



Standard Evaluation Framework

for physical activity interventions

noo

National Obesity
Observatory

noo.org.uk

Acknowledgements

The National Obesity Observatory (NOO) would like to acknowledge the input of a large number of people who provided valuable comments during the development of this document.

Authors (NOO)

Nick Cavill, Kath Roberts,
Harry Rutter

Editor (NOO)

Di Swanston

Consultees

Emma Adams, Loughborough University
Beelin Baxter, Department of Health
Maria Bryant, University of Leeds
Janet Cade, University of Leeds
Adrian Coggins, NHS West Essex
Nick College, Wigan Leisure and Culture Trust
Ashley Cooper, University of Bristol
Lindsey Dugdill, University of Salford
Louisa Ells, University of Teeside/National Obesity Observatory
Charlie Foster, University of Oxford
Jo Foster, Macmillan
Oliver Francis, University of Cambridge
Claire Glazzard, Kent Community Health NHS Trust
Melvyn Hillsdon, University of Exeter
Susan Jebb, Medical Research Council
Sonia McGeorge, Loughborough University
Claire Ramwel, Blackburn with Darwen Borough Council
Debra Richardson, Independent consultant
Sara Roberts, Cornwall and Isles of Scilly PCT
Jamie Sadler, East End Health
Pinki Sahota, Leeds Metropolitan University
Dawn Scott, NHS North of Tyne
Paul Stonebrook, Department of Health
Gareth Stratton, Swansea University
Madeleine Thomas, NHS Derbyshire County
Kay Thomson, Sport England
Sarah Worbey, NHS Milton Keynes

NOO is also grateful to all those who contributed to the original Standard Evaluation Framework for weight management interventions, on which this publication is based.

NOO was established to provide a single point of contact for wide-ranging authoritative information on data, evaluation and evidence related to obesity, overweight, underweight and their determinants. The Observatory works with a range of organisations and supports policy makers and practitioners involved in obesity and related issues.

Front cover illustration by Michael Heasman

Delivered by NOO on behalf of the Public Health Observatories in England.

September 2012

Contents

➤ 1. Introduction	4
➤ National Obesity Observatory resources on evaluation	4
➤ Introduction to this document	4
➤ What does this document aim to do?	4
➤ What does this document not aim to do?	5
➤ Target audiences	6
➤ What is physical activity?	6
➤ Why do we need a SEF for physical activity interventions?	6
➤ 2. Principles of evaluation	7
➤ 3. Physical activity: selecting and measuring outcomes	9
➤ Selecting physical activity outcomes	9
➤ Options for measuring physical activity	10
➤ Measuring outcomes	11
➤ Potential sources of error and bias	12
➤ Use of existing datasets	13
➤ Further resources on physical activity assessment	13
➤ 4. The SEF for physical activity interventions	15
➤ 5. Explanatory notes	17
➤ Part one: programme details	17
➤ Part two: evaluation details	21
➤ Part three: demographics of individual participants	22
➤ Part four: baseline data	24
➤ Part five: follow-up data (impact evaluation)	25
➤ Part six: process evaluation	26
➤ Part seven: analysis and interpretation	27
➤ 6. Conclusions	29
➤ Appendices	30
➤ Appendix 1: ONS categories for ethnic group monitoring in England	30
➤ Appendix 2: Examples of factors that may be modified to increase physical activity	31
➤ Appendix 3: Physical activity indicators in the 2012 Public Health Outcomes Framework	32
➤ Glossary of terms	34
➤ References	36

1. Introduction

National Obesity Observatory resources on evaluation

The National Obesity Observatory (NOO) is committed to improving the quality and quantity of evaluations of interventions related to obesity, overweight, underweight and their determinants. In order to do this, we have produced a number of resources that provide information and support to practitioners with an interest in evaluation:

- the '*Standard Evaluation Framework (SEF) for weight management interventions*'¹
- the '*Standard Evaluation Framework (SEF) for dietary interventions*'²
- the '*Standard Evaluation Framework (SEF) for physical activity interventions*' (this document)³
- a Collection of Resources on Evaluation (CoRE) which provides additional guidance and case studies⁴
- an evaluation data collection tool⁵
- a searchable database of interventions⁶
- guidance on '*Measuring diet and physical activity in weight management interventions*'^{7,8}

Figure 1 (on page 5) is a flow chart to guide readers to the most relevant resources to assist with evaluations.

Introduction to this document

This document builds on the *Standard Evaluation Framework (SEF) for weight management interventions*, published by NOO in April 2009.¹ It takes the principles described in the original SEF and applies them to physical activity interventions.

This SEF contains a list of 'essential' and 'desirable' criteria for data required for a comprehensive and robust evaluation. Essential criteria are the minimum data and information recommended to perform a basic evaluation of an intervention that sets out to increase physical activity. Desirable criteria are additional data that would improve the quality of an evaluation, and enhance understanding about what has been achieved and the processes that have taken place during the intervention.

A glossary of terms is available on page 34.

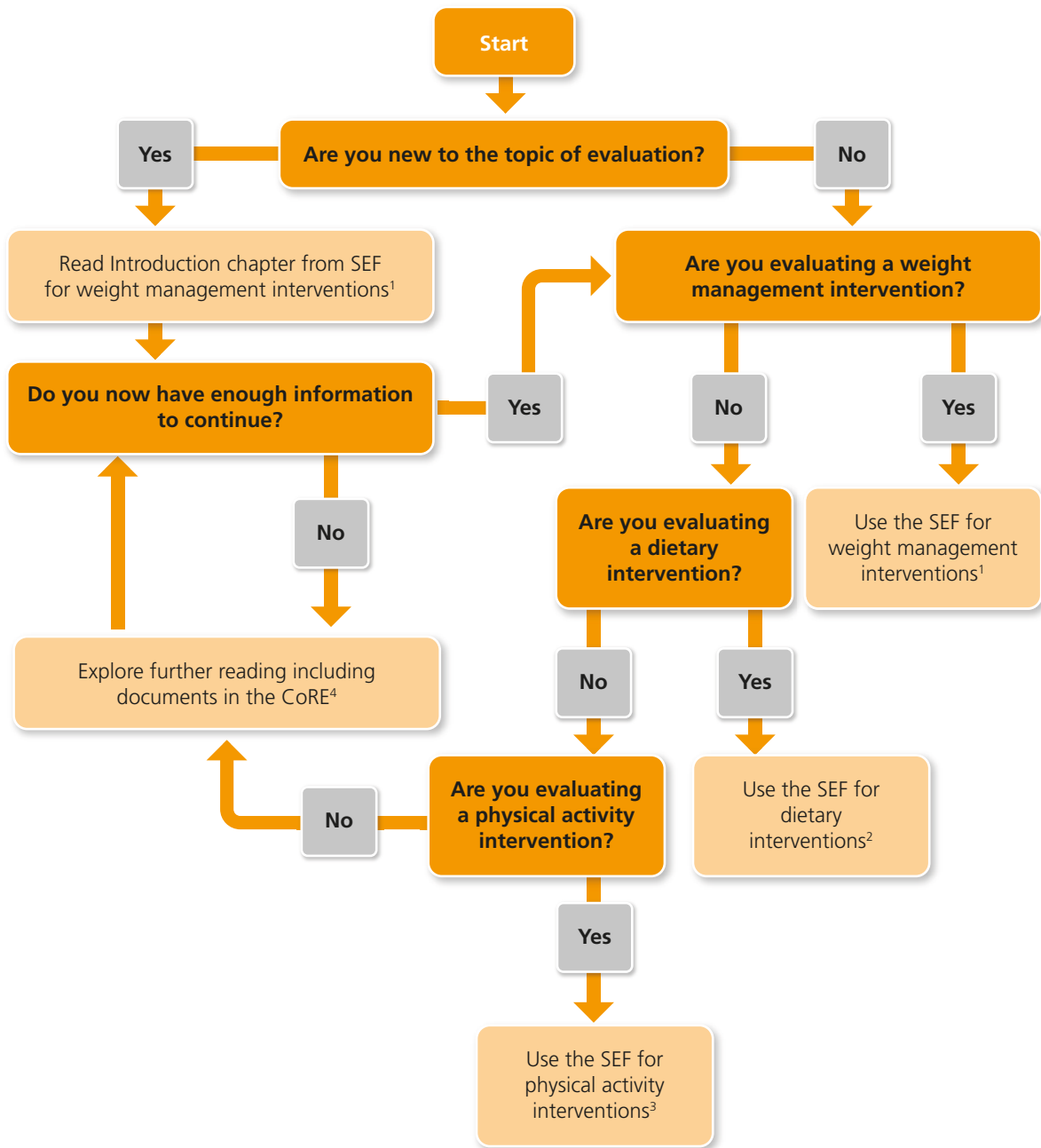
What does this document aim to do?

The SEF for physical activity aims to describe and explain the information that should be collected in any evaluation of an intervention that aims to increase participation in physical activity. It is aimed at interventions that work at individual or group level, not at population level. It provides detailed, specific guidance on the following areas:

1. How to identify appropriate physical activity outcomes for evaluating different types of intervention.
2. How to define suitable measures for different types of physical activity outcome.
3. How to approach the challenges of assessing and measuring physical activity and energy expenditure.

An accompanying SEF is available with a specific focus on interventions designed to influence diet and nutrition.²

Figure 1: Flow chart to show which NOO evaluation resource to use



What does this document not aim to do?

- Provide guidance that is intended to support the evaluation of medical interventions, such as surgery or medications.
- Provide guidance on the evaluation of broader community-level programmes, or programmes that promote physical activity through changing the physical environment. Such programmes present additional evaluation challenges that are beyond the scope of this document.
- Provide an introduction to the concepts of evaluation; such an introduction is contained in Chapter 1 of the *SEF for weight management interventions*.¹

Target audiences

The target audiences for this document are:

- commissioners or managers of weight management or obesity prevention interventions with a physical activity element
- commissioners or managers of physical activity interventions
- physical activity or sport and leisure leads in local authorities
- commissioners or managers of active travel projects
- practitioners running physical activity or active travel projects.

What is physical activity?

Physical activity is defined as 'any bodily movement produced by skeletal muscles that results in energy expenditure'.⁹ The term encompasses: exercise; sport; recreation; walking, cycling and other active transport; occupational activity and domestic activity such as gardening. This evaluation framework should therefore be useful to a wide range of professionals involved in planning and evaluating physical activity projects and interventions.

Why do we need a SEF for physical activity interventions?

There are many public health interventions that have an element of physical activity behavioural change. However, few of these interventions are evaluated.¹⁰ Additionally, there is a need for more rigorous evaluation of local interventions, particularly those relating to overweight and obesity.¹¹

Evaluations of public health interventions are often poorly designed, use inappropriate measures, do not report on health outcomes, and tend to focus disproportionately on process measures such as attendance and participant satisfaction.¹¹ Variation in evaluation practice also makes it difficult to compare the impacts of interventions. The wide variety of physical activity interventions also presents challenges for evaluation, as interventions such as sports promotions, exercise programmes or active travel initiatives are likely to require the assessment of different health outcomes and the use of different data collection methods.

This document aims to help address these issues by providing a framework to guide the design and implementation of evaluations of interventions with physical activity outcomes.

2. Principles of evaluation

Evaluation is about judging the value of an activity and assessing whether or not it has achieved what it set out to do. In public health and health promotion, an evaluation determines the extent to which a programme has achieved its objectives, and will assess how different processes have contributed to this.¹² Evaluation is a critical aspect of good project management, and should be a component of projects of any size.

The *SEF for weight management interventions* contains an introduction to evaluation principles and methods.¹

When designing an evaluation, the single most important principle is to ensure that the evaluation focuses on the objectives of the programme.

The two basic evaluation questions:

1. What are the objectives of your project?
2. Will the evaluation measure whether these objectives have been achieved?

The objectives determine the outcomes of the intervention and what should be measured – known as ‘outcome measures’ or ‘indicators’.

The main focus of an evaluation should be to assess whether the primary outcomes that the project is trying to achieve have been met. Secondary outcomes are things that are of interest, and may add to the knowledge generated by the evaluation. It may be of particular interest to collect secondary outcome data relating to the mechanisms by which the intervention is expected to work. Primary and secondary outcomes are illustrated in the following two scenarios.

Scenario 1

If an active travel project has a core objective of increasing the proportion of people walking to work, then it should measure the change in the proportion of people who have walked to work, irrespective of details such as speed, distance or changes to other forms of activity. This would be the primary outcome measure. However, participation in total physical activity would be a valuable secondary outcome measure.

If the project focused on achieving its outcomes by reducing people’s perceived barriers to walking to work, then these might also be measured as secondary outcomes. For instance, the proportion of people who agree with the statement: ‘I don’t have the time to walk to work,’ should be measured before and after the intervention.

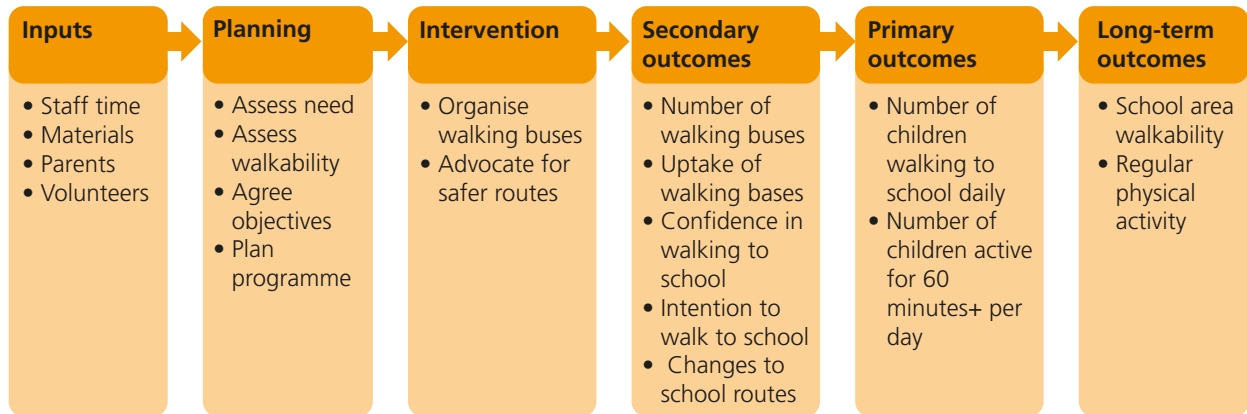
Scenario 2

A sports outreach project has a core objective to increase the number of children aged 9–13 from a specific community who play mini-rugby twice a week. The project does not have any objectives relating to total physical activity. In this case, the evaluation should measure any change in the proportion of the target audience who participate in mini-rugby at least twice a week as the primary outcome measure. Other measures – such as total physical activity or satisfaction with the programme – would be interesting, but not essential to the evaluation. These would be secondary outcome measures or indicators.

A logic model can help to identify primary and secondary outcome indicators. Logic models describe the relationship between each element in a project or intervention, and the likely direction of change. They can be useful in describing and explaining what is expected to happen in a project, providing a mechanism to check that the appropriate indicators have been selected and the project is likely to achieve its objectives.

Figure 2 provides an example of a logic model for a ‘Walk to School’ project. It shows how the outcomes are related closely to the precise nature of the intervention. For example, whilst the primary outcome is the number of children who walk to school daily, a further objective of the project is to organise walking buses; the secondary outcome would therefore be uptake of the walking bus.

Figure 2: Example logic model for a ‘Walk to School’ programme, showing stages of evaluation (adapted from US Physical Activity Evaluation Handbook¹³)



3. Physical activity: selecting and measuring outcomes

Selecting physical activity outcomes

An outcome indicator should be selected on the basis of how well it measures the objectives of the intervention. However, there are other factors that are important in choosing an indicator:

- selected outcome(s)
- level of data required (individual, group or population)
- characteristics of the individuals, group or population
- timeframe of interest
- available time and resources (including expertise).

If change in physical activity is the primary outcome there are a number of types of physical activity outcomes that may be the focus of the intervention. The acronym FITT is used to describe the four main ways to classify physical activity:

- frequency
- intensity
- time (duration)
- type.

Any type of activity may change in terms of frequency (for example, to walk to work more often), intensity (walk to work at the same frequency but walk faster to increase intensity), and duration (walk to work at the same frequency and walking pace but take a slightly different route to increase the time spent walking each trip). It is also possible to classify physical activity as a combination of frequency, intensity and duration, and this is often referred to as the 'volume of physical activity'. Being clear about the element of physical activity that is the focus of the intervention will help in the choice of instrument used to measure change. Examples of intervention objectives that focus on the different elements of physical activity are as follows:

- changes in total physical activity, aiming to increase the volume of physical activity of any type
- changes in the frequency of specific physical activity behaviours, such as increasing the number of days that children engage in sport at school
- changes in the frequency and intensity of physical activity, such as increasing the number of older adults who report walking at a brisk or fast pace on at least five days of the week
- changes in the duration (time) of specific physical activities, such as increasing the proportion of primary school children who engage in active play for at least 30 minutes per day
- changes in determinants of activity behaviours, such as improving attitudes to physical activity, increasing knowledge about physical activity and exercise, improving self-efficacy towards physical activity.

Intervention objectives and outcomes should be agreed early in the development of the project, along with the proposed evaluation design. When agreeing these outcomes, it is important to think about the purpose of the evaluation, how the results will be used and what decisions might be made as a result.

In addition, in order to be evaluated, outcomes need to be measurable. Therefore, the practicalities of collecting the required outcome data should be considered at the project design stage, including the

choice of instrument used to measure change. Measurement of physical activity can be challenging; more detail is provided in the accompanying NOO reviews of measurement of diet and physical activity.^{7,8} The next section presents the main options for measuring physical activity in evaluations.

Options for measuring physical activity

Option 1: a measure of specific type of physical activity

As mentioned previously, the key measure for an evaluation should focus on the main outcome that the project is trying to achieve. Secondary outcomes are things that are of interest, and may add to the knowledge generated by the evaluation.

If the intervention focuses on increasing a specific type of activity, such as walking or cycling, this should be the main outcome measure. It is not always necessary to measure total physical activity, although this may improve understanding of the impact of the intervention – for example, if it is thought that people might increase one type of physical activity and compensate by reducing another.

Examples of outcome measures of specific types of activity are shown in Table 1.

Table 1: Examples of outcome measures of specific types of activity

Objective	Intervention	Examples of outcome measures
Increase walking to school	'Walk to School' promotion programme	Proportion of total trips to school taken on foot
		Proportion of children who say walking is usual mode of travel to school
Increase swimming	Free swimming vouchers for children	Number of free swim sessions taken by children
		Proportion of children who swim once a week or more
Increase leisure walking	Walk promotion using pedometers	Minutes spent on leisure walks per week
Increase total walking	Counselling	Number of steps walked per week (measured with pedometer)

In each of these examples measurement of total physical activity would add detail to the evaluation but is not essential. However, if an evaluation relies on the measure of a single mode of activity, it is important to note that this would not reflect total physical activity. For example, measuring an increase in trips along a bike path after environmental improvements might seem like an indication of project success. However, it may be that existing cyclists are using that path instead of another (so their total cycling has not increased); or it may be that the improved path encourages new cyclists, who give up going to the gym (again, total activity would not have increased). Restricting measurement to one mode of physical activity may therefore preclude the drawing of conclusions about the health benefits of the observed changes in physical activity.

Option 2: measure of total physical activity

As previously discussed above, the primary outcome measure used in any evaluation will depend on the objectives of the programme. However, a good option in any evaluation of a physical activity programme is a measure of total physical activity. This is because some aspects of activity might be counterbalanced by others (for example, a programme promoting walking might mean that people stop cycling). Total physical activity is best measured by continuous variables such as total kcal energy expended, or total number of minutes spent at specific levels of energy expenditure such as hours per week of moderate intensity activity (usually expressed in MET-minutes).

A measure of total physical activity is a better option if the study is able to follow up individuals throughout the course of the intervention. This will enable the change in total volume of physical activity or energy expenditure to be calculated.

Option 3: proportion achieving recommended physical activity levels

There is a dose-response relationship between the amount (frequency, intensity, duration) of physical activity and most health outcomes.¹⁴ This means that more physical activity will generally lead to greater benefits. However, many programmes are framed in the context of guidelines such as the Chief Medical Officers’ recommendations for physical activity,¹⁵ and may focus on trying to help people reach a threshold of activity, such as the proportion of the participants in the programme achieving 150 minutes of moderate intensity physical activity per week (or the equivalent of 75 minutes of vigorous activity). In a walking study, with monitoring by pedometer, a common threshold measure used is 10,000 steps a day.¹⁶

If the objective of an intervention is to increase the proportion of people achieving recommended levels of activity, this would require a different type of outcome measure from those measuring total physical activity (described in Option 2 above). Examples of interventions and appropriate outcome measures are presented in Table 2:

Table 2: Examples of outcome measures of total activity and proportion of participants achieving recommended levels

Objective	Intervention	Examples of measure	Examples of outcome
Increase physical activity	12-week primary care counselling programme	Mean number of occasions being active per week	Average increase of 1.5 episodes of activity per week ¹⁷
	Walking promotion programmes	Time spent walking	Mean increase in walking of between 30–60 minutes per week ¹⁸
Increase proportion active at recommended levels	10-week primary care programme	Proportion active at recommended levels	11% increase in the proportion of participants active for at least 150 minutes of moderate/vigorous activity per week after six months ¹⁹

It is important to note that if different samples are used for baseline and follow-up, it is harder to assume that any changes were due to the intervention. For example, in the last example in Table 2 above, while we know there was an 11% increase in the proportion active overall, we do not know that the increase is as a result of the intervention – there may be other factors that are confounding the outcome. Also, the overall mean increase may mask other changes: some people may be more active, others less. If we are able to follow-up individuals we would have a clearer understanding of who had changed and who had not.

Measuring outcomes

Measurement of physical activity is complex and presents significant challenges, particularly in a public health setting where there are likely to be restrictions on the time and tools available for measurement. In an academic research setting, an investigator may be able to:

- gather data on participation in different activities from detailed and lengthy questionnaires
- use objective measures such as accelerometers (which measure body movement in two or three planes) or global positioning systems (which measure where the user is at any given time and hence can estimate movement)

> CONTENTS

- employ the 'gold-standard' method of taking estimates of total energy expenditure from doubly labelled water.^{a, 20}

Such objective methods are far more likely to provide reliable assessments of energy expenditure than self-completion questionnaires. However these methods are unlikely to be feasible for a public health intervention in free-living adults, without additional assistance from academics or other specialists experienced in using objective measures and analysing and interpreting the data.

In the context of public health, self-report methods are commonly used to collect data on physical activity behaviour, from which estimates of energy expenditure can be derived. All self-report methods of physical activity assessment have advantages and disadvantages, and it is important to consider these before deciding on a suitable evaluation measure.

The appropriate method of measurement for assessment of physical activity in an evaluation will depend on the objectives of the intervention as well as the need for individual-level or population/cohort/group-level data, the characteristics of the population/cohort/group, the timeframe of interest and the resources (including expertise) available.²¹

Objective measurement of walking and cycling

If an intervention focuses only on walking, then pedometers provide an objective measure that can be used easily in a public health setting, provided the model of pedometer used has been validated (assessed as providing valid and reliable measures compared to a 'gold-standard' measure). Cheap pedometers may be useful for motivation or as promotional devices but they are unlikely to provide accurate measures of physical activity.²² Some models have seven-day memories. If the unit has no memory, participants will need to record their steps each day, which may reduce the quality of the data.

Cycle computers could be used in a similar way to measure distance and time cycled, although there is little published evidence on the validity of cycle computers.

Potential sources of error and bias

It is important to understand that any method of assessing physical activity only provides an estimate of a complex behaviour. Therefore, wherever possible, it is important to identify and quantify sources of error in order to minimise them through careful study design or analysis. This is particularly important when estimating total energy expenditure from participation in individual activities. All tools should ideally have been previously tested for reliability and validity. Reliability refers to how consistently an instrument or tool will measure something on two or more separate occasions.²³ Validity is the extent to which an instrument genuinely records what it is intended to measure.²⁴

Self report of physical activity is subject to a number of types of bias, these include:

- recall bias (where a respondent's memory influences their response)
- bias caused by lack of compliance (where certain people might refuse to answer some questions)
- 'social desirability' bias (in which people may be more likely to provide responses that they think are acceptable, rather than those that represent their real opinions or behaviour).

The choice of measure will, to an extent, dictate the data collection method employed. The usual method of estimating the proportion of people meeting guidelines using self-assessment questionnaires is to assign energy expenditure values to every activity, and use this to calculate total time at moderate intensity

.....
a Doubly labelled water provides an accurate, objective measure of energy expenditure.

> CONTENTS

or above. But this can be complex and time-consuming, and may be unrealistic for many evaluations. However, there are questionnaires available that can produce categorical measures of activity relatively easily. One simple physical activity questionnaire that has been well validated is the short EPIC physical activity questionnaire. It can be analysed to place respondents into four categories of activity.²⁵ It should, however, be noted that it has been validated for ranking participants according to their level of activity, not as an absolute estimate of their level of activity.

A shorter method is to use global self-completion questionnaires, but these are generally less precise than a physical activity item-specific (type-specific) measure. An example is the single-item physical activity questionnaire developed to assess respondents' eligibility for physical activity interventions.²⁶ This is very short and easy to use, and asks respondents on how many days in the last week they have been active for 30 minutes or more. The questionnaire is a valid screening tool to determine whether respondents are sufficiently active to benefit their health, when compared to data from accelerometers.²⁷ Such questionnaires might offer practical solutions to evaluators looking for easy-to-use tools, but it is important to bear in mind their limitations, notably their lack of criterion validity,^b and lack of sensitivity to measure the small changes commonly observed in public health interventions.

Use of existing datasets

It is also worth considering whether existing datasets can be used to evaluate changes in key indicators at local level. A good example is the Active People Survey, which collects data annually on participation in sport and physical activity that is undertaken for at least 30 minutes per occasion. The data are available online²⁸ and the sample size is large enough to allow analysis at local authority level. However there are disadvantages of using data such as these: (the questions cannot be changed); sport is the focus; the sample size (although large) is not large enough to allow sub group analysis (such as age or sex) or exploration of geographical areas smaller than a local authority; there is a time lag between fieldwork and publication that may be too long for local level evaluations.

Further resources on physical activity assessment

NOO has conducted a rapid review of the most appropriate tools for measuring physical activity (in terms of energy expenditure) in local-level weight management interventions. *'Measuring diet and physical activity in weight management interventions'*⁷ identifies a shortlist of practical and validated questionnaires for the assessment of physical activity and diet. It recommends three questionnaires for measuring children's physical activity, and four for adults. The shortlist is based on best available evidence and highlights the strengths and limitations of each questionnaire. Copies of the questionnaires are also available online.⁸ The questionnaires generally ask respondents to recall the amount of time spent in particular activities, or in categories of activity (such as moderate intensity activity). These questionnaires are therefore recommended for the measurement of activity in interventions in which total energy expenditure is the primary outcome.

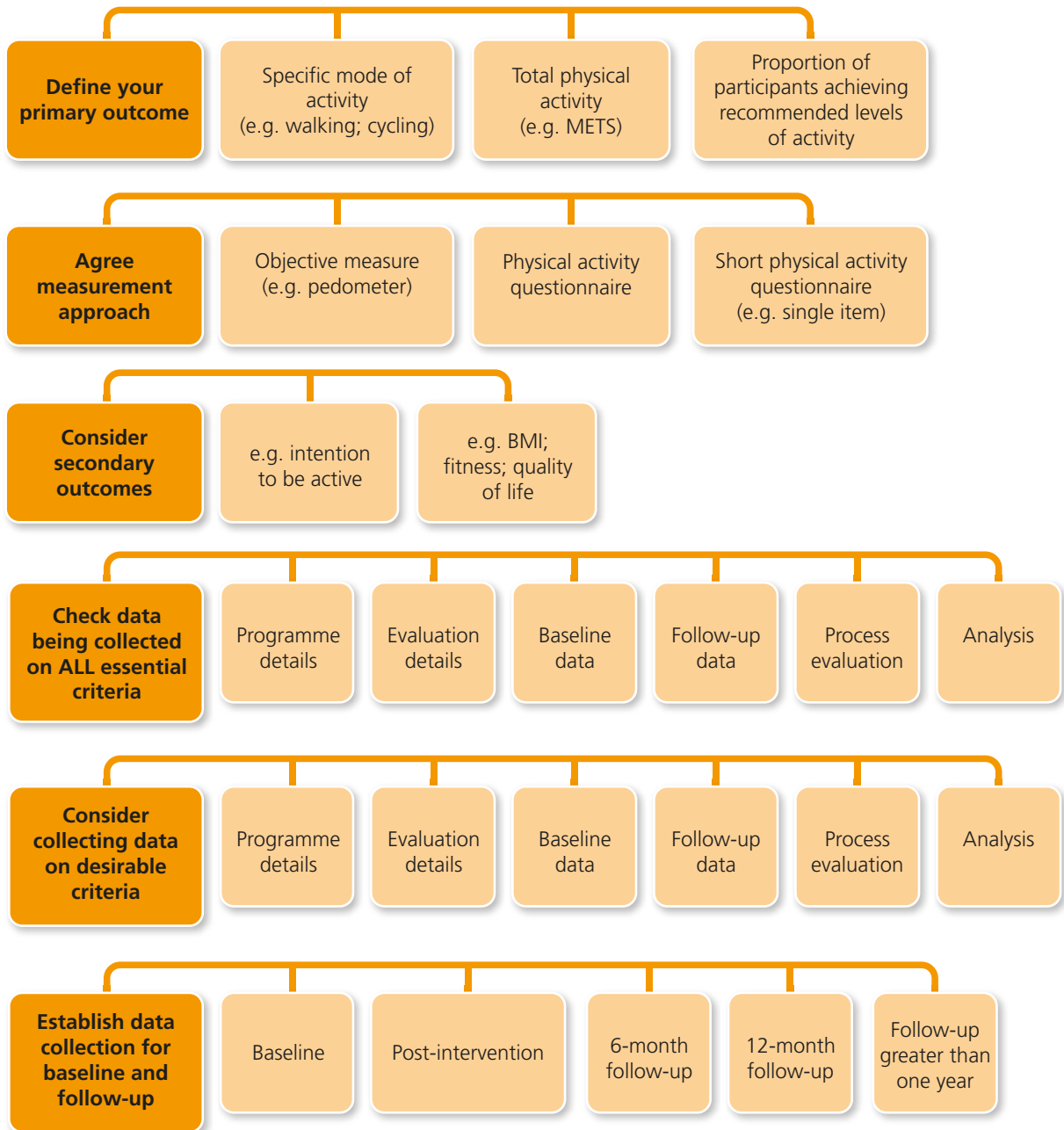
In addition, the Medical Research Council (MRC) has produced an online *'Diet and Physical Activity Measurement Toolkit'* that provides advice to researchers on selecting appropriate physical activity measurement methods. The tool provides details about the pros and cons of each method and links to publications and instruments.²¹

All tools have some limitations, and these should be acknowledged in any report so that the reader can make his or her own judgement about the validity of the results.

Figure 3 shows the process that should be followed when planning an evaluation – from defining outcomes to the 12-month follow-up.

.....
b Criterion validity describes how well the tool's measures compare to those considered 'gold standard', such as measures from doubly labelled water or accelerometers.

Figure 3: Flow chart (and examples) for data collection and analysis



4. The SEF for physical activity interventions

This section of the document presents evaluation criteria necessary to undertake a comprehensive and robust evaluation. Essential criteria are the minimum recommended data and information required to evaluate a physical activity intervention. Desirable criteria are additional data that would enhance the evaluation. Section 5 contains supporting information on each criteria and also provides guidance on data collection.

Table 3: Evaluation criteria

Click on each criteria to link to the relevant supporting information and guidance.

	ESSENTIAL	DESIRABLE
Part one: programme details		
➤ 1. Title or name of intervention		
➤ 2. Aims and objectives (including primary and secondary outcomes)		
➤ 3. Rationale for the intervention (including theoretical basis and logic model)		
➤ 4. Contact details		
➤ 5. Commissioner(s) of the intervention and sources of funding		
➤ 6. Intervention timescale (exposure, quantity and duration)		
➤ 7. Intervention delivery dates		
➤ 8. Duration of funding (including dates)		
➤ 9. Location and setting		
➤ 10. Description of intervention: <ul style="list-style-type: none"> • target population • content • delivery method • deliverer • quality assurance mechanisms • assessment of potential unintended consequences 		
➤ 11. Method of recruitment and referral		
➤ 12. Participant admission/exclusion criteria		
➤ 13. Participant consent mechanism (and ethical approval)		
➤ 14. Equipment and resources required		
➤ 15. Core staff competencies (and training required)		
➤ 16. Incentives for attendance		
➤ 17. Detailed breakdown of costs		
➤ 18. Cost of the intervention per participant		
➤ 19. Cost to the participant		
➤ 20. Relevant policy context		
➤ 21. Details of health needs assessments		
➤ 22. Details of equality impact assessments		
➤ 23. Declaration of interest		
Part two: evaluation details		
➤ 24. Type of evaluation and evaluation design		
➤ 25. Methods and timings of data collection		

	ESSENTIAL	DESIRABLE
Part three: demographics of individual participants		
➤ 26. Age		
➤ 27. Sex		
➤ 28. Ethnicity		
➤ 29. Disability		
➤ 30. Measure of socio-economic status		
➤ 31. Additional information (such as marital status, medical history and parental status)		
➤ 32. Details of parental physical activity and weight (for interventions aimed at children)		
Part four: baseline data		
➤ 33. Measure(s) of physical activity behaviour		
➤ 34. Measure(s) of correlates of physical activity (if relevant to intervention objectives)		
➤ 35. Other outcome measures (such as BMI; physiological measures; quality of life)		
Part five: follow-up data (impact evaluation)		
➤ 36. Follow-up data on key physical activity measures (minimum of three follow-up points, including at one year)		
➤ 37. Follow-up data on key physical activity measures over a greater term than one year		
➤ 38. Follow-up data on correlates of physical activity behaviour (if collected at baseline)		
➤ 39. Follow-up data on other secondary outcome measures		
Part six: process evaluation		
➤ 40. Number of participants invited		
➤ 41. Number of participants recruited		
➤ 42. Number of participants attending each session or contact point		
➤ 43. Number of participants at each follow-up point, including final stage		
➤ 44. Reasons for opt-out (where relevant)		
➤ 45. Description of what was actually delivered and details of any unexpected outcomes		
➤ 46. Participants' satisfaction with the intervention		
➤ 47. Plans for sustainability		
Part seven: analysis and interpretation		
➤ 48. Summary of results compared to baseline (for primary and secondary outcomes)		
➤ 49. Details of any further analyses and statistical methods used		
➤ 50. Limitations and generalisability		
➤ 51. Recommendations and changes to future projects		
➤ 52. Dissemination of learning and findings		

5. Explanatory notes

This section provides explanatory notes for each of the criteria listed in Section 4. These notes are designed to help with the collection and recording of information required to complete the SEF. Particular Issues for consideration are highlighted and discussed.

Part one: programme details

1. Title or name of intervention

ESSENTIAL

Record of the name or title of the intervention, for example: 'Be Active Be Healthy.'

2. Aims and objectives (including primary and secondary outcomes)

ESSENTIAL

What does the intervention aim to do? What is the intended impact on the long-term physical activity behaviours and health of the participants? What are the actions that will bring about this outcome?

3. Rationale for the intervention (including theoretical basis and logic model)

DESIRABLE

In developing and evaluating interventions, it is essential to have a clear understanding of the rationale for the intervention. This should provide an explanation as to how and why the intervention is supposed to achieve its objectives and includes the processes that will take place before any final changes in behaviour.

Being clear about the rationale for the intervention will also help to clarify or reinforce what the outcomes and appropriate measures should be. The use of 'logic models' (Figure 2) can be helpful in clarifying a rationale for an intervention with clear inputs, processes, outputs and outcomes.

For example, a project may aim to increase regular walking among older adults. The rationale for this project may be that a high proportion of older people said they would like to walk but didn't have anyone to walk with. The project would therefore aim to put on guided walks and put people in touch with each other so they could continue walking when the project had finished.

In some cases, interventions may be based on: existing scientific theories, such as behaviour change theory; guidance such as that from the National Institute of Health and Clinical Excellence (NICE); or scientific evidence from peer-reviewed journals.

4. Contact details

ESSENTIAL

List the key people involved in the intervention's planning, delivery and evaluation. This should include full contact information and details of staff positions (as individuals may change during or following the intervention).

5. Commissioner(s) of the intervention and sources of funding

ESSENTIAL

How is the intervention funded and who has commissioned it? For example: 'Funding is provided by the Department of Health and the intervention was commissioned by the local authority.'

6. Intervention timescale (exposure, quantity and duration)

ESSENTIAL

For how long does the intervention run? How many sessions, episodes or events are/were delivered? For example: 'The intervention is delivered in 20 two-hour sessions, twice a week for ten weeks.'

7. Intervention delivery dates**ESSENTIAL**

This includes dates for the initial recruitment of organisations (for example, GPs) and participants, first point of contact and any follow-ups.

8. Duration of funding (including dates)**DESIRABLE**

What are the start and completion dates for the overall programme? The intervention may be run a number of times throughout the duration of the overarching programme.

9. Location and setting**ESSENTIAL**

Where is the intervention taking place? It could be in a GP surgery, school or community centre, or online. If it takes place in several settings, they should all be included here. It may be useful to add a description of any transport that is provided for participants to attend.

10. Description of the intervention**ESSENTIAL**

This section should include all the information about planned elements of the intervention, what the intervention intends to achieve, how it will do it, and the timing details of each element. Details should be provided, where applicable, of the following:

- Target population – what is the intervention’s target population? From which population are the participants recruited? For example: ‘Children aged 7-13 from the Anytown area who are overweight or obese’, or ‘Bangladeshi women living in a specific community.’
- Content – what is the intervention aiming to deliver, and how? For example: ‘The intervention will provide physical activity advice and ideas for ways to be active. Each session will be one hour in duration, once a week for 12 weeks. At weeks 4, 8 and 12, individuals will be offered a session of motivational interviewing.’
- Delivery method – for example, one-to-one sessions, face-to-face group sessions, by telephone, online or via the media.
- Deliverer – for example, an exercise professional, school nurse, health trainer or local media.
- Quality assurance mechanisms – what mechanisms are in place to ensure that the intervention is being delivered in the way in which it was planned? For example, spot checks, audits carried out by an external assessor or participants’ satisfaction questionnaires. It is useful to provide details of any relevant health checks, risk assessments and Criminal Records Bureau checks if applicable. Also include details of how any adverse events will be recorded and addressed.
- Assessment of any potential unintended consequences of the intervention – for example, if the programme is successful in increasing cycling to school, there may be a demand for more bike parking.

11. Method of recruitment and referral**ESSENTIAL**

How have participants been recruited to the intervention? What percentage of those that are eligible have been recruited? Is there a referral process or is it self-selecting? For example, are participants referred by a GP or are leaflets and posters used to advertise in GP surgeries?

Give brief details here of any sampling processes that are undertaken. Is there any targeting of particular groups by, for example, advertising of the intervention in certain communities or at specific locations? The method by which participants are recruited should be taken into account when conducting the evaluation. For example, a self-selecting group of participants may be more motivated than referred participants. Provide details here of the percentage of those that are eligible who have actually been recruited.

12. Participant admission/exclusion criteria**ESSENTIAL**

Some projects may have specific criteria which participants are required to meet in order for them to be eligible to participate. Participants who have been referred or have self-referred should meet pre-defined criteria. For example, the target population may be inactive children aged 7–13 from a particular location. This should then inform the specific design of the intervention: for example, children with particular or complex needs, such as those with learning or educational difficulties, significant or serious obesity-related morbidity, or a suspected underlying medical cause of obesity may need more intensive or specific types of therapy.

13. Participant consent mechanism (and ethical approval)**DESIRABLE**

It is important that participants give their consent to data being collected from them as part of the evaluation. The appropriate mechanism for gaining participant consent will vary for different groups of participants and according to the nature of the intervention. For example:

- those able to give their own consent
- those with parental responsibility giving consent on behalf of a child or young person under the age of 16
- those who lack the capacity to give consent.

Policy guidance on seeking consent from different groups, including people unable to consent for themselves, is available from the National Research Ethics Service website.²⁹

The Data Protection Act 1998 must be adhered to when collecting personal data from individuals, and a data protection statement should be given to participants before any personally identifiable data are collected. It should explain exactly which personal data are being held, why, where they will be held and who will have access to the data. This is particularly important when collecting sensitive data such as ethnicity and socio-economic status. More information about the requirements of the Data Protection Act 1998 can be found on the National Archives website.³⁰

Some evaluation projects may be classed as research and so will require ethical approval from the National Research Ethics Service. Its leaflet, *'Defining Research'*, provides guidance on distinguishing between research, audit, service evaluation and public health surveillance.³¹

14. Equipment and resources required**ESSENTIAL**

What equipment or resources are required to run the intervention? For example, is it necessary that a gym or exercise facility is available and, if so, what equipment is needed? Can the facility accommodate population groups with specific requirements (such as people with physical limitations or specific needs)? Does the intervention require any equipment for measuring outcomes, such as pedometers or scales to weigh participants?

15. Core staff competencies (and training required)**ESSENTIAL**

How are those delivering the intervention recruited? What are the core skills needed by everyone involved in delivering the intervention? For example, does the intervention require the involvement of a registered exercise professional? Other personal skills such as communication or facilitation may also be important.

Do those staff delivering the intervention need to be trained in certain aspects of the intervention such as a fitness qualification, counselling, childcare and youth work? Should they be trained to a specific level?

16. Incentives for attendance**ESSENTIAL**

Are any incentives provided to either recruit or retain participants and, if so, what are they? If incentives are used, it is important to record any evidence of their potential impact. This is important as incentives may influence the effectiveness of the intervention and the sustainability of any behaviour change.

17. Detailed breakdown of costs**DESIRABLE**

A detailed breakdown of the cost of the intervention is important for an economic analysis of the entire intervention and to judge if it is good value for money. It also enables commissioners to judge whether sufficient resources are available to run the intervention.

Take into account costs incurred during the planning and set-up stages, as well as during the delivery and evaluation. Some examples of input costs are: staff time; transport; venue hire; equipment; publicity; incentives. It is especially important to factor in 'invisible' costs. For example, a room in a local authority leisure centre may be available free of charge as part of a partnership agreement. However, this resource needs to be taken into account so that, if the intervention is repeated, financial resources can be accurately planned.

NICE has a number of costing templates available that may assist with the accurate estimate of costs.³²

18. Cost of the intervention per participant**ESSENTIAL**

This describes the total cost of the project (point 17 above) divided by the total number of people who have received the intervention. It should be based on real data where possible, with any estimates or assumptions clearly documented. Costs should be calculated on the basis of the cost per person receiving the full 'dose' of the programme at follow-up – that is, recruitment, participation and completion of the intervention. However, it should also take account of the costs associated with non-completers. So, for example, if a project spent a total of £10,000 on a project that recruited 100 people to a walking course, but only 50 completed the course, then the cost per participant would be $£10,000/50 = £200$.

19. Cost to the participant**ESSENTIAL**

Costs to the participant should be recorded. This would include charges made to the participant for any part of the intervention (such as in some exercise referral schemes), the cost of any necessary equipment purchased by the participant (such as clothing and shoes), the cost of travel or loss of potential earnings while taking part in the intervention.

20. Relevant policy context**DESIRABLE**

It may be useful to show whether an intervention is part of a specific local strategy, or whether it addresses needs identified in the Joint Strategic Needs Assessment (JSNA).³³ Does it address national priorities or indicators such as those in the Public Health Outcomes Framework (PHOF)?³⁴

Relevant indicators in the PHOF include:

- 2.13i: proportion of adults achieving at least 150 minutes of physical activity per week in accordance with UK CMO recommended guidelines on physical activity
- 2.13ii: proportion of adults classified as 'inactive'
- 1.16: percentage of people using green space for exercise or health reasons.

21. Details of health needs assessments**DESIRABLE**

Has a health needs assessment been undertaken that identifies a need for the intervention or service? Information to support the rationale for the intervention may also be available from other sources. For example, data relating to health inequalities and gaps in service provision may already be available from local policies such as JSNAs.

22. Details of equality impact assessments**DESIRABLE**

Public bodies have a duty to undertake equality impact assessments (EIAs) under race, sex and disability equality legislation.³⁵ It is useful to provide an intervention's EIA as part of its overall evaluation. It can provide valuable information if particular outcomes are witnessed in different groups. An EIA provides a systematic way of ensuring legal obligations are met. It is also a 'practical way of examining new and existing policies and practices to determine what effect they may have on equality for those affected by the outcomes'.³⁶

23. Declaration of interest**ESSENTIAL**

This covers any potential conflicts of interest and is particularly important if the evaluation is funded by an agency that could be perceived to have influence over the results for commercial reasons.

NICE has produced a clear statement covering different categories of potential conflicts of interest that should be declared, including pecuniary interests (where a financial payment or other benefit has been received) and non-pecuniary interests (where someone may have publicly expressed a clear opinion on the intervention in question, and this may influence their impartiality).³⁷

In general it is best to declare any potential conflicts, even if they do not appear to be important. Perceived conflicts of interest do not necessarily mean the intervention should not go ahead as planned; it may be acceptable to state how potential conflicts are going to be avoided.

Part two: evaluation details**24. Type of evaluation and evaluation design****ESSENTIAL**

The way in which an evaluation is designed to collect data, and the methods by which data may be compared with any control population are very important considerations. For example, does the evaluation use mainly qualitative or quantitative data? Guidance is provided in Section 3 of this document, and Section 2.7 in the 'SEF for weight management interventions' provides a more detailed explanation of evaluation designs.¹

It is also important to consider who is conducting the evaluation. Is it being evaluated by the same people running the intervention, or is it being externally, independently evaluated? An independent evaluation is always preferable as it is more likely to produce an objective view of the success or otherwise of the intervention.

25. Methods and timings of data collection**ESSENTIAL**

Does the evaluation use qualitative or quantitative data? Which data collection tools are used – questionnaires, diaries, interviews, focus groups? At what stages of the intervention are the follow-up points, for example, week six, month three, year one? How are the data collected at each of these contact points, for example, in person at an event or session, by telephone, by home visit or online?

Part three: demographics of individual participants

It is important to collect individual-level data on each participant in the intervention; this is typically collected at baseline. Data may be collected in the following categories:

26. Age

ESSENTIAL

It is essential to record the age of all participants in the intervention. In most cases this means recording the actual age of the participant. If age data are collected in categories (such as 0-4, 5-9), then these categories must also be used in the analysis.

27. Sex

ESSENTIAL

Record the sex of all participants. This is useful to identify whether the intervention tends to be more effective for males or females, or if uptake of the intervention varies by sex.

28. Ethnicity

ESSENTIAL

It is standard practice in public health interventions to record the ethnic origin of participants. If the intervention is targeted at a specific ethnic group, then a record of ethnic origin is essential for screening participants for eligibility. If the intervention is not targeted in this way, the record still provides important information to understand the extent to which response to, and effectiveness of, the intervention may vary between different ethnic groups. Particular ethnic groups also have higher than average prevalence of obesity, and lower levels of participation in physical activity.³⁸

In addition, there is a legal requirement to carry out ethnic monitoring. The Race Relations (Amendment) Act 2000 requires public bodies, including local authorities, primary care trusts and their partners to take account of race equality in policy making and service delivery. Ethnic monitoring is a way of identifying potential discrimination and whether policies promoting equality of opportunity and good relations between different racial groups are being implemented. For further information, see the Equalities and Human Rights Commission's website.³⁹ Ethnic category codes for England are defined by the Office for National Statistics.⁴⁰ These categories are listed in Appendix 1.

29. Disability

ESSENTIAL

It is standard practice in healthcare interventions to record the disability status of participants. The Equality Act 2010 defines a disabled person as: 'Someone who has a physical or mental impairment that has a substantial and long-term adverse effect on their ability to carry out normal day-to-day activities.' Detailed guidance on the interpretation of this definition can be found on the DirectGov website.⁴¹

The Public Sector Equality Duty came into force under the Equality Act 2010 and requires all public bodies, including NHS trusts and boards, to actively look at ways of ensuring that people with disabilities have equal access to service provision. It may be particularly important to consider this duty when an intervention has been delivered by an external provider. More information is available on the Public Sector Equality Duty website.⁴²

In the context of physical activity, monitoring of disability, co-morbidities (such as diabetes) or long-term illness is important, as physical or learning difficulties in adults and children can confound the outcome of an intervention. In many cases interventions will need to be adapted specifically to meet the physical capacity of participants. It is also important to consider learning disabilities: children with learning or physical difficulties may be at a greater risk of obesity,⁴³ and are likely to have lower capacity for physical activity.

30. Measure of socio-economic status**ESSENTIAL**

Participation in physical activity can be influenced by a range of factors related to socio-economic status. It is useful to record socio-economic status in some form, and to monitor uptake of interventions by different socio-economic groups to ensure the intervention is not systematically excluding or favouring any particular groups through their design, delivery, recruitment or referral methods.

There are standard indicators of socio-economic status, such as the National Statistics Socio-economic Classification (NS-SEC). However, collecting and analysing the data needed for this classification is likely to be too complex for most public health interventions. A more appropriate approach would therefore be to consider proxy variables for socio-economic status that may be more relevant. Such proxy variables should relate closely to the purpose and nature of the intervention. It is important to ask the question: 'What social factors are likely to affect uptake of the intervention?', and think of appropriate ways to measure these factors.

For example, with an intervention focusing on children's outdoor play, it may be useful to collect information on social factors that might influence children's propensity to play outside – such as access to a garden, or the presence of a TV in a child's bedroom. Other simplified proxy measures of socio-economic status include:

- Educational attainment

This could use the Census categories; the age at leaving full-time education; or having a higher degree.

- Housing tenure

The 2001 and 2011 Census split tenure into ten categories; these could be simplified into 'owner-occupied' and 'not owner-occupied'.

- Free school meals status

Analysis of these data can be complex, and it may be necessary to seek specialist help. Local public health analysts or researchers may be able to assist with this type of analysis.

31. Additional information including marital status, medical history and parental status**DESIRABLE**

Additional information can enhance the evaluation and provide an indication of possible confounding factors when analysing results:

Marital status/cohabitation

Data on marital status/cohabitation is a useful measure as there is evidence to indicate that the BMIs of married people tend to be correlated, and that men and women may increase weight when married.⁴⁴ It can therefore be useful not only to collect data on whether the subject is married (or cohabiting), but also the activity levels of the subject's partner.

Medical history

This can include factors such as a participant having an existing clinical condition that is known to be linked with lack of physical activity, such as type II diabetes. Relevant medical history information can be gathered from a screening questionnaire prior to participants enrolling on a project.

Parental status

This can be useful as the presence of children in a household may have an influence on the adults' levels of physical activity.

32. Details of parental physical activity/weight (for evaluations of interventions aimed at children)

DESIRABLE

There is some evidence that children are more likely to be active if one or both parents are active,⁴⁵ so this information may provide useful context. Parents could be asked to complete physical activity questionnaires alongside their children as part of the evaluation. Parents could also record their diet and weight status. More detail on measurement methods is included in the NOO document, *'Measuring diet and physical activity in weight management interventions'*.⁸

Part four: baseline data

It is extremely important to collect baseline data before an intervention begins. If an intervention commences without baseline data it will not be possible to accurately assess its impact.

33. Measure(s) of physical activity behaviour

ESSENTIAL

Baseline measures of physical activity behaviour should be made before the intervention begins. Details on selecting and measuring appropriate outcomes can be found in Section 3 (pages 9–14).

34. Measure(s) of correlates of physical activity (if relevant to the intervention objectives)

DESIRABLE

The vast majority of physical activity programmes aim primarily to increase levels of physical activity, so some measure of activity should be the primary outcome measure. Most interventions will be based on a theory or assumption about how the intervention will work. For example, it may be thought that a 'Walk to Work' intervention should try to increase motivation and confidence in walking to work because it is assumed that motivation and confidence are determinants of behaviour, and that improving them may increase the likelihood of positive behaviour change. Therefore, as well as measuring the primary outcome (walking to work), a secondary outcome measure could be taken of these other correlates.

For instance, there is evidence that physical activity behaviour is related to levels of self-efficacy (a person's belief that he or she is capable of doing something, such as taking up a new activity). So an intervention may focus on modifying self-efficacy to increase physical activity. If so, it would be helpful to be able to determine if there were changes in self-efficacy (secondary outcome measure) as well as any changes in physical activity (the primary outcome).

An example is given in the logic model of a 'Walk to School' project (Figure 2, page 8). In this example, data are collected on the pupils' intention to walk to school, and levels of confidence in walking to school. This information will identify whether the walking intervention has changed confidence or intention to walk, as well as whether or not it has increased walking.

A number of reviews have explored the correlates of physical activity for adults,⁴⁶ children and adolescents,⁴⁷ and for the environment.^{48,49} Examples of correlates that may be suitable for modification through physical activity projects are shown in Appendix 2. However, it is essential to think carefully about which factors are relevant to a specific intervention.

35. Other outcome measures (such as BMI, other physiological measures, quality of life)

DESIRABLE

Finally, it is worth considering other outcome measures in addition to participation in physical activity. Such outcomes should only be measured if they are relevant to the aims and objectives of the project.

Body mass index (calculated from height and weight) is listed in the *SEF for weight management interventions*¹ as an essential criterion, as it is recommended that all weight management interventions

should monitor body weight as part of their evaluation. However, with physical activity interventions it is only a desirable criterion, as interventions may aim to increase physical activity without changing body weight.

It is also important to note that body weight may not change dramatically in the course of a physical activity intervention, particularly if dietary intake remains stable (or increases) and the volume of physical activity undertaken is relatively low.⁵⁰

Other physiological measures are sometimes taken to monitor physiological changes due to increased physical activity. These include blood pressure; cholesterol or assessments of cardiorespiratory fitness. These are beyond the scope of this document, as it is unlikely that most evaluations of physical activity interventions would have the resources to take physiological measurements. If such measures are taken, however, they should adhere to appropriate clinical standards.

Quality of life has been shown to be related to physical activity.⁵¹ It can be measured using questionnaires, the most common of which is SF-36, a multi-purpose, short-form health survey with only 36 questions. It provides a profile of functional health and well-being scores as well as psychometrically-based physical and mental health summary measures.⁵² Permission for use of the SF-36 is required from its copyright holders.

Part five: follow-up data (impact evaluation)

36. Follow-up data on key physical activity measures (minimum of three follow-up points, including at one year)

ESSENTIAL

Part four describes the minimum data that should be collected at baseline, before an intervention begins. These key measures of physical activity behaviour should ideally be collected at least three times after the intervention has taken place. Typically this will be at the end of the intervention, then six and 12 months after the intervention has completed. Although measures from two data points can give some indication of change, it is not sufficient to describe a trend.

Also, the impact is likely to be greatest at completion of an intervention. To measure at this point only might give an artificially inflated indication of the intervention's impact. Behaviour change needs to be sustained to have significant health impacts, therefore it is important to check if behaviours that change at 12 weeks or six months are maintained in the longer term.

Where data on additional outcome measures have been collected (such as psycho-social health outcomes), these should also be collected at defined follow-up points.

It is important to collect follow-up data from as many members of the original sample as possible. The WHO Europe *'Good Practice Appraisal Tool'* recommends that follow-up is performed with a representative sample of the target group that includes more than 80% of the original intervention participants.⁵³

37. Follow-up data on key physical activity measures over a greater term than one year

DESIRABLE

Ideally follow-up data should be collected over a period greater than 12 months. See point 36 above.

38. Follow-up data on correlates of physical activity behaviour (if collected at baseline)

DESIRABLE

See points 34 and 36.

39. Follow-up data on other secondary outcome measures

DESIRABLE

See points 35 and 36.

Part six: process evaluation**40. Number of participants invited****ESSENTIAL**

An important aspect of evaluating an intervention is to determine the flow of participants through the project. Accurate monitoring of participant numbers is often a weakness of published evaluations.⁵⁴ It is important to distinguish between interventions offered to the whole community and those offered only to a select group. How are participants initially recruited? How many individuals were invited to attend the intervention? If invitation /recruitment is via GP referrals or referral from a school nurse, state the number of people who were advised to attend the intervention, and state if the invitation was via leaflets, posters or other publicity.

Recruitment information can be easily set out in a flow chart describing the numbers of participants at each stage.⁵⁴ Where possible, the number of participants invited should be categorised by age, sex, ethnicity and socio-economic position – to identify any possible inequality introduced during the invitation process. This is also the case for points 41–44.

41. Number of participants recruited**ESSENTIAL**

How many people applied to join the intervention or accepted the invitation? How many were screened? How many participants were found to be eligible for the intervention? How many actually enrolled?

42. Number of participants attending each session or contact point**ESSENTIAL**

How many participants attended each exposure, episode, session or contact point? For example, if an intervention is run twice a week for ten weeks, how many participants attended each of these sessions?

43. Number of participants at each follow-up point, including final stage**ESSENTIAL**

How many participants attended each of the follow-up stages for data collection? If high numbers of people drop out of the intervention, this can lead to biased findings. Some research studies define 'completion' as 75 per cent attendance of the programme/ intervention.⁵⁵

44. Reasons for opt-out (where relevant)**ESSENTIAL**

Whilst opt-out information is not always easy to gather, it is important to understand why participants drop out of an intervention. It is particularly important if the intervention is to be run more than once, or as part of a rolling programme. Opt-out information can be collected in a number of ways and is similar to gathering information on participants' overall satisfaction with the intervention (see point 46).

45. Description of what was actually delivered and details of any unexpected outcomes**DESIRABLE**

Those delivering an intervention should keep a record of what was actually delivered. Such information may be collected via intervention diaries or logs of activities. The log is useful to identify if the planned intervention was changed in any way. It is also helpful to explain any changes observed in collected data, as these can be linked to project activity.

It is useful to note whether there were any unexpected side effects or outcomes from the intervention. For example, did participants take up one type of physical activity but stop doing another during the project? Unexpected outcomes do not necessarily have to be negative and there may be unanticipated positive health outcomes.

46. Participants' satisfaction with the intervention**ESSENTIAL**

Satisfaction questionnaires are frequently used as part of evaluations. It is important to find out if participants are satisfied or dissatisfied with the way in which an intervention is delivered, or unhappy with an element of the overall design of the intervention. Understanding which elements of an intervention were not popular will help in the planning of future programmes. Dissatisfaction may also lead to non-attendance and the intervention would then be less likely to achieve its defined outcomes.

When collecting data on participants' satisfaction, it should be noted that it is often very difficult to glean unbiased opinions from participants if there have been problems and difficulties. Therefore, any research of this nature should be carefully and sensitively conducted. To identify strengths and weaknesses of the intervention, it can be more useful to use qualitative methods such as focus groups or semi-structured interviews. It is not advisable for the deliverer of the intervention to carry out the evaluation. Participants may feel more able to be honest with a person they have not previously encountered as part of the delivery team.

47. Plans for sustainability**DESIRABLE**

Consider whether plans have been made to ensure participants are offered the opportunity to continue with the intervention in some way – perhaps through a follow-up group or referral to a permanent project. This will help the intervention's effects to be sustained over time. There may be resource implications for this type of long-term planning, and such information should be included in the evaluation.

Part seven: analysis and interpretation

No matter how good the quality of the collected data, it is vital that the data are correctly analysed and interpreted, otherwise they may not produce useful, robust information and feedback about the intervention. It is beyond the scope of this document to provide detailed guidance on data analysis. Readers should either seek the advice of a statistician or analyst, or consult published information on study design and statistical methods.

48. Summary of results compared to baseline (for primary and secondary outcomes)**ESSENTIAL**

The most basic type of analysis is to show whether primary and secondary outcomes have changed over the course of the intervention. The method for analysing and presenting results from the evaluation will depend on the study design. This in turn will determine the degree of confidence in the results. Section 2.7 of '*SEF for weight management interventions*' describes the main study designs used for evaluations and appropriate analysis methods.¹

49. Details of any further analyses or statistical methods used**DESIRABLE**

It is beyond the remit of this document to detail the statistical methods that could be used in the analysis of collected data. However when comparing outcome data with baseline, or with data from other similar studies it is important to consider statistical significance: the extent to which we can be confident that a result did not occur by chance.

The use of confidence intervals (CIs) is a good way of presenting a level of certainty with data. CIs describe a range of values where there is 95 per cent confidence that the observed outcome is the true value. So, for example, there may be a mean change of 15 minutes in time spent being active following an intervention, with 95 per cent confidence intervals stated as 13 to 17 minutes per day. This means there is a 95 per cent likelihood that the true amount of activity is between 13 and 17 minutes.

50. Limitations and generalisability**ESSENTIAL**

The generalisability of the intervention is the likelihood that the results of the intervention would be reproduced if the intervention were carried out with another group, or in the whole population. When assessing generalisability, it is important to take into account: the target population, nature of the intervention, nature of sampling and recruitment methods, length of follow-up, and settings and resources needed.

The more tightly controlled the intervention, the more likely it is that the intervention can be described so that it can be reproduced in another setting or population.

51. Recommendations and changes to future projects**ESSENTIAL**

Finally, it is important to 'take a step back' from the project and consider what should be done differently next time. Are there key recommendations to pass on to those running similar projects in the future?

52. Dissemination of learning and findings**DESIRABLE**

It is important to share the findings of your evaluation to help build the evidence base for what works, and what doesn't work, to increase physical activity levels. This can be done through formal academic routes such as published papers and conference presentations, as well as feedback to the funders and partners.

Please share your results with NOO. We have developed an evaluation data collection tool for weight management, physical activity and dietary interventions, based on the fields presented in this and other SEF frameworks. Submitting your data will enable others to learn from your experiences and also provide you with an audit record of your intervention.

NOO also hosts a national searchable database of weight management, physical activity and diet interventions. The evaluation data collection tool and the searchable database of interventions can both be found on the NOO website at www.noo.org.uk/core.

6. Conclusions

We hope this document has been helpful to those commissioning, running and evaluating physical activity interventions. Our aim has been to establish the beginnings of a consensus on the minimum data that need to be collected for an evaluation to be effective and to help raise standards. This will hopefully contribute towards the long-term aim of advancing knowledge of what is effective in combating the rise in obesity in this country.

We welcome any comments on this document: please email info@noo.org.uk

Declaration of interest statement: Nick Cavill, one of the authors of this document, runs a public health consultancy that sells pedometers as part of its business. This document was reviewed by an independent advisory group, and the sections on pedometers were specifically checked by NOO directors to avoid any potential conflict of interest.

Appendices

Appendix 1: Office for National Statistics categories for ethnic group monitoring in England^{c,40}

White

English/Welsh/Scottish/Northern Irish/British

Irish

Gypsy or Irish Traveller

Any Other White Background

Mixed/Multiple ethnic groups

White and Black Caribbean

White and Black African

White and Asian

Any other Mixed/multiple ethnic background

Asian/Asian British

Indian

Pakistani

Bangladeshi

Chinese

Any other Asian background

Black/African/Caribbean/Black British

African

Caribbean

Any other Black/African/Caribbean background

Other ethnic group

Arab

Any other ethnic group

.....
c At the time of publication many NHS data sources may not collect the full range of categories above. In particular 'Gypsy or Irish Traveller' and 'Arab'. These are new categories included in the 2011 Census.

Appendix 2: Examples of factors that may be modified to increase physical activity

Personal and social factors

- enjoyment of exercise
- expectation of benefits
- intention to exercise
- self-efficacy
- self-motivation
- stage of change
- social support from friends/peers/spouse/family
- barriers to exercise
- lack of time
- perceived effort.

Environmental factors

- school facilities for physical activity
- neighbourhood facilities for physical activity
- traffic safety
- availability of physical activity equipment (associated with vigorous physical activity/sports)
- connectivity of trails (associated with active commuting)
- availability, accessibility and convenience of recreational facilities.

Appendix 3: Physical activity indicators in the 2012 Public Health Outcomes Framework³⁴

PHOF number 1.16 Utilisation of green space for exercise/health reasons

Rationale	Inclusion of this indicator is recognition of the significance of accessible green space as a wider determinant of public health. There is strong evidence to suggest that green spaces have a beneficial impact on physical and mental wellbeing and cognitive function through both physical access and usage.
Final indicator available from	Now
Indicator definition	<p>The indicator definition is ready.</p> <p>1.16 Percentage of people using green space for exercise/health.</p> <p>Numerator: Number of people reporting that they have taken a visit to green space for health or exercise over the previous seven days.</p> <p>Denominator: Total number of respondents to survey.</p>
Data source	<p>The data source is ready.</p> <p>Monitor of Engagement with the Natural Environment (MENE) survey.</p> <p>Data is fully available at England and regional level. It will be possible to produce upper tier local authority level data from this source.</p>
Reporting	Results from the MENE survey (annual report and monthly updates) are published by Natural England: http://www.naturalengland.ork.uk/ourwork/enjoying/research/monitor/#statistics

PHOF number 2.13 Proportion of physically active and inactive adults

Rationale	<p>Lack of sufficient physical activity costs the NHS over £1bn per year – £6.5bn per year to the wider economy – and is one of the top few risk factors for premature mortality. The need for physical activity has become particularly high profile since the publication of the UK CMO guidelines and in the context of the 2012 legacy.</p> <p>Physical activity provides important health benefits across the life course. Participation in sport and active recreation during youth and early adulthood can lay the foundation for lifelong participation in health-enhancing sport and wider physical activity.</p>
Final indicator available from	January 2012
Indicator definition	<p>The indicator definition is ready.</p> <p>2.13i Proportion of adults achieving at least 150 minutes of physical activity per week in accordance with UK CMO recommended guidelines on physical activity.</p> <p>Numerator: Number of adults (16+) doing at least 150 minutes of at least moderate intensity physical activity per week in bouts of 10 minutes or more*.</p> <p>2.13ii Proportion of adults classified as “inactive”.</p> <p>Numerator: Number of adults (16+) who do less than 30 minutes of moderate intensity physical activity per week in bouts of 10 minutes or more*.</p> <p>Denominator: Population of adults (aged 16+)</p> <p>*Based on 2011 CMO report “Start active, stay active”.</p>
Data source	<p>The data source is ready.</p> <p>Sport England’s Active People Survey.</p> <p>Work is under way to align questions with the new physical activity guidelines – data collection to commence in January 2012.</p>
Reporting	Active People Survey data relating to sport and wider physical activity (the data source for this indicator) will be published for national and local levels at six monthly intervals.

Glossary of terms

Aim

A broad statement of intent setting out the purpose of the project. For example: 'The project aims to reduce obesity through increasing participation in regular physical activity in schools.'

Correlates

Measures that are related to the outcome of interest. For example, the outcome may be increased walking to school; a correlate may be confidence in crossing busy roads. Interventions may then attempt to modify this correlate (for example, by providing a school crossing warden).

Evaluation

Establishing whether a project has achieved what it set out to do.

Focus group

A group of people who discuss an issue, led by a researcher. Focus groups generate qualitative data. They are sometimes used in preference to individual interviews, as some researchers believe that a group dynamic will reveal more about people's beliefs and attitudes.

Impact evaluation

This focuses on whether a project has met its aims and objectives. This might be in terms of health outcomes, such as obesity, or impacts on health behaviours, such as increased physical activity or improved diet.

Interviews

A discussion between a researcher and subject(s), usually using a script or pre-designed list of questions, prompts and topics. Interviews can be face-to-face or by phone, structured (with fixed questions) or semi-structured (where discussion can be more flexible).

Monitoring

Collection of routine data that helps one assess whether projects are proceeding to plan.

Objective

A statement of exactly what the project is trying to achieve. For example: 'To increase the proportion of children who are active for an hour or more each day.'

Outcome

A visible or practical result, effect or product. It highlights the change or impact a project will have on the target population. For example: 'An average reduction of 2cm in waist circumference among project participants within six months.'

Outcome measure or indicator

A measure of something which demonstrates a change in a particular outcome. For example, the number of children who increase the time spent active each week.

Outputs

Outputs of the project or activities that occur, through the use of the resources in the project. For example: 'Ten community walks in local parks.'

Process evaluation

Evaluation which focuses on the processes used throughout an intervention. It aims to identify why the project met or did not meet its aims and objectives; what went right and what went wrong; what can be learnt for future projects.

Qualitative research

Qualitative research provides information that is often reported in narrative form or which is based on descriptive information, such as diaries, open-ended responses to questions and field notes. Qualitative data are usually analysed using non-statistical methods.

Quantitative research

Research which measures and analyses observations to produce numerical or statistical data.

Reliability

The consistency of a measure. A measure is said to have a high reliability if it produces consistent results under consistent conditions.

Target population

The people the project aims to reach. These may be segmented by a number of factors including age, sex and social class.

Validity

The validity of a tool is the degree to which it measures what it claims to measure. For example, the extent to which a self-report physical activity questionnaire reflects actual behaviour.

References

1. Roberts K, Cavill N, Rutter H. Standard Evaluation Framework for Weight Management Interventions. Oxford: National Obesity Observatory, 2009.
2. Roberts K, Cavill N, Rutter H. The Standard Evaluation Framework for dietary interventions London: National Obesity Observatory, 2012.
3. Cavill N, Roberts K, Rutter H. Standard Evaluation Framework for Physical Activity Interventions. London: National Obesity Observatory, 2012.
4. National Obesity Observatory. Collection of Resources on Evaluation (CoRE) Oxford: National Obesity Observatory, 2012.
5. National Obesity Observatory. Evaluation data collection tool. Oxford: National Obesity Observatory, 2011.
6. National Obesity Observatory. Searchable Database of Interventions Oxford: National Obesity Observatory, 2012.
7. Richardson D, Cavill N, Ells L, Roberts K. Measuring Diet and Physical Activity in Weight Management Interventions: a Briefing Paper. Oxford National Obesity Observatory, 2011.
8. Richardson D, Cavill N, Ells L, Roberts K. Supplement: Measuring Diet and Physical Activity in Weight Management Interventions. Oxford National Obesity Observatory, 2011.
9. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports* 1985;100(2):126-31.
10. National Institute for Health and Clinical Excellence. Obesity: Guidance on the Prevention, Identification, Assessment and Management of Overweight and Obesity in Adults. London: NICE, 2006.
11. Katikireddi SV, Higgins M, Bond L, Bonell C, Macintyre S. How evidence based is English public health policy? *BMJ* 2011;343:d7310.
12. Nutbeam D, Bauman A. *Evaluation in a Nutshell: a Practical Guide to the Evaluation of Health Promotion Programs*. Sydney: McGraw Hill Publishers, 2006.
13. US Department of Health and Human Services. Physical Activity Evaluation Handbook. Atlanta: US Department of Health and Human Services, 2002.
14. Department of Health. At Least Five a Week: Evidence on the Impact of Physical Activity and its Relationship to Health. London: Department of Health, 2004.
15. Department of Health. Start active, Stay Active: a Report on Physical Activity from the Four Home Countries' Chief Medical Officers. London: Department of Health 2011.
16. Tudor-Locke C, Bassett DJ. How many steps/day are enough? Preliminary pedometer indices for public health. *Sports Med* 2004;34(1):1-8.
17. Stevens W, Hillsdon M, Thorogood M, McArdle D. Cost-effectiveness of a primary care based physical activity intervention in 45-74 year old men and women: a randomised controlled trial. *British Journal of Sports Medicine* 1998;32(3):236-41.
18. Ogilvie D, Foster C, Rothnie H, Cavill N, Hamilton V, Fitzsimons C, et al. Interventions to promote walking: systematic review. *BMJ* 2007;334(7605):1204.

> CONTENTS

19. Isaacs AJ, Critchley JA, Tai SS, Buckingham K, Westley D, Harridge SD, et al. Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. *Health Technol Assess* 2007;11(10):1-165, iii-iv.
20. Rutishauser I H E, Black A E. Measuring Food Intake. In: Gibney M J, Kok F J, editors. *Introduction to Human Nutrition*. Oxford: Blackwell Science, 2002.
21. Medical Research Council. Diet and Physical Activity Measurement Toolkit. Cambridge: MRC, 2012.
22. De Cocker K, Cardon G, De Bourdeaudhuij I. Validity of the inexpensive Stepping Meter in counting steps in free living conditions: a pilot study. *British Journal of Sports Medicine* 2006;40(8):714-6.
23. Gibson RS. *Principles of nutritional assessment*. 2nd ed. New York: Oxford University Press, 2005.
24. Block G, Hartman AM. Issues in reproducibility and validity of dietary studies. *The American journal of clinical nutrition* 1989;50(5 Suppl):1133-8; discussion 231-5.
25. The INTERACT consortium. Validity of a short questionnaire to assess physical activity in 10 European countries. *European journal of epidemiology* 2011;27(1):15-25.
26. Milton K, Bull FC, Bauman A. Reliability and validity testing of a single-item physical activity measure. *British Journal of Sports Medicine* 2011;45(3):203-8.
27. Milton K, Clemes S, Bull F. Can a single question provide an accurate measure of physical activity? *British Journal of Sports Medicine* 2012 (epub ahead of publication).
28. Sport England. Active People Survey. London: Sport England, 2012.
29. National Research Ethics Service. Consent Guidance 2012. <http://www.nres.nhs.uk/applications/guidance/> (accessed 24 July 2012)
30. National Archives. Data Protection Act 1998, 1998. <http://www.nationalarchives.gov.uk/information-management/legislation/data-protection.htm> (accessed 24 July 2012)
31. National Research Ethics Service. Defining Research London, 2009. http://www.red.mmu.ac.uk/documents/res_files/ethics/NRES_leaflet_Defining_Research-2.pdf (accessed 24 July 2012)
32. National Institute for Health and Clinical Excellence. Costing template London, 2006.
33. Department of Health. Joint Strategic Needs Assessment. London: Dept of Health. 2007.
34. Department of Health. Healthy lives, healthy people: Improving outcomes and supporting transparency. London: Department of Health, 2012.
35. Department of Health. Equality Impact Assessment: Summary Tool and Guidance for Policy Makers. London: Department of Health, 2008.
36. NHS Employers. Equality Analysis and Equality Impact Assessments London, 2011.
37. National Institute for Health and Clinical Excellence. A Code of Practice for Declaring Interests and Resolving Conflicts. London: NICE, 2006.
38. Swanton K. Healthy Weight, Healthy Lives: A toolkit for Developing Local Strategies. London, 2007.
39. Equalities and Human Rights Commission. Our Equality Information, 2012.

40. Office for National Statistics categories for ethnic group monitoring, 2011. <http://www.ons.gov.uk/ons/guide-method/measuring-equality/equality/ethnic-nationalityreligion/ethnic-group/index.html> (accessed 6 August 2012)
41. Direct Gov. Disability and the Equality Act 2010, 2010. http://www.direct.gov.uk/en/disabledpeople/rightsandobligations/disabilityrights/dg_4001068 (accessed 24 July 2012)
42. Office for Disability Issues. Public Sector Equality Duty, 2010. <http://odi.dwp.gov.uk/disabled-people-and-legislation/disability-equality-duty-and-impact-assessments.php> (accessed 24 July 2012)
43. Ells LJ, Shield JP, Lidstone JS, Tregonning D, Whittaker V, Batterham A, et al. Teesside Schools Health Study: body mass index surveillance in special needs and mainstream school children. *Public Health* 2008;122(3):251-4.
44. The NS, Gordon-Larsen P. Entry into romantic partnership is associated with obesity. *Obesity* 2009;17(7):1441-7.
45. Lawman HG, Wilson DK. A Review of Family and Environmental Correlates of Health Behaviors in High-Risk Youth. *Obesity* 2012; 20(6):1142-57.
46. Trost SG, Owen N, Bauman AE, Sallis JF, Brown W. Correlates of adults' participation in physical activity: review and update. *Medicine and Science in Sports and Exercise* 2002;34(12):1996-2001.
47. Craggs C, Corder K, van Sluijs EM, Griffin SJ. Determinants of change in physical activity in children and adolescents: a systematic review. *American journal of preventive medicine* 2011;40(6):645-58.
48. de Vet E, de Ridder DT, de Wit JB. Environmental correlates of physical activity and dietary behaviours among young people: a systematic review of reviews. *Obesity reviews : an official journal of the International Association for the Study of Obesity* 2011;12(5):e130-42.
49. Wendel-Vos W, Droomers M, Kremers S, Brug J, van Lenthe F. Potential environmental determinants of physical activity in adults: a systematic review. *Obesity reviews : an official journal of the International Association for the Study of Obesity* 2007;8(5):425-40.
50. Jakicic JM, Clark K, Coleman E, Donnelly JE, Foreyt J, Melanson E, et al. American College of Sports Medicine position stand. Appropriate intervention strategies for weight loss and prevention of weight regain for adults. *Medicine and Science in Sports and Exercise* 2001;33(12):2145-56.
51. Bize R, Johnson JA, Plotnikoff RC. Physical activity level and health-related quality of life in the general adult population: a systematic review. *Preventive Medicine* 2007;45(6):401-15.
52. Ware JE. SF-36 Health Survey Update., 2011. See <http://www.sf-36.org/tools/SF36.shtml/> (accessed 24 July 2012)
53. World Health Organization for Europe. Good Practice Appraisal Tool for Obesity Prevention Programmes, Projects, Initiatives and Interventions. Copenhagen: WHO Europe 2011.
54. Foster CE, Brennan G, Matthews A, McAdam C, Fitzsimons C, Mutrie N. Recruiting participants to walking intervention studies: a systematic review. *The international journal of behavioral nutrition and physical activity* 2011;8:137.
55. Hughes AR, Stewart L, Chapple J, McColl JH, Donaldson MD, Kelnar CJ, et al. Randomized, controlled trial of a best-practice individualized behavioral program for treatment of childhood overweight: Scottish Childhood Overweight Treatment Trial (SCOTT). *Pediatrics* 2008;121(3):e539-46.

Title	Standard Evaluation Framework for physical activity interventions
Authors	Nick Cavill Kath Roberts Harry Rutter
Reviewers	Advisory Group – see page 2
Editor	Di Swanston
Publication Date	September 2012
Target Audience	<ul style="list-style-type: none"> • Commissioners or managers of weight management or obesity prevention interventions with a physical activity element. • Commissioners or managers of physical activity interventions. • Physical activity or sport and leisure leads in local authorities. • Commissioners or managers of active travel projects. • Practitioners running physical activity or active travel projects.
Description	This document describes and explains what information should be collected in any evaluation of an intervention that aims to increase participation in physical activity
How to cite	Cavill N, Roberts K, Rutter H. Standard Evaluation Framework for physical activity interventions. Oxford: National Obesity Observatory. 2012.
Contact	National Obesity Observatory www.noo.org.uk info@noo.org.uk
Electronic location	www.noo.org.uk/core
Copyright	© National Obesity Observatory, 2012

noo

National Obesity
Observatory

noo.org.uk

NOO is delivered by Solutions for Public Health.