluation (56) was only reported in four studies but was in all cases used as a standalone framework rather than in combination with other frameworks. Fourteen studies reported applying more than one framework (Table 3-6). The frameworks most frequently reported as being used in combination with others were Saunders and Joshi's (149) and Steckler and Linnan's (148) process evaluation frameworks. Both these frameworks provide a similar step-wise approach to process evaluation. The MRC guidance on process evaluation (59) and logic models (116) were also reported in several studies, both as a standalone framework and in combination with other frameworks.

Table 3-4 Evaluation frameworks reported within the 69 studies

Named Framework	Number of studies reporting
RE-AIM (47)	27
Developing a process evaluation plan (149)	12
Process evaluation for public health (148)	10
MRC Guidance on evaluation of complex interventions (35)	8
MRC Guidance on process evaluation (59)	8
Logic Model (116)	7
Realist Evaluation (56)	4
Precede-Proceed (114)	3
Intervention Mapping (115)	2
Outcome Model (140)	2
CDC Framework (120)	1
Evaluation: a Systematic Approach (192)	1
Model of Implementation (193)	1
WHO Process Evaluation Workbook (194)	1
Swiss Model for Outcome Classification (195)	1
Concepts in process evaluation (150)	1

Note: 14 papers referred to more than one of these frameworks informing the evaluation.

Study characteristics

Study characteristics are shown in the supplementary material (Additional File 3.1). The frameworks have been used in a wide range of contexts and circumstances. Most of the criteria used to describe the interventions were clearly specified, and there was good agreement in the sample validated independently. The study population was reported in all studies; 37 studies (54%) reported interventions targeting children or young adults, 24 (35%) targeted adults, and five (7%) targeted older people. The remaining three (4%) studies did not specify an age group but implied the intervention was targeted at multiple population groups or the general public. Relevant details of demographic and/or health status of target populations were also described fully in studies where this was relevant: interventions targeting populations with or at increased risk of diabetes, the metabolic syndrome or heart disease; low socio-economic groups; and women or men only. Details of the included population were reported variously as sample size, participants recruited, or the number of intervention sites. Intervention setting was described in all studies; 28 (40%) were implemented in schools (including pre-schools), 13 (20%) in health care settings, four (6%) in the workplace, and 24 (35%) in other community settings (e.g. youth groups, churches). All studies provided some description of the intervention components (i.e. activities delivered), although the level of detail was variable. For example, most studies described specific activities delivered (e.g. walking, dance, counselling, staff training, online tools), whilst fewer studies provided details of who delivered the intervention, the mode of delivery, the dose, or modifications to the delivery of the intervention. Most studies were multi-component and described several activities delivered together. Training (n=50, 72%), education (n=47, 68%) and enablement (n=42, 61%) were the most frequently reported intervention functions stated in the studies. Studies less frequently reported modelling (n=12, 17%), incentivisation (n=9, 13%), environmental restructuring (n=9, 13%) and persuasion (n=4, 6%).

Additional File 3.1 shows the data we extracted related to the study objectives, study design and outcomes reported. Study designs included quantitative, qualitative and mixed-methods studies, controlled trials, quasi-experimental, case studies and hybrid designs. Thirty-five (51%) studies were described as a process evaluation and 15 (22%) as an outcome evaluation. In addition to physical activity outcomes, a range of secondary outcomes were reported: 52 (75%) reported on various implementation measures e.g. reach, dose, fidelity and maintenance; 14 (20%) reported outcomes related to anthropometric measures; and 15 (22%) reported details of participant demographics. Only nine (13%) studies reported outcome measures related to quality of life and only five (7%) reported on economic or cost analysis.

Appraisal of the quality of reporting on the use of evaluation frameworks

Table 3-5 shows the data extracted on the use and reporting of an evaluation framework for studies referring to a single framework, and Table 3-6 shows the data for studies referring to more than one framework. The level of agreement for the validation of data extracted for these items was 80%. Six studies mentioned a framework but did not state that the evaluation was informed by it. These included one study that provided a logic model but made no reference to this other than in the figure caption (196), and four studies that mentioned the MRC guidance on evaluating complex interventions and one that mentioned the MRC guidance on process evaluation of complex interventions but did not explicitly state that the study was informed by these guidance documents (197-201) (four of these were companion studies relating to the same intervention). In three (4%) further studies the description lacked sufficient clarity to determine whether the study was intended to be based on the reported framework or not; for example these referred to the formulation of a logic model but did not describe the evaluation and outcomes as being based on the logic model (202-204). The remaining 60 (87%) studies all stated that the evaluation was informed by one or more specified framework. However, based on the extracted data on how studies had reported framework components, how these had been applied and how the results linked to the framework components, we identified only 51 (74%) of the studies as being explicitly based on the reported framework.

Forty-four studies (64%) referred to the framework(s) in the introduction, while thirty-six (52%) provided a description of the framework components. Fifty-three (77%) reported outcomes linked to relevant framework components, the remaining sixteen (23%) studies provided no evidence of how the outcomes reported were linked to the framework components. Only 26 (38%) studies provided detailed descriptions consistently across all of the indicators; this included 13 that used RE-AIM, three that used realist evaluation, two that used the MRC guidance on process evaluation, and two that used Saunders and Joshi's process evaluation framework. Four studies (205-208) that had applied frameworks in combination also consistently reported details of the frameworks and their use across all indicators. Twenty-nine studies (42%) described strengths or limitations, whilst only 17 (25%) described adaptations in how the framework had been used.

Table 3-5 Appraisal of use and reporting of an evaluation framework in studies using a single evaluation framework

Framework(s)	Intervention name	First author and publication	Framework stated in introduction	Framework components described	Study stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
CDC Framework (120)	WAVE	Meng (209)	No	No	Yes	Yes	✓		Not sure	No	No
Developing a Process-Evaluation Plan (149)	APAN	Blackford (210)	No	Yes	Yes	Yes		✓	All	Yes	Yes
	Exercise Counselling	McCarthy (211)	Yes	Yes	Yes	Yes		✓	At least one	No	No
	NECaSP	Curry (212)	No	No	Yes	Not sure		✓	Not sure	Yes	Yes
	PACES	Webster (213)	Yes	Yes	Yes	No	✓		Not sure	No	No
	ToyBox-study	De Craemer (214)	Yes	Yes	Yes	Yes		✓	At least one	Yes	Yes
Evaluation: a Systematic Approach (192)	FLEX	Wright (215)	No	Yes	Yes	Yes		✓	Not sure	Yes	No
Logic Model	Girls Active	Harrington (196)	No	No	No	No	✓		Not sure	No	No
	GOTR	Ullrich-French (204)	No	No	Not sure	Yes	✓		Not sure	Yes	No
	Healthy Start	Chow (216)	No	No	Yes	Not sure	✓		At least one	No	Yes
	School–Community Linked PA	Griffiths (217)	No	No	Yes	Yes		✓	At least one	No	Yes
Model of Implementation (193)	MAGNET	Burkart (218)	Yes	Yes	Yes	Yes		✓	Not sure	No	No
MRC Guidance for Development & Evaluation of	Action 3.30	Jago (197)	No	No	No	No	✓		Not sure	No	No
	BGDP	Jago (198)	No	No	No	Yes	✓		At least one	No	No
	BGDP	Sebire (199)	Yes	No	No	No	✓		At least one	No	No

Table 3-5 Appraisal of use and reporting of an evaluation framework in studies using a single evaluation framework

Framework(s)	Intervention name	First author and publication	Framework stated in introduction	Framework components described	Study stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
Complex Interventions (35)	GoActive	Corder (219)	Yes	No	Yes	No	✓		Not sure	No	No
	Movement as Medicine	Avery (220)	Yes	Yes	Yes	Yes		✓	At least one	No	Yes
	STAND	Biddle (221)	Yes	No	No	No	✓		None	No	No
MRC Process Evaluation of Complex Interventions (59)	BGDP	Sebire (201)	Yes	No	Yes	Not sure	✓		At least one	Not sure	Yes
	BGDP	Sebire (222)	No	No	No	No	✓		None	No	No
	LPAW	Lefler (223)	Yes	No	Yes	No	✓		None	No	No
	PACE-UP	Furness (224)	Yes	Yes	Yes	Yes		✓	At least one	No	No
	We Act	Bonde (225)	Yes	Yes	Yes	Yes		✓	All	No	No
Outcome Model (140)	Healingo Fit	Dadaczynski (226)	No	No	Yes	Yes		✓	At least one	No	No
PRECEDE PROCEED (227)	SPACE	Tucker (228)	No	No	Yes	No	✓		Not sure	No	No
Process Evaluation for Public Health (148)	Group fitness	Sofija (229)	No	Yes	Yes	Yes		✓	All	No	No
	PAC	Matthews (230)	Yes	No	Yes	Yes	✓		At least one	No	No
RE-AIM (47)	5-As	Galaviz (231)	No	Yes	Yes	Yes		✓	All	No	No
	ACTIVE	Christian (232)	No	No	Yes	Yes		✓	All	No	No
	CHAM JAM	Reznik (233)	No	Yes	Yes	Yes		✓	All	No	No
	COMMUNICATE	Kamada (183)	No	No	Yes	Yes		✓	At least one	No	No
	Enhance®Fitness	Kohn (234)	Yes	No	Yes	Yes		✓	At least one	No	No

Table 3-5 Appraisal of use and reporting of an evaluation framework in studies using a single evaluation framework

Framework(s)	Intervention name	First author and publication	Framework stated in introduction	Framework components described	Study stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
	Enhance®Fitness	Petrescu-Prahova (235)	Yes	Yes	Yes	Yes		✓	At least one	Yes	Yes
	FAN	Wilcox (236)	Yes	No	Yes	Not sure		✓	At least one	No	No
	FitEx & ALED	Harden (61)	Yes	Yes	Yes	Yes		✓	At least one	No	Yes
	Guided Walking	Baba (237)	Yes	Yes	Yes	Yes		✓	All	No	Yes
	HKOS	Economos (238)	Yes	Yes	Yes	Yes		✓	All	Yes	No
	Healthy Start-Départ Santé	Ward (239)	Yes	Yes	Yes	Yes		✓	All	No	No
	Healthy Together	Jung (240)	Yes	Yes	Yes	Yes		✓	All	No	No
	IMIL	Allar (241)	Yes	Yes	Yes	Yes		✓	At least one	Yes	Yes
	ManUp	Caperchione (242)	Yes	Yes	Yes	Yes		✓	All	No	Yes
	PAFES	Gonzalez-Viana (243)	Yes	Yes	Yes	Yes		✓	All	No	Yes
	Promotora Community Health Program	Schwingel (244)	Yes	Yes	Yes	Yes		✓	All	Yes	Yes
	RCP & ACP	Paez (245)	Yes	Yes	Yes	Yes		✓	All	Yes	Yes
	Sport England funded project	Koorts (246)	Yes	No	Yes	Yes		✓	All	Yes	Yes
	Stair Climbing	Bellicha (247)	No	Yes	Yes	Yes		✓	At least one	Yes	Yes

Table 3-5 Appraisal of use and reporting of an evaluation framework in studies using a single evaluation framework

Framework(s)	Intervention name	First author and publication	Framework stated in introduction	Framework components described	Study stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
	STEPs & LET US Play	Beets (248)	No	No	Yes	Yes		✓	At least one	Not sure	Yes
	STEPs & LET US Play	Beets (249)	No	No	Yes	Yes		✓	At least one	Yes	Yes
	SAGE	Lee (250)	Yes	Yes	Yes	Yes		✓	All	No	Yes
	TAME health	Lewis (251)	Yes	No	Yes	Yes		✓	All	No	Yes
	Walking Works	Adams (252)	Yes	Yes	Yes	Yes		✓	All	No	Yes
Realist Evaluation (56)	CBHEPA	Herens (253)	Yes	Yes	Yes	Yes		✓	All	Not sure	Yes
	Local Authority Sport & PA	Daniels (254)	Yes	Yes	Yes	Yes		✓	All	Yes	Yes
	Local Environment Model	Willis (255)	Yes	Yes	Yes	Yes		✓	All	No	No
	Project SoL	Mikkelsen (256)	No	Yes	Yes	No	✓		Not sure	No	No

Table 3-6 Appraisal of use and reporting of the use of evaluation frameworks in studies using multiple frameworks

No. of frameworks & references	Intervention name	First author & publication	Framework stated in introduction	Framework components described	Stated as based on framework	Outcomes linked to components	Framework mentioned	Explicitly based on framework	No. of components linked	Adaptations described	Limitations described
2 (116, 149)	WWPP	Fournier (257)	Yes	No	Yes	Yes		✓	At least one	No	No
2 (114, 149)	SPACE	Driediger (258)	Yes	No	Yes	Yes		✓	At least one	No	No
2 (148, 149)	IDEFICS	Verloigne (259)	Yes	No	Yes	Not sure	✓		Not Sure	No	Yes
2 (148, 149)	PA for grandparent	Young (205)	Yes	Yes	Yes	Yes		✓	At least one	No	No
2 (148, 149)	It's LiFe!	Verwey (260)	No	Yes	Yes	Yes		✓	At least one	No	Yes
2 (116, 195)	Classes in Motion	Grillich (202)	No	No	Not sure	Yes	✓		At least one	No	No
2 (35, 116)	ENGAGE-HD	Quinn (203)	Yes	No	Not sure	Not sure	✓		Not Sure	No	No
2 (35, 47)	Move for Well-being in School	Smedegaard (206)	Yes	Yes	Yes	Yes		✓	All	Yes	No
2 (59, 148)	WAVES	Griffin (261)	Yes	No	Yes	Yes		✓	At least one	Yes	Yes
2 (148, 194)	Walk Well	Matthews (262)	No	Yes	Yes	Yes		✓	At least one	No	No
3 (47, 148, 149)	BeweegKuur	Berendsen (263)	Yes	Yes	Yes	Yes		✓	All	Yes	Yes
3 (47, 59, 115)	Workplace intervention for Nurses	Torquati (264)	Yes	No	Yes	Yes		✓	All	Yes	Yes
3 (140, 148, 149)	SLIMMER	van Dongen (208)	Yes	Yes	Yes	Yes		✓	All	No	Not sure
5 (114-116, 148, 150)	SHAPES	Saunders (265)	Yes	No	Yes	Yes		✓	At least one	No	Yes

Discussion

The extent to which evaluation frameworks have been used and reported

This is the first systematic review that has attempted to comprehensively assess the use of evaluation frameworks within evaluations of physical activity interventions. We identified 292 evaluation studies of interventions in which physical activity was the primary goal, published between 2015 and the date of our search. Only 69 (23%) of these studies reported using an evaluation framework; within these 16 different frameworks were mentioned. Given that we previously identified 68 published evaluation frameworks that could be used to facilitate evaluation of physical activity interventions (85), our findings highlight that evaluation frameworks are under-used and/or under-reported. Their limited use suggests missed opportunities to apply frameworks to guide evaluation and reporting in intervention studies. For example, despite recommendations in several guidance documents to use logic models to support intervention development and evaluation (29, 59, 93), logic models were only referred to in seven of the studies, and their application was poorly reported. None of the studies reported using any frameworks that have been developed specifically for use in physical activity programme evaluation such as the SEF for physical activity interventions (93). This may be explained by its more limited guidance on process evaluation, given that 51% of the studies were a process evaluation and 75% reported implementation measures. The SEF was developed for use in a UK practice context and may therefore be less likely to be used in a research led intervention than a real-world programme evaluation. Its absence from any of the studies in this review suggests not just a limited use made of it but also highlights the gap between research and practice and the challenges of reporting real-world evaluations in the scientific literature. The more frequent use and reporting of RE-AIM may be because it provides guidance on both outcome and process evaluation components. However, its use may also be influenced by its greater exposure within the literature.

Framework use, choice of framework and the quality of reporting is likely to be influenced by the intervention's context and circumstances in which they are used. Many of the studies (n=35, 51%) were process evaluations and it therefore follows that the most frequently reported frameworks were process evaluation frameworks. However, we found that a range of frameworks were used across different intervention types, contexts and study designs. This suggests that many evaluation frameworks are widely applicable and the decision to use and report a framework is more critical than the choice of which framework to use.

The quality of reporting with regards to how frameworks were used

Our checklist of indicators (Table 3-3) enabled us to appraise the quality of use and reporting of evaluation frameworks. There was considerable variation in the quality of reporting of framework use (Tables 3-5 and 3-6). Whilst some studies did report the framework and how it had been used consistently across all indicators in our checklist, others were less consistent in the quality of reporting and some only mentioned a framework without specifying the details of its use. In some studies, the evaluation was reported as being informed by a framework even where there was little evidence of the evaluation being based on it.

Studies tended to be poorer at describing framework components and adaptations or limitations in how these had been used, whilst links between outcome measures and framework components were more clearly described. For example, those which applied just one or more framework's components, rather than all the components, provided very little explanation or rationale for these adaptations. Publishing constraints can mean that reporting an evaluation study fully requires companion papers or supplementary files (35). However, where this was done, we found that there was often inconsistency in reporting the use of frameworks across the different reported elements e.g. (197-199, 201, 213, 222, 225, 266). More detailed and consistent reporting of the framework components and how these have been applied would help those trying to understand the intervention effectiveness fully.

It is inevitable that some frameworks lend themselves to better quality reporting. For example, studies using RE-AIM and Realist evaluation provided a more consistent report of their use across all indicators. RE-AIM is a structured framework; whilst Realist evaluation is a methodological approach, it too provides a guiding framework to facilitate a systematic evaluation and as such has been referred to as a framework within this paper. Both RE-AIM and Realist evaluation have a clear set of components that are relevant to both process and outcomes; they are therefore applicable to a range of evaluation objectives and can be used to identify appropriate data sources. Many of the studies using RE-AIM provided a full description of the components, an explanation of how these linked to data sources, and used the framework components to structure the reporting of findings. In this way the framework facilitated both a systematic evaluation and consistent reporting. RE-AIM was the most frequently used framework. There is a body of literature on how RE-AIM has been developed and used over time (267), and examples of its application. This may have helped to build a better understanding of how its components are defined and how they can be linked to data sources. Some of the less structured guidance documents, for example the MRC guidance on the development and evaluation of complex interventions (35), were used more loosely as a framework, particularly in studies that used more than one framework in combination. This does not necessarily equate to a poorer quality

evaluation. However, we suggest those studies drawing on several frameworks and general guidance documents would benefit from a more detailed reporting of how these have been used to assist the reader in understanding which intervention components are reported on, and why. Whilst there is variability in the quality of reporting of how frameworks have been used, this review does highlight that evaluation frameworks can, when used appropriately, facilitate a systematic evaluation, and that studies that use a framework can facilitate systematic reporting of the evaluation process and outcomes.

Despite recommendations on the importance of fully reporting contextual factors and intervention components, and guidance within the frameworks to facilitate this (35, 59), our review supports previous review findings (88, 189) that the reporting of intervention components is variable, with wider effects (e.g. quality of life and costs) and wider contextual factors (e.g. dose, intervention modifications) being particularly poorly reported. The Behaviour Change Wheel was developed to characterise intervention types and identify behaviour change techniques as 'active ingredients' to improve the reporting and synthesis of evidence of what works in different populations and settings (38). Yet we found ambiguity in the way in which studies reported intervention functions. It is noteworthy that intervention function was the item where we initially had most disagreement in the data extraction validation process and we would argue that clearer specification, or mapping of intervention functions against behaviour change techniques, would make them more useful in characterising interventions. Poor reporting of intervention components and types limits their comparability and transferability.

If evaluation studies are to contribute to an evidence base on which policymakers, practitioners and researchers can draw to inform the development and implementation of interventions, both the framework and intervention components need to be more clearly defined and documented. Clear, consistent and full reporting of interventions and their evaluation is essential to ensure that critical evidence gets shared and used to develop understanding of causal mechanisms, contextual factors and good practice (41). This is vital to allow resources and efforts to address public health issues, such as increasing physical activity, to be focused on effective and efficient intervention components.

Where frameworks are used, their application to guide the full evaluation process from planning to reporting can improve the quality of reporting of their use. A focus on evaluation at the design and development stages of interventions and a clear understanding of the purpose of the evaluation can help to ensure outcome measures are linked to framework components. However, there is a need to improve understanding of how framework and intervention components are defined. Training and documentation can play a role, but more consistent and precise reporting within the scientific literature is needed. Our set of indicators (see Table 3-3) can be used to guide

the reporting of framework use. Those reporting an evaluation study can apply the indicators as a checklist to provide a clear and consistent description of how framework components have been applied across all stages of the evaluation. Reviewers and journal editors can also play a role in using the checklists available to appraise evaluation reports.

Strengths and limitations

The strengths of this study are that we developed a comprehensive checklist of indicators to appraise the use and reporting of evaluation frameworks, based on a widely accepted coding scheme designed to assess the use and reporting of theory (191). Our checklist and its use as a guide to data extraction was piloted and developed iteratively, and agreed by all authors. This enabled us to review the use and reporting of different frameworks.

Limitations of our study include the fact that some studies may use frameworks or framework components in a way that is implied but not explicitly stated, and we acknowledge that this may have led to underrepresentation of the full use made of evaluation frameworks. A more detailed assessment of evaluation studies against each specific framework's components may have provided greater insight into the limitations or fidelity of use and reporting of frameworks. This was not practical to do within a single review of multiple evaluation frameworks. Extracting details of outcome measures (findings) and intervention characteristics for all physical activity evaluation studies may have enabled a fuller appraisal of the quality of the studies and a comparison between those using and those not using an evaluation framework. This may have provided further insights on the impact of using evaluation frameworks on the quality of the evaluation study, however this was beyond the scope of this review.

Conclusion

Despite the use of evaluation frameworks being advocated to improve the rigour of evaluation studies, frameworks are underused and reported inconsistently in many studies. Applying an evaluation framework to inform both the evaluation and reporting of physical activity intervention studies facilitates a more systematic evaluation study. However, intervention and framework components need to be more precisely and consistently defined and documented to help improve the quality of reporting. Variability in the quality of reporting limits the comparability and transferability of evidence. This means that critical evidence that could be used to inform interventions to support the health of the population is not making it into the public domain. The indicators we developed enabled us to appraise the use and reporting of a range of different evaluation frameworks within evaluations of physical activity interventions. These indicators can be used by those reporting an evaluation to guide them in developing a systematic

evaluation report, and by reviewers and journal editors to appraise evaluation studies that have reported the use of an evaluation framework.

There is a growing appreciation of the value of evaluation of 'real-world' interventions to provide practice-relevant evidence. These interventions are recognised as difficult to evaluate, and use of evaluation frameworks may be particularly useful in facilitating practice-based evaluation. Chapter 4 explored the use of a standard evaluation framework, amongst other evaluation strategies in an applied context.

Chapter 4. Use of evaluation guidance in practice - exploring influences on evaluation practice within a national physical activity programme

Introduction

Chapter 3 showed that appropriate use of an evaluation framework can improve the quality of an evaluation study and facilitate a systematic evaluation report. However, the review also highlighted limited use and reporting of frameworks within published evaluation studies of physical activity interventions, and variability in the quality of reporting of framework use. This suggested that there may have been missed opportunities to apply frameworks to facilitate more robust reporting, and that this likely limits the comparability and transferability of evidence. This means that critical evidence that could be used to inform evidence-based decisions to support and improve the health of the population may not be making it into the public domain.

Chapter 3 was based on published studies, though. In evidence-based public health, practice-relevant evidence is critical to increasing the likelihood that evidence will be taken up and used to inform policy and practice decisions. Practice-based evaluation, in other words evaluation of real-world interventions, can address that challenge. However, the complexity of real-world interventions makes robust and rigorous evaluation difficult. Whilst the use of evaluation frameworks may be particularly important to facilitate systematic evaluation in applied contexts, simply providing guidance may not be sufficient.

Chapters 4 and 5 explored influences on evaluation practice in an applied context by undertaking a case study of a national physical activity programme, Sport England's Get Healthy Get Active (GHGA) programme. Sport England funded 33 physical activity projects through the GHGA programme. Applying a collective case study approach allowed the use of multiple projects and data sources to develop an in-depth, multi-faceted understanding of influences on evaluation and evidence-based practices within real world interventions. Chapter 4 reports on framework analysis of documents and data from stakeholder interviews to compare findings within and across projects and the programme. The findings in Chapter 4 suggested that partnerships were a key influence on evaluation practices and highlighted a need to better understand the relationships between partnership characteristics, processes and practices to support evaluation, dissemination and evaluation use. Chapter 5 therefore explored these themes in more detail by combining network analysis with thematic analysis to develop a conceptual model of effective partnership working.

Background

Interventions to increase physical activity are a core part of public health policy and practice (1, 2, 13, 39), yet the complexity of public health interventions, which are often multi-component and multi-sectoral, inevitably leads to complexity in terms of their implementation and evaluation (268, 269). Nevertheless, it is essential that we understand if and how these interventions are effective and act upon this evidence if we are to meet targets for increasing physical activity at the population level, including the World Health Organization Global Action Plan target for a 15% reduction in physical inactivity by 2030 (1).

Evidence-based public health aims to ensure that decisions and interventions are based on sound evidence to safeguard and improve the health of the population. Appropriate evaluation is central to the generation of this evidence (14, 23, 50, 270). One of the key challenges is to generate practice-relevant evidence, where external validity and adoption into routine practice may be more likely (61, 270, 271). Evaluation of 'real-world' interventions, implemented as part of normal service delivery or in practice-based settings rather than in a research environment, provides an opportunity to address this challenge. However, this type of evaluation requires careful selection of approaches that are appropriate and feasible within real-world contexts (35, 127, 272).

Much progress has been made within the field of public health evaluation in the last two decades, and we have a better understanding of the challenges. Examples include limitations in expertise, capacity, and resources within normal service delivery to conduct evaluation, too much focus on operational objectives and outputs, and barriers to knowledge translation (11, 48-51). As our understanding of the challenges to evaluation has developed, so too has the guidance available. This includes guidance on methodological approaches, such as theory-based or realist evaluation (56, 57), recommendations for good practice (8, 11, 14, 35, 58, 59), and specific frameworks to facilitate systematic evaluation (47, 93, 120). The application of frameworks and logic models are now commonly recommended to guide the evaluation and reporting of physical activity interventions. However, our own systematic review of evaluation frameworks showed limited use and/or reporting of frameworks in evaluation studies of physical activity interventions (86). The reasons for this remain unclear.

Further to the concerns regarding the limited use of frameworks, additional gaps remain in our understanding of how to improve evaluation. Previous reviews of health promotion programmes have highlighted a need for a greater consideration of programme theory (157), investment and planning for evaluation (50), and a need for multi-level strategies that involve multiple stakeholders (11, 50, 51). Collaboration with independent experts in evaluation, such as through research-practice partnerships, is recommended as an approach to improve the quality of evaluation, build capacity for evaluation (8, 11, 48, 50, 51), and improve the use of evidence to

inform programme development (61). However, our understanding of the effectiveness of these strategies in practice remains limited (9, 51, 54, 61).

There is a need for research to develop a better understanding of how different factors interact to influence evaluation practice (51). Lack of insight into these influences may lead to variability in the quality of evaluation and reporting, which limits the generation and use of critical evidence to inform interventions and decisions to improve population health.

In this study, we report the findings of a case study of Sport England's 'Get Healthy Get Active' (GHGA) programme (273) to explore evaluation practices, and influences on practice, in an applied context. Sport England is the agency in England with primary responsibility for developing grassroots sports and increasing physical activity across England (274). The GHGA programme was chosen as our case study as it was specifically designed to build an evidence base for the role of sport in increasing physical activity, improving health and reducing health inequalities (275); evaluation was therefore a key element of the programme. The GHGA programme exemplifies multi-sectoral and multi-component approaches within public health (2). We explored the relationships between organisational structures and processes, and evaluation practice. Although we focus on a national programme to increase physical activity, the aim was to produce research findings that were applicable to other health-promotion interventions, particularly those operating in multi-sectoral public health contexts.

Objectives

1. To identify the logic of the programme and explore the relationships between programme and project aims.
2. To explore influences on evaluation practices, including requirements to use a standardised evaluation framework and specific data collection methods.
3. To appraise whether the programme was effective in generating high quality generalisable evidence that enabled it to meet its aims.
4. To formulate and discuss implications for the effective commissioning and evaluation of public health interventions.

Method

The GHGA programme

Through the GHGA programme Sport England funded 33 physical activity projects, 31 projects within two funding rounds and two invited projects, which were delivered between 2013 and 2018 to communities and population groups across England. For clarity, we refer to the GHGA intervention as 'the programme' and local, funded interventions as 'projects'. Projects were

developed, implemented and evaluated in partnership with Local Authorities, charities, Clinical Commissioning Groups and evaluation partners.

The programme provided an opportunity to explore evaluation practices, and to appraise whether strategies intended to facilitate project evaluation were effective. Sport England put in place several funding requirements to support evaluation. All projects were required to engage an independent evaluation partner, either an academic organisation or consultant. Projects were also required to use validated evaluation tools. This included the use of the Standard Evaluation Framework for physical activity interventions (SEF) (93) to guide project evaluation, the Single Item Physical Activity Measure (276), a validated tool to screen participants for eligibility for physical activity interventions, and the International Physical Activity Questionnaire (IPAQ) (277) to measure physical activity at baseline and follow-up.

Study design

We applied a collective case study design (83), using documentary analysis and semi-structured interviews, to conduct an in-depth analysis of multiple sources of evidence from a range of physical activity projects funded by GHGA. The purpose of a collective case study was to provide an in-depth, multi-faceted understanding of evaluation practices in a real-world context, using multiple data sources to increase the internal validity (83). Ethical approval was received from the University of East Anglia Faculty of Medicine and Health Sciences Reseach Ethics Committee (REF: 201718 – 133) (see Appendix 2).

Sampling and data collection for the documentary analysis

Agreement to conduct the research was gained from Sport England (Appendix 3). We conducted initial screening of documents provided by Sport England or published on their website, such as the "Project Summaries", to develop an overview of projects and to identify the lead organisation for each project. Each of the organisations responsible for the 31 projects in the two funding rounds were contacted and asked to share the final project evaluation report along with documents related to the funding application and intervention planning if available. Contact was initially made by email and then by telephone up to three times. All documents were given a unique code to de-identify them prior to importing them into NVivo 12 Pro for analysis.

Sampling and data collection for the semi-structured interviews

For the interviews, we applied purposive sampling to select stakeholders who were involved in the development, delivery or evaluation of the GHGA programme and projects. This included stakeholders with a role in the national programme and the project lead of each organisation who had shared an evaluation report. We applied snowball sampling to identify additional stakeholders, such as evaluation partners and project facilitators. Each stakeholder was contacted

up to three times via email or telephone and invited to participate in an interview. We continued sampling until we were confident that the sample was representative of projects across the two funding rounds, and different types of lead organisation, evaluation partnership, and stakeholder role. All participants provided written consent prior to participating in the interview (Appendix 4 and 5 provide a copy of the Participant Information Sheet and Consent Form).

We used semi-structured interviews to ensure we obtained data in relation to the objectives yet allow flexibility that may elicit richer data. An interview guide was developed to facilitate practitioner reflection and allow clarification of findings from the documentary analysis. The guide was piloted with one practitioner, however using semi-structured interviews allowed us to be responsive to emerging findings and refine the questions throughout the data collection period in an iterative approach. The guide consisted of 13 open ended questions that explored practitioners' experiences of the evaluation process, influences on evaluation, barriers and facilitators, and dissemination activities (provided in Appendix 6).

The interview guide was sent to participants in advance to provide them with prompts for reflection prior to the interview. Interviews were conducted face-to-face, by Skype or telephone. One participant communicated their responses via email. Interviews were conducted by the lead author (JF) between May and December 2019 and lasted an average of 46 minutes (range 25-86 min). Interviews were audio recorded and transcribed verbatim. All transcripts were sent to participants to check and provide the opportunity to add additional comments or clarification. Transcripts were given a unique numerical identifier to de-identify them before being imported into NVivo12 Pro.

Analysis of documents and interview data

To understand the programme aims and logic (objective one) we analysed Sport England's organisational documentation related to programme design, funding and monitoring, to develop a logic model and pathway diagram. These were refined through interviews and consultation with key stakeholders at Sport England to ensure that our interpretation and representation of the programme was accurate.

To address objectives two and three we applied Framework Analysis (278, 279). We combined deductive (*a priori)* and inductive (emergent) approaches to conduct thematic analysis of the documents and interview data. Initial categories and codes were identified *a priori.* These included codes related to the use and reporting of the SEF criteria, the single-item physical activity measure and the IPAQ. The SEF provides a structured framework to support project design, evaluation and reporting; the 52 criteria included in the SEF are intended to provide guidance on the information required to undertake a comprehensive and robust evaluation (93). The criteria

are grouped into seven sections (Table 4-1). We used these criteria as codes to guide data extraction and anaylsis, and provide a systematic approach to summarise the projects and their evaluation. Other codes identified *a priori* were informed by our interview guide and research objectives, for example influences on evaluation design, barriers and facilitators, and dissemination. Through repeated reading and familiarization with the data emergent codes were added, for example reference to additional evaluation methods such as logic models and case studies. The codes were reviewed and organised into categories and sub-themes (by JF) to develop the coding framework and were iterated and agreed with all authors.

Table 4-1 Summary of criteria included in the Standard Evaluation Framework for Physical Activity Interventions (SEF)

SEF sections	Criteria	Examples of criteria included
1 Programme details	16 essential	Aims, timescales, location and setting, description, recruitment, costs, resources
	7 desirable	Rationale, policy context, health needs assessment
2 Evaluation details	2 essential	Evaluation design, methods and timing of data collection
3 Demographics of participants	5 essential	Age, sex, ethnicity, disability, socio-economic status
	2 desirable	Additional information
4 Baseline data	1 essential	Measures of physical activity
	2 desirable	Correlates of physical activity, other outcomes
5 Follow up data	1 essential	Physical activity at ≥ 3 time points
	3 desirable	Physical activity > 1 year, correlates of physical activity, other outcomes
6 Process evaluation	6 essential	Participant numbers invited, recruited, attending, at follow up, satisfaction
	2 desirable	Unexpected outcomes, sustainability plans
7 Analysis and interpretation	3 essential	Summary of results, limitations and generalisability, recommendations
	2 desirable	Details of analysis, dissemination

We extracted data from NVivo12 Pro into a final analytical framework matrix to systematically synthesise the data by cases and codes. Using the framework we analysed themes by individual cases (funded projects), across different data sources (documents and interviews), and across the whole data set (representing the programme). To explore how evaluation practices had been applied and documented, and to identify influencing factors, we combined data from the documentary anaysis with data from the interviews.

The findings are presented as a narrative synthesis. Firstly, we present the programme's aim and logic, and then describe how these compare to project aims and characteristics (objective 1). We then present key themes identified as influences on evaluation practices (objective 2). To appraise whether the programme aim of generating evidence had been met (objective 3), we summarise

the reported outputs and outcomes from the project and programme evaluation, and map these against the intended outcomes. Finally, we formulate and discuss implications for effective commissioning and evaluation of health promotion interventions (objective 4) within the discussion.

Results

The Case study sample

In addition to the programme-level documents provided by Sport England, representatives from 23 out of 31 (74%) projects shared documents, including the final evaluation reports. These documents formed our sample for the documentary analysis. Lead organisations of two projects declined to share reports, and the leads of the remaining projects did not respond, of which two organisations were known to be no longer in operation.

Thirty-five stakeholders participated in an interview, including stakeholders with a role in the development, management or evaluation of the national programme (n=5), and stakeholders with a role in the design, delivery and/or evaluation of one or more local projects (n= 31). Some stakeholders had held more than one position with differing roles in the programme and projects. The interview sample was representative of 16 different projects; six from the first funding round and 10 from the second round.

Objective 1: To identify the logic of the programme and explore the relationships between programme and project aims

The rationale for the programme and its evaluation is shown in a logic model (Figure 4-1). A pathway diagram (Figure 4-2) shows the contextual factors influencing the programme. The programme was described as a response to a review commissioned by Sport England that highlighted the limited evidence base for the role of sport in tackling inactivity (280), and to government strategies that sought to increase participation in sport and physical activity among the least active adults (281, 282). Stakeholders involved in the programme's design highlighted the desire to build evidence that could support the commissioning of sport interventions to improve physical activity and health. One programme-level stakeholder explained:

> "*The reason why we did it the way we did it, was because of the lack of the evidence base … so when somebody else does a systematic review we are hoping that there will be at least 33 papers that will come up, if not more, to help answer that question in future*". (stakeholder 1)

Table 4-2 summarises the aims and key characteristics of the projects. Whilst the primary aim of all projects aligned to the programme aims, projects also reported various secondary aims and objectives. Projects were delivered by a range of organisations and cross-sector partnerships in a

range of locations and settings to diverse population groups. Several included multiple components and/or delivery pathways.

The pathway diagram (Figure 4-2) shows changes in organisational structures and strategies, as well as organisational learning, which influenced programme processes and practices across the two funding rounds. A key factor was the shift to Local Authority Health and Well-being Boards and Clinical Commissioning Groups being made accountable for Public Health commissioning in England from 2013, which informed an additional funding requirement for projects to address local needs and gain approval from Local Health and Well-being Boards in Round Two; a change which is reflected in the target populations and objectives of those projects.

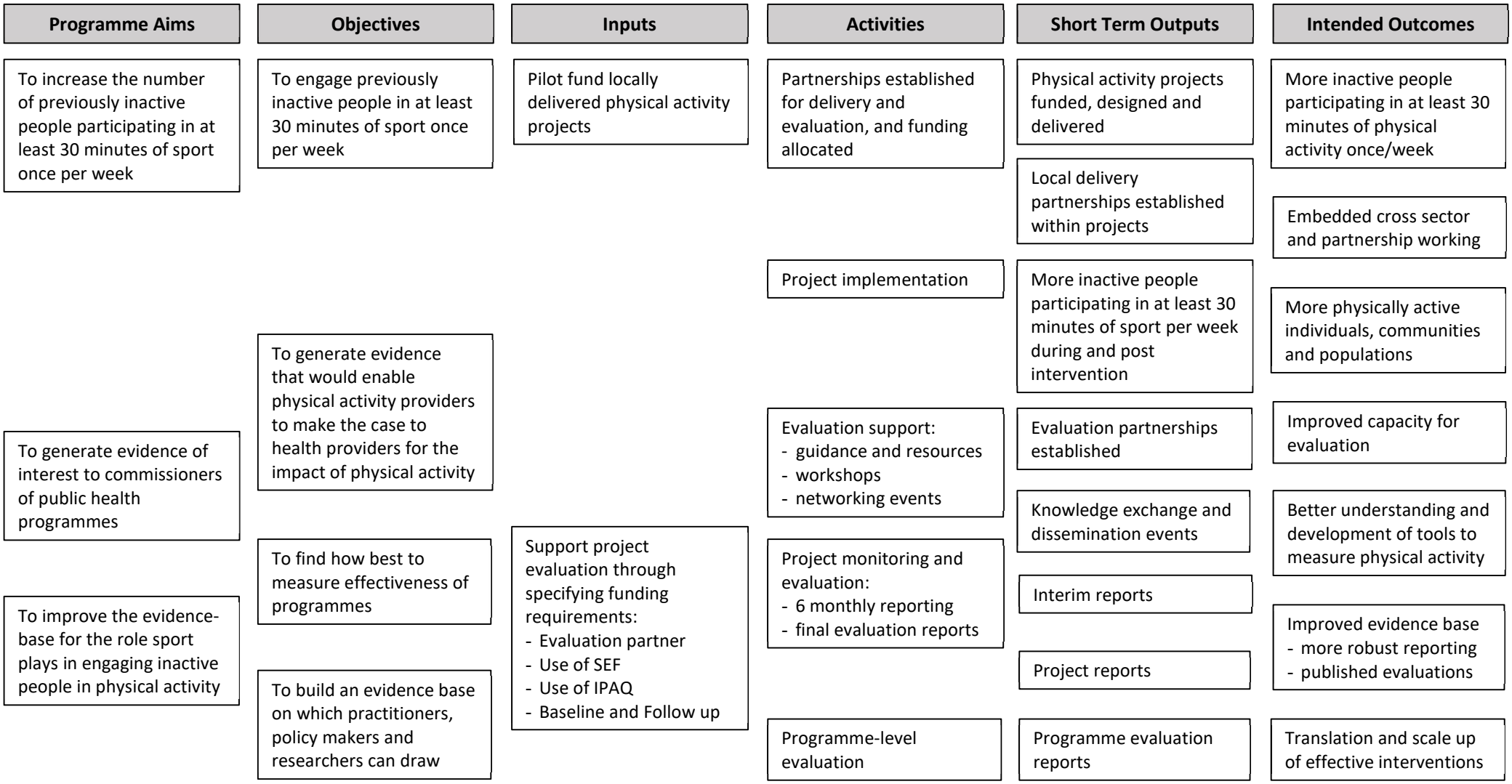

Figure 4-1 Logic Model for the Get Healthy Get Active (GHGA) programme

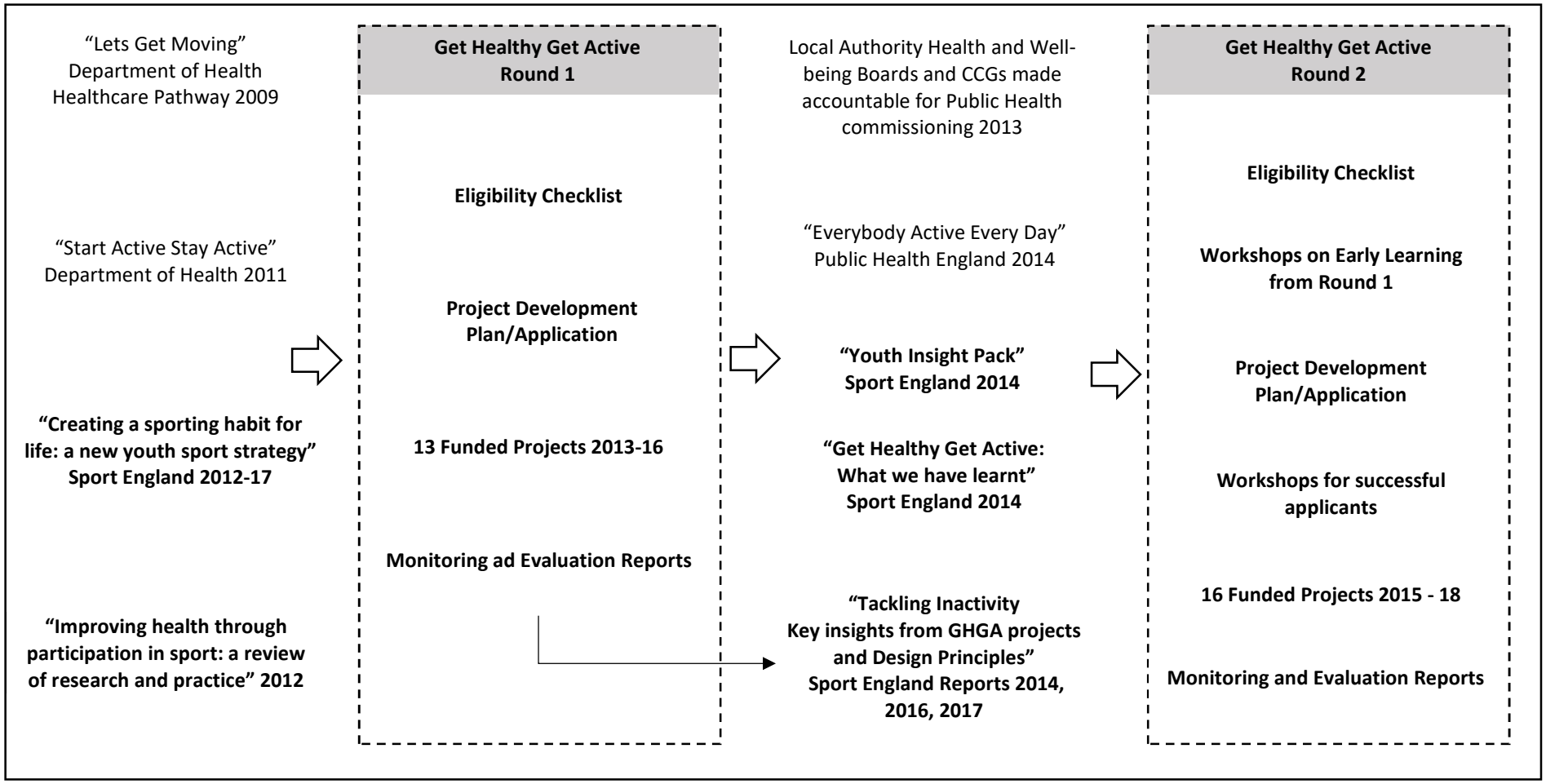

Figure 4-2 Pathway diagram of the Get Healthy Get Active (GHGA) programme

Notes: Round One was originally referred to as Get Healthy Get into Sport, Normal text shows external documents and influences on the programme e.g. Start Active Stay Active (281), Everybody Active Every Day (2), Bold text shows documents published or commissioned by Sport England and steps in the GHGA programme e.g. Sport England Strategy 2012-17 (282), Improving health through participation in sport (280), Get Healthy Get Active: What we have learnt (275), Tackling Inactivity (275, 283)

Table 4-2 Summary of the reported programme and project characteristics, aims and objectives

Project	Lead Organisation	Evaluation Partner	Location and Setting	Target Population	Aims and Objectives
GHGA	Sport England	In-house and independent consultants	NA	Inactive people aged 14 years and over	To encourage inactive adults to increase their physical activity by participating in sport, and build the evidence base
1-01	County Sports Partnership	University Partner	County-wide community settings	Inactive adults aged 16 years and over	How inactive adults can be recruited into sport and PA; How sport can be used to engage inactive adults in PA; Assess the impact and cost-effectiveness
1-02	University	University Led	CCG area, sport and leisure settings	Inactive people with hypertension, suspected or pre-hypertension or high-normal blood pressure	Whether sports-based referral for exercise would be effective compared to traditional gym-based projects; Whether a self-help web-based tool would add any additional benefit
1-03	University	University Led	Metropolitan borough, community settings	Inactive people	To design and deliver innovative community sports for health projects in different local contexts; Evaluate the design, outcomes, processes and costs of the project.
1-04	County Sports Partnership	University Partner	County-wide	Sedentary people at excess risk of cardiovascular disease and Type 2 diabetes	To describe the demographic details and impact of the project on self-reported and objectively measured physical activity; To gain insights into the experiences of participants and deliverers
1-05	County Sports Partnership Network	University Partner	National workplaces	Inactive employees	To develop a package of interventions to engage people in PA in workplaces; Assess the effectiveness of the project on increasing sport and PA and on business outcomes; Understand factors associated with using the workplace to engage the inactive in sport and PA
1-06	County Sports Partnership	University Partner	City and County districts, community settings	Inactive people living in target areas	To develop and test a community model for engaging inactive individuals in sport and PA; Assess whether one-to-one mentoring influences experiences and adherence to participation in sport and physical activity; Explore influences of engagement of family and friends; Explore wider benefits; Explore impact of engaging volunteers

Table 4-2 Summary of the reported programme and project characteristics, aims and objectives

Project	Lead Organisation	Evaluation Partner	Location and Setting	Target Population	Aims and Objectives
1-07	Charity	Evaluation Consultant	Geographical Health regions across UK	People Living with Cancer	Understand how the pathway has been implemented; Assess the extent to which delivery is in line with the ideal model; Explore efficacy of the interventions, scalability of the pathway, processes for best practice delivery, and impact of the pathway on service users and their families
1-08	County Sports Partnership	University Partner	County-wide, leisure settings	Referrers of inactive people (various health services)	To help individuals meet recommended levels of physical activity, based on the Lets Get Moving pathway
1-09	County Council	University Partner	County-wide, community settings	Inactive adults with long-term health conditions: cancer, cardiovascular disease, type II diabetes, mental health and from deprived communities	To establish the effectiveness of the project at increasing and sustaining PA of inactive individuals; Establish the effectiveness of tailoring interventions to specific population groups; Understand the mechanisms by which outcomes were reached and identify good practice and difficulties
1-10	Not-for-profit association	Not Stated	City and County-wide, GP surgeries	Individuals 18-75 years with a BMI between 28-35 resident in the catchment of participating surgeries	To provide an overarching assessment of the project and its impact upon participation in sporting sessions and physical activity levels
1-11	Borough Council group	University Partner	Metropolitan borough	Inactive people aged 14 and over, with a BMI of 28 or more	To help people get fit and lose weight by taking up sport; Evaluate effects of a community sports referral project compared with standard community exercise referral
2-01	County Sports Partnership	University Partner	County-wide, sheltered housing and care homes	Residents aged 65 years and over in sheltered housing and care home sites	To promote physical activity among residents in group homes with the aim of normalising physical activity
2-02	Not-for-profit association	University Partner	County districts	Inactive people over 16 years, living in target areas, one or more risk factors for cardiovascular disease and/or mild to moderate mental health problems	To support inactive adults to become more active and to work with Primary Health Care as a primary route of referral; Assess the measurable change on PA, general health and wellbeing; Understand how the project worked

Table 4-2 Summary of the reported programme and project characteristics, aims and objectives

Project	Lead Organisation	Evaluation Partner	Location and Setting	Target Population	Aims and Objectives
2-03	City Council	Evaluation Consultant	City areas, community settings	Pregnant and post-pregnant women	To increase the activity levels of pregnant and post-pregnant women
2-04	County Sports Partnership	University Partner	County-wide, leisure and community settings	People with drug and alcohol related problems	To encourage active and healthier lifestyles for adults recovering from drug and alcohol misuse
2-05	Borough Council	University Partner	Metropolitan borough, community settings	Inactive people with a high risk of developing type 2 diabetes, aged 47-74 years	To show the impact of a targeted sport and PA project on helping prevent or reduce the onset of type 2 diabetes and risk factors, for high risk adults; Assess differences across demographic categories; Assess if peer support can impact on someone increasing (and maintaining) PA; Assess differences in GP- and self-referred
2-06	Borough Council	University Partner	County-wide	Inactive people with a long-term condition: Cardiac Phase IV, Chronic Heart Failure, Stroke, Cancer, Lower Back Pain, Chronic Obstructive Pulmonary Disease and Falls Prevention	To support individuals with long term conditions to become and stay more physically active; To understand how effective the project was in providing condition specific support via PA pathways for seven long-term conditions, cost effectiveness, and the process of delivering the programme
2-07	Borough Council	University Partner	Metropolitan borough	Older adults	To engage inactive older adults in PA at least once a week for 30 minutes; Evaluate project effectiveness on older adults' physical activity, sedentary behaviour and self-reported health indicators
2-08	District Council	University Partner	District, leisure and community settings	Inactive, hypertensive, pre-diabetic, diabetic or overweight/obese people	To engage individuals in sport and PA through collaborative working between general practice and community leisure services; Understand the population impact; Understand Reach, Effectiveness, Adoption, Implementation and Maintenance

Table 4-2 Summary of the reported programme and project characteristics, aims and objectives

Project	Lead Organisation	Evaluation Partner	Location and Setting	Target Population	Aims and Objectives
2-09	Not-for-profit association	University Partner	Metropolitan borough, community settings	Residents	To support and empower residents to lead healthier lives, to be more active and lose/maintain a healthy weight
2-10	University	University Led	City-wide	Young people (14-25yrs), working adults and older adults (65+), and those with an identified health risk through smoking or obesity	To put in place a city-wide (whole systems) approach to tackling physical inactivity; Investigate changes in PA awareness and behaviour in response to the implementation of a consortium-led, multi-agency, person-centred behaviour change project
2-11	County Council Public Health	Evaluation Consultant	County-wide, leisure and community settings	Inactive people in the County	To enable inactive people to engage with sporting activities to lower rates of physical and mental ill-health and to reduce public expenditure related to preventable illness; Evaluate how implementation has improved outcomes and experiences for participants, including improvements in quality of life, health and well-being
2-12	Not-for-profit association	University Partner	City-wide	Inactive men and women (aged 26-75) who already had type 2 diabetes or were pre-diabetic or were at high risk of type 2 diabetes	To engage target population in a community-based sport and PA intervention to increase PA, enhance health and wellbeing and facilitate the management of disease symptoms

Objective 2: Influences on evaluation practices

We identified five main themes describing factors that influenced evaluation practices: (1) programme and project design; (2) evaluation design; (3) partnerships; (4) resources; and (5) organisational structures and systems. Examples of how various factors within these themes can act as barriers or facilitators to evaluation are shown in Table 4-3, and explored further below. The data highlighted the complex inter-connections between influences, and how many influences can act as both facilitators and barriers depending on the project characteristics and context.

1. Programme and project design

Evaluation was shaped by the programme and project design. The choice and use of evaluation and data collection methods within projects was determined by programme and project objectives and outcomes of interest. However, these also needed to be adapted to the contexts and characteristics of the projects. Within this theme we identified four sub-themes of important influences on evaluation: timescales, participant demographics, settings, and implementation.

Timescales were seen as a barrier to data collection and to formative work. For example, short lead-in times impacted participant recruitment, ability to pilot evaluation methods, and to develop and embed data collection systems. Stakeholders noted that it took time to build relationships with delivery partners and to recruit participants. Timescales related to funding, project conclusion and outcome review were also felt to be a barrier to project sustainability. For example, stakeholders commented:

> *"the main thing was that lead in time, and I think the second thing is that it takes time to set up the project especially in these hard to reach communities and I think you can't underestimate how much time it takes to build those relationships with the participants, community groups, with the referrers…so it is how we can move away from that two to three years funding cycle, with the reality that it probably takes a year to two years to build relationships in the community and then you are taking that intervention away."* (stakeholder 15)

> *"I think there was sometimes a lack of time to actually pilot test some of the data collection instruments and processes because the projects are under pressure to start delivering as quickly as possible. And if we had had that time we might have maybe done things differently or refined things before we actually started to ensure it all went smoothly."* (stakeholder 21)

Participant demographics also influenced the outcomes of interest and how data were collected. Stakeholders described the importance of adapting data collection methods, project design and activities, to facilitate recruitment and data collection with specific demographic groups.

Project locations, settings and contexts, including resource availability and accessibility for participants, further impacted recruitment, implementation and response rates. The need for flexibility and adaptability was a recurring theme. This was linked to changes to projects during implementation, such as: staffing and promotional material; adding or tailoring activities and engagement opportunities; and refining eligibility criteria or referal processes. Flexibility in both project and evaluation implementation were described as essential to facilitate data collection, whilst also being a potential barrier to the generalisability of outcomes.

Table 4-3 Summary of influences on evaluation practice

Influence	Examples of how these can act as barriers or facilitators
Programme and project design	
Timescales	Lead in time, delivery and funding cycles influence opportunities for relationship building, recruitment, piloting methods and formative evaluation. Scheduling and duration of delivery sessions influence resource availability and capacity for data collection.
Participant demographics	Participant demographics influence recruitment and data collection, capacity for self reporting, response rates, outcomes of interest, requirements for different outcome measures and need for adaptations to data collection methods (impacts standardisation and generalisability).
Settings	Location, facilities and resource availability influence recruitment, response rates and data collection.
Implementation	Tailoring and adaptability in project and evaluation implementation can facilitate recruitment, participant engagement and response rates, but limit standardisation.
Evaluation design	
Standardised data collection	Facilitates consistency of reporting and comparability, however use in diverse project contexts and participant groups limits generalisability. Increases research-practice tensions, data collection burden and impacts response rates. Choice of tools, appropriateness to participants, and ease or difficulty of implementation influence data collection and outcomes.
Standard Evaluation Frameworks	Evaluation frameworks and guidance facilitate more consistent evaluation and reporting of required evaluation criteria and outcomes of interest. Variability in how criteria are applied and reported can act as a barrier to generalisability and quality of data. Limitations in guidance included in frameworks used can lead to variability in the quality of evaluation and reporting of specific evaluation components.
Use of non-required evaluation methods	Use of non-required evaluation components is dependent on knowledge, experience and priorities of project stakeholders, e.g. the value placed on qualitative methods influenced the inclusion of qualitative methods. Limitations in the specified requirements to address objectives drives inclusion of additional methods. Limitations in guidance, understanding of methods and capacity to conduct qualitative research influences the quality of analysis and reporting. Pilot and formative evaluation facilitates development, testing and embedding of evaluation approaches and data collection systems, intermediate evaluation facilitates learning, adaptation and improvement. These are dependent on timescales, regular reporting and feedback processes. Adaptability and flexibility facilitates ability to be responsive to needs, to improve participant and stakeholder engagement with evaluation processes, and to improve response rates and quality of data collection.
Resources	
Staffing	Staff expertise, experience, capacity, buy-in for evaluation, and how roles and responsibilities are defined influence evaluation processes, project sustainability, knowledge management and dissemination.
Funding level	Funding for evaluation, including staffing and partnership working, is a major influence on evaluation practice.

Table 4-3 Summary of influences on evaluation practice

Influence	Examples of how these can act as barriers or facilitators
	Differing levels of funding and the proportion allocated to evaluation, position of decisions for this at local or national level, and timescales of funding cycles influence evaluation practices.
Time	Time impacts the choice of evaluation methods, and the capacity for data collection and evaluation processes.
Equipment/facilities	Influences project activities, recruitment, implementation, and data collection methods, including opportunities for use of innovative methods.
Partnerships	
Essential partners/roles and responsibilities	Definning roles and responsibilities of delivery, funding and evaluation partners for evaluation processes is a key factor. Capacity for evaluation and success of partnership working is dependent on costs, funding, resources, and the nature of the partnership.
Stakeholder priorities, objectives and expectations	Differing partner priorities and expectations can lead to research-practice tensions. Approaches to balance research objectives, policy priorities and practicalities of what will work in real-world and in budget are required. Strategies to manage expectations are needed.
Expertise, experience, capacity	Prior experience, knowledge and training of stakeholders influence evaluation design, choice of methods, innovation and implementation. Research-practice partnerships can improve evaluation through access to expertise, skills and experience, and access to additional resource for implementing evaluation and data collection.
Relationships and Communication	Close relationships between partners are key. Local partnerships increase opportunities to observe and understand local project needs and facilitate relationship building. Available, approachable and adaptable partners enable open and trusting relationships, regular comminication, opportunities for stakeholders to challenge, learn from each other, find solutions and make decisions collaboratively. Appropriate language facilitates relationship building (jargon busting).
History of partnership, embeddedness	Continuity of relationships facilitates understanding of local project evaluation priorities, helps to embed processes, which can help mitigate effects of limited lead-in times, piloting and insight phases. Arms-length or transactional relationships act as barriers.
Organisational structures, systems and processes	
Funding systems and requirements	Clearly defined, agreed and communicated funding requirements act as facilitators to evaluation and use of evidence. Funding cycles and time scales for reporting and review can limit learning from evaluation, dissemination and project sustainability. Understanding future commissioning needs facilitates evaluation planning and implementation to ensure practice-relevant evidence is collected.
Staffing structures	Clearly defining roles and responsibilities of staff, volunteers and partners is vital to successful partnership working, project implementation and evaluation processes.

Table 4-3 Summary of influences on evaluation practice

Influence	**Examples of how these can act as barriers or facilitators**
	Key staff that have capacity and/or responsibility for co-ordinating processes, relationships and practices can be essential for the success of a project and its evaluation. These may be embedded in the staff structure as an evaluation officer, or an external partner that champions evaluation. Highly mobile workforce and employment contracts linked to short funding cycles act as a barrier to continuity of partnerships, relationships, and organisational learning, but as a facilitator to inter-organisational learning.
Systems for oversight, monitoring and communication	Information and support from funders, essential to guide project planning, but also to make use of feedback from intermediate monitoring and evaluation. Service level agreements help to define and agree roles, responsibilites, objectives and outputs, but can limit adaptability and flexibility. Steering groups (project boards or operational groups) enable sharing of good practice, open dialogue and support. Regular meetings that include evaluation feedback facilitates evaluation process. Challenges remain to ensure decisions are transferred between strategic and operational stakeholders, and that actions agreed are followed up.
rocesses for capacity building and knowledge exchange	Training to build capacity, knowledge and gain buy-in is essential, especially where data collection is dependent on delivery staff. Workshops and networking opportunities facilitate knowledge exchange across projects, partners and wider audiences.
Data management systems	Effective data management systems facilitate data collection and management, participant engagement and project implementation. Developing, agreeing and embedding systems that meet the needs of practitioners and researchers is essential, but has implications for resources such as time, staffing and budgets. System development and use needs to consider implications for data security policies and practices, reliability, flexibility, integration with existing service delivery systems and needs, standardisation to allow reporting and comparison between partners, projects and programme.
Wider external influences	Embedding project and evaluation into existing service delivery offers opportunities for efficiencies, e.g. shared resources, staffing economies and use of existing infrastructure such as data management systems. Embedding in existing service delivery can also facilitate project sustainability. Evolving policies, strategies, commissioning priorities and knoweldge development interact to influence priorities for funding, project and evaluation objectives, reporting and dissemination, and use made of evidence. Multi-sectoral, multi-component projects or localised delivery and evaluation can lead to fragmentation of projects across organisations and locations, which can act as a barrier to standardised approaches to evaluaton, knowledge exchange and use of evidence.
Organisational culture and embeddedness of evaluation	Organisational culture and a history of evaluation and partnership working within organisations can increase opportunities for integrating evaluation and project design, improve the skills base, capacity and buy-in to evaluation process and practices and facilitate the embedding of evaluation.

2. Evaluation design

Evaluation design was shaped primarily by the requirements to use standardised data collection tools and a standard evaluation framework. In addition to these required elements, projects reported on a wide range of study designs, evaluation methods, and data collection tools, as shown in Table 4-4. As one stakeholder explained:

> *"There was a big influence there in terms of consistency across the projects across the country … Sport England were a big influence in terms of the IPAQ and the things that they were asking for, but we also had the additional secondary questions that we added into the evaluation that were very much around what do we need locally to evidence that this works ... I know that a lot of the academic studies included a process evaluation, but that wasn't a direct output that Sport England were expecting, or they didn't dictate that."* (stakeholder 6)

To illustrate how the application and reporting of required and optional evaluation methods influenced the evaluation in practice these elements are discussed below.

2.1 Use of standardised tools

Sport England recommended using the Single Item Measure (276) to identify inactive participants for eligibility. Sixteen projects reported using this tool. Two projects did not refer to any screening tool, whilst four mentioned using alternative screening tools (Table 4-4). There was variability in how eligibility criteria were applied, and in the use made of the Single Item Measure; for example four projects used it to assess changes in physical activity over time. Stakeholders reflected on differences in how eligibity criteria and screening tools were applied as a challenge to recruitment and comparability across projects.

Projects were also required to use the IPAQ to collect baseline and follow-up measures. Twenty-two projects reported using IPAQ-short form or IPAQ-E (developed for older people), whilst one project had agreement to use an alternative tool, the Scottish Physical Activity Questionnaire (SPAQ). Sport England also recommended using a single question to assess sport participation; which ten projects referred to.

The use of standardised tools in real-world settings and with specific demographic groups was identified as a key challenge. In particular, stakeholders emphasised the negative effect of data collection burden on recruitment and response rates, and in turn on generalisability. For example, stakeholders described the following challenges in using the IPAQ:

> *"One of the biggest challenges is taking validated questions and looking at the practicality of implementing them in the community."* (stakeholder 15)

"They were a fairly lengthy questionnaire for the type of people we were working with and it led to a real reduction in numbers. The evaluation led to the reduction in numbers. The reduction in numbers was because of the way the evaluation was working but to make the evaluation effective we needed more people. So it was a bit of a vicious circle." (stakeholder 19)

2.2 Use and reporting of the Standard Evaluation Framework

The purpose of including the use of the essential SEF criteria as a funding requirement was to facilitate standardised evaluation and reporting. According to one programme-level stakeholder its strength was in the guidance on reporting contextual factors that would allow Sport England to *"understand what works, for who and how; or what doesn't." (stakeholder 1)*

Eleven (48%) of the evaluation reports, specifically stated that the evaluation was guided by the SEF. Eleven reports did not refer to any evaluation framework, and one referred to the RE-AIM framework (47) as guiding the evaluation.

Reporting of the SEF criteria was variable. Tables 4-5 and 4-6 summarise which projects reported on the criteria related to programme details and participant demographics. All projects gave a detailed description of their aims and objectives, recruitment methods, location and setting, and reported on age and gender. Those that targeted specific population groups described these in detail. Quality assurance mechanisms, potential unintended consequences, and costs were reported on by fewer projects. The rationale for the intervention, relevant policy context and health needs assessment were not always differentiated. The SEF recommends the use of a logic model, yet just five reports (22%) provided this.

All projects reported on the timing of data collection at baseline and follow-up. Whilst there was some variation in how impact data were reported, all projects reported on change in self-reported physical activity across time points. Seven (30%) projects reported a comparison of outcomes between intervention and control groups or across demographic, disease-risk, referral or service pathway sub-samples. Details of statistical tests used to analyse physical activity measures and the rationale for their use were reported fully, whilst sixteen (70%) projects reported on limitations and generalisability and ten (44%) reported on how findings were disseminated.

The SEF provides more limited guidance on process evaluation (Table 4-1). Participant numbers were reported variably based on attendance at at least one session, completion of a 10 or 12 week course, or registration at one-off events or online. One project provided a flow diagram of participant numbers with reasons for drop out. Fourteen (61%) projects combined exit survey and interview data to report on participant satisfaction. Nineteen (83%) projects reported on plans for sustainability. One project included this as a research objective to explore features that may lead

to sustainable delivery models. Five (22%) projects described how the delivery model had been developed with sustainability in mind.

Table 4-4 Study design and data collection methods included in project evaluation

Methods		Project Codes:	1-01	1-02	1-03	1-04	1-05	1-06	1-07	1-08	1-09	1-10	1-11	2-01	2-02	2-03	2-04	2-05	2-06	2-07	2-08	2-09	2-10	2-11	2-12	%
Physical Activity Measurement	IPAQ short		x	x	x		x	x		x	x	x	x		x	x	x	x			x	x	x	x	x	78
	IPAQ-E													x					x	x						13
	SPAQ								x																	4
	Stanford 7 day recall					x																				4
	Sport participation question		x		x		x	x	x		x	x									x			x		43
	Objective measure (accelerometer)					x						x		x										x		17
	Borg scale																								x	4
Screening	Single Item Measure		x		x	x	x	x		x		x		x	x	x	x		x			x		x	x	x	70
	PARQ								x				x					x				x				17
	General Practice Physical Activity Questionnaire			x																						4
Self-report Surveys	Cancer Physical Activity SEF								x																	4
	Health Related Quality of Life (e.g.EQ-5D-5L)		x		x				x	x				x	x				x							30
	Kemp Quality of Life Scale																			x						4
	Warwick Edinburgh Mental Wellbeing Scale						x							x	x			x			x					22
	Functional Assessment of Chronic Illness Therapy								x																	4
	General Self-Efficacy (GSE) scale								x															x		9
	Wellbeing (e.g. Adolescent Wellbeing Scale)				x		x				x		x													17
	WHO-5 Well-being Index																							x		4
	RAND SF32			x									x													9
	Loneliness Questionnaire												x	x												9
	Motivation Questionnaire												x													4
	Fear of Falling Visual Analogue Scale													x												4
	Life satisfaction scale		x			x																				9
	Cantril Self-Anchoring Striving Scale																x									4
	Mediators of sport or physical activity			x			x												x			x	x			22
	Other self-reporting health status or behaviours					x	x	x								x				x		x	x			30
	Feedback/satisfaction survey		x	x						x												x				17
Other	Attendance		x		x			x										x			x		x			26
	Costs, resource use, programme records		x		x				x	x				x	x				x		x			x		39
	Objective measures (e.g. anthropometric, health)			x										x				x			x	x				22
	Interviews, Focus groups		x	x	x		x	x	x		x	x		x	x	x	x	x	x	x	x		x	x	x	83
	Ethnographic/observation				x				x					x		x										17

Table 4-5 Summary of project reporting on SEF criteria related to programme details

Project Codes:	1-01	1-02	1-03	1-04	1-05	1-06	1-07	1-08	1-09	1-10	1-11	2-01	2-02	2-03	2-04	2-05	2-06	2-07	2-08	2-09	2-10	2-11	2-12	%
SEF mentioned	X				X	X	X	X	X			X					X	X	X			X		48
1. Intervention title	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
2. Aims & objectives	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
3. Rationale for the intervention	X	X	X			X	X	X	X	X	X	X			X	X	X	X	X	X	X	X	X	83
4. Contact details	X	X	X		X	X	X		X		X	X	X						X		X	X		57
5. Commissioners/sources of funding	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	96
6. Intervention timescale	X	X	X		X	X	X	X	X	X	X	X	X	X			X	X	X	X	X	X	X	87
7. &/or 8. Delivery or funding dates	X		X		X	X	X	X	X	X		X	X	X	X	X	X	X	X		X	X	X	83
9. Location & setting	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
10a. Target population	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X		X	91
10b. Content	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X		X		X	X	X	X	87
10c. Delivery method	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	91
10d. Deliverer	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X		X	X	X	X	91
10e. Quality assurance mechanisms							X	X	X	X			X			X					X			30
10f. Potential unintended consequences			X						X															9
11. Method of recruitment & referral	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	100
12. Admission/inclusion criteria	X	X	X		X	X		X	X	X			X	X		X			X			X	X	61
13. Consent /ethical approval	X	X	X		X	X	X		X	X		X			X	X	X	X	X		X	X	X	74
14. Equipment & resources	X		X		X	X	X		X	X		X	X				X		X					48
15. Core staff competencies/training	X		X		X	X	X	X	X	X		X	X	X	X	X	X		X		X		X	74
16. Incentives for attendance	X		X	X	X	X	X	X	X	X	X		X					X			X			57
17. Detailed breakdown of costs	X		X				X	X				X	X				X		X			X		39
18. Costs per participant	X						X	X		X		X	X						X		X	X		39
19. Cost to the participant	X	X			X	X	X		X	X	X		X		X		X	X						52
20. Relevant policy context	X		X				X	X	X	X	X	X	X		X	X	X	X	X	X	X		X	74
21. Health needs assessment			X	X		X		X					X			X	X		X	X		X	X	48
22. Equality impact assessments																								0
23. Declaration of interest																								0

Table 4-6 Summary of project reporting on SEF criteria related to participant demographics

Participant Demographics	1-01	1-02	1-03	1-04	1-05	1-06	1-07	1-08	1-09	1-10	1-11	2-01	2-02	2-03	2-04	2-05	2-06	2-07	2-08	2-09	2-10	2-11	2-12	%
Age	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	100
Sex	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	100
Ethnicity	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x		x	x	x	91
Disability	x	x	x	x	x	x	x	x	x	x	x	x		x	x				x		x	x	x	78
Socio-economic status	x	x	x	x	x	x	x	x	x		x		x	x	x		x	x	x		x	x	x	83
Additional information e.g. health status	x	x	x		x	x	x				x								x			x		39

2.3 Use and reporting of optional evaluation components

Table 4-4 shows that projects included a range of additional self-report surveys. Nineteen (83%) of the projects conducted interviews and/or focus groups to provide additional understanding and insights about how the projects worked and were received. The choice and use of these methods was influenced by project level stakeholders' priorities and expertise, but also limitations in the required tools to generate evidence in relation to evaluation objectives.

Several stakeholders reflected on the value of qualitative methods to answer questions about the project, for example:

> *"there's certain cohorts of people we work with where it's really hard to collect robust evaluation and actually it's the qualitative that matters and the process. I'd like to see a lot more investment in process evaluation because I think at the moment at this time of system changes, so much transformation going on in the health system, and it's the processes that are important."* (stakeholder 6)

> *"I think for us some of the most important information came from the qualitative side."* (stakeholder 15)

Twelve projects provided a separate section or report described as either a process or qualitative evaluation. There was variability in how qualitative methods were applied, analysed and reported. For example, some simply mentioned thematic analysis, whilst others provided details of the coding and method of reporting. Four projects combined different data sources to explore project impementation and contextual factors, whilst eight reported on data as case studies of individual participants, organisations or delivery pathways.

3. Resources

Resources, including staff, time, funding, equipment and facilities, were a major influence on evaluation as shown in Table 4-3. In particular, the availability and use of resources illustrates how the context and characteristics of each project can affect how factors interact and can act as both facilitators and barriers. For example staffing was essential for data collection and evaluation, and depended on the roles, responsibilities and capacity of partners, which in turn were dependent on organisational staffing structures, funding levels and time-scales. Stakeholders from some projects regarded the level of funding as enabling a more rigorous evaluation process than is often possible within real-world interventions, whilst stakeholders from other projects highlighted limited funding as a barrier to their ability to resource the evaluation.

4. Partnerships

Partnerships shaped the nature of project evaluations. All projects were required to have an independent evaluation partner, and were developed and implemented through working with a

range of delivery and funding partners. Evaluation partners were central to the evaluation design. Whilst some stakeholders reflected on differing objectives, priorities and understanding between research and practice as potential sources of tension, most highlighted access to expertise, and in some cases access to additional resources for evaluation as a benefit.

Variation in the responsibilities, priorities and capacities of staff employed by delivery organisations and evaluation partners was thought to have impacted the evaluation design and process. Delivery staff were seen as essential to recruitment and managing data collection. Defining responsibilities, communication, and training were seen as vital to build capacity, and to get buy-in to the evaluation process. As shown in Table 4-3, the nature of the relationships and history of the partnerships were key influences. For example, close relationships and local partnerships enabled regular communication, and facilitated relationship building and sustainable partnerships, whereas arms-length relationships were described as barriers to successful partnerships and evaluation.

5. Organisational structures, systems and processes

We identified seven sub-themes of influences related to organisational structures, systems and processes: funding systems; staffing structures; systems for communication, monitoring and oversight; processes for capacity building and knowledge exchange; data management systems; wider external influences; and organisational culture and embeddedness of evaluation (Table 4-3).

Several of these factors are inter-connected, and also underpin factors identifed within the other main themes. For example, whilst defining roles and responsibilities early in the project was essential to successful partnership working and evaluation, this was dependent on appropriate funding and staffing structures. High staff turnover was mentioned as a challenge to evaluation in nine of the reports, and by eighteen of the stakeholders interviewed. Stakeholders felt this was linked to short funding cycles and contracts, and to have negatively influenced continuity, the capacity for evaluation and dissemination. In particular, stakeholders felt that delays in staff recruitment added to the challenges associated with short lead in times; and early departure of staff influenced dissemination and use of evidence. Having a central co-ordinator who could act as a conduit between partner organisations was seen as critical to successful project evaluation in several cases.

As shown in Table 4-3, various structures and systems that can act as facilitators to evaluation were identified. Examples include: steering groups and service level agreements to enable regular and formal communication and oversight; training and knowledge exchange to build capacity; and data management systems and processes to integrate evaluation within normal service delivery.

Stakeholders reflected on the potential for efficiencies from integrated systems and processes, but also on the considerable time and resource implications of developing these and the difficulties in implementing them across multiple project partners and/or components.

A key underpinning theme was the importance of systems to facilitate monitoring, oversight and communication throughout the project planning, implementation and evaluation cycle. However stakeholders reflections on their experiences of these were variable. For example, service level agreements were seen as critical to agreeing and defining responsibilities in some projects, and as limiting flexibility in others. Many stakeholders reflected on the value of networking and knowledge exchange events facilitated by the funding agency, whilst others commented on a lack of such oportunities as a limitation:

> *"We found the workshops that they held, … actually to get the GHGA projects in a room together was really useful and because you could share the issues that you were having and people understood and you could share ideas and realize how people have overcome them."* (stakeholder 24)

> *"They were really good at that side of things, they would bring us in and then different projects would speak each time on different topic areas that we would cover in workshop scenarios, that was really good. They did that really well … I think Sport England could make a lot more of the network than they do in terms of avoiding that duplication of effort and resources."* (stakeholder 6)

> *"I never had a chance to talk to anyone else who was doing any of the other evaluations so there was never that kind of network and support which I think it might have been quite useful to have had."* (stakeholder 28)

Variability in communication and involvement of stakeholders in networking across different projects appears to have limited the opportunity for a more consistent approach to wider scale knowledge exchange and use of evidence. Some stakeholders also identified a need for organisational structures that enabled forward planning and closer working with local services to ensure that evaluation and evidence generation met future commissioning requirements.

Objective 3: Appraisal of whether the programme was effective in generating high quality generalisable evidence that enabled it to meet its aims

Figure 4-3 provides a summary of project and programme outputs mapped against the intended outcomes included in the logic model (Figure 4-1). Two separate evaluation consultancies were commissioned to produce summary reports from Round One and Round Two respectively. At the time of writing, only the reports following Round One were available (275, 284) these reported numbers of participants engaged in the programme, changes in numbers of participants identified

as active or inactive, and case studies of individual projects. Stakeholders at programme and project levels acknowledged the challenges of pooling large data sets from multi-component, multi-sectoral projects due to diverse project designs, settings and participant demographics, and variability in response rates, secondary outcomes, and in how outcome measures were analysed and reported:

> *"It was good to specify a measure to get the consistency across all the programmes, I guess the quality of that data collection probably varied quite a lot across different projects, depending on who did the data collection and how it was done."* (stakeholder 21)

One programme level stakeholder commented on the need to accept flexibility in how projects applied the specified requirements but that this:

> *"created a number of challenges at programme level, when you try to pull it all together."* (stakeholder 1)

Programme level stakeholders reported that findings had informed the development of resources to support project and service design and evaluation (173, 283, 285), and that several project reports had been included in subsequent reviews of practice (286, 287). In total nine projects disseminated findings through published articles in academic journals, eleven through publicly available reports, and nine through conference presentations. Five stakeholders mentioned plans for publishing articles, but identified a lack of time or time lag between end of project and publication as a challenge.

Project level stakeholders felt the need for knowledge exchange activities and reporting methods that were more appropriate to a wider audience, including local stakeholders and commissioners. Stakeholders involved in projects that had been showcased through best practice projects and conferences saw it as an important way of valueing the project and disseminating findings. Other stakeholders, who had not been involved seemed less aware of dissemination activities beyond what they were doing locally, and were keen to know more about how findings from across the programme were being shared. For example, stakeholders commented:

> *"I think it is a constant frustration that I have, that there is a huge amount of knowledge that gets built up and then never gets shared."* (stakeholder 31)

> *"I don't think out of all those projects across the whole network, that was really shared with people. So I think we got to hear more about it because we were part of it. I think where they have done one or two things more recently where they do try and bring people back together where they are all working on similar types of project and I think that's really valuable but I still think they can do a lot more to then share that with the wider network."* (stakeholder 30)

Whilst there was limited understanding amongst some project level stakeholders of how the reports were received, used or shared at the programme level, many described project evaluation as influencing practices, project sustainability or partnerships locally. One programme-level stakeholder commented on learning and capacity building remaining at a project or person level, and fragmentation of projects across multiple organisations, limiting the ability to influence at scale.

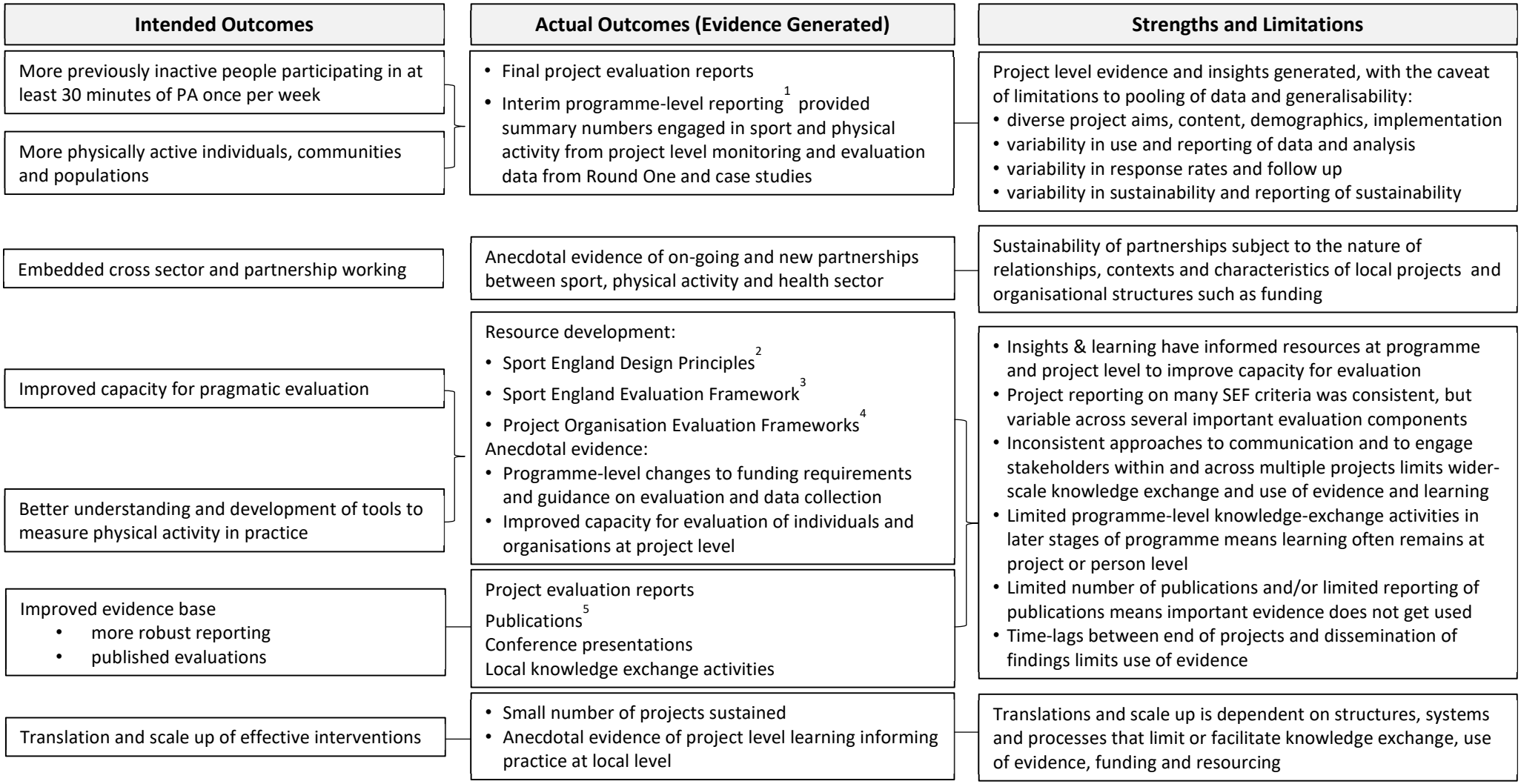

Figure 4-3 Evidence generated from the Get Healthy Get Active programme mapped against the intended outcomes

Notes: [1]Get Active Get Healthy, what we have learned so far (275), Tackling Inactivity (284), [2]Design Principles (283), [3]Sport England Evaluation Framework (173), [4]Hertfordshire Evaluation Framework (288), [5]Examples of publications include (65, 289-298)

Discussion

The GHGA programme included physical activity projects with a wide range of secondary aims, partnerships, participant groups, settings, and project and evaluation designs. Despite the variability in projects, we identified common influences on evaluation practices that act as facilitators or barriers depending on the context and how they interact within a project. Multiple factors influence programme implementation and evaluation in real-world interventions (11, 51). This is especially true in multi-sectoral and multi-component programmes such as GHGA. This makes gauging the role of any one factor difficult. Accordingly, our findings highlight the importance of understanding the interactions between influences on evaluation practices and, in particular, the implications for commissioning and evaluation of interventions. Whilst our focus is on physical activity interventions, the findings are applicable to other interventions, particularly those operating in multi-agency public health contexts.

A frequent criticism of real-world evaluation has been that evaluation is approached as an "add on" to intervention design and implementation, and that insufficient attention is given to evaluation during intervention planning (11, 50). Previous studies of health promotion programmes have also identified barriers such as limited investment for evaluation, and differing value placed on evaluation by stakeholders (14, 50, 52, 299). Within the GHGA programme these barriers were largely overcome by the specification of evaluation as a funding requirement at the outset of the programme. Our study showed the vital role that commissioners play in influencing evaluation practice through resourcing and demands for evaluation, and more critically, in providing appropriate guidance and support, and how they value different forms of evidence.

Stakeholders' understanding of what counts as evidence, and their use of appropriate evaluation methods, are recognised challenges of conducting real-world evaluation (14, 68, 92, 300, 301). Evaluation in an applied context often requires a balance to be found between scientific rigour and pragmatism, internal and external validity, and standardisation and adaptability (8, 14). It can be a challenge to balance differing stakeholder priorities for evidence. The value of combining systematic and flexible approaches (164, 302, 303), and applying theory based approaches (56, 57, 304) to evaluate the variability within complex interventions is well recognised. Standardised requirements for evaluation of funded projects can facilitate a systematic approach to evaluation and improve the consistency of reporting. This may be particularly important within multi-project programmes like GHGA, which are designed and funded nationally but delivered and evaluated through local projects. We have previously argued that appropriate use of an evaluation framework to guide evaluation and reporting can improve the quality of an evaluation study (86). Use of a framework can also facilitate identification and agreement of evaluation objectives and

methods between stakeholders (46). Logic models are commonly recommended to identify objectives, inputs, contextual factors and outcomes to help explain an intervention's theory or rationale (8, 59, 116, 178); their use is also recommended in the SEF (93). Qualitative or mixed methods are also advocated to help explain quantitative findings, and generate evidence about project implementation, programme theory or causal mechanisms (35, 59, 157, 303). Despite putting in place specific evaluation requirements, there was considerable variation in how important evaluation components were applied and reported. Components that were reported in detail, such as project descriptions and participant demographics, reflected the more detailed guidance of these components in the evaluation framework applied. Gaps in the evaluation reports highlighted limitations in the guidance provided in the SEF and the field generally on important evaluation components, and limited the ability to compare or generalise findings across projects. Further guidance or training is needed to improve the evaluation and reporting of specific components, in particular qualitative methods, process evaluation, economic evaluation, logic models, and data analysis. We argue that specifying evaluation requirements alone is insufficient. The context-specific nature of influences within diverse projects makes it more critical to implement processes that facilitate collaborative decision making to select, agree and apply the most appropriate methods to generate the evidence required and valued, rather than specifying standardised data collection across heterogenous projects.

Evaluation partnerships were a strong influence on evaluation. Many of the benefits of partnership working that we identified in this study, such as access to expertise, capacity building, and efficiencies from shared resources or integrated systems were also found in other studies (11, 50, 51, 61). We also suggest that partnerships can bring greater opportunities for evaluation to be tailored to the needs of individual projects and stakeholders, and to enable a more flexible and innovative evaluation approach. However, the effectiveness of partnerships were dependent on the nature of the relationships, the embeddedness and continuity of partnerships, and on organisational structures and systems. In line with other studies, we also found partnerships to be context specific, and changeable (65). For funders and partners to initiate and embed processes and systems that facilitate partnerships and that retain benefits of partnership working beyond a project's lifetime, it is essential that we develop a better understanding of the influences of, and on, partnership working.

Our appraisal of the extent to which the programme had generated evidence to achieve its aims (Figure 4-3) identified several resources and publications resulting from the programme, but showed that dissemination and use of evidence remains a challenge. At this stage, questions remain as to how useful local project evaluation has been in addressing the programme aim to

build an evidence-base that would inform scale up of effective interventions or translation to other settings. The programme sits within a system of evolving national and local policies, strategies and priorities, and knowledge base (Figure 4-2). Our findings highlight the importance of rapid feedback to ensure that evidence and insights are disseminated and used to inform policy and practice. Further, we show the importance of thinking forward to the next cycle of project planning and funding to ensure that relevant evidence is generated and used beyond the project. Systems that enable collaboration in the early stages of evaluation planning to identify and agree types of evidence needed and stakeholder engagement throughout the project lifespan are essential. In addititon, systems are needed that minimise time lags between project end and dissemination and facilitate knowledge transfer between and beyond projects and partners. The role of research partners is critical in bringing practice-relevant studies to publication (61), and reviewers and editors also have a role in this. Our study showed that funders and practitioners have a vital role in facilitating and contributing to knowledge-exchange activities. Multi-sectoral and multi-component projects, particularly where projects and evaluation are locally designed and implemented, need appropriate processes and systems to facilitate flows of information between all stakeholders. Without this, fragmentation of projects can lead to fragmentation of learning across organisations and individual stakeholders. In line with other studies (11, 48, 51), we show that cross-sector partnerships and networks appear to offer opportunites to improve knowledge-management and dissemination. Further research is needed to understand their value and how these can be implemented and embeded to help close current gaps in the evidence-based practice cycle.

Our findings have highlighted the important influences of differing stakeholder demands for evaluation, and resources for evaluation, in shaping the design and implementation of intervention evaluation. More critically, it showed the important influence of the underpinning organisational structures and systems, and the complex interactions between influences that act as facilitators or barriers to good practice, even when measures to address known challenges are put in place. Previous studies have identified a need for multi-level strategies to improve evaluation and for more research to understand these (11, 51); this study supports this view. We argue that stakeholders need to work together to understand, develop and implement systems to enable: (i) collaborative decision making; (ii) synergies between data needed for project delivery, participant engagement, accountability, research and evaluation; and (iii) timely knowledge transfer and dissemination. It is vital to improve our understanding of how influences interact to facilitate or limit good practice within evaluation. This will enable structures and systems to be developed and implemented that capitalise on factors acting as facilitators and that address

barriers, and help to ensure that effective interventions are adopted, and that ineffective interventions or unnecessary research are avoided.

Strengths and limitations

A key strength of this study is that we combined data from multiple sources, including evaluation reports and documents from 23 physical activity projects and from the programme as a whole, and data from 35 stakeholder interviews. A further strength is our use of a rigorous and transparent methodology to extract and analyse the data. The logic model that we imputed from the documents was based on the programme aims, objectives and intended outputs reported, and implied outcomes, and was further refined through consultation and interviews with key stakeholders.

There are several limitations of the study. Time lags between end of project delivery and publication mean that our appraisal of the evidence generated could not include the final programme summary evaluation that has been commissioned, and we may have missed additional publications from individual projects. The retrospective nature of the study limited the use of a more ethnographic approach. This may also have contributed to a lower response rate from project organisations and our ability to obtain documents related to project planning and the funding application. This time line also limited our ability to adopt a more collaborative approach to agree the theory of the programme as represented on the logic model.

Conclusion

We identified multiple influences on evaluation practice that can act as barriers and facilitators to good practice. These influences are context-specific and operate through a complex set of interactions. It is vital that commissioners, researchers and practitioners engaged in intervention evaluation or with an interest in improving evaluation and the generation of high-quality evidence, develop a better understanding of these influences and implement appropriate systems and processes to support good practice. Critically, organisational structures, systems and processes are needed to: (i) build and retain individual and organisational capacity for evaluation; (ii) enable collaborative and flexible decision making to identify and agree the most appropriate evaluation objectives, methods and types of evidence; and (iii) improve the transfer of knowledge and insights between stakeholders. This is critical to close current gaps in the evidence-based practice cycle, and ensure that relevant evidence is generated and used in a timely manner.

The findings highlighted the important role of the various partners in evaluation and dissemination. To improve practice it is essential to develop a better understanding of how partnerships and networks can be initiated and sustained. Chapter 5 applied network and

thematic analysis to gain a deeper understanding of the relationships between partners, processes and practice. This is critical to develop recommendations for how organisational structures, systems and networks can be implemented to support and improve evidence-based practice.

Chapter 5. Partnerships, collaborations and networks for evaluation: their use in facilitating evaluation, dissemination and evaluation use

Introduction

Chapter 4 revealed complex inter-connections between influences on evaluation practice and dissemination. Whilst the requirement for projects to use an evaluation framework and standardised data collection may have facilitated a systematic evaluation approach, the case study highlighted the important roles of different stakeholders in designing, conducting and using the evaluation. However, differences in the nature of the relationships suggested these may influence how partnerships can act as a facilitator or barrier to evidence-based practice.

There has been recent growth in multi-agency partnerships and networks to develop, implement and evaluate public health interventions, such as physical activity programmes. These include research-practice partnerships which bring together researchers and practitioners. Such partnerships can potentially provide opportunities for multi-sectoral approaches to address complex health behaviours. However, gaps remain in our understanding of influences on partnership working and their effectiveness, and the association between partnerships and the capacity to conduct and use evaluation. Chapter 5 therefore explored in more detail specific themes related to partnership working and evaluation use within the data generated through the interviews conducted as part of the case study. This aimed to advance understanding of how partnership working can best be implemented to improve evidence-based practice.

Background

As our understanding of the wider determinants of health behaviours has grown, there has been an increasing appreciation and understanding of the need for multi-agency and multi-component approaches to address complex public health challenges such as increasing physical activity (40). Examples include interventions that aim to address multiple influences on behaviour through adopting a range of modes of delivery and intervention functions, such as environmental restructuring alongside education. As a result, there has been an expansion of cross-sector and inter-organisational partnerships in intervention development, implementation, and evaluation. These include collaborations between physical activity providers and health organisations (65, 254, 298). In parallel, demands for evidence-based interventions have driven increasing interest in research-practice partnerships, which bring together researchers and practitioners with responsibility for programme delivery. These partnerships provide opportunities for collaborative

approaches to address complex health behaviours and to understand the implementation and effectiveness of complex interventions.

Evidence-based public health seeks to ensure that interventions are based on sound evidence, and involves three fundamental elements: evaluation, dissemination, and evaluation use (9, 270). Evaluation is defined as assessment of an activity, project or programme; it aims to provide accountability and facilitate learning for future practice (305). Dissemination is the process of communicating findings in ways that will facilitate their use in practice (66), and knowledge exchange is central to this. Following Alkin and King's conceptual model (55, 71), the term 'evaluation use' includes both use of evidence generated (findings use) and the effects of being involved in evaluation (process use). Their typology of evaluation use provides a framework to differentiate between the source or stimulus for use (findings or process), and how it has been used, for example: to inform direct actions (instrumental use); in improving knowledge or changing attitudes (conceptual use); or justifying decisions and actions (symbolic use) (55). Alkin and King also note that a broad definition of evaluation use incorporates the influence of an evaluation on wider systems (71). For consistency with the evaluation literature, we have applied these terms in our descriptions of evaluation use.

Research-practice partnerships (referred to as 'partnerships' hereafter) have been advocated as an approach to facilitate evidence-based practices (50, 61, 62, 65). Engagement of practitioners and policy makers in an evaluation can improve understanding amongst researchers of what evidence is relevant and valued for decision making in a real-world context, whilst engagement of research partners can bring knowledge and expertise to help identify and implement appropriate and innovative evaluation methods, and improve the rigour of evaluation (50, 64). Further, dissemination and evaluation use can be improved by research partners' understanding of the appropriateness of evidence for academic publication (61); this has the potential to increase the likelihood that evidence is taken up and used to inform policy and practice decisions. Yet despite this potential, there is currently limited evaluation and evidence use undertaken in organisations responsible for the design and delivery of health interventions, leading to little institutional learning and unnecessary cycles of programme re-invention. A key challenge is that we do not understand well how these partnerships can be shaped and implemented to improve practice.

Studies that have explored partnership working within physical activity and health promotion interventions have identified several benefits and challenges (50-52, 61, 62, 64, 65). Benefits include the generation of practice-relevant evidence, capacity-building, improved implementation of evidence-based practices, and access to additional funding and resources. Challenges include differing evaluation priorities and objectives, time scales, and organisational systems and cultures

(9, 49). Different stakeholders' demands for evaluation, the value they place on different forms of evidence, and how partners interact to implement appropriate evaluation methods within certain contexts influences the capacity to conduct and use evaluation (71, 73). Indeed, models of evaluation and evaluation use have focused on capacity building at the organisational level (73, 74, 306). Building on this, Labin et al.'s integrative evaluation capacity model (307) highlights the importance of collaborative processes.

Previous studies (49, 51), including our own work reported on in Chapter 4 (87), have highlighted the complex interconnections between influences on partnership working and evaluation practices. These have identified limitations in the empirical evidence and gaps in our understanding of organisational structures and processes within multi-agency partnerships and networks (51, 64). Questions remain unanswered regarding the influences on partnership working and their effectiveness, the value of being involved in partnerships to different stakeholders, and how partnerships and networks may influence evaluation, dissemination, and evaluation use (65, 72, 73). If organisations are to initiate and implement collaborative practices that are effective and sustainable, research that takes an inter-disciplinary approach is needed to understand evaluation practices and information flow between partners (51, 73, 74, 306).

To address these gaps, we explored the experiences and perceptions of stakeholders that were involved in partnerships to develop, implement and/or evaluate a national physical activity programme. The Get Healthy Get Active (GHGA) programme (308) was designed and funded by Sport England, the agency in England with primary responsibility for developing grassroots sports and getting more people active (274). Through the GHGA programme Sport England funded a portfolio of 33 projects, 31 projects within two funding rounds and two invited projects. Projects were delivered to communities across England between 2013 and 2018 (273). Projects were designed, implemented, and evaluated through various multi-agency partnerships (87).

All projects funded through the GHGA programme had the shared aim to increase physical activity in the most inactive adults and to generate evidence of the role of sport in improving physical activity and health. Projects differed in their target populations, secondary objectives, and approaches to partnership working and project implementation. The programme was chosen for this study, firstly as it exemplifies the multi-agency and partnership approach increasingly prevalent in health promotion interventions, and secondly because all lead organisations of funded projects were required by Sport England to engage an independent evaluation partner.

Objectives

1. To identify the partners involved in the evaluation of a multi-agency intervention, and the roles of these partners.

2. To explore how different stakeholders perceived and described the partnerships and their influence on evaluation.

3. To explore how different stakeholders involved in evaluation partnerships described the use made of the evaluation by themselves, their organisations or partners.

4. To apply the findings from objectives one to three, to develop a conceptual model of how descriptions of partnerships and networks may be associated with knowledge exchange and the capacity to do and use evaluation.

Method

This Chapter used data collected for the case study reported on in Chapter 4. We combined network and thematic analysis to describe the network of partners, and to identify themes within the data from the semi-structured interviews related to stakeholders' experiences and perceptions of those partnerships, the evaluation process, and evaluation use. We then adopted an inter-disciplinary approach to draw on concepts of evaluation use and organisational systems (55, 71, 73, 74) to help interpret our findings. Ethical Approval was obtained from the University of East Anglia Faculty of Medicine and Health Sciences Research Ethics Committee (REF:201718-133) (Appendix 2). Permission to conduct the research was received from Sport England (Appendix 3).

Study sample

We combined purposive and snowball sampling to identify stakeholders involved in the design and/or evaluation of projects or the overall programme. Organisations and stakeholders named as either the project lead or evaluation lead were identified from evaluation reports and documentation that had been shared with us. We contacted stakeholders directly via email or telephone to invite them to participate in an interview. Participants were asked during the interview to suggest other partners that they felt it would be useful for us to interview. We continued sampling until we had a sample that was representative of projects across the two funding rounds of the programme, different organisation types and stakeholder roles. Some stakeholders had multiple roles within the projects and programme, for example some evaluators were involved in evaluating more than one project, and some were involved at both project and

programme levels. Table 5-1 shows the final sample, which included 35 stakeholders from 16 projects and the GHGA programme.

Table 5-1 Sample of interview participants according to their role in the programme or projects

		Participants according to role	
Programme Component	**Projects**	**Delivery Organisation**	**Evaluation Organisation**
GHGA Programme	N/A	3 Sport England Staff	2 Evaluation Consultants
Round 1	6 Projects	5 Project Leads 2 Managers 1 Delivery Staff	5 Evaluation Leads
Round 2	10 Projects	8 Project Leads 5 Managers 2 Delivery Staff	8 Evaluation Leads

Data collection

Thirty-five interviews were conducted and audio recorded by the lead author (JF) between May and December 2019. Interviews lasted an average of 46 minutes (range 25-86 minutes). The topic guide was sent to participants in advance. This included questions that asked them to reflect on their experiences of partnership working and its influence on the evaluation, and their perceptions about how the evaluation had been used by themselves or their organisation(s) (see Appendix 6). Interviews took place over Skype, telephone or face to face, and one respondent responded via email. Interviews were transcribed verbatim and given a unique identifier to de-identify stakeholders, and then uploaded into the NVivo12Pro software for analysis.

Data analysis

To identify partners involved in the project and programme evaluation (objective one) we applied principles of network analysis. Firstly, we coded each interview transcript and each project as a separate "case" within NVivo12Pro. Secondly, we coded any named individuals, groups or organisations that were mentioned in the content of the transcripts as being involved in the programme or project evaluation as additional 'cases'. To de-identify individuals and organisations each of these was also given a unique number.

Details of the projects, individuals and organisations were then exported into an Excel spreadsheet for further analysis. To preserve anonymity, individuals were grouped at the organisational level. These were coded as organisational types to describe the key attributes of each partner; for ease of interpretation these were then grouped into broader sector-based categories (Health, Sport, University and Other). 'Other' included public, private, and third-sector organisations. Each "case" was also coded by role (Funder, Lead Organisation, Evaluator, Delivery

Partner, External Partner). The code “delivery partner” included any partners engaged in project recruitment, implementation or evaluation that were identified as playing a role in the evaluation; “external partner” included those identified as being connected, but not directly involved, in the project or programme evaluation.

Within Excel we created a spreadsheet to identify the connections between partners from reported descriptions of the projects and the interview data. This information was imported into UCINET (309) to generate a visual representation to describe the network of partners included in our sample, and their role in the project and/or programme evaluation.

To explore how different stakeholders described their experiences of partnership working, the nature of those partnerships, and their influence on evaluation and evaluation use (objectives 2 and 3), we applied thematic coding to the interview data. Initial codes were identified *a priori*, informed by our research objectives and the interview schedule. Other themes were identified iteratively through the processes of repeated familiarization, coding and recoding. Codes were reviewed and organised into categories (by JF) to develop the draft coding framework, which was then discussed and agreed by all authors (Table 5-2). Framework analysis was used to compare across and between stakeholder types and projects.

Table 5-2 Coding framework for the thematic analysis of interview data

Key Themes	Sub-themes
Partnership characteristics	Roles and responsibilities
	Nature of the relationship
	Continuity
	Engagement
	Communication
Evaluation use	Use of Findings
	Use of Process
	Instrumental (direct action)
	Intervention maintenance
	Informing local decisions
	Informing national decisions
	Capacity building
	Catalyst for change
	Developing partnerships
	Initiating
	Embedding

To explore how the descriptions of partnerships may be associated with knowledge exchange and the capacity to do and use evaluation (objective 4), we drew on concepts of evaluation use and organisational systems to help interpret our findings from the network and thematic analysis, and to develop a conceptual map. This was refined and agreed through discussion with all authors.

Results

The results are presented within four sections, reflecting the four objectives of the research. Firstly, we describe the partners involved in the project and/or programme evaluation. Secondly, we describe how stakeholders described the partnerships and how partnership working influenced the evaluation. Thirdly, we provide a synthesis of how stakeholders described their use of the evaluation. Lastly, we explore how descriptions of partnerships and networks may be associated with knowledge exchange and the capacity to do and use evaluation, and present and explain our conceptual model of the flow of information and processes between partners.

1. The partners involved in the project and/or programme evaluation

Figures 5-1 and 5-2 show the partners involved in programme and project evaluation. Partners are grouped and colour coded by the categories used to describe their main role within the partnership (funder, lead organisation, evaluation partner and delivery partners). The sectors used to group the organisational types (sport, health, university and other) are shown by symbol shape. The purpose of these maps is to illustrate the complexity of the network, rather than to drill down to examine the complexity in detail. They serve as a descriptive tool on which to base the exploration of the characteristics of the partnerships and discussion of influences on partnership working and their effectiveness.

Figure 5-1 shows the formal partners reported to have been involved in the delivery and evaluation of the 16 projects and the programme. Projects brought together a range of private, public, and voluntary organisations and individuals from different sectors to facilitate recruitment and implementation. Most involved partnerships between: (i) sport and physical activity providers such as County Sports Partnerships, leisure centres, National Governing Bodies, and community-based clubs and individuals; (ii) partners from the health sector such as public health teams and primary care, and (iii) Local Authorities. Eleven of the projects engaged a university evaluation partner and two engaged evaluation consultants. Three projects were university led, and each of these also led the project evaluation (shown as Lead & Evaluator in Figure 1). Sport England engaged two consultancies to conduct summative evaluations of the overall programme following rounds one and two. Figure 5-1 shows that within each project-based group of partners, the project lead organisation is the central link between partners. It also shows two cases where there are connections between projects via a common evaluation partner. These connections represent flows of information. The dotted lines represent where boundaries exist between the key partner types and show how these intersect the connecting lines and potentially interrupt flows between partners.

Figure 5-2 shows the wider network of both formal partners and informal connections between individuals and organisations identified from the interview data. This reveals a more complex set of relationships, with connections between individuals and groups that transcend project and organisational boundaries within the network and appear as additional networks nested within the overall programme network. The additional partners include charities, local services, and community-based groups, mentioned by stakeholders as essential partners in project evaluation. Stakeholders described the role that these partners played in recruitment, undertaking baseline and follow up data collection, and building relationships with participants, which in turn enhanced response rates. Stakeholders from two projects also mentioned links to additional universities that supported, but did not lead, the project evaluation.

Figure 5-2 also shows (in red) external partners that were not directly involved in the project or programme evaluation but that were mentioned as influencing either the evaluation methods adopted, dissemination or evaluation use. This included individuals and organisations that informed programme-level decisions about project evaluation design, organisations connected by movement of staff between them, and organisations involved in dissemination activities.

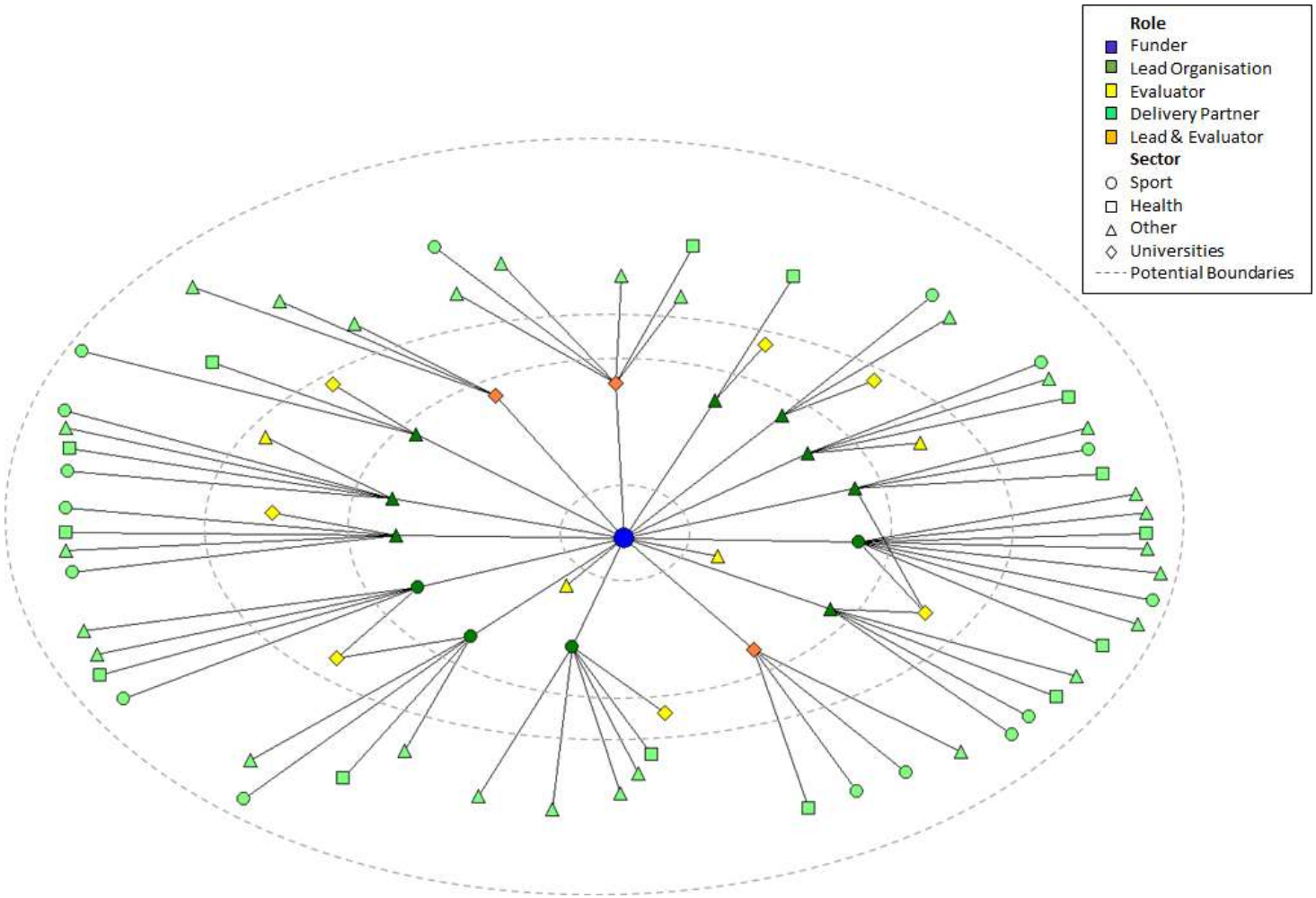

Figure 5-1 The network of reported partners

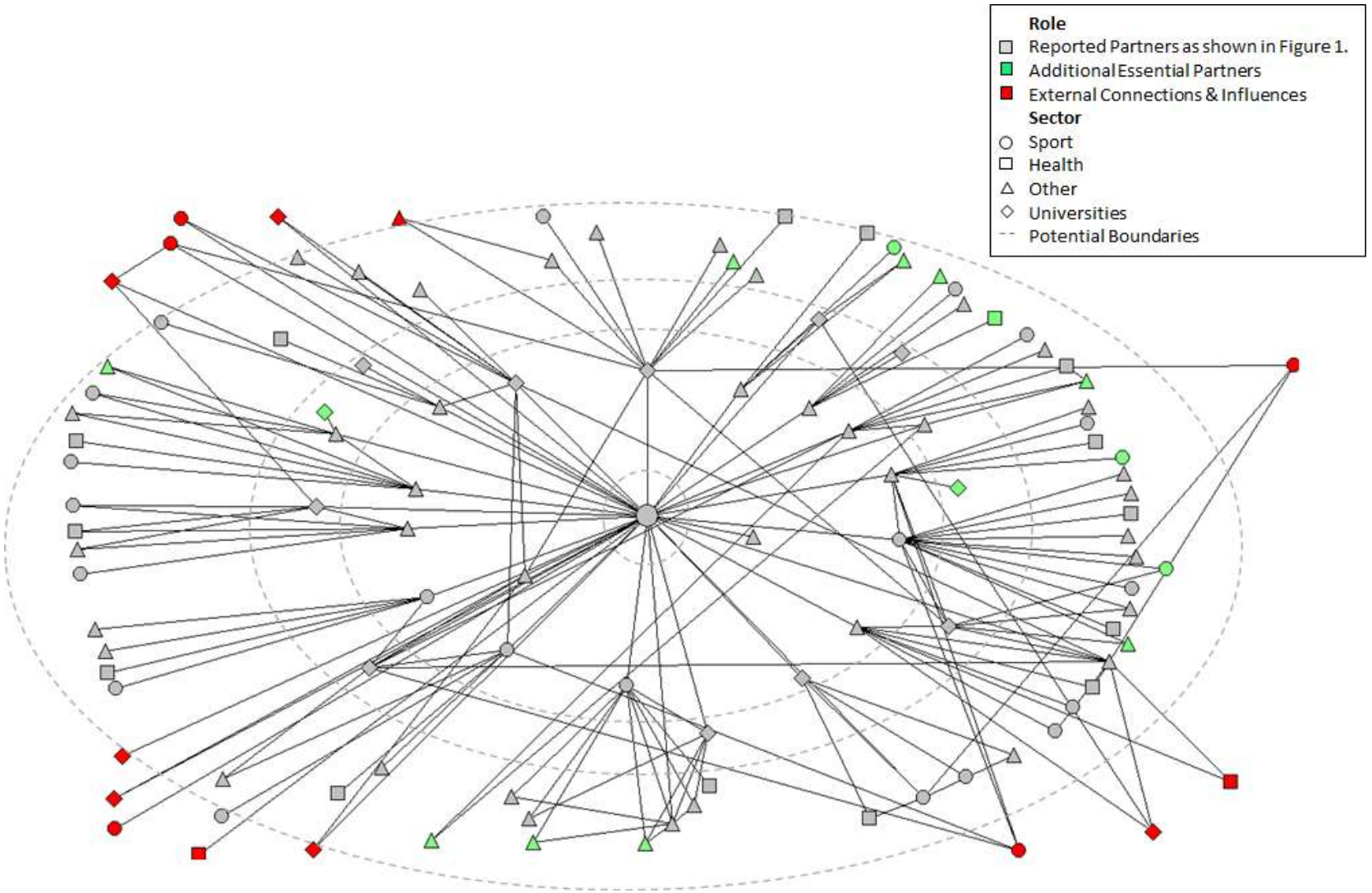

Figure 5-2 The network, showing additional partners and relationships identified from the interviews

2. Partnership characteristics and their influence on evaluation

Partnerships were described by their roles and responsibilities (as applied in our categorisation in Figures 5-1 and 5-2) but were more fully described by the nature of the relationships, collaboration, and communication. How stakeholders described each of these four themes, and how these were perceived as facilitators or barriers to partnership working and evaluation, are summarised below, and explored in more detail in Table 5-3.

Table 5-3 Partnership characteristics, and facilitators and barriers to successful partnerships and evaluation as perceived by stakeholders

Partnership characteristics	Facilitating influences	Challenging influences
Roles and Responsibilities		
Conduit and Co-ordinators	Key stakeholders that act as a bridge between partners to co-ordinate and manage relationships and activities	Staffing structure, funding and resourcing does not always facilitate a co-ordinating role
Leadership/Driving force	Having partner(s) that can act as the architect for the project, relationships and evaluation	Staffing structure, time and resource are needed
Expert/Adviser	Evaluation expertise is a valued source of advice and guidance	Tensions between evaluation rigour and pragmatic approaches require recognition of and value placed on differing perspectives and approaches to ensure evaluation works in practice
Data collector and/or Recruitment	Critical resource and capacity for successful evaluation	Requires understanding and agreement of roles and responsibilities, training, capacity building and buy-in to evaluation processes
Relationships		
Building good working and close relationships is key	Accessible, approachable and adaptable partners are vital to build close, open, honest relationships and trust, and facilitate candid discussions and collaboration	Building relationships and trust is critical, but takes time
Adaptability	Adaptability facilitates pragmatic approaches to evaluation and problem solving	Evaluation rigour can be seen as limiting adaptability, and impacting negatively on delivery objectives
Local relationships	Local relationships facilitate relationship building and regular communication	Geographically distanced partnerships negatively influence relationship building and partnership working
Reciprocal relationships	Reciprocal relationships and shared understanding of expectations and mutual benefits are important for collaboration	Disconnect or tensions between partners and perceptions of a lack of interest may arise from a lack of understanding of expectations, targets, priorities and pressures
Collaboration		

Table 5-3 Partnership characteristics, and facilitators and barriers to successful partnerships and evaluation as perceived by stakeholders

Partnership characteristics	Facilitating influences	Challenging influences
Collaborative	Recognition of value in bringing differing perspectives together is vital to the evaluation and getting buy-in from partners	Transactional relationships negatively impact buy-in and engagement of partners in evaluation
Level of engagement	Hands-on approach and engagement with activities and partners is critical for developing an understanding of the project, ensuring data collection, building relationships, getting buy-in and embedding processes	Hands-off partnerships negatively impact partnerships and evaluation. Time and effort are needed to build trust with delivery staff and participants prior to data collection
Prior connections, previous collaboration	Established relationships facilitate shared understanding and early collaboration to develop an evaluation plan that works for all	Newly formed partners require time to build relationships and understand needs for the project and evaluation
Embeddedness	Embedded partnerships, mature relationships, better understanding of how "evaluation ready" the organisation is, greater engagement with evaluation; embedding all partners, including evaluators, in project management structures facilitates regular communication and collaboration	Where partners were not embedded time was needed at the start of projects to build relationships and to agree roles and priorities for the evaluation
Continuity of relationships	Early collaboration enables partners to influence evaluation design and integration of evaluation into project implementation, continuity of staffing facilitates consistency of approaches, relationships and communication	Short funding cycles and staffing structures do not always facilitate early collaboration or continuity; staff turnover (late starts, early departures) impact continuity even where the organisational partnership is maintained
Commitment	Commitment from all partners is essential	Tensions where not all stakeholders were committed to the evaluation, where it was seen to interfere with delivery, or where evaluators had differing priorities
Communication		
Regular communication	Facilitates engagement, review, knowledge exchange and shared understanding, mechanisms are needed for formal and informal communication	Challenges of sustaining active participation by different partners and through different stages of planning, implementation and reporting can limit ongoing evaluation, feedback, adaptation and evaluation use
Appropriate communication	Two-way dialogue, bringing the right people together and use of appropriate language to enable shared learning is critical	Tensions between partners can arise from differences in understanding of terminology, language, and differing priorities. Collaboration requires differing perceptions and voices to be respected and valued

There was variability in the way stakeholders described their own experiences of partnerships, but consistency in the way they described strengths and weaknesses of partnerships. Within the four themes, we identified key processes and partnership characteristics critical to effective partnership working and evaluation. Based on these processes and characteristics, we formulated an ABC for effective partnership working: A. Approachable, adaptable, and accessible partners, B. Building relationships and building capacity, and C. Communication, collaboration, and continuity. These cut across the themes, and highlight the relationships between processes, such as communication and building capacity, and characteristics that influence these processes, such as the approachability, accessibility and adaptability of partners.

2.1 Roles and responsibilities

Stakeholders reflected on the importance of understanding, agreeing and valuing the differing roles and responsibilities of partners, and also the benefits from partners bringing different skills and expertise to facilitate evaluation:

> *"You need to be able to draw on a number of different skills. I think the beauty of having the University involved in this project is that you can draw on expertise quite quickly."* (Project Lead and Evaluator)

In other projects stakeholders reflected on the challenges and tensions between partners, and the need for a shared understanding of expectations and the value placed on differing perspectives and approaches to ensure evaluation works in practice:

> *"The biggest challenge at the time, was the interest from the evaluation partner, and not understanding the bigger picture … and how we wanted to show that we were having a big impact. It was too much of a facts and figures focus."* (Project Lead)

> *"There is a disconnect between them* [evaluators and practitioners], *and there still remains to be a disconnect but I think it's just trying to appreciate as best you can each other's roles really, especially for the first year of this project that really didn't happen."* (Evaluator)

Others reflected on the importance of key partners acting as a conduit or bridge to facilitate partnership working, and to co-ordinate and manage relationships and activities:

> *"I do think that academia has different outputs and objectives to policy and practice. Having an understanding and being able to be a bit of a bridge between the two was important."* (Project Lead)

> *"Everyone is driving towards the same thing, but they have to do it in different ways because they are either contractually bound, or they are limited by their resources, and so that partnership network was essential. That community of sport and physical activity network was*

a central way in which we could have debates and discussions but crucial in that partnership was the role I took. You need an architect really to pull that together." (Project Lead and Evaluator)

2.2 Relationships

Relationships in which partners found each other to be accessible, approachable, and adaptable were described as essential to facilitating open and honest conversations, and to enabling capacity building and collaborative approaches. Stakeholders recognised that building relationships, trust and capacity required time and investment:

"The partnerships that were really key were myself with the project lead, project coordinator and program manager, we had really good working relationship… having the key relationships with them was useful. Also, we had to have really good relationships with those who were actually delivering the intervention or the programme and exercise, having good relationships with them was absolutely essential for enabling data to be captured." (Evaluator)

"I think the partnership comes down to an investment of time into building it and a mutual benefit in doing it. We put a lot of time and energy into the development of the relationship, and we even now do try to touch base regularly. Collaboration is very different to working in partnership… it really takes time to embed if you think about building trust, respect, honesty and I think we have built on a lot of those. So it is a very open, honest, transparent relationship." (Project Lead)

Relationships with the funding partner were described variably across projects and between different partners within projects. Experiences of the relationship with the funder were also felt to have changed over the course of the programme's life cycle. Nine participants (representing delivery partners, lead organisations and evaluators) commented on the supportive relationship between themselves and the funding organisation. Stakeholders also referred to the important role that Sport England played in bringing projects and partners together through knowledge exchange events to facilitate capacity building and shared learning. Nevertheless, nine participants (representing delivery staff and evaluators) described the relationship as transactional, and commented on limited opportunities for communication or engagement in knowledge exchange and feedback.

2.3 Collaboration

Collaboration was thought to be facilitated by early and ongoing engagement of evaluators in the project implementation and of delivery staff in the evaluation. This was described as mutually beneficial; evaluators developed a better understanding of the project and needs of stakeholders,

whilst delivery staff were more likely to buy-in to the evaluation processes when time was taken to train and explain the purpose, methods, and importance of evaluation.

> *"You can't just tack it on, you need to be there from the start and to be involved. When everyone gets their opportunity to give their thoughts and ideas everyone is engaged, and that makes a big difference. People can see what they're going to get out of the evaluation, it makes a better experience for everyone, and then measures get completed. Without which you don't have an evaluation. When everyone's bought into the process, that's when it works."* (Evaluator)

Where there had been a prior connection, or working relationship, participants reflected on this facilitating a closer partnership. Local partnerships were thought to enable closer relationships, more regular communication and engagement with project activities, and better understanding of local needs and priorities. The findings also highlighted the influence on continuity of organisational structures and processes, such as funding and staffing. Stakeholders described late project starts and early staff departures within project teams as a challenge to building relationships, and to planning, agreeing, and implementing evaluation practices:

> *"That consistency, which is always difficult, people do leave, continuity really helps if you can get it, in terms of relationship."* (Project Lead)

> *"There were changes in the clinical team, changes in the council team, changes from the delivery teams, and changes in the evaluation team, that's really hard if you've not got the good relationships there. ...Since the evaluation got published there's been a ton of changes in staffing again, I do wonder if it was still the same leads from the beginning whether that would have been more broadly disseminated."* (Evaluator)

2.4 Communication

Communication was described as a key process to facilitate knowledge exchange, and in turn to build capacity to both do and use evaluation. Communication that was regular, timely and appropriate was seen as critical to effective partnership working, whether between funders, delivery staff, project leads or evaluators. Limitations in communication and feedback in the later stages of the programme, and particularly following final reporting, were identified as barriers to knowledge exchange and evaluation use:

> *"It would have been useful to have a little bit of communication when the report was submitted."* (Evaluator)

3. Evaluation use

We identified the following themes related to evaluation use: findings or process use, instrumental use (direct action), capacity building, being a catalyst for change, and initiating and embedding partnerships. There was consistency in the way stakeholders with differing roles within and across projects described their experiences and perceptions of evaluation use.

Stakeholders described their experiences holistically. For example, they did not always differentiate between findings use or process use, or between engagement in partnership working or the evaluation itself. Project and programme stakeholders described how the evaluation as a whole had been used to enable the project or elements of the project to continue, and to inform approaches used in subsequent projects, or in future commissioning activities:

> *"We have massively used it as a way of trying to develop better tools that will measure and do what we want ….which has certainly built on the experiences not just of this programme but across the whole organisation and how we support other organisations."* (Funding organisation)

> *"It made the biggest difference to how we tackle and move towards tackling inactivity locally, and so that is not necessarily about the evaluation process but it is the impact and outcome of that whole learning from the evaluation. … The legacy of the project has carried on, it has had a massive impact on the physical activity strategy."* (Project Lead and Evaluator)

> *"We have secured further funding and this was probably a part of it, but that was halfway through, not the end evaluation report."* (Project Lead)

> *"Through that we've got a three-year contract to deliver activities as part of a different project…we wouldn't have got that without the GHGA project and the evaluation, the evidence that we had from that." (Project Lead)*

These observations illustrate the instrumental use of evaluation, but also the value of concurrent evaluation and intermediate feedback, rather than purely summative evaluation and evidence generation. Some stakeholders commented on their own limited understanding of how the evaluation had been used at the programme level:

> *"I don't know how useful the evaluation has been, in terms of the report which we submitted"* (Evaluator)

Capacity building was more explicitly linked to process use, and was identified as increasing knowledge, skills and attitudes. Stakeholders described their learning from the experiences of being involved in evaluation processes and of being exposed to different evaluation approaches and methods. Where formal training was mentioned, this related to training for programme

delivery or data collection methods. Developing a better understanding of the purpose and importance of evaluation, and gaining buy-in from all partners, were seen as critical for successful evaluation, and to bring about changes to evaluation practices during and after the project. Stakeholders described their learning as a catalyst for changing practices, and, in five projects, for changing staffing structures, with the creation of insight and evaluation officer roles being embedded into organisations.

> *"The learning has transferred across to other projects, the importance of capturing really good quality evaluation. We have developed evaluation resources and run training sessions for organizations locally to share our learning with the sector. I would say this project was the catalyst."* (Project Lead)

> *"It has been huge; it shapes much of what I do on a day to day basis and probably the same for the other people here. Embedding that evaluation, that partnership working across everything we do, I think that's crucial."* (Delivery Partner)

Stakeholders at both the local project and national programme levels also reflected on the value of initiating cross-sector partnerships, opening doors for conversations, and developing networks.

> *"One of the big things that came out of the project was the steering group that was set up at the start, that has led to more and more partners coming round the table and that is because people were hearing about it and wanted to be involved in the project and they were bringing their own projects and their own ideas to the table as well, so certainly evidence from my point of view that that was leading to more partnership working locally."* (Evaluator)

> *"I think it has been quite significant but isn't necessarily that easy to quantify or that tangible. One of the effects of GHGA has been this much closer partnership between sporting and some of the health partners, and Public Health England nationally. I think having the evaluation arrangements, for all their imperfections, were probably more rigorous than we would have had historically and has been helpful in getting some of that buy in and engagement with health and wellness … through evidence, but also through relationships and wider political changes, a shift has happened … and I think the evaluation has been relevant to winning some of that support or some of that shift." (Funding Organisation)*

One stakeholder reflected on the value of relationships with the wider network evident in Figure 5-2:

> *"from my own personal relationships, I still have those networks … that is how I get most of my information, and find out the best things to be doing."*

4. A conceptual model of the relationships between partnerships, processes and partnership characteristics that facilitate evaluation, dissemination and evaluation use

To address the final objective we applied the findings from the previous objectives to develop our conceptual model of how partnerships and networks may be associated with knowledge exchange, and the capacity to do and use evaluation, this is shown in Figure 5-3. Informed by the network maps (Figures 5-1 and 5-2), the model shows partners with differing roles within the projects, the programme and external to the programme, connected by arrows. Figure 5-2 revealed groups of connections and partners which transcended project and programme boundaries, through having differing roles, connections to external partners, or staff mobility. They can be viewed as smaller networks nested within the overall network. These connections represent important opportunities for information flow between partners and across the network, but also where alignment of processes along connecting lines is required to facilitate effective partnership working.

As in Figures 5-1 and 5-2, our model shows the groups of partners separated by boundaries (dotted lines) which represent potential interruption to flows of information and barriers to alignment of processes. For example, differences in priorities, organisational structures, and a lack of a common language between evaluators and practitioners can act as barriers to communication, collaboration and building capacity. Differences in organisational structures and cultures can influence time lags in engaging staff, agreeing evaluation processes and in communicating and providing feedback.

Thematic analysis highlighted key processes that are interlinked, such as communication, building relationships and knowledge exchange, and how these are essential to build capacity to both do and use evaluation. These are shown in boxes on the left of our model spanning the connections between partners. Our analysis also showed how partnership characteristics can negatively or positively influence these partnership processes, and in turn influence evaluation, dissemination and evaluation use. Partnership characteristics identified as important influences on the success of the partnerships and the evaluation, are shown within the boxes on the right of our model. Within the boxes we highlight (in bold) the processes and characteristics that we identified from the thematic analysis as critical to effective partnership working (the ABC). Our model illustrates how boundaries may act as barriers and close, effective relationships as facilitators. The model can be used to understand, and implement, approaches to support partnership working.

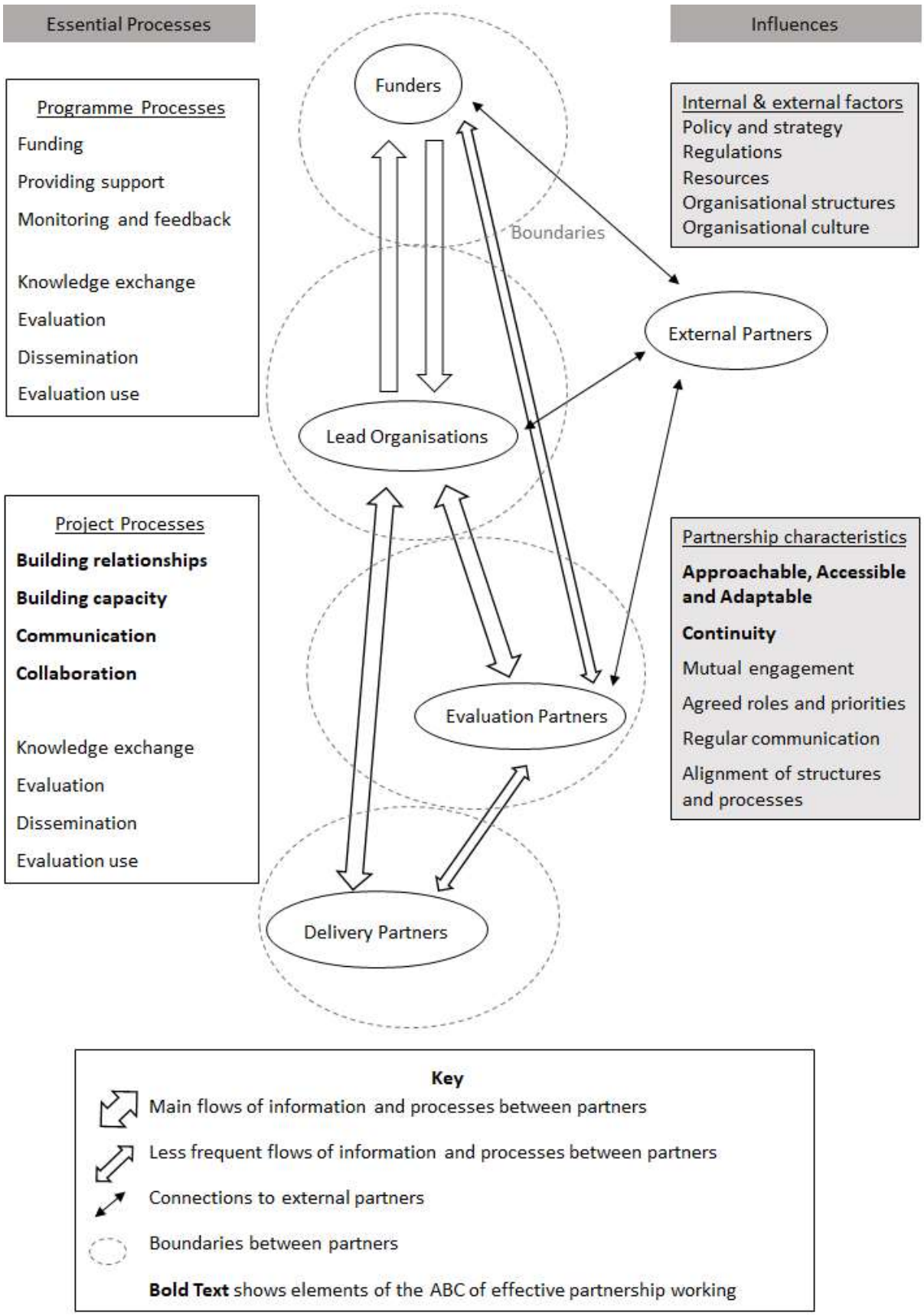

Figure 5-3 Conceptual model of the flow of information between partners and networks, and the relationships between processes, partnership characteristics and practices to support evaluation, dissemination and evaluation use

Discussion

We identified a complex network of partners that were involved in or influenced programme and project evaluation. By combining network analysis with framework analysis, we have shown how partnership characteristics can influence the flow of information and alignment of processes between partners, and how this in turn influences evaluation, dissemination and evaluation use. We have developed a conceptual model to help visualise this. Our model builds on concepts within previous models of evaluation and evaluation use that focused on capacity building for evaluation at an organisational level (73, 74, 306, 307), and on concepts of partnership working (61, 62, 65). Through the model, we have highlighted important elements of partnerships and networks, and how these are essential for collaborative evaluation activities, quality evaluation, knowledge exchange, shared learning, and evaluation use. Compared to previous models, our study offers a deeper understanding of the roles of different partners within multi-agency interventions in evaluation, and how characteristics of partnerships and networks can be shaped to positively influence evaluation processes and practices.

Network analysis revealed a complex set of formal and informal relationships, and groups of more, or less, connected partners within wider programme and external networks. These connections are essential for knowledge exchange; they provide the potential for building capacity and professional development to improve evidence-based practices for individual and organisational partners. Communication to support multi-directional flows of information between partners is crucial. Our findings showed that communication and knowledge exchange were critical to evaluation, and to the use of both evaluation findings and process in multi-agency interventions.

Through the thematic analysis we identified important benefits of research-practice partnerships, such as access to expertise, improved evaluation rigour, generation of practice relevant evidence, and capacity building, that support findings from previous studies (51, 61, 65). We also identified key processes and partnership characteristics that were critical to successful partnership working and evaluation. For example, appropriate and regular communication, and early mutual engagement were essential to facilitate effective collaboration, communication and capacity building. Close relationships in which stakeholders were, and were seen to be, approachable, accessible, and adaptable were important. Continuity of partnerships facilitated these processes. By including these in our model as an ABC of effective partnership working for evidence-based practices, we highlight their importance so that practitioners, funders and evaluators can take steps to address these when engaging in partnership-based evaluation. Funders and commissioners play a crucial role through their requirements for evaluation and the information and support they provide for project evaluation, knowledge exchange, and feedback. They need

to implement organisational structures and processes that support (i) initiation and continuity of relationships and practices, (ii) alignment of processes to minimise barriers to information flows across boundaries between partners, and (iii) development of systems to support knowledge exchange and capacity building.

In line with previous studies, our findings highlighted the context-specific and changeable nature of partnerships (65), and the complex inter-connections between influences on partnerships and practices within multi-agency public health interventions (49, 51, 310). To facilitate information flow across boundaries there needs to be alignment of organisational structures and systems, time scales and communication approaches. By identifying where boundaries may exist within the network, where they may limit information flow, where time-lags may occur, and where knowledge may be lost or gained, our model helps explain the relationships between partnership working and processes fundamental to evidence-based practice. Staff movement represents the potential for both loss and gain of learning and capacity from organisations. The net effect depends on their role and position within the network, but funding and organisational structures that minimise staff loss are vital. Knowledge exchange via informal or personal connections for stakeholders in the wider network was also important. We suggest there may be added value in releasing the intrinsic value of these "hidden communities of practice" by developing organisational structures and processes that systemise networking and embed knowledge sharing practices. Communities of practice in health settings offer opportunities for capacity building and knowledge exchange to support professional development (311). To realise these benefits of networking, and to make these accessible to all stakeholders at any stage in the evidence-based practice cycle, there is a need for sustainable networks to bring researchers, policy makers and practitioners together and act as a conduit for knowledge exchange, advice, and professional development. Both the research community and those with responsibility for strategy, policy and practice decisions have a role to play in facilitating this at the local, national or inter-national level.

In a similar vein to previous conceptual models of evaluation and evaluation use (73, 307) we offer our model as a contribution to what we see as an ongoing enquiry and conversation to improve evaluation and evidence-based practices in multi-agency public health interventions. We have drawn on concepts from organisational learning and systems to help interpret our findings, rather than applying specific theories. For example, we have described clusters of connected partners at the project level, nested in wider networks operating at the programme level and with partners external to the programme, and have identified boundaries as potential barriers to the flow of information and alignment of processes, much like those described in systems theories.

We have not however delved more deeply into systems thinking or communities of practice, and highlight these as areas that would add value in further research.

Strengths and limitations

An important strength of this study was the support we received from Sport England to conduct the study, and the access to and participation from stakeholders at all levels of the GHGA programme. Another strength is our use of empirical evidence and inter-disciplinary approaches to inform our analysis and development of the conceptual model. This has enabled us to develop a novel view that builds on and integrates current understanding of partnerships and networks with an understanding of evaluation and evaluation use. There are limitations in our approach. The full extent of formal, and especially informal networks, is likely under-represented, due to the retrospective nature of the data collection process, and the grouping at organisational and sector level which was essential for anonymity. In future studies of this nature, a more systematic, prospective method of data collection for the network analysis would be beneficial.

Conclusion

Partnerships and networks represent a complex set of informal and formal relationships that have the potential to positively influence evidence-based practice. Our conceptual model highlights key processes and characteristics of partnerships that facilitate evaluation, dissemination, and evaluation use, the three fundamental steps in evidence-based practice. The model, highlights the importance of relationships and communication to facilitate the flow of information between partners and the network, and where there are potential barriers between partners. Based on the ABC of effective partnership working for evidence-based practices this research has identified key processes and influences as critical components in evidence generation and knowledge exchange: A. Approachable, adaptable, and accessible partners, B. Building relationships and capacity, and C. Communication, collaboration, and continuity. The model can be used by funders, practitioners, and evaluators engaged in multi-agency interventions and research-practice partnerships to identify important processes and influences that can shape the success of partnership working and evaluation practices. If partners are to realise the benefits of partnerships and networks, it is essential that they understand and implement these, and invest time, resources and effort to develop the structures and systems to support them

Chapter 6. Discussion and conclusions - the use, usability and usefulness of strategies to support evaluation and evidence-based public health

6.1 Introduction

This concluding chapter of the thesis summarises the key findings, provides a discussion of how these contribute to knowledge, and considers their implications for research and practice. It then provides a reflection on the research approach and methods applied in the thesis, including a discussion of their strengths and limitations. This is followed by a personal reflection on the experiences of conducting the research and writing the thesis. Lastly, it provides some concluding comments.

6.2 Summary of findings

This thesis explored strategies to support and improve evaluation practices and public health evidence-based practice, taking a programme of interventions to promote physical activity as a case study. The research has highlighted the complex inter-connected influences on practice. It has considered in detail two particular strategies that are recommended to support and improve practice (46, 51, 60-62) - the use of evaluation frameworks and research-practice partnerships.

The scoping review (Chapter 2) identified seventy-one evaluation frameworks, more than in any previous reviews (63), and highlighted the extensive range of guidance available. In considering the use of frameworks (Chapters 3 and 4) the thesis has shown that such frameworks can improve the quality of evaluation when applied appropriately, but that there is considerable variability in their use and reporting. Indeed, the systematic review (Chapter 3) highlighted that frameworks are under used and that there is much heterogeneity in the way intervention and evaluation components are described within published evaluation studies. As reported in Chapter 4, variability in the use and reporting of evaluation frameworks was evident even when a standardised framework and method of data collection was stipulated as a requirement of funding. This limits the comparability and transferability of evaluation studies and their findings, and suggests that providing evaluation guidance is not, of itself, sufficient to ensure the generation of high quality, generalisable evidence, and its' dissemination and use.

The typology and mapping of evaluation frameworks presented in Chapter 2 and the checklist of indicators to appraise the reporting of frameworks within studies (Chapter 3) aim to address some of these challenges. The mapping and typology can be used by stakeholders to identify and agree frameworks relevant to their needs, whilst the indicators can be used to facilitate systematic reporting of how a framework was used. These tools can also be used by researchers to identify

where to focus their efforts to address the gaps in the guidance provided by existing frameworks, and reviewers to appraise the quality of reporting in evaluation studies. However, the findings from the research suggest that evaluators may need more support to improve real-world evaluation and evidence-based practices.

Inter-agency and research-practice partnerships are the other key strategy that the research reported in this thesis explored in detail. Chapters 4 and 5 showed that such partnerships can improve the quality of an evaluation and help to build capacity to conduct and use evaluation. Yet, the nature of relationships and the characteristics of the partnerships were found to be context-specific and changeable, and this influenced the success of partnerships and evaluation. Critically, the thesis highlighted that within inter-agency interventions, researchers, funders, policy makers, practitioners, and reviewers all have a vital role to play in shaping structures, systems and processes to support and improve evaluation, dissemination and evaluation use. Such networks are important to ensure that relevant evidence is generated and makes its way into the public domain in a timely manner, and to help close current gaps in the evidence-based practice cycle. Findings from Chapter 4 illustrated how organisational structures and systems are needed to: build and retain individual and organisational capacity for evaluation; facilitate collaborative and flexible decision making; and improve the transfer of knowledge and insights between stakeholders.

The conceptual model presented in Chapter 5 shows the relationships between partners, processes and partnership characteristics, and how these can be shaped to positively influence evaluation, dissemination, and evaluation use. Key elements of effective partnership working highlighted in the model include: having approachable, adaptable, and accessible partners; building relationships and capacity to conduct and use evaluation; and communication, collaboration and continuity. This model can be used to guide understanding and implementation of structures, systems, and processes to facilitate effective partnerships and evaluations.

By exploring in detail specific strategies to support evaluation and evidence-based practice, and to appraise their use, usability and usefulness, it is hoped that this thesis has made a meaningful contribution to the evolving understanding of and improvement in evaluation and evidence-based practices within public health contexts, including interventions to promote physical activity. The implications of the key findings for research and practice are discussed in more detail in the following sections.

6.2 Discussion and implications for research and practice

Evaluation, dissemination and evaluation use are fundamental processes to facilitate evidence-based public health, as depicted in Figure 1-1. These processes are essential to bring about

sustainable change and meet targets to increase physical activity levels and improve the health of individuals and populations. This thesis has shown that strategies, such as the use of evaluation frameworks and research-practice partnerships and networks, can facilitate evidence-based practice when implemented appropriately. However, it has also shown that there remains considerable variation in the use of these strategies, despite the value placed on evidence-based practice by both the research and practice communities, and the plethora of guidance and recommendations that have been developed. If we are to ensure that effective evaluation is undertaken, relevant evidence is generated, and evidence is shared and used to inform decisions, strategies to encourage and support good practice need to be understood and implemented. The use of frameworks and research-practice partnerships, which are two strategies that this research explored, are discussed in the following two sections of this chapter.

6.2.1. The use, usability and usefulness of evaluation frameworks and guidance

Evaluation frameworks are intended to facilitate a systematic approach to evaluation (46). The reasons for their limited use and reporting in evaluation studies remains unclear, though is likely influenced by a range of factors. The abundance and breadth of frameworks has previously been suggested as a barrier to evaluators being able to find and make use of the most appropriate guidance (54, 63). Given the importance of generating practice-relevant evidence and building an evidence base on which decision makers can draw, it is important to direct practitioners and evaluators to appropriate guidance, and researchers to where further guidance is needed. Reviews of frameworks (63, 69, 101) and tools to signpost to relevant frameworks, such as those presented within the scoping review (85), are just one solution that can help to address this challenge.

Unfortunately, providing and signposting to the guidance will not necessarily guarantee uptake or effective use of the guidance. There are several factors that can influence use of evaluation guidance. These include resourcing, organisational constraints, capacity of evaluators, and other contextual barriers that have been discussed previously in the evaluation research literature (49, 51, 52). The demands for evaluation and value placed on differing forms of evidence by different stakeholders are driving forces for evaluation recommendations, and for the methods used. Where this is determined by the funder, as in the case study used here, this can lead to tensions and difficulties in implementing standardised methods consistently across differing projects and contexts. As both the scoping review and case study showed, there are strengths and gaps in the available guidance. Strengths that were identified included strong guidance on process evaluation and using logic models, which reflects the growth in appreciation of the importance of these evaluation components to gain a detailed description of the intervention under evaluation and to understand it's causal mechanisms. The gaps identified reflect areas that have been expanding or

are emerging as areas of interest within evaluation and implementation research; for example, guidance on participatory evaluation, sustainability, and evaluation of wider non-health and unanticipated outcomes. Nevertheless, it should be noted that the guidance on evaluation is continually developing. For example, there has been an expansion of frameworks developed by and for practitioners in the last few years, such as Sport England's Evaluation Framework (173), and the West of England Academic Health Science Network's toolkit to support commissioning of health and care services (177). What is perhaps most crucial, is that efforts are focused on developing guidance or frameworks that are: fit for purpose; available and accessible by a range of stakeholders; and applicable to a range of evaluation needs, including emerging needs. This may increase the usability and usefulness of the guidance to different users, and therefore increase use. However, to promote the uptake and use of guidance, more is needed than just better guidance. Powell et al. (20) in their work on the Expert Recommendations for Implementing Change (ERIC) project provide a list of seventy-three discrete strategies that can be used to support and improve implementation, such as developing academic partnerships, and core definitions that can be used to improve the reporting of implementation strategies. This supports the view presented in this thesis, of implementation as a complex and context specific endeavour.

Improving the quality and consistency of reporting, as well as increasing the number of published studies, has been a key focus within implementation research (10, 20). Indeed, considerable progress has been made in developing guidance, checklists and statements to improve the reporting of interventions. Examples include the Medical Research Council's guidance on the development and evaluation of complex interventions and on natural experiments (35, 127), statements on Strengthening the Reporting of Observational Studies (STROBE) (185, 186), and Standards for Reporting Implementation Studies (StaRI) (312). The rationale behind them is to provide tools to improve the reporting of studies, and so enhance translation and adoption of effective interventions (312). In a similar vein, in the systematic review undertaken here (Chapter 3) the checklist of indicators was presented as a tool to enable appraisal of the use and reporting of evaluation frameworks, for example by reviewers, and to improve the reporting of the use made of a framework. Its suggested use is intended to be alongside the guidance and statements mentioned above, as an additional tool to improve the completeness and transparency of evaluation studies, with the aim of improving the likelihood that relevant evidence is disseminated and taken up. However, improving dissemination is just one element of the evidence-based practice cycle, and dissemination through publication just one strategy. Within practice and policy decision making, other avenues of dissemination may be equally critical to

sharing effective interventions and good practice. An example includes the role of research networks in facilitating dissemination, which is considered in the following section.

The thesis findings have highlighted some key challenges in implementing strategies for evaluation and evidence-based practice. Firstly, the complex inter-connecting and context specific influences on evaluation and evidence-based practices suggest that simply providing guidance, and stipulating requirements for evaluation and reporting are not enough (20). There is an ongoing need for training and capacity building to support stakeholders to conduct and use evaluation. Secondly, public health interventions are complex and operate in an evolving political, organisational, economic, and environmental context, where priorities for evaluation, and the relevance and value of evidence changes. Thirdly, evaluation and evidence-based practices also occur in an evolving knowledge base. For example, there is a growing appreciation and understanding of the importance of multi-agency, multi-component and whole-systems approaches to address public health priorities (36, 40, 77, 313). The World Health Organization's call for multi-sectoral and multi-component approaches to achieve their targets to reduce physical inactivity amongst the population is just one example of the value now placed on these approaches to support intervention design, implementation and evaluation (1). Understanding and implementing such approaches and networks will require new and innovative evaluation methods and strategies (10, 16, 298). Indeed it may be more useful to implement pragmatic evaluation approaches that facilitate collaborative decision making to identify, agree and apply the most appropriate methods to generate the evidence required, and that can be adapted in response to emerging findings, than to provide or specify standardised frameworks such as those considered in this thesis.

6.2.2 The role of partnerships and networks in supporting and improving evidence-based practices

Research partnerships and networks are increasingly recognised as a valuable strategy to improve practice. They bring together researchers and practitioners with responsibility for programme delivery to collaborate in the design, implementation and evaluation of interventions. Multi- or inter-agency collaborations offer opportunities to adopt inter-disciplinary methods to understand and implement multi-component interventions. Further, research-practice partnerships offer flexibility that can allow adaptation and refinement of evaluation methods in response to emerging findings, changing stakeholder and programme needs, and changing demands for types of evidence during an intervention's development and delivery phases. This can be particularly important for practitioners to recognise the value of the evaluation to the intervention.

Evidence presented in this thesis has shown that research-practice partnerships can improve the rigour and relevance of evaluation, help build capacity to conduct and use evaluation, and improve use of evidence to inform policy and practice, as suggested in previous studies (8, 11, 48, 50, 61). The combination of researchers' knowledge, research expertise and resources, with practitioners' understanding of constraints and priorities in evaluating real-world interventions, can help when making decisions about pragmatic evaluation, and improve the likelihood that practice relevant evidence is generated. The research presented here highlighted that effective collaboration and communication between stakeholders throughout the stages of intervention planning, implementation, and evaluation were essential to allow flexible, responsive and innovative evaluation. The flow of information between partners is dependent on effective communication strategies, whilst wider networks of partners, including funders, play a crucial role in enabling timely and appropriate forms of communication. The use of the Sport England case study revealed that communication in the early and final stages of a programme is often a challenge, and that improvements are needed to better support this. Examples include longer term funding, structures that embed partnerships and staff with relevant expertise within organisations, and systems to promote knowledge exchange. Critically, the thesis has shown that key processes and characteristics of partnership working and relationships may be central to success. These processes and partnership characteristics need to be understood and applied by stakeholders when initiating and implementing partnerships and networks. Funders, researchers and practitioners all have a role to play in shaping these if the benefits of research-practice partnerships and networks are to be realised.

Evaluation is only one element in the evidence-based cycle; dissemination is a critical process linking evaluation and evaluation use. This requires effective organisational structures and systems to facilitate timely and appropriate communication between partners. Without this, even the most robust evaluation will not make it into the public domain. Reviewers and journal editors have a role to play in enabling practice-relevant evidence to reach publication, whilst research networks that facilitate the flow of knowledge between researchers, funders and practitioners may be even more critical in communicating relevant evidence to inform policy and practice decisions.

The thesis also highlighted gaps in our understanding of the complex interconnections between influences on practice, and in particular the need for more inter-disciplinary research to explore the role that research-practice partnerships and networks play in knowledge exchange and building capacity to support evidence-based practices. The thesis findings support a growing body of work that recommends further research that draws on theories of organisational learning (51, 87), communities of practice (311, 314, 315) and systems theories (16, 77, 316) to help explain

how organisational structures, networks, and systems can be developed and implemented to facilitate evaluation and evidence-based practices in multi-agency public health interventions.

6.3 Implications for research and practice

Public health interventions are developed within an evolving knowledge base and context. A shifting political, economic, environmental and social context drives requirements for evaluation and influences the value placed on differing forms of evidence. What counts as evidence changes within this context. As we move to a more system wide focus for interventions that can integrate systems, environments, and services to bring about sustained systemic change (36, 40, 313), evaluators, practitioners and decision makers need to: ask the right questions of an evaluation; identify, agree and apply appropriate methods; and disseminate findings in a timely manner to inform intervention implementation.

The research presented in this thesis has highlighted the complex interconnections between influences on evidence-based practice, and the highly context specific nature of these influences. In line with the Expert Recommendations for Implementing Change (ERIC) (20), this thesis has argued that it is not simply a case of identifying discrete recommendations and providing specific guidance, but that multiple methods and strategies may be needed. Organisations should implement strategies to facilitate innovative and flexible approaches. This will allow them to respond to evolving demands for evidence, and to evolving understanding of evaluation, dissemination, and implementation practice.

As shown in the case study, there is an inextricable link between the demands for evaluation and value placed on different forms of evidence by stakeholders as well as the use of methods and strategies to conduct evaluation, which in turn influences the usefulness of an evaluation to different stakeholders. It is vital that commissioners, practitioners, researchers, and other decision makers who are involved in intervention development, implementation and evaluation develop a better understanding of influences on practice, and put in place organisational structures, systems, and processes to better support evidence-based practice. For example, this research has highlighted the importance of applying and embedding strategies that facilitate capacity building, communication, collaboration and continuity of relationships. Without this, there is a risk that valuable resources are spent on less effective interventions, whilst critical evidence about what works does not reach the public domain.

Funders of interventions need to consider their requirements for evaluation to ensure they are appropriate to enable the generation of high quality, relevant evidence, and that systems to support evaluation and knowledge exchange are embedded. For example, funding cycles, timescales for reporting, and strategies to enable knowledge exchange are crucial. It is also

essential that organisations build and retain individual and organisational capacity to conduct and report on intervention evaluation, and to use evidence generated. The thesis has highlighted the potential for research-practice partnerships and networks to positively influence evidence-based practice. Practice-based and research organisations need to explore how they can initiate and embed such partnerships within their work, and how these can be shaped to allow for collaborative and flexible approaches that are mutually beneficial to those involved.

Researchers need to respond to the challenges to develop and implement appropriate methods and guidance to support emerging evaluation needs. The thesis has identified gaps in the current guidance where researchers can most usefully focus efforts to develop further guidance, such as sustainability and wider non-health and unanticipated outcomes. However, as discussed in this last chapter of the thesis, there is a growing appreciation of systems approaches to address public health concerns. Systems approaches can provide a useful framework, and tools, to help understand the complex relationships between influences on interventions and how these change over time (40, 313, 317). Nevertheless, there is a need to develop a better understanding of how these approaches can be applied within public health evaluations (16). Multi-agency partnerships and networks, including research-practice partnerships, provide opportunities to draw on differing perspectives and inter-disciplinary approaches to understand complex public health interventions and evaluation. Further research is recommended to more fully understand how networks and communities of practice work to improve knowledge exchange and shared learning, and how these may be implemented to achieve shared goals to improve health (311, 314).

6.4 Reflections on the research approach and its strengths and limitations

A key strength of the approach adopted for this thesis was the engagement with researchers and practitioners. This included consultation with stakeholders to conduct the scoping review, to ensure that the list of evaluation components against which frameworks were mapped was comprehensive, and that the results would be of interest and value to both practitioners and researchers. Working with stakeholders engaged in the design, implementation and evaluation of Sport England's 'Get Healthy Get Active' programme was also essential to allow a collective case study approach to be applied. Such working enabled the focus for the case study to be agreed, and provided access to multiple sources of data from a representative sample of the funded projects, and from differing stakeholder perspectives. This was important to ensure that the programme's rationale and evaluation was understood and represented appropriately in the thesis.

As outlined in the introduction to the thesis (section 1.4) and the conceptual framework developed to guide the research (Figure 1-2), a pragmatic approach was adopted to identify and apply the most appropriate methods to address each of the research questions. The conceptual

framework provided a useful point of reference to reflect on the research process and methods, and to appraise their strengths and limitations in addressing the research questions and in meeting the intended outputs and outcomes. The discussion of some of the strengths and limitations of the methods used illustrates the challenge of balancing methodological decisions in pragmatic research.

Firstly, applying a collective case study, and the choice of programme for this, was important. This enabled multiple data sources from the programme and 23 locally delivered physical activity interventions to be explored with multiple methods. Nevertheless, there were challenges; the case study was conducted largely retrospectively, so at the time of data collection the projects included were in their final few months of delivery or already completed. Whilst this was important to be able to access final evaluation reports for each of the projects as a key data source, it limited the ability to take a more ethnographic or participatory approach that may have provided more insights into the experiences and reflections of stakeholders during the evaluation process. It may also have limited the number of projects and stakeholders who agreed to participate in the research and to share documents. In addition, time lags between interventions ending and publication of programme level evaluation and any peer reviewed articles arising from project evaluations meant these could not be included in the data collection and analysis.

The retrospective nature of the case study, and also time since the searches for the scoping and systematic reviews were conducted means that, as with much research, many of the findings reported within the evidence synthesis and recommendations made may have since been implemented. This thesis is not an endpoint though, and seeks to contribute to the continuous harvesting of experiences and effects that are relevant to an ongoing discussion of evaluation, dissemination and evaluation use.

There is a growing recognition of the value of using qualitative research, multiple methods and case studies in health research to understand interventions holistically (10, 83, 84). Adopting a qualitative approach was essential to explore influences on practice in detail, and to gain insights into stakeholders' descriptions and reflections of their experiences. This allowed a deeper understanding of practices, and how influences can act as barriers or facilitators to evaluation and evidence-based practice. However, this was perhaps at the expense of taking a quantitative or meta-evaluation approach that may have provided a more systematic assessment of whether the evaluation strategies required by the funder had been effective in achieving the programme's aim of generating high quality, generalisable evidence.

Case studies are often criticised as a research approach, in that they are based on a specific entity or phenomenon and as such can be limited in their generalisability (83, 318). It is important to understand the purpose of a case study and how this informs the epistemological approach (319).

This was critical to its use in this thesis. The purpose here was to explore evaluation and evidence-based practices and influences on practice within the real-world contexts in which they occur. The methods applied therefore allowed a critical and interpretivist approach to be combined to understand similarities and differences in cases and perspectives, to draw generalisations where appropriate, and to understand the wider contextual factors that shaped practices within the cases. This differs from the use of case studies within many of the evaluation reports appraised as part of the case study, where the approach was mostly descriptive, and the purpose described by stakeholders more akin to promotion than evaluation. This distinction is noteworthy, and important to avoid the 'misuse' of the case study as an evaluative tool. The systematic and transparent analysis and reporting within the case study, the use of multiple data sources, framework analysis (278, 279) and directed content analysis (84) were critical to address some of the challenges associated with qualitative research (318) and case studies (83).

The thesis has been presented as a set of linked studies, in which each chapter applied appropriate methods to address a specific set of research questions. Each chapter built on the findings of the previous one to contribute to the overall aim of the thesis to improve understanding and implementation of evaluation and evidence-based practices in public health, and physical activity interventions in particular. This iterative approach dictated the sequential nature of the research which limited the ability to adopt a more deductive approach and for research elements to be conducted simultaneously. However, this is also a strength of the thesis as it allowed a progressive narrowing of the focus to explore emergent aspects of the research in more detail. This had implications for the research and the experience of conducting the research, which are discussed below in the personal reflection.

6.5 Personal reflection of the candidate

The starting point for the thesis was the aim to improve evaluation in practice, and the approach taken has primarily been informed by an applied research perspective. One frustration has been the retrospective nature of the research, and a feeling that this limited the opportunity to take a more action-research and participatory approach. Engaging with stakeholders to gather data, whether that was negotiating for access to documents or conducting interviews, helped to mitigate that in part. However, working within an applied context but operating outside of that context also lent itself to a questioning of the purpose and value of the research. One of the benefits of the emergent research approach has been that the chapters that have contributed to the overall thesis could also be presented as independent pieces; this meant that three papers have been published, and one more is under review. This was important to be able to recognise the value of the research, and was critical for motivation.

As is often the case in qualitative research, managing the extensive bodies of data that were included in each of the reviews and the case study brought challenges, both in terms of the time needed to process and analyse the data, and in identifying (or accepting) what was important to include and what could comfortably be left out within the narrative synthesis of the thesis. The emergent and sequential nature of the thesis may also have contributed to this. The identification and selection of an appropriate programme for the case study took time and was inevitably somewhat opportunistic. The scoping review was the first piece of work to be completed. With hindsight the focus of the scoping review could perhaps have been narrowed from physical activity and dietary change interventions to physical activity specifically at an earlier stage, which would have aligned more closely with the rest of the thesis.

Within this last chapter the growing interest in multiple methods and systems approaches to better understand health interventions has been discussed. Recent calls for the adoption of systems approaches to understand the wider context in which public health interventions occur have emphasised a pragmatic approach to identify which tools, theories and methods are most useful in particular contexts and to answer particular questions (16, 77, 313). Reflecting on the research undertaken with this perspective in mind, three things present. Firstly, the scoping review was a useful starting point to understand the breadth of frameworks and tools available; several of those included in the review, such as Participatory Impact Pathway Analysis (123), and Developmental Evaluation (78), are mentioned in the literature as relevant to the application of systems thinking in health (77). Secondly, the progressive narrowing of the focus of interest that allowed a more detailed consideration of partnership working, and the use of network mapping in Chapter 5, were important to bring the discussion in line with emerging issues in public health evaluation. Thirdly, the use of multiple methods and software in this thesis, including: NVivo 12 Pro to support the content and framework analysis; Excel spreadsheets to process large data sets; Ucinet to produce network maps; and PRISMA statements to support the reporting of systematic and scoping reviews, were all essential to address the research objectives, and importantly provided an opportunity to develop skills and experience that will be applicable in future research.

It has been crucial to keep at the heart of the thesis the understanding that the findings are more broadly applicable than physical activity interventions, and that insights can be applied to evaluation and evidence-based practices in any domain that operates in similar multi-agency contexts. This was presented in the conceptual framework (Figure 1-2), and is important to reflect on. The value of the research to Sport England as a key stakeholder in this research remains to be established. Findings have been shared in order to gather comments on the final drafts of chapters relating to the case study, but beyond that opportunities to disseminate findings to

stakeholders involved in the individual projects, or the wider network of researchers and practitioners involved in promotion of physical activity, have not arisen thus far. Two conferences for which I had abstracts accepted were postponed, and preliminary discussions with Sport England regarding dissemination of findings to stakeholders are on hold due to the Covid-19 pandemic. The publications arising from the study provide a useful contribution to the literature on which practitioners and researchers can draw; however, it is important in applied research that findings are also disseminated appropriately to a wider range of stakeholders.

Other challenges that are perhaps worthy of note, have been those of working on a single project over a long duration and independently. Opportunities to engage in delivering teaching activities, participating in training and professional development, and undertaking additional research activities were essential to sustain momentum in this research. In addition, the findings and experiences have been used to write a successful research grant application, and to inform a four-month long project in which I applied network analysis and collaborative approaches to explore research activities within a local authority. I include this here, as this has helped me to realise the value of the thesis, and the potential for its contribution to practice, to addressing University requirements to show impact, and critically to addressing the underlying aim of the thesis which was to contribute to closing the research-practice gap. It is only in the final stages that the value of the research has been really appreciated.

6.5 Concluding comments

To bring about sustainable change and meet targets to improve the health of individuals and populations, evaluation, dissemination and evidence uptake are essential processes in the evidence-based public health cycle. In practice, complex inter-connected influences can act as barriers and facilitators to the effective implementation of these three processes. This thesis sought to improve understanding and implementation of evidence-based practice in multi-agency public health interventions, taking the evaluation of physical activity interventions as a case study. It is hoped it has made an important contribution to practice, by signposting to appropriate evaluation guidance, and by identifying examples of good practice and of where improvements are needed to better support practice.

The contextual and changeable nature of interventions and their evaluation will always impact the use, usability and usefulness of strategies to support evaluation and evidence-based practice. Critically, stakeholders with an interest in conducting or using evaluation need to understand and implement organisational structures, systems and processes to support and improve evaluation, dissemination and evaluation use. These need to be shaped to ensure that individual and organisational capacity for evaluation is built and retained, and to encourage collaboration and communication between stakeholders. This is critical to close current gaps in the evidence-based

practice cycle and to ensure that relevant evidence is generated and used in a timely manner. Without this, critical evidence that could be used to inform interventions to support the health of the population will not make it into the public domain.

By identifying the gaps in understanding and in the evaluation guidance, it is hoped that the thesis stimulates further conversations and research to improve understanding of the relationships between research and practice, between evaluation and evaluation use, and between the structures, systems and processes that influence evaluation practices, dissemination, and evaluation use. Overall it is hoped that this body of work has contributed to the underlying aim of helping to close the research-practice gap.

Appendices

Appendix 1. Glossary of key terms

Community	Interventions that are both place-based and where people share goals or affinity (13, 43).
Community-centred	Approaches that promote relationships, mobilise local assets, and strengthen community capacities are more than simply community-based (13, 43).
Context	A general term that includes diverse internal and external factors that may influence an intervention and/or its evaluation.
Dissemination	The process of communicating findings in ways that will facilitate their use in practice (66).
Evaluation	S*ystematic examination and assessment of the features of an initiative and its effects, in order to produce information that can be used by those who have an interest in its improvement or effectiveness*. (19), p3.
Evaluation component	Individual elements encompassed within evaluation; for example elements of process or outcome evaluation.
Evaluation framework	Any structured guidance that facilitates a systematic evaluation of the implementation or outcomes of an intervention.
Evaluation use	The use of evidence generated from evaluation, and the effects of being involved in evaluation (55, 71), related terms include:
Findings use	The use of evidence generated from an evaluation.
Process use	The effects of being involved in evaluation.
Instrumental use	The use of evaluation for direct action.
Conceptual use	The use of evaluation in changing attitudes or improving knowledge.
Symbolic use	The use of evaluation to justify decisions or actions.
Evidence-based practice	The use of evidence to inform and improve future decision making.
Generic framework	One that is intended for use across a range of contexts, settings and sectors, as opposed to one that has been developed for use in a specific context or field.

Health promotion	Interventions that adopt methods to enable people to improve their health or well-being (32).
Implementation	The act of carrying an intention into effect, which in health research can be policies, programmes, or individual practices (collectively called interventions) (320).
Implementation research	T*he scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and, hence, to improve the quality and effectiveness of health services and care.* (17) (p.1).
Intervention	A general term that encompasses a broad spectrum of components, and includes interventions developed within a research context as well as those developed within a practice-based context.
Intervention function	Broad categories to define the general means by which an intervention might change behaviour (Education, Enablement, Persuasion, Coercion, and Incentivisation, Modelling, Environmental Restructuring) (38, 190).
Multi-component	Interventions that have several elements, such as different modes of delivery or intervention functions.
Multi- or inter-agency	Bringing together stakeholders from different groups or organisations within the same sector or different sectors, for example health charities, public health teams, or health researchers.
Multi-sectoral	Bringing together different sectors, for example health and sports sectors.
Practice-based evidence	The knowledge and insights generated from evaluation of 'real-world' interventions.
Practitioner	Those involved in decisions and actions related to intervention development, delivery and evaluation from a practical standpoint, including funders, commissioners, policy makers and intervention facilitators or delivery staff.
Pragmatic evaluation	An approach that seeks to balance the need for pragmatism within service delivery with demands for evaluation rigour.
Programme	Real-world interventions that represent a group of related projects.

Public Health Intervention	Interventions that seek to modify socio-ecological determinants of health, for example to bring about behaviour change, to address non-communicable health outcomes (30).
Real-world/Practice-based	Interventions that are part of normal service delivery or delivered in a practice setting, rather than within a research setting.
Researcher	Those primarily engaged in research and evaluation from an academic standpoint.
Setting	The physical, geographical, or organisational space in which an intervention is implemented.
Systems thinking	A field of enquiry and practice aimed at seeing how things are connected to each other within a notion of a wider 'whole entity' (77).
Translational research	A research approach that explores which evidence and knowledge-transfer strategies are used within specific policy and programmes to understand the relationship between evidence generation and evidence-based policy and practice (68).
Use	The action of using something or the state of being used for a purpose.
Usability	The degree to which something is easy to use.
Usefulness	The quality or degree of being effective, indicating the value to the user.

Appendix 2 Research ethics approval

Faculty of Medicine and Health Sciences Research Ethics Committee

Judith Fynn MED

Research & Innovation Services
Floor 1, The Registry
University of East Anglia
Norwich Research Park
Norwich, NR4 7TJ

Email: fmh.ethics@uea.ac.uk

Web: www.uea.ac.uk/researchandenterprise

16 October 2018

Dear Judith

Project title: Exploring facilitators and barriers to good practice in intervention evaluation: A case study of the Sport England 'Get Healthy Get Active' programme

Reference : 201718 - 133

The amendments to your above proposal have been considered by the Faculty Research Ethics Committee and we can confirm that your proposal has been approved.

Please could you ensure that any further amendments to either the protocol or documents submitted are notified to us in advance and also that any adverse events which occur during your project are reported to the Committee. Please could you also arrange to send us a report once your project is completed.

Approval by the FMH Research Committee should not be taken as evidence that your study is compliant with GDPR and the Data Protection Act 2018. If you need guidance on how to make your study GDPR compliant, please contact your institution's Data Protection Officer.

Yours sincerely,

Professor
M J
Wilkinson
Chair
FMH Research Ethics Committee

CC Supervisor

Project officer & REN project code (if we have this information)

Appendix 3 Gatekeepers consent to conduct the research from Sport England

Permission to conduct research on Sport England's Get Healthy Get Active programme

TW

Toby Wood <Toby.Wood@sportengland.org>

Reply

Wed 30/05/2018 09:56

To:

Judith Fynn (MED - Student)

Cc:

Darcy Hare <Darcy.Hare@sportengland.org>;
Tom Burton <Tom.Burton@sportengland.org>

Hi Judith,

Permission to conduct research on Sport England's Get Healthy Get Active (GHGA) programme

This is to confirm that Sport England gives permission for the research team at UEA, led by Judith Fynn a PhD candidate, to conduct the research on the Get Healthy Get Active (GHGA) programme. We have been given the opportunity to see and comment on the protocol. We have agreed to support the research team by contacting the leads for the funded projects within the GHGA programme to provide them with information about the research project and to invite them to participate in interviews with the lead researcher.

Toby Wood

Project and Relationship Manager

T: 07920295281

M: 07920295281

F: 01132 422 189

E: Toby.Wood@sportengland.org

Appendix 4 Participant Information Sheet for the case study

Participant Information Sheet (PIS)

Study title: Exploring facilitators and barriers to good practice in intervention evaluation: A case study of the Sport England 'Get Healthy Get Active' programme.

You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Please ask if there is anything that is not clear or if you would like more information. If you have questions, contact details are at the end of this information sheet.

What is the purpose of the study?

Sport England's Get Healthy Get Active programme has funded, in partnership with Local Authorities, Charities and Clinical Commissioning Groups, a portfolio of physical activity projects, and aims to evaluate these to provide evidence of how sport can contribute to decreasing inactivity and improving public health. A team of researchers from the University of East Anglia are conducting interviews with practitioners involved in projects funded by the Sport England programme to identify examples of good practice and to gain a better understanding of how systems and practices support or limit evaluation and evidence-based practice. The study aims to contribute to our understanding of best practice and to make recommendations to support intervention evaluation and sharing of evidence generated from interventions.

Why have I been chosen?

Sport England's Get Healthy Get Active programme has been identified as the focus for the research. As a person with a role in this programme or one of the funded projects within the programme your views are important to this study and we would like to invite you to take part.

Do I have to take part?

You do not have to take part in the study if you do not want to. If you do want to take part in the study, we attach a copy of the consent form which we ask you to sign and return to us, please. You will be given a copy of the consent form to keep. If you decide to take part you are still free to withdraw from the study at any time up to the point of analysis without giving a reason. If you decide to withdraw after the interview, the process for withdrawal will be explained at the end of the interview, and any data collected during the interview will not be included in the study. A decision to withdraw at any time, or a decision not to take part, will not affect you in any way.

What does taking part involve?

If you agree to take part, you will be asked to participate in an interview with the researcher. Interviews will be conducted face to face, or via Skype or telephone, by mutual arrangement with participant(s), and will last a maximum of 60 minutes. Interviews will be recorded and typed-up. If you would like to, you will be able to see a copy of the transcript and/or reported analysis to check that it is a true representation of what was said.

What are the possible disadvantages and benefits of taking part?

We do not believe there are any disadvantages or risks in taking part in the study other than your time taken to take part in the interview. In terms of benefits, by taking part you will have the chance to reflect on the evaluation and reporting process and your own involvement. We are interested in your experiences and views, and by taking part you will be contributing to our understanding of good practice in evaluation and evidence-based practice, along with facilitators and barriers to good practice. Understanding gained from the study is intended to help develop recommendations for best practice.

Will my taking part in the study be kept confidential?

All information which is collected during the course of the research, including information relating to the Get Healthy Get Active programme and individual funded projects, and individual's opinions or comments will be kept strictly confidential so that only the researcher(s) carrying out the research will have access to such information.

Participants should note that things you say during interview may be included in reports as direct quotes or as summaries in an anonymised form. By agreeing to participate in this project, you are consenting to the retention and publication of information gathered. Any quotes that have the potential to be recognisable will be shared with you and your consent sought for their inclusion in reports and publications. You will be able to withdraw direct quotes from the reports up to the point at which they are submitted for publication. At the end of the interview you will be provided with information for the withdrawal process and dates up to which you can either fully withdraw or withdraw quotes that may be recognisable.

What will happen to the results of the research study?

Understandings gained will be shared with Sport England and partner organisations through the final report and presentations. Findings may also be shared with a wider audience through publication in academic journals and by presenting at conferences as well as a chapter in a PhD thesis being produced by the lead researcher, Judith Fynn.

Who is organising the research?

The research is being conducted by a team at the Norwich Medical School at the University of East Anglia (UEA). The research is funded by the university as well as the Centre for Diet and Activity Research (CEDAR). The study is being led by Judith Fynn with Professor Andy Jones, Dr Wendy Hardeman, Dr Karen Milton and Dr Charlotte Salter.

Has the project been approved on ethical grounds?

This study has been reviewed and approved by an independent group of people as part of the Faculty of Medicine and Health Sciences Research Ethics Committee, University of East Anglia which protects your safety, rights, wellbeing and dignity. The design and management of the research has taken account of GDPR requirements to ensure compliance.

Complaints Procedure

If you have any complaints about the study you can contact in the first instance:

Professor Andy Jones, Norwich Medical School, University of East Anglia, Norwich, NR4 7TJ. Tel: 01603 593127, a.p.jones@uea.ac.uk

Who may I contact for further information?

If you would like more information about the research before you decide whether or not you would be willing to take part, please email in the first instance: Judith Fynn (j.fynn@uea.ac.uk), Norwich Medical School, University of East Anglia, Norwich, NR4 7TJ

Appendix 5 Consent Form for participants in the case study

CONSENT FORM

Title of Project: Exploring facilitators and barriers to good practice in intervention evaluation: A case study of the Sport England 'Get Healthy Get Active' programme.

Name of Researcher: Judith Fynn

	Please initial box
1. I confirm that I have read and understand the information sheet dated 18/09/2018 (version 2) for the above study. I have had the opportunity to consider the information and ask questions.	☐
2. I understand that my participation is voluntary and that I am free to withdraw fully from the study at any time up to the point of analysis and to withdraw quotes that may be recognisable up to the point of submission for publication, without giving any reason and without my legal rights being affected.	☐
3. I understand that the information collected will be used to support other research in the future, and may be shared anonymously with other researchers.	☐
4. I agree to take part in the above study.	☐

_____ _____ _____

Name of Participant Date Signature

_____ _____ _____

Name of Researcher Date Signature

1 for participant; 1 for researcher

Appendix 6 Topic guide for semi-structured interviews

General Contextual 1. Please can you tell me about your role in the project ?
Project Evaluation 2. Can you tell me about your experiences of being involved in the project evaluation, please? 3. What do you feel were the main factors that influenced how the evaluation was designed and implemented? Prompts: Any specific requirements, evidence, tools, or frameworks? 4. What do you feel were the main strengths and weaknesses of the methods or approaches used in the project evaluation? Prompts: What worked well?/ What worked less well, any challenges?/ Any examples? 5. How useful do you feel the evaluation was? Prompts: Any examples of how it was used?/ Any examples of challenges to it being useful? 6. Please could you tell me more about any systems or organisational structures that were put in place to support project evaluation? How effective were these? 7. Reflecting back, is there anything you feel you would have been done differently in evaluating the project? 8. Do you have any thoughts or suggestions for what is needed to support project evaluation?
Partnership working: 9. Please can you tell me more about any partners involved in the evaluation, and the roles they played? Prompts: Who do you see as essential partners? / Had you worked together before? Who did what? How was this decided? 10. Can you tell me more about your experiences of working with partners as part of the evaluation? Prompts: What works well? (facilitators) / What works less well? (any barriers or challenges?) / Anything you would do differently regarding partnership working?

11. How would you describe the processes, information or support for evaluation from partners? Prompts: From the funder, within your organisation, other partners?
Evaluation reporting/knowledge sharing: 12. Please can you tell me more about your experiences of how the evaluations were reported or shared? 13. Do you have any thoughts on how projects could be better supported to share knowledge gained from evaluation?
Other questions e.g. Specific follow up on observations from the evaluation report. Is there anything else you would like to tell me, that you feel you have not yet had the opportunity to discuss? Is there anyone else involved in the project and/or it's evaluation that you feel it may be useful for me to speak to?

References

1. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. Geneva: World Health Organization; 2018.
2. Public Health England. Everybody Active, Every Day: An evidence-based approach to physical activity. London: Public Health England; 2014.
3. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet. 2012;380(9838):219-29.
4. World Health Organization. Global health observatory data repository. Geneva: World Health Organization; 2018.
5. Bauman AE. Updating the evidence that physical activity is good for health: an epidemiological review 2000–2003. Journal of Science and Medicine in Sport. 2004;7(1):6-19.
6. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. The Lancet. 2012;380(9838):247-57.
7. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1· 6 million participants. The Lancet Child & Adolescent Health. 2020;4(1):23-35.
8. Pettman TL, Armstrong R, Doyle J, Burford B, Anderson LM, Hillgrove T, et al. Strengthening evaluation to capture the breadth of public health practice: ideal vs. real. Journal of Public Health. 2012;34(1):151-5.
9. Bowen S, Zwi AB. Pathways to "evidence-informed" policy and practice: a framework for action. PLoS Med. 2005;2(7):e166.
10. Ogilvie D, Bauman A, Foley L, Guell C, Humphreys D, Panter J. Making sense of the evidence in population health intervention research: building a dry stone wall. BMJ Global Health. 2020;5(12):e004017.
11. Lobo R, Petrich M, Burns SK. Supporting health promotion practitioners to undertake evaluation for program development. BMC Public Health. 2014;14:1315.
12. Smith BJ, Rissel C, Shilton T, Bauman A. Advancing evaluation practice in health promotion. Health Promotion Journal of Australia. 2016;27(3):184-6.
13. Public Health England. A Guide to Community-Centred Approaches to Health and Well-being Full Report PHE publications gateway number: 20147112014711. London: Public Health England; 2015.
14. Li V, Carter SM, Rychetnik L. Evidence valued and used by health promotion practitioners. Health Education Research. 2015(2):193.
15. Moore G, Audrey S, Barker M, Bond L, Bonell C, Cooper C, et al. Process evaluation in complex public health intervention studies: the need for guidance. Journal of Epidemiology & Community Health. 2015;68(2):101-2.
16. McGill E, Er V, Penney T, Egan M, White M, Meier P, et al. Evaluation of public health interventions from a complex systems perspective: a research methods review. Social Science & Medicine. 2021;Jan 11:113697.
17. Eccles MP, Mittman BS. Welcome to implementation science. 2006. p. 1-3.
18. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychology. 2015;3(1):32.
19. World Health Organization. Health Promotion Evaluation: Recommendations to Policy-Makers. Geneva: World Health Organization; 1998.
20. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implementation Science. 2015;10(1):1-14.
21. Public Health England. From Evidence to Action: Opportunities to protect and improve the nation's health. London: Public Health England; 2014.

22. Allen P, Jacob RR, Lakshman M, Best LA, Bass K, Brownson RC. Lessons learned in promoting evidence-based public health: perspectives from managers in state public health departments. Journal of Community Health. 2018;43(5):856-63.
23. Learmonth AM. Utilizing research in practice and generating evidence from practice. Health Education Research. 2000;15(6):743-56.
24. Brownson R, Fielding J, Green LW. Building Capacity for Evidence-Based Public Health: Reconciling the Pulls of Practice and the Push of Research. Annual Review of Public Health. 2018;39(1):27-53.
25. Ammerman A, Smith TW, Calancie L. Practice-based evidence in public health: improving reach, relevance, and results. Annual Review of Public Health. 2014;35:47-63.
26. Honeycutt S, Hermstad A, Carvalho ML, Arriola KRJ, Ballard D, Escoffery C, et al. Practice to Evidence: Using Evaluability Assessment to Generate Practice-Based Evidence in Rural South Georgia. Health Education & Behavior. 2017;44(3):454-62.
27. Centers for Disease Control and Prevention Program Performance and Evaluation Office. Introduction to Program Evaluation for Public Health Programs: A Self-Study Guide [Internet]. U.S.: Centers for Disease Prevention U.S. Department of Health & Human Services; 2019 [updated 2019; cited 2020 Feb 8]. Available from: https://www.cdc.gov/eval/framework/index.htm
28. Brownson RC, Gurney JG, Land GH. Evidence-based decision making in public health. Journal of Public Health Management and Practice. 1999;5:86-97.
29. Bauman A, Nutbeam D. Evaluation in a Nutshell. Australia: McGraw-Hill Education 2013.
30. Bauer G, Davies JK, Pelikan J, Noack H, Broesskamp U, Hill C. Advancing a theoretical model for public health and health promotion indicator development: proposal from the EUHPID consortium. European Journal of Public Health. 2003;13(suppl_3):107-13.
31. Dahlgren G, Whitehead M. European strategies for tackling social inequities in health: Levelling up Part 2. Studies on social and economic determinants of population health, No. 3. Copenhagen: World Health Organization Regional Office for Europe 2007.
32. Eriksson M, Lindström B. A salutogenic interpretation of the Ottawa Charter. Health Promotion International. 2008;23(2):190-9.
33. Rogers PJ. Using programme theory to evaluate complicated and complex aspects of interventions. Evaluation. 2008;14(1):29-48.
34. Glouberman S, Zimmerman B. Complicated and complex systems: what would successful reform of Medicare look like? Romanow Papers. 2002;2(Discussion Paper No. 8):21-53.
35. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: The new Medical Research Council guidance. British Medical Journal. 2008;337:a1655.
36. Rutter H, Savona N, Glonti K, Bibby J, Cummins S, Finegood DT, et al. The need for a complex systems model of evidence for public health. Lancet. 2017;390(10112):2602-4.
37. Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. Bmj. 2008;336(7656):1281-3.
38. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. Implementation Science. 2011;6(1):42.
39. Sport England. Towards an Active Nation. London: Sport England; 2016.
40. Rutter H, Cavill N, Bauman A, Bull F. Systems approaches to global and national physical activity plans. Bulletin of the World Health Organization. 2019;97(2):162.
41. Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implementation Science. 2009;4(1):40.
42. Bertotti M, Jamal F, Harden A. Connected Communities: A review of conceptualisations and meanings of 'community'within and across research traditions: a meta-narrative approach. London, UK: Arts and Humanities Research Council; 2012.

43. National Institute for Health and Care Excellence (NICE). NICE Guideline: Community engagement: improving health and wellbeing and reducing health inequalities. London: National Institute for Health and Care Excellence; 2016.
44. Pellegrinelli S. What's in a name: Project or programme? International Journal of Project Management. 2011;29(2):232-40.
45. Weiss CH. The interface between evaluation and public policy. Evaluation. 1999;5(4):468-86.
46. Wimbush E, Watson J. An Evaluation Framework for Health Promotion: Theory, Quality and Effectiveness. Evaluation. 2000;6(3):301-21.
47. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. American Journal of Public Health. 1999;89(9):1322-7.
48. Francis LJ, Smith BJ. Toward Best Practice in Evaluation: A Study of Australian Health Promotion Agencies. Health Promotion Practice. 2015;16(5):715-23.
49. Schneider CH, Milat AJ, Moore G. Barriers and facilitators to evaluation of health policies and programs: Policymaker and researcher perspectives. Evaluation and Program Planning. 2016;58:208-15.
50. Brug J, Tak NI, Te Velde SJ. Evaluation of nationwide health promotion campaigns in the Netherlands: an exploration of practices, wishes and opportunities. Health Promotion International. 2011;26(2):244-54.
51. Schwarzman J, Bauman A, Gabbe B, Rissel C, Shilton T, Smith B. Organizational determinants of evaluation practice in Australian prevention agencies. Health Education Research. 2018;33(3):243-55.
52. Habicht JP, Victora CG, Vaughan JP. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. International Journal of Epidemiology. 1999;28(1):10-8.
53. Raphael D. The question of evidence in health promotion. Health Promotion International. 2000;15(4):355-67.
54. Denford S, Lakshman R, Callaghan M, Abraham C. Improving public health evaluation: a qualitative investigation of practitioners' needs. BMC Public Health. 2018;18(1):190.
55. Alkin MC, King JA. The Historical Development of Evaluation Use. American Journal of Evaluation. 2016;37(4):568-79.
56. Pawson R, Tilley N. Realistic Evaluation. London: Sage Publications Ltd; 1997.
57. Brousselle A, Buregeya J-M. Theory-based evaluations: Framing the existence of a new theory in evaluation and the rise of the 5th generation. Evaluation. 2018;24(2):153-68.
58. Smith RD, Petticrew M. Public health evaluation in the twenty-first century: time to see the wood as well as the trees. Journal of Public Health. 2010;32(1):2-7.
59. Moore G, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process Evaluation of Complex Interventions: UK Medical Research Council (MRC) Gudiance. London: MRC Population Health Science Research Network; 2015.
60. Aarestrup AK, Jørgensen TS, Due P, Krølner R. A six-step protocol to systematic process evaluation of multicomponent cluster-randomised health promoting interventions illustrated by the Boost study. Evaluation and Program Planning. 2014;46:58-71.
61. Harden SM, Johnson SB, Almeida FA, Estabrooks PA. Improving physical activity program adoption using integrated research-practice partnerships: an effectiveness-implementation trial. Translational Behavioral Medicine. 2017;7(1):28-38.
62. Nyström ME, Karltun J, Keller C, Gäre BA. Collaborative and partnership research for improvement of health and social services: researcher's experiences from 20 projects. Health Research Policy and Systems. 2018;16(1):1-17.
63. Denford S, Abraham C, Callaghan M, Aighton P, De Vocht F, Arris S. A review of Grey and academic literature of evaluation guiance relevant to public health interventions. BMC Health Services Research. 2017;17:643.

64. Page-Reeves J, Regino L. Community–University Health Research Partnerships: Challenges and Concrete, Plain Language Strategies for Community Engagement in Research. Anthropology in Action. 2018;25(2):1-12.
65. Mansfield L. Resourcefulness, reciprocity and reflexivity: the three Rs of partnership in sport for public health research. International Journal of Sport Policy and Politics. 2016;8(4):713-29.
66. Wilson PM, Petticrew M, Calnan MW, Nazareth I. Disseminating research findings: what should researchers do? A systematic scoping review of conceptual frameworks. Implementation Science. 2010;5(1):91.
67. Tabak RG, Khoong EC, Chambers DA, Brownson RC. Bridging research and practice: models for dissemination and implementation research. American Journal of Preventive Medicine. 2012;43(3):337-50.
68. Rychetnik L, Bauman A, Laws R, King L, Rissel C, Nutbeam D, et al. Translating research for evidence-based public health: key concepts and future directions. Journal of Epidemiology and Community Health. 2012;66(12):1187-92.
69. Milat A, Li B. Narrative review of frameworks for translating research evidence into policy and practice. Public Health Research & Practice. 2017;27(1):e2711704.
70. Milat AJ, King L, Bauman AE, Redman S. The concept of scalability: increasing the scale and potential adoption of health promotion interventions into policy and practice. Health Promotion International. 2013;28(3):285-98.
71. Alkin MC, King JA. Definitions of evaluation use and misuse, evaluation influence, and factors affecting use. American Journal of Evaluation. 2017;38(3):434-50.
72. King JA, Alkin MC. The centrality of use: Theories of evaluation use and influence and thoughts on the first 50 years of use research. American Journal of Evaluation. 2019;40(3):431-58.
73. Cousins JB, Goh SC, Elliott CJ, Bourgeois I. Framing the capacity to do and use evaluation. New Directions for Evaluation. 2014;2014(141):7-23.
74. Amo C, Cousins JB. Going through the Process: An Examination of the Operationalization of Process Use in Empirical Research on Evaluation. New Directions for Evaluation. 2007;116:5-26.
75. Berman J, Smyth R. Conceptual frameworks in the doctoral research process: A pedagogical model. Innovations in Education and Teaching International. 2015;52(2):125-36.
76. Norris E, Finnerty AN, Hastings J, Stokes G, Michie S. A scoping review of ontologies related to human behaviour change. Nature Human Behaviour. 2019;3(2):164-72.
77. Peters DH. The application of systems thinking in health: why use systems thinking? Health Research Policy and Systems. 2014;12(1):1-6.
78. Patton MM. Developmental Evaluation Applying Complexity Concepts to Enhance Innovation and Use. New York: The Guildford Press; 2011.
79. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implementation Science. 2010;5:69.
80. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology. 2005;8(1):19-32.
81. Tricco AC, Lillie E, Zarin W, K.K. OB, Colquhoun H, Levac D. PRISMA Extension for Scoping Reviews (PRISMAScR): Checklist and Explanation. Annals of Internal Medicine. 2018;169(7):467-73.
82. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 2010;8(5):336-41.
83. Crowe S, Cresswell K, Robertson A, Huby G, Avery A, Sheikh A. The case study approach. BMC Medical Research Methodology. 2011;11(1):100.
84. Hsieh H-F, Shannon SE. Three approaches to qualitative content analysis. Qualitative Health Research. 2005;15(9):1277-88.

85. Fynn JF, Hardeman W, Milton K, Jones AP. A scoping review of evaluation frameworks and their applicability to real-world physical activity and dietary change programme evaluation. BMC Public Health. 2020;20(1):1-16.
86. Fynn JF, Hardeman W, Milton K, Murphy J, Jones A. A systematic review of the use and reporting of evaluation frameworks within evaluations of physical activity interventions. International Journal of Behavioral Nutrition and Physical Activity. 2020;17(1):1-17.
87. Fynn JF, Hardeman W, Milton K, Jones A. Exploring influences on evaluation practice: A case study of a national physical activity programme. International Journal of Behavioral Nutrition & Physical Activity. 2021;In Press.
88. Hanson S, Jones A. Missed opportunities in the evaluation of public health interventions: a case study of physical activity programmes. BMC Public Health. 2017;17(1):674.
89. National Institute for Health and Care Excellence (NICE). Obesity: Working with local communities, PH42 November 2012, updated June 2017. London: National Institute for Health and Care Excellence; 2012.
90. HM Government. Sporting Future: A New Strategy for an Active Nation. London: Cabinet Office; 2015.
91. Lee RG, Garvin T. Moving from information transfer to information exchange in health and health care. Social Science & Medicine. 2003;56:449-64.
92. Milstein B, Wetterhall S. A Framework Featuring Steps and Standards for Program Evaluation. Health Promotion Practice. 2000;1(3):221-8.
93. Cavill N, Roberts K, Rutter H. Standard Evaluation Framework for Physical Activity Interventions. Oxford: National Obesity Observatory; 2012.
94. Cavill N, Roberts K, Rutter H. Standard Evaluation Framework for Dietary Interventions. Oxford: National Obesity Observatory; 2012.
95. Roberts K, Cavill N, Rutter H. Standard Evaluation Framework for Weight Management Interventions. London: National Obesity Observatory; 2009.
96. Nilsen P. Making sense of implementation theories, models and frameworks. Implementation Science. 2015;10:53-79.
97. Davies JK, Sherriff N. The gradient in health inequalities among families and children: A review of evaluation frameworks. Health Policy. 2011;101(1):1-10.
98. Chambers AH, Kolbe A, Murphy K. Designs and methods used in published Australian health promotion evaluations 1992-2011. Australian and New Zealand Journal of Public Health. 2015;39(3):222-6.
99. Lim MSC, Wright CJC, Carrotte ER, Pedrana AE. Reach, engagement, and effectiveness: a systematic review of evaluation methodologies used in health promotion via social networking sites. Health Promotion Journal of Australia. 2016;27(3):187-97.
100. Escoffery C, Lebow-Skelley E, Udelson H, Böing EA, Wood R, Fernandez ME, et al. A scoping study of frameworks for adapting public health evidence-based interventions. Translational Behavioral Medicine. 2018;9(1):1-10.
101. Abraham C, Denford S, Callaghan M. Evaluation in Health and Wellbeing: Guidance Summaries [Internet]. London: Public Health England; 2018 [updated 7 August 2018; cited 2021 Feb 8]. Available from: https://www.gov.uk/government/publications/evaluation-in-health-and-wellbeing-guidance-summaries/evaluation-in-health-and-well-being-guidance-summaries
102. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Information & Libraries Journal. 2009;26(2):91-108.
103. Judd J, Frankish CJ, Moulton G. Setting standards in the evaluation of community-based health promotion programmes— a unifying approach. Health Promotion International. 2001;16(4):367-80.
104. Vatcharavongvan P, Hepworth J, Marley J. The application of the parallel track model in community health promotion: A literature review. Health and Social Care in the Community. 2013;21(4):352-63.

105. Dauphinee WD. The role of theory-based outcome frameworks in program evaluation: Considering the case of contribution analysis. Medical Teacher. 2015;37(11):979-82.
106. Ory MG, Altpeter M, Belza B, Helduser J, Zhang C, Smith ML. Perceived utility of the RE-AIM framework for health promotion/disease prevention initiatives for older adults: A case study from the U.S. evidence-based disease prevention initiative. Frontiers in Public Health. 2015;2:143.
107. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Education Quarterly. 1988;15(4):351-77.
108. Kramer L, Schwartz P, Cheadle A, Rauzon S. Using Photovoice as a Participatory Evaluation Tool in Kaiser Permanente's Community Health Initiative. Health Promotion Practice. 2013;14(5):686-94.
109. Davis LA, Morgan SE, Mobley AR. The Utility of the Memorable Messages Framework as an Intermediary Evaluation Tool for Fruit and Vegetable Consumption in a Nutrition Education Program. Health Education and Behavior. 2015;43(3):321-7.
110. Smith LT, Johnson DB, Lamson E, Sitaker M. A framework for developing evaluation tools used in Washington State's Healthy Communities projects. Preventing Chronic Disease. 2006;3(2):9.
111. Lubans DR, Lonsdale C, Cohen K, Eather N, Beauchamp MR, Morgan PJ, et al. Framework for the design and delivery of organized physical activity sessions for children and adolescents: Rationale and description of the 'SAAFE' teaching principles. International Journal of Behavioral Nutrition and Physical Activity. 2017;14(1):1-11.
112. Angeles RN, Dolovich L, Kaczorowski J, Thabane L. Developing a Theoretical Framework for Complex Community-Based Interventions. Health Promotion Practice. 2014;15(1):100-8.
113. Laverack G, Labonte R. A planning framework for community empowerment goals within health promotion. Health Policy and Planning. 2000;15(3):255-62.
114. Green LW, Kreuter MW. Health Promotion Planning: An Educational and Environmental Approach 2nd edition. Mountain View, California: Mayfield Publishing Co.; 1991.
115. Bartholomew LK, Parcel GS, Kok G, Gottlieb NH, Schaalma HC, Markham CC, et al. Planning Health Promotion Programs: An Intervention Mapping Approach. California: Jossey-Bass Inc.; 2006.
116. W K Kellogg Foundation. WK Kellogg Foundation Logic Model Development Guide. Battle Creek Michigan: WK Kellogg Foundation; 2004.
117. Poland B, Krupa G, McCall D. Settings for health promotion: an analytic framework to guide intervention design and implementation. Health Promotion Practice. 2009;10(4):505-16.
118. Vanderkruik R, McPherson ME. A Contextual Factors Framework to Inform Implementation and Evaluation of Public Health Initiatives. American Journal of Evaluation. 2017;38(3):348-59.
119. Crosby R, Noar SM. What is a planning model? An introduction to PRECEDE-PROCEED. Journal of Public Health Dentistry. 2011;71(SUPPL. 1):S7-S15.
120. Centers for Disease Control and Prevention (CDC). Framework for program evaluation in public health. 1999;48(RR-11).
121. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: new guidance. Medical Research Council; 2008.
122. Moore G, Audrey S, Barker M, Bond L, Bonell C, Cooper C, et al. Process evaluation of complex interventions: Medical Research Guidance. British Medical Journal. 2015;350:h1258(2):101-2.
123. Douthwaite B, Alvarez S, Thiele G, Mackay R. Participatory Impact Pathway Analysis: A practical method for project planning and evaluation. The Institutional Learning and Change (ILAC); 2008. Contract No.: Brief 17 May 2008.
124. Douthwaite B, Alvarez S. Participatory Impact Pathways Analysis Manual [Internet]. 2018 [updated 29 September 2018; cited 2021 Feb 8]. Available from: http://pipamethodology.pbworks.com/w/page/70345685/PIPA%20Manual

125. Abram JK, Hand RK, Parrott JS, Brown K, Ziegler PJ, Steiber AL. From the Academy: What Is Your Nutrition Program Missing? Finding Answers with the Guide for Effective Nutrition Interventions and Education (GENIE). Journal of the Academy of Nutrition and Dietetics. 2015;115:122-30.
126. Hand RK, Abram JK, Brown K, Ziegler PJ, Parrott JS, Steiber AL. Development and Validation of the Guide for Effective Nutrition Interventions and Education (GENIE): A Tool for Assessing the Quality of Proposed Nutrition Education Programs. Journal of Nutrition Education and Behavior. 2015;47(4):308-16.
127. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. Journal of Epidemiol Community Health. 2012;66(12):1182-6.
128. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, et al. Using natural experiments to evaluate population health interventions - guidance for producers and users of evidence [Internet]. Glasgow: Medical Research Council; 2011 [updated 2011; cited 2021 Feb 8]. Available from: https://mrc.ukri.org/documents/pdf/natural-experiments-guidance/
129. Themessl-Huber M, Lazenbatt Anne A, Taylor J. Overcoming health inequalities: A participative evaluation framework fit for the task. Journal of The Royal Society for the Promotion of Health. 2008;128(3):117-22.
130. Goldsmith R, Harris S. Thinking Inside the Box: The Health Cube Paradigm for Health and Wellness Program Evaluation and Design. Population Health Management. 2013;16(5):291-5.
131. South J, Fairfax P, Green E. Developing an assessment tool for evaluating community involvement. Health Expectations. 2005;8(1):64-73.
132. World Health Organization Regional Office of Europe. Good Practice Appraisal Tool for Obesity Prevention Programmes, Projects, Initiatives and Interventions. Copenhagen: World Health Organization; 2011.
133. Israel BA, Cummings KM, Dignan MB, Heaney CA, Perales DP, Simons-Morton BG, et al. Evaluation of Health Education Programs: Current Assessment and Future Directions. Health Education Quarterly. 1995;22(3):364-89.
134. Poland BD. Knowledge development and evaluation in, of and for Healthy Community initiatives. Part I: Guiding principles. Health Promotion International. 1996;11(3):237-47.
135. Hawe P, Dageling D, Hall J. Evaluating Health Promotion: A Health Workers Guide. Sydney: Maclennan & Petty Pty Ltd; 1990.
136. Patton M. Utilization-focused evaluation. 3rd edition ed. London SAGE; 1997.
137. Brännström IA, Persson LÅM, Wall SI. Towards a framework for outcome assessment of health intervention: Conceptual and methodological considerations. European Journal of Public Health. 1994;4(2):125-30.
138. Bartholomew LK, Parcel GS, Kok G. Intervention mapping: a process for developing theory and evidence-based health education programs. Health Education Behavior. 1998;25(5):454-63.
139. Sanderson C, Haglund BJA, Tillgren P, Svanström L, Östenson CG, Holm LE, et al. Effect and stage models in community intervention programmes; and the development of the model for management of intervention programme preparation (MMIPP). Health Promotion International. 1996;11(2):143-56.
140. Nutbeam D. Evaluating Health Promotion—Progress, Problems and solutions. Health Promotion International. 1998;13(1):27-44.
141. Goodman RM. Principles and tools for evaluating community-based prevention and health promotion programs. Journal of Public Health Management and Practice. 1998;4(2):37-47.
142. Kegler MC, Twiss JM, Look V. Assessing community change at multiple levels: The genesis of an evaluation framework for the California Healthy Cities Project. Health Education and Behavior. 2000;27(6):760-79.
143. Wallerstein N, Polascek M, Maltrud K. Participatory evaluation model for coalitions: The development of system indicators. Health Promotion Practice. 2002;3(3):361-73.

144. Wallerstein N. A participatory evaluation model for healthier communities: Developing indicators for New Mexico. Public Health Reports. 2000;115(2-3):199-204.
145. Rootman I, Goodstadt M, Hyndman B, McQueen DV, Potvin L, Springett J, et al. Evaluation in health promotion: principles and perspectives Geneva: WHO Regional Office Europe; 2001.
146. Harris R, Hardman E. A formative model of service evaluation. Journal of Clinical Excellence. 2001;3(2):69-73.
147. O'Connor-Fleming ML, Parker E, Higgins H, Gould T. A framework for evaluating health promotion programs. Health Promotion Journal of Australia. 2006;17(1):61-6.
148. Steckler AB, Linnan L. Process evaluation for public health interventions and research. San Francisco, California: Jossey-Bass; 2002.
149. Saunders RP, Evans MH, Joshi P. Developing a Process-Evaluation Plan for Assessing Health Promotion Program Implementation: A How-To Guide. Health Promotion Practice. 2005;6(2):134-47.
150. Baranowski T, Stables G. Process Evaluations of the 5-a-Day Projects. Health Education & Behavior. 2000;27(2):157-66.
151. Beery WL, Senter S, Cheadle A, Greenwald HP, Pearson D, Brousseau R, et al. Evaluating the legacy of community health initiatives: A conceptual framework and example from the California Wellness Foundation's health improvement initiative. American Journal of Evaluation. 2005;26(2):150-65.
152. Chinman M, Imm P, Wandersman A, Kaftarian S, Neal J, Pendleton K, et al. Using the getting to outcomes (GTO) model in a statewide prevention initiative. Health Promotion Practice. 2001;2(4):302-9.
153. Butterfoss FD, Francisco VT. Evaluating Community Partnerships and Coalitions with Practitioners in Mind. Health Promotion Practice. 2004;5(2):108-14.
154. Jolley G, Lawless A, Hurley C. Framework and tools for planning and evaluating community participation, collaborative partnerships and equity in health promotion. Health Promotion Journal of Australia. 2008;19(2):152-7.
155. Kruger DJ, Morrel-Samuels S, Davis-Satterla L, Harris-Ellis BJ, Slonim A. Developing a Cross-Site Evaluation Tool for Diverse Health Interventions. Health Promotion Practice. 2010;11(4):555-61.
156. Brandstetter S, Curbach J, Lindacher V, Rueter J, Warrelmann B, Loss J. Empowerment for healthy nutrition in German communities: A study framework. Health Promotion International. 2017;32(3):500-10.
157. Jolley G. Evaluating complex community-based health promotion: Addressing the challenges. Evaluation and Program Planning. 2014;45:71-81.
158. Dunne A, Scriven A, Howe A. Research partnership and knowledge transfer in the development of a generic evaluation toolkit for health promotion interventions. International Public Health Journal. 2011;3(2):253.
159. Wagemakers A, Koelen MA, Lezwijn J, Vaandrager L. Coordinated action checklist: A tool for partnerships to facilitate and evaluate community health promotion. Global Health Promotion. 2010;17(3):17-28.
160. Masso M, Quinsey K, Fildes D. Evolution of a multilevel framework for health program evaluation. Australian Health Review. 2017;41(3):239-45.
161. Mantziki K, Renders CM, Westerman MJ, Mayer J, Borys JM, Seidell JC. Tools for a systematic appraisal of integrated community-based approaches to prevent childhood obesity. BMC Public Health. 2018;18(1):1-12.
162. Hoddinott P, Allan K, Avenell A, Britten J. Group interventions to improve health outcomes: A framework for their design and delivery. BMC Public Health. 2010;10(1):1-9.
163. Grant A, Treweek S, Dreischulte T, Foy R, Guthrie B. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. Trials. 2013;14:15.

164. Wagemakers A, Vaandrager L, Koelen MA, Saan H, Leeuwis C. Community health promotion: A framework to facilitate and evaluate supportive social environments for health. Evaluation and Program Planning. 2010;33(4):428-35.
165. W.K. Kellogg Foundation. The Step-by-Step Guide to Evaluation Battle Creek, Michigan: W.K. Kellogg Foundation; 2017.
166. National Institute for Health and Care Excellence (NICE). Behaviour Change: individual approaches Public Health Guideline [PH49]. London: NICE; 2014. Available from: https://www.nice.org.uk/guidance/ph49/chapter/1-recommendations#recommendation-15-monitor-behaviour-change-interventions
167. Taylor M, Purdue D, Wilson M, Wilde P. Evaluating Community Projects: A practical guide. York: Joseph Rowntree Foundation; 2005. Available from: https://www.jrf.org.uk/report/evaluating-community-projects-practical-guide.
168. The Centre for the Advancement of Community Based Public Health. An Evaluation Framework for Community Health Programs. Durham, NC.: The Centre for the Advancement of Community Based Public Health; 2000.
169. Dugdill L, Stratton G. Evaluating sport and physical activity interventions: a guide for practitioners. Salford: University of Salford; 2007.
170. Health System Intelligence Project (HSIP). The Health Planner's Toolkit. Ontario HSIP; 2006. Available from: http://www.ontla.on.ca/library/repository/mon/22000/283845.pdf.
171. National Health Service (NHS) Health Scotland. LEAP for Health: Learning, Evaluation and Planning. Edinburgh: Health Scotland; 2003. Available from: http://www.healthscotland.com/uploads/documents/308-LEAP_for_health.pdf.
172. US Department of Health and Human Services. Physical Activity Evaluation Handbook. Atlanta GA: US Department of Health and Human Services, Centers for Disease Control and Prevention; 2002.
173. Sport England. Evaluation Framework: Measurement and Evaluation Guidance Tools and Resources [Internet]. London: Sport England; [date unknown; cited 2021 Feb 8] Available from: https://evaluationframework.sportengland.org/
174. Better Evaluation. Better Evaluation Rainbow Framework and Planning Tool [Internet]. Australia: Better Evaluation Ltd; 2014 [updated 2014; cited 2021 Feb 8]. Available from: http://www.betterevaluation.org
175. Leeman J, Sommers J, Vu M, Jernigan J, Payne G, Thompson D, et al. An evaluation framework for obesity prevention policy interventions. Preventing Chronic Disease. 2012;9:11_0322.
176. Community Toolbox. Community Toolbox: Evaluating Community Programs and Initiative [Internet]. Kansas: Centre for community health and development of University of Kansas; 2018 [updated 2018; cited 2021 Feb 8]. Available from: https://ctb.ku.edu/en/evaluating-initiative
177. West of England Academic Health Science Network, BNSSG Research and Evidence Team, NIHR CLAHRC West. Evaluation Works: A toolkit to support commissioning of health and care services [Internet]. Bristol: West of England Academic Health Science Network, BNSSG Research and Evidence Team, NIHR CLAHRC West; 2016 [updated 2016; cited 2021 Feb 8]. Available from: http://www.nhsevaluationtoolkit.net/
178. Cavill N, Roberts K, Ells L. Evaluation of weight management, physical activity and dietary interventions: an introductory guide. Oxford: Public Health England; 2015.
179. HM Treasury. The Magenta Book: Guidance for Evaluation. London: HM Treasury; 2011.
180. Ontario Agency for Health Protection and Promotion (Public Health Ontario), Snelling S, Meserve A. Evaluating health promotion programs: introductory workbook [Internet]. Toronto, ON: Queen's Printer for Ontario; 2016 [updated 2016; cited 2021 Feb 8]. Available from: https://www.publichealthontario.ca/en/erepository/Evaluating_health_promotion_programs_workbook_2016.pdf

181. Prevention and Population Health Branch. Evaluation framework for health promotion and disease prevention programs [Internet]. Melbourne: Victorian Government Department of Health and Human Services; 2011 [updated 2011; cited 2021 Feb 8] Available from: https://www2.health.vic.gov.au
182. Pérez-Escamilla R, Segura-Pérez S, Damio G. Applying the Program Impact Pathways (PIP) evaluation framework to school-based healthy lifestyles programs: Workshop Evaluation Manual. Food and Nutrition Bulletin. 2014;35:S97-S107.
183. Kamada M, Kitayuguchi J, Abe T, Taguri M, Inoue S, Ishikawa Y, et al. Community-wide intervention and population-level physical activity: A 5-year cluster randomized trial. International Journal of Epidemiology. 2018;47(2):642-53.
184. Van Hoye A, Vandoorne C, Absil G, Lecomte F, Fallon C, Lombrail P, et al. Health enhancing physical activity in all policies? Comparison of national public actors between France and Belgium. Health Policy. 2019;123(3):327-32.
185. Knottnerus A, Tugwell P. STROBE--a checklist to Strengthen the Reporting of Observational Studies in Epidemiology. Journal of Clinical Epidemiology. 2008;61(4):323.
186. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Journal of Clinical Epidemiology. 2008;61(4):344-9.
187. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. Bmj. 2014;348:g1687.
188. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Annals of Behavioral Medicine. 2013;46(1):81-95.
189. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: a systematic review of use over time. American Journal of Public Health. 2013;103(6):e38-e46.
190. Steinmo S, Fuller C, Stone SP, Michie S. Characterising an implementation intervention in terms of behaviour change techniques and theory: the 'Sepsis Six'clinical care bundle. Implementation Science. 2015;10(1):111.
191. Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. Health Psychology. 2010;29(1):1-8.
192. Rossi PH, Lipsey MW, Henry GT. Evaluation: A systematic approach. California: Sage publications; 2018.
193. Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. American Journal of Community Psychology. 2008;41(3-4):327-50.
194. World Health Organization (WHO). Process Evaluation Workbook 4. Geneva: WHO; 2001.
195. Spencer B, Broesskamp-Stone U, Ruckstuhl B, Ackermann G, Spoerri A, Cloetta B. Modelling the results of health promotion activities in Switzerland: Development of the Swiss model for outcome classification in health promotion and prevention. Health Promotion International. 2008;23(1):86-97.
196. Harrington DM, Davies MJ, Bodicoat DH, Charles JM, Chudasama YV, Gorely T, et al. Effectiveness of the 'Girls Active' school-based physical activity programme: A cluster randomised controlled trial. International Journal of Behavioral Nutrition & Physical Activity. 2018;15(1):1-18.
197. Jago R, Sebire SJ, Davies B, Wood L, Banfield K, Edwards MJ, et al. Increasing children's physical activity through a teaching-assistant led extracurricular intervention: Process evaluation of the action 3:30 randomised feasibility trial. BMC Public Health. 2015;15(1):1-15.
198. Jago R, Edwards MJ, Sebire SJ, Tomkinson K, Bird EL, Banfield K, et al. Effect and cost of an after-school dance programme on the physical activity of 11-12 year old girls: The Bristol Girls

Dance Project, a school-based cluster randomised controlled trial. International Journal of Behavioral Nutrition and Physical Activity. 2015;12(1):1-15.
199. Sebire SJ, .Jago R, Banfield K, Edwards MJ, Campbell R, Kipping R, et al. Results of a feasibility cluster randomised controlled trial of a peer-led school-based intervention to increase the physical activity of adolescent girls (PLAN-A). International Journal of Behavioral Nutrition and Physical Activity. 2018;15(1):1-13.
200. Biddle SJ, Edwardson CL, Gorely T, Wilmot EG, Yates T, Nimmo MA, et al. Reducing sedentary time in adults at risk of type 2 diabetes: process evaluation of the STAND (Sedentary Time ANd Diabetes) RCT. BMC Public Health. 2017;17(1):80.
201. Sebire SJ, Kesten JM, Edwards MJ, May T, Banfield K, Tomkinson K, et al. Using self-determination theory to promote adolescent girls' physical activity: Exploring the theoretical fidelity of the Bristol Girls Dance Project. Psychology of Sport and Exercise. 2016;24:100-10.
202. Grillich L, Kien C, Takuya Y, Weber M, Gartlehner G. Effectiveness evaluation of a health promotion programme in primary schools: a cluster randomised controlled trial. BMC Public Health. 2016;16:679.
203. Quinn L, Trubey R, Gobat N, Dawes H, Edwards RT, Jones C, et al. Development and delivery of a physical activity intervention for people with huntington disease: Facilitating translation to clinical practice. Journal of Neurologic Physical Therapy. 2016;40(2):71-80.
204. Ullrich-French S, Cole AN, Montgomery AK. Evaluation development for a physical activity positive youth development program for girls. Evaluation and Program Planning. 2016;55:67-76.
205. Young T, Sharpe C. Process evaluation results from an intergenerational physical activity intervention for grandparents raising grandchildren. Journal of Physical Activity and Health. 2016;13(5):525-33.
206. Smedegaard S, Brondeel R, Christiansen LB, Skovgaard T. What happened in the 'Move for Well-being in School': A process evaluation of a cluster randomized physical activity intervention using the RE-AIM framework. International Journal of Behavioral Nutrition & Physical Activity. 2017;14(1):1-11.
207. Berendsen BA, Hendriks MR, Verhagen EA, Schaper NC, Kremers SP, Savelberg HH. Effectiveness and cost-effectiveness of'BeweegKuur', a combined lifestyle intervention in the Netherlands: rationale, design and methods of a randomized controlled trial. BMC Public Health. 2011;11(1):1-9.
208. van Dongen EJD, G.Jansen, S. C.Ter Beek, J.Huijg, J. M.Leerlooijer, J. N.Hiddink, G. J.Feskens, E. J.Haveman-Nies, A. Process evaluation of a randomised controlled trial of a diabetes prevention intervention in Dutch primary health care: the SLIMMER study. Public health nutrition. 2016;19(16):3027-38.
209. Meng Y, Wong SS, Manore MM, Patton-López M. WAVE~Ripples for Change Obesity Two-Year Intervention in High School Soccer Players: Process Evaluation, Best Practices, and Youth Engagement. Nutrients. 2018;10(6):711.
210. Blackford K, Lee A, James AP, Waddell T, Hills AP, Anderson AS, et al. Process evaluation of the Albany Physical Activity and Nutrition (APAN) program, a home-based intervention for metabolic syndrome and associated chronic disease risk in rural Australian adults. Health Promotion Journal of Australia. 2017;28(1):8-14.
211. McCarthy MM, Dickson VV, Katz SD, Sciacca K, Chyun DA. Process evaluation of an exercise counseling intervention using motivational interviewing. Applied Nursing Research. 2015;28(2):156-62.
212. Curry WB, Dagkas S, Wilson M. Newham's every child a sports person (NECaSP): A summative process evaluation of a school- and community-based intervention in east London, United Kingdom. Journal of Physical Activity and Health. 2016;13(10):1124-31.
213. Webster CA, Weaver RG, Egan CA, Brian A, Vazou S. Two-year process evaluation of a pilot program to increase elementary children's physical activity during school. Evaluation and Program Planning. 2018;67:200-6.

214. De Craemer M, Verloigne M, De Bourdeaudhuij I, Androutsos O, Iotova V, Moreno L, et al. Effect and process evaluation of a kindergarten-based, family-involved cluster randomised controlled trial in six European countries on four- to six-year-old children's steps per day: the ToyBox-study. International Journal of Behavioral Nutrition & Physical Activity. 2017;14:1-16.
215. Wright CM, Chomitz VR, Duquesnay PJ, Amin SA, Economos CD, Sacheck JM. The FLEX study school-based physical activity programs - Measurement and evaluation of implementation. BMC Public Health. 2019;19(1):1-12.
216. Chow AF, Leis A, Humbert L, Muhajarine N, Engler-Stringer R. Healthy start – Départ santé: A pilot study of a multilevel intervention to increase physical activity, fundamental movement skills and healthy eating in rural childcare centres. Canadian Journal of Public Health. 2016;107(3):e312-e8.
217. Griffiths LA, Griffiths MA. Evaluation of a School–Community Linked Physical Activity Intervention Targeting 7- to 12-Year-Olds: A Sociocultural Perspective. American Journal of Health Education. 2019;50(2):112-26.
218. Burkart S, St. Laurent CW, Alhassan S. Process evaluation of a culturally-tailored physical activity intervention in African-American mother-daughter dyads. Preventive Medicine Reports. 2017;8:88-92.
219. Corder K, Brown HE, Schiff A, van Sluijs EM. Feasibility study and pilot cluster-randomised controlled trial of the GoActive intervention aiming to promote physical activity among adolescents: outcomes and lessons learnt. BMJ open. 2016;6(11):e012335.
220. Avery L, Charman SJ, Taylor L, Flynn D, Mosely K, Speight J, et al. Systematic development of a theory-informed multifaceted behavioural intervention to increase physical activity of adults with type 2 diabetes in routine primary care: Movement as Medicine for Type 2 Diabetes. Implementation Science. 2016;11(1):1-21.
221. Biddle SJ, Edwardson CL, Wilmot EG, Yates T, Gorely T, Bodicoat DH, et al. A randomised controlled trial to reduce sedentary time in young adults at risk of type 2 diabetes mellitus: project STAND (Sedentary Time ANd Diabetes). PLoS One. 2015;10(12):e0143398.
222. Sebire SJ, Edwards MJ, Kesten JM, May T, Banfield KJ, Bird EL, et al. Process evaluation of the Bristol girls dance project. BMC Public Health. 2016;16(1):1-13.
223. Lefler LL, .Jones S, Harris B. Key Strategies for Physical Activity Interventions Among Older Women: Process Evaluation of a Clinical Trial. American Journal of Health Promotion. 2018;32(3):561-70.
224. Furness C, Howard E, Limb E, Cook DG, Kerry S, Wahlich C, et al. Relating process evaluation measures to complex intervention outcomes: findings from the PACE-UP primary care pedometer-based walking trial. Trials. 2018;19(1):1-8.
225. Bonde AH, Stjernqvist NW, Sabinsky MS, Maindal HT. Process evaluation of implementation fidelity in a Danish health-promoting school intervention. BMC Public Health. 2018;18(1):1-10.
226. Dadaczynski K, Schiemann S, Backhaus O. Promoting physical activity in worksite settings: Results of a German pilot study of the online intervention Healingo fit. BMC Public Health. 2017;17(1):1-9.
227. Green J, Tones K. Health Promotion: Planning and Strategies. 3rd edition ed. London: SAGE; 2010.
228. Tucker P, Vanderloo LM, Johnson AM, Burke SM, Irwin JD, Gaston A, et al. Impact of the Supporting Physical Activity in the Childcare Environment (SPACE) intervention on preschoolers' physical activity levels and sedentary time: A single-blind cluster randomized controlled trial. International Journal of Behavioral Nutrition and Physical Activity. 2017;14(1):1-11.
229. Sofija E, Plugge M, Wiseman N, Harris N. 'This is the beginning of the new me': Process evaluation of a group fitness intervention to promote wellbeing in formerly homeless individuals. BMC Public Health. 2018;18(1):1-14.

230. Matthews L, Kirk A, McCallum M, Mutrie N, Gold A, Keen A. The feasibility of a physical activity intervention for adults within routine diabetes care: a process evaluation. Practical Diabetes. 2017;34(1):7-12a.
231. Galaviz KI, Estabrooks PA, Ulloa EJ, Lee RE, Janssen I, Lopez y Taylor J, et al. Evaluating the effectiveness of physician counseling to promote physical activity in Mexico: an effectiveness-implementation hybrid study. Translational Behavioral Medicine. 2017;7(4):731-40.
232. Christian D, Todd C, Hill R, Rance J, Mackintosh K, Stratton G, et al. Active children through incentive vouchers - evaluation (ACTIVE): a mixed-method feasibility study. BMC Public Health. 2016;16(1):1-10.
233. Reznik M, Wylie-Rosett J, Kim M, Ozuah PO. A classroom-based physical activity intervention for urban kindergarten and first-grade students: A feasibility study. Childhood Obesity. 2015;11(3):314-24.
234. Kohn M, Belza B, Petrescu-Prahova M, Miyawaki CE. Beyond Strength: Participant Perspectives on the Benefits of an Older Adult Exercise Program. Health Education and Behavior. 2015;43(3):305-12.
235. Petrescu-Prahova M, Belza B, Kohn M, Miyawaki C. Implementation and Maintenance of a Community-Based Older Adult Physical Activity Program. Gerontologist. 2016;56(4):677-86.
236. Wilcox S, Saunders RP, Kaczynski AT, Forthofer M, Sharpe PA, Goodwin C, et al. Faith, Activity, and Nutrition Randomized Dissemination and Implementation Study: Countywide Adoption, Reach, and Effectiveness. American Journal of Preventive Medicine. 2018;54(6):776-85.
237. Baba CT, Martins Oliveira I, Ferreira Silva AE, Vieira LM, Cerri NC, Florindo AA, et al. Evaluating the impact of a walking program in a disadvantaged area: using the RE-AIM framework by mixed methods. BMC Public Health. 2017;17:1-11.
238. Economos CD, Anzman-Frasca S, Koomas AH, Bakun PJ, Brown CM, Brown D, et al. Dissemination of healthy kids out of school principles for obesity prevention: A RE-AIM analysis. Preventive Medicine. 2019;119:37-43.
239. Ward S, Chow AF, Humbert ML, Belanger M, Muhajarine N, Vatanparast H, et al. Promoting physical activity, healthy eating and gross motor skills development among preschoolers attending childcare centers: Process evaluation of the Healthy Start-Depart Sante intervention using the RE-AIM framework. Evaluation and Program Planning. 2018;68:90-8.
240. Jung ME, Bourne JE, Gainforth HL. Evaluation of a community-based, family focused healthy weights initiative using the RE-AIM framework. International Journal of Behavioral Nutrition and Physical Activity. 2018;15(1):1-16.
241. Allar I, .Jones E, Elliott E, Kristjansson A, Taliaferro A, Mann M, et al. The perceived impact of i am moving, i am learning on physical activity and family involvement: A preliminary investigation. American Journal of Health Behavior. 2017;41(6):683-92.
242. Caperchione CM, Duncan M, Kolt GS, Vandelanotte C, Rosenkranz RR, Maeder A, et al. Examining an Australian physical activity and nutrition intervention using RE-AIM. Health Promotion International. 2016;31(2):450-8.
243. Gonzalez-Viana A, Violan Fors M, Castell Abat C, Rubinat Masot M, Oliveras L, Garcia-Gil J, et al. Promoting physical activity through primary health care: the case of Catalonia. BMC Public Health. 2018;18(1):1-7.
244. Schwingel A, Gálvez P, Linares D, Sebastião E. Using a Mixed-Methods RE-AIM Framework to Evaluate Community Health Programs for Older Latinas. Journal of Aging and Health. 2017;29(4):551-93.
245. Paez DC, Reis RS, Parra DC, Hoehner CM, Sarmiento OL, Barros M, et al. Bridging the gap between research and practice: an assessment of external validity of community-based physical activity programs in Bogotá, Colombia, and Recife, Brazil. Translational Behavioral Medicine. 2015;5(1):1-11.

246. Koorts H, Gillison F. Mixed method evaluation of a community-based physical activity program using the RE-AIM framework: Practical application in a real-world setting Energy balance-related behaviors. BMC Public Health. 2015;15(1):1-10.
247. Bellicha A, Kieusseian A, Fontvieille A-M, Tataranni A, Copin N, Charreire H, et al. A multistage controlled intervention to increase stair climbing at work: effectiveness and process evaluation. International Journal of Behavioral Nutrition and Physical Activity. 2016;13:1-9.
248. Beets MW, Weaver RG, Brazendale K, Turner-Mcgrievy G, Saunders RP, Moore JB, et al. Statewide dissemination and implementation of physical activity standards in afterschool programs: Two-year results. BMC Public Health. 2018;18(1):1-14.
249. Beets MW, Weaver RG, Turner-McGrievy G, Saunders RP, Webster CA, Moore JB, et al. Evaluation of a statewide dissemination and implementation of physical activity intervention in afterschool programs: a nonrandomized trial. Translational Behavioral Medicine. 2017;7(4):690-701.
250. Lee RE, Parker NH, Soltero EG, Ledoux TA, Mama SK, McNeill L. Sustainability via Active Garden Education (SAGE): results from two feasibility pilot studies. BMC Public Health. 2017;17:1-11.
251. Lewis ZH, Ottenbacher KJ, Fisher SR, Jennings K, Brown AF, Swartz MC, et al. The feasibility and RE-AIM evaluation of the TAME health pilot study. International Journal of Behavioral Nutrition and Physical Activity. 2017;14:1-15.
252. Adams EJ, Chalkley AE, Esliger DW, Sherar LB. Evaluation of the implementation of a whole-workplace walking programme using the RE-AIM framework. BMC Public Health. 2017;17:1-18.
253. Herens M, Wagemakers A, Vaandrager L, van Ophem J, Koelen M. Contexts, Mechanisms, and Outcomes That Matter in Dutch Community-Based Physical Activity Programs Targeting Socially Vulnerable Groups. Evaluation and the Health Professions. 2017;40(3):294-331.
254. Daniels JE, Bell B, Horrocks C. Capturing the realities of sports programmes: systematic 'messiness'? International Journal of Sport Policy and Politics. 2018;10(4):779-94.
255. Willis CE, Reid S, Elliott C, Rosenberg M, Nyquist A, Jahnsen R, et al. A realist evaluation of a physical activity participation intervention for children and youth with disabilities: What works, for whom, in what circumstances, and how? BMC Pediatrics. 2018;18(1):1-15.
256. Mikkelsen BE, Bloch P, Reinbach HC, Buch-Andersen T, Winkler LL, Toft U, et al. Project SoL—A community-based, multi-component health promotion intervention to improve healthy eating and physical activity practices among danish families with young children part 2: Evaluation. International Journal of Environmental Research and Public Health. 2018;15(6):1097-112.
257. Fournier B, Manon P, Johanne F, Nathalie B, Lorthios-Guilledroit A, Marie-Ève M. Development and implementation of a community-based pole walking program for older adults. Activities, Adaptation and Aging. 2019;43(1):1-22.
258. Driediger M, Vanderloo LM, Burke SM, Irwin JD, Gaston A, Timmons BW, et al. The Implementation and Feasibility of the Supporting Physical Activity in the Childcare Environment (SPACE) Intervention: A Process Evaluation. Health Education and Behavior. 2018;45(6):935-44.
259. Verloigne M, Ahrens W, De Henauw S, Verbestel V, Marild S, Pigeot I, et al. Process evaluation of the IDEFICS school intervention: Putting the evaluation of the effect on children's objectively measured physical activity and sedentary time in context. Obesity Reviews. 2015;16(Supplement 2):89-102.
260. Verwey R, van der Weegen S, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. Process evaluation of physical activity counselling with and without the use of mobile technology: A mixed methods study. International Journal of Nursing Studies. 2016;53:3-16.
261. Griffin TL, Clarke JL, Lancashire ER, Pallan MJ, Adab P. Process evaluation results of a cluster randomised controlled childhood obesity prevention trial: the WAVES study. BMC Public Health. 2017;17:1-13.

262. Matthews L, Mitchell F, Stalker K, McConnachie A, Murray H, Melling C, et al. Process evaluation of the Walk Well study: a cluster-randomised controlled trial of a community based walking programme for adults with intellectual disabilities. BMC Public Health. 2016;16(1):1-11.
263. Berendsen BAJ, Kremers SPJ, Savelberg HHCM, Schaper NC, Hendriks MRC. The implementation and sustainability of a combined lifestyle intervention in primary care: mixed method process evaluation. BMC Family Practice. 2015;16(1):1-12.
264. Torquati L, Kolbe-Alexander T, Pavey T, Leveritt M. Changing Diet and Physical Activity in Nurses: A Pilot Study and Process Evaluation Highlighting Challenges in Workplace Health Promotion. Journal of Nutrition Education and Behavior. 2018;50(10):1015-25.
265. Saunders RP, Pfeiffer K, Brown WH, Howie EK, Dowda M, O'Neill JR, et al. Evaluating and Refining the Conceptual Model Used in the Study of Health and Activity in Preschool Environments (SHAPES) Intervention. Health Education and Behavior. 2017;44(6):876-84.
266. Stjernqvist NW, Sabinsky M, Morgan A, Trolle E, Thyregod C, Maindal HT, et al. Building school-based social capital through 'We Act - Together for Health' – a quasi-experimental study. BMC Public Health. 2018;18(1):1-13.
267. Harden SM, Strayer TE, 3rd, Smith ML, Gaglio B, Ory MG, Rabin B, et al. National Working Group on the RE-AIM Planning and Evaluation Framework: Goals, Resources, and Future Directions. Front Public Health. 2019;7(390):1-7.
268. Datta J, Petticrew M. Challenges to evaluating complex interventions: a content analysis of published papers. BMC Public Health. 2013;13(568):1-18.
269. Petticrew M, Rehfuess E, Noyes J, Higgins JP, Mayhew A, Pantoja T, et al. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. Journal of Clinical Epidemiology. 2013;66(11):1230-43.
270. Brownson RC, Roux AVD, Swartz K. Commentary: generating rigorous evidence for public health: the need for new thinking to improve research and practice. Annual Review of Public Health. 2014;35:1-7.
271. Glasgow RE, Emmons KM. How Can We Increase Translation of Research into Practice? Types of Evidence Needed. Annual Review of Public Health. 2007;28(1):413-33.
272. Northridge ME, Metcalf SS. Enhancing implementation science by applying best principles of systems science. Health Research Policy and Systems. 2016;14(1):1-8.
273. Sport England. New £5 million drive to help tackle inactivity [Internet] London: Sport England; 2014 [updated 2014; cited 2021 Feb 8] Available from https://www.sportengland.org/news/new-5-million-drive-to-help-tackle-inactivity
274. Sport England. Why we're here [Internet]. London: Sport England; 2021 [updated 2021; cited 2021 Feb 8] Available from: https://www.sportengland.org/why-were-here#buildinganactivenation-5887.
275. Cavill N, Adams E, Gardner S, Ruane S. Tackling Inactivity What we know: Key insights from our Get Healthy Get Active pilots. London: Sport England; 2016.
276. Milton K, Bull F, Bauman A. Reliability and validity testing of a single-item physical activity measure. British Journal of Sports Medicine. 2011;45(3):203-8.
277. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Medicine & Science in Sports & Exercise. 2003;35(8):1381-95.
278. Richie J, Spencer L. Qualitative data analysis for applied policy research. Analysing qualitative data. Edited by: Bryman A, Burgess B. 1994. London: Routledge.
279. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Medical Research Methodology. 2013;13(1):1-8.
280. Cavill N, Richardson D, Foster C. Improving health through participation in sport: a review of research and practice. Oxford: British Heart Foundation Health Promotion Research Group University of Oxford; 2012.

281. Department of Health. Start Active, Stay Active: A report on physical activity from the four home countries' Chief Medical Officers. London: HM Government; 2011.
282. Sport England. Creating a sporting habit for life: a new youth sport strategy. London: Department for Digital, Culture, Media & Sport; 2012.
283. Sport England. Tackling Inactivity: The Design Principles [Internet]. London: Sport England; 2016 [updated 2016; cited 2021 Feb 8]. Available from: https://sportengland-production-files.s3.eu-west-2.amazonaws.com/s3fs-public/tackling-inactivity-design-principles.pdf
284. Sport England. Tackling Inactivity. London: Sport England; 2016.
285. Sport England. Applying Behaviour Change Theories: Real World Examples from The Get Healthy Get Active Programme. London: Sport England; 2016.
286. Varney J, Lawson R, Williams T, Copeland R, Brannan M, Lane A, et al. Moving at scale: Promising practice and practical guidance on evaluation of physical activity programmes in the UK. London: UK Active; 2018.
287. Beedie C, Mann S, Copeland R, Domone S. Identifying what works for local physical inactivity interventions. London: Public Health England; 2014.
288. University of Hertfordshire Higher Education Corporation. Hertfordshire Evaluation Framework [Internet]. Hertfordshire: Hertfordshire Sports Partnership; 2017 [updated 2017; cited 2021 Feb 8]. Available from: https://sportinherts.org.uk/app/uploads/2018/05/Hertfordshire-Evaluation-Framework.pdf
289. Mansfield L, Kay T, Anokye N, Fox-Rushby J. Community Sport and the Politics of Aging: Co-design and Partnership Approaches to Understanding the Embodied Experiences of Low-Income Older People. Frontiers in Sociology. 2019;Feb 12;4:5.
290. Anokye N, Mansfield L, Kay T, Sanghera S, Lewin A, Fox-Rushby J. The effectiveness and cost-effectiveness of a complex community sport intervention to increase physical activity: an interrupted time series design. BMJ Open 2018;8(12):e024132.
291. Mansfield L, Anokye N, Fox-Rushby J, Kay T. The Health and Sport Engagement (HASE) Intervention and Evaluation Project: protocol for the design, outcome, process and economic evaluation of a complex community sport intervention to increase levels of physical activity. BMJ open. 2015;5(10):e009276.
292. Burke A, Jones A, Hughes R, Player E. From evidence to practice: Developing best practice guidelines for the delivery of activities to people living with moderate to advanced dementia using a pragmatic observational study. Dementia. 2020:1471301220957805.
293. Buffin J. The Benefits of the Commuinity Sports Initiative: An evaluation of a 6 month pilot programme to enhance recovery among problematic drug and alcohol users. Lancashire: University of Central Lancashire; 2014.
294. Sanders GJ, Roe B, Knowles ZR, Kaehne A, Fairclough SJ. Using formative research with older adults to inform a community physical activity programme: Get Healthy, Get Active. Primary Health Care Research & Development. 2019;20(e60):1-10.
295. Sanders G. Get Healthy Get Active: prevention is better than care [dissertation]. Ormskirk, UK: Edge Hill University; 2018.
296. Adams EJ, Musson H, Watson, A, Mason, L. Bright spots, physical activity investments that work: Workplace Challenge. British Journal of Sports Medicine. 2018;52(16):1026-8.
297. Atchinson R, Frith G, Roden A, Copeland RJ, Reece LJ. Active for Health Rotherham: be active to stay healthy. British Journal of Sports Medicine. 2019;53(16):1036-7.
298. Cavill N, Richardson D, Faghy M, Bussell C, Rutter H. Using system mapping to help plan and implement city-wide action to promote physical activity. Journal of Public Health Research. 2020;9(3):1759.
299. Jolley GM, Lawless AP, Baum FE, Hurley CJ, Fry D. Building an evidence base for community health: a review of the quality of program evaluations. Australian Health Review. 2007;31(4):603-10.

300. Rychetnik L, Frommer M, Hawe P, Shiell A. Criteria for evaluating evidence on public health interventions. Journal of Epidemiology Community Health. 2002;56:119-27.
301. Koenig G. Realistic evaluation and case studies: stretching the potential. Evaluation. 2009;15(1):9-30.
302. Lipsey MW, Cordray DS. Evaluation methods for social intervention. Annual Review of Psychology. 2000;51(1):345-75.
303. Bauman A, Nutbeam D. Planning and evaluating population interventions to reduce noncommunicable disease risk–reconciling complexity and scientific rigour. Public Health Research and Practice. 2014;25(1):e2511402.
304. Stame N. Theory-based evaluation and types of complexity. Evaluation. 2004;10(1):58-76.
305. World Health Organization. WHO Evaluation Practice Handbook. Geneva: World Health Organization; 2013.
306. Preskill H, Boyle S. A multidisciplinary model of evaluation capacity building. American Journal of Evaluation. 2008;29(4):443-59.
307. Labin SN, Duffy JL, Meyers DC, Wandersman A, Lesesne CA. A research synthesis of the evaluation capacity building literature. American Journal of Evaluation. 2012;33(3):307-38.
308. Sport England. Get Healthy Get Active: learn more about our initial investments into tackling inactivity from 2012-2016. Retrieved; 2017.
309. Borgatti SP, Everett MG, Freeman LC. UCINET for Windows: Software for Social Network Analysis Version 6.705. 2020.
310. Tremblay M-C, Richard L, Brousselle A, Beaudet N. How can both the intervention and its evaluation fulfill health promotion principles? An example from a professional development program. Health Promotion Practice. 2013;14(4):563-71.
311. McKellar KA, Pitzul KB, Juliana YY, Cole DC. Evaluating communities of practice and knowledge networks: a systematic scoping review of evaluation frameworks. Ecohealth. 2014;11(3):383-99.
312. Pinnock H, Barwick M, Carpenter C, Eldridge S, Grandes G, Griffiths C, et al. The Standards for Reporting Implementation studies (StaRI) statement. BMJ. 2017;356:i6795.
313. Egan M, McGill E, Penney T, Anderson de Cuevas R, Er V, Orton L, et al. NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. Part 1: Introducing Systems Thinking. London: National Institute of Health Research School for Public Health Research; 2019.
314. Kothari A, Boyko JA, Conklin J, Stolee P, Sibbald SL. Communities of practice for supporting health systems change: a missed opportunity. Health Research Policy and Systems. 2015;13(1):1-9.
315. Kothari A, Hamel N, MacDonald JA, Meyer M, Cohen B, Bonnenfant D. Exploring Community Collaborations: Social Network Analysis as a Reflective Tool for Public Health. Systemic Practice and Action Research. 2014;27(2):123-37.
316. Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. American Journal of Community Psychology. 2009;43(3-4):267-76.
317. Egan M, Penney T, Anderson de Cuevas R, Er V, Orton L, White M, et al. NIHR SPHR Guidance on Systems Approaches to Local Public Health Evaluation. Part 2: What to consider when planning a systems evaluation. London: National Institue of Health Reseach School for Public Health Research; 2019.
318. Graneheim UH, Lindgren B-M, Lundman B. Methodological challenges in qualitative content analysis: A discussion paper. Nurse Education Today. 2017;56:29-34.
319. Yin RK. Case study research: design and methods 3rd ed., Ser. Applied social research methods series, v. 5 ed. Thousand Oaks, California: Sage Publications; 2003.
320. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Implementation Research: what it is and how to do it. BMJ. 2013;20(347):f6753.