Public Health England

Protecting and improving the nation's health

Standard Evaluation Framework for Weight Management Interventions

February 2018 (revised)

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. We do this through world-leading science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. We are an executive agency of the Department of Health and Social Care, and a distinct delivery organisation with operational autonomy. We provide government, local government, the NHS, Parliament, industry and the public with evidence-based professional, scientific and delivery expertise and support.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

Prepared by: Louisa Ells, Kath Roberts and Nick Cavill

© Crown copyright 2017
You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0. To view this licence, visit OGL. Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Published March 2018
PHE publications
gateway number: 2017812

PHE supports the UN
Sustainable Development Goals

Acknowledgements

This document is an update of the original ‘Standard Evaluation Framework for weight management interventions’ published by the former National Obesity Observatory in 2009. This original document was written by Kath Roberts, Nick Cavill and Harry Rutter, with input from Louisa Ells and Hywell Dinsdale.

Public Health England’s Risk Factors Intelligence Team would like to acknowledge contributions from the following people:

Contributors

Jamie Blackshaw, Vicki Coulton, Tim Chadbourn and the PHE evaluation group, Mary Gatineau , Shireen Mathrani , Clare Perkins, Harry Rutter, Alison Tedstone , Claire O’Malley Lina Toliekyte and Jennifer Logue.

Consultees

Steve Cummins, Adrian Coggins, James Creaghan, Claire Glazzard, Zoe Helman, Jane Hynes, Stuart King, Scott Lloyd, Sam Montel, Carolyn Pallister, Duncan Radley, Claire Ramwell, Natalie Randell, Emma Strachan, Carol Weir.

Contents

Resources for the evaluation of weight management interventions

Public Health England (PHE) is committed to improving the quality and quantity of evaluations of weight management interventions.

This document provides an update to the 2009 Standard Evaluation Framework for weight management interventions, and is accompanied by a number of other useful evaluation resources:

- evaluation of public health interventions: an introductory guide (2015)
- the Magenta Book: HM Treasury guidance on what to consider when designing an evaluation (2011)
- an archive of evaluation resources (archived in 2017), and the forthcoming PHE online evaluation resource suite.
- capturing data: a tool to collect and record adult weight management intervention data (2017)
- capturing data: a tool to collect and record child weight management intervention data (2017)
- key performance indicators: tier 2 weight management interventions for adults (2017)

Introduction

This document is an update of the 'Standard Evaluation Framework (SEF) for weight management interventions', (2009) which was originally published by the National Obesity Observatory, and is now widely used across England.[1]

It has been updated as a result of feedback from practitioners in the field following a consultation exercise, and to provide support for the Guides to Commission Tier 2 Weight Management Interventions for Children, Families and Adults (2017).[2]

It has also been informed by findings from a project led by the University of Glasgow, funded by the Chief Scientists Office (CSO) Scotland, to develop core outcome measures for lifestyle weight management programmes by expert consensus[a]. The CSO project once completed (July 2018) will also provide consensus on measurement tools to support data collection.

This document 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 a weight management intervention. 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.

In this document, the term 'weight management intervention' refers to any service that explicitly sets out to manage or reduce body weight (including the primary prevention of weight gain).

What does this document aim to do?

The SEF for weight management interventions aims to describe and explain the information that should be collected in any evaluation of a weight management intervention.

This document contains guidance on evaluating interventions which have reduction in body weight, body mass index (BMI), BMI-z or centile as a primary outcome. Separate standard evaluation frameworks have been developed for programme which specifically focus on physical activity[3] and diet[4] outcomes.

[a] CGA/17/08 - Developing core outcome measures for lifestyle weight management programmes by expert consensus (unpublished findings were kindly provided by Dr Jennifer Logue).

The SEF provides a list of data collection criteria and supporting guidance for collecting high quality information that supports the evaluation of weight management interventions across England. It is aimed at interventions that work at individual or group level, not at population level.

What does this document not aim to do?

Provide guidance that is intended to support the evaluation of medical interventions, such as bariatric surgery or medication.

Provide guidance on the evaluation of broader societal-level interventions, digital interventions or interventions that promote weight management through changes to the physical environment. Whilst this document may help inform these types of intervention, it is acknowledged that they present additional evaluation challenges that are beyond the scope of this document.

Provide detailed advice on ethics, information governance and how to seek external funding support for intervention evaluations.

Provide an introduction to the concepts of evaluation, as this is provided in a separate document: Evaluation of public health interventions: an introductory guide (2015).[5]

Who is the target audience?

The target audiences for this document are:

- commissioners of weight management interventions
- obesity leads in local authorities
- practitioners running weight management interventions
- evaluators of weight management interventions

Why do we need a SEF for weight management?

The evidence base for the effectiveness of weight management interventions has grown in recent years. However, there remains a need to continue to obtain high quality evidence on effective weight management interventions.[6] While interventions are being commissioned by a variety of organisations, data concerning the relative ‘success’ of the interventions are often patchy and inconsistent. There is a need for rigorous evaluation of local weight management interventions, particularly those that can be applied as part of routine care[7] and those that specifically reduce the health inequalities in obesity prevalence.[8]

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.[9] Such evaluation practice makes it hard to compare the impact of interventions, and understand which interventions are more effective, and for which populations.

In evidence-based medicine, randomised controlled trials (RCTs) are usually considered to be the 'gold standard' for a scientifically robust assessment of whether an intervention is effective. The Consolidated Standards of Reporting Trials (CONSORT) statement provides guidelines for how RCTs should be reported for evaluation purposes.[10] An extension to the CONSORT statement provides similar guidance about the reporting of 'pragmatic' trials which are intended to inform decisions about whether an intervention works in 'normal' practice.[11] The CONSORT 2010 statement (item 5) is further supported by the publication of the 12 item TIDieR checklist, which aims to improve intervention description reporting, and hence reproducibility.[12]

However, RCTs are not always practical or appropriate when evaluating public health interventions. For example, RCTs typically evaluate a specific intervention and its effects on specified outcomes, within a defined population group, and dedicated timeframe and budget. By contrast, public health interventions often have multiple elements, involve complex partnerships, have restricted timeframes, resources and budget, and include participants with complex needs and circumstances who may be less likely to participate in a formal trial. Therefore where RCTs are not possible, it is important to explore other evaluation methods to ensure interventions are fit for purpose, and learning is shared. Remember, good evaluation planning starts at the same time as the intervention is being developed or commissioned.

The Medical Research Council's 'Developing and evaluating complex interventions' includes non-experimental methods and complex interventions outside the health service.[13] It recognises that there are methodological and practical constraints to carrying out complex interventions that need to be considered when evaluating an intervention. The Transparent Reporting of Evaluations with Nonrandomised Designs (TREND) statement has general guidance and a checklist, similar to the CONSORT statement, for the evaluation of non-RCTs and other types of research design.[14]

The document 'Evaluation of public health interventions: an introductory guide' provides an introduction to evaluation principles and methods[5] and the Magenta book provides an introduction to the standards of good practice in conducting evaluations, and the issues faced when undertaking evaluations of projects, policies, and the delivery of intervention.[15]

The Consort Statement, Medical Research Council guidance, TIDieR checklist and TREND statement have all been consulted and referenced in the development of this document.

This document does not set out to be over-prescriptive or stifle innovation. Evaluation is not a fixed or stand alone process: it needs to be flexible and adaptable to the needs of the intervention being appraised. This document provides an evidence-based framework that will help to improve the quality of evaluation, and the development of relevant policies.

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, health promotion and health improvement, an evaluation determines the extent to which an intervention has achieved its objectives and assesses how different processes have contributed to this.9

Evaluation is a critical aspect of good project management, and should be an integral component of any intervention, irrespective of size, and should involve all key stakeholders.

Typically a good evaluation comprises of:

- a process evaluation (principally, was the intervention implemented as intended, where there any issues with the implementation?)
- impact evaluation (principally, what changes occurred following the intervention implementation: did it make a difference, were there any unintended consequences?)
- economic evaluation (principally, a comparison of the intervention benefits and costs)

It is important to note that whilst this framework supports the collection of basic cost data to facilitate a simple cost effectiveness analysis (ie the cost of the intervention implementation in relation to the impact generated eg cost per unit outcome), the primary focus is on the process and impact evaluations. Evaluators are advised to seek specialist advice from a health economist at the start of the evaluation planning process, if a more advanced cost-benefit analysis is required, to ensure all required data is captured during the evaluation. PHE has also developed a weight management economic assessment tool for adults,[16] which may be helpful.

The 2 basic evaluation questions

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

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

Primary and secondary outcomes

The main focus of an evaluation should be to assess whether the primary outcomes that the intervention is trying to achieve have been met. Secondary outcomes are items 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. This is illustrated in the following scenario:

Scenario

An adult weight management intervention has the objective of helping participants to lose weight. Weight reduction is the primary outcome. The primary outcome measures or indicators might be expressed as the proportion who have lost some weight since the beginning the intervention, or may be quantified (for example the proportion of participants who have lost 5% or more of their starting weight).

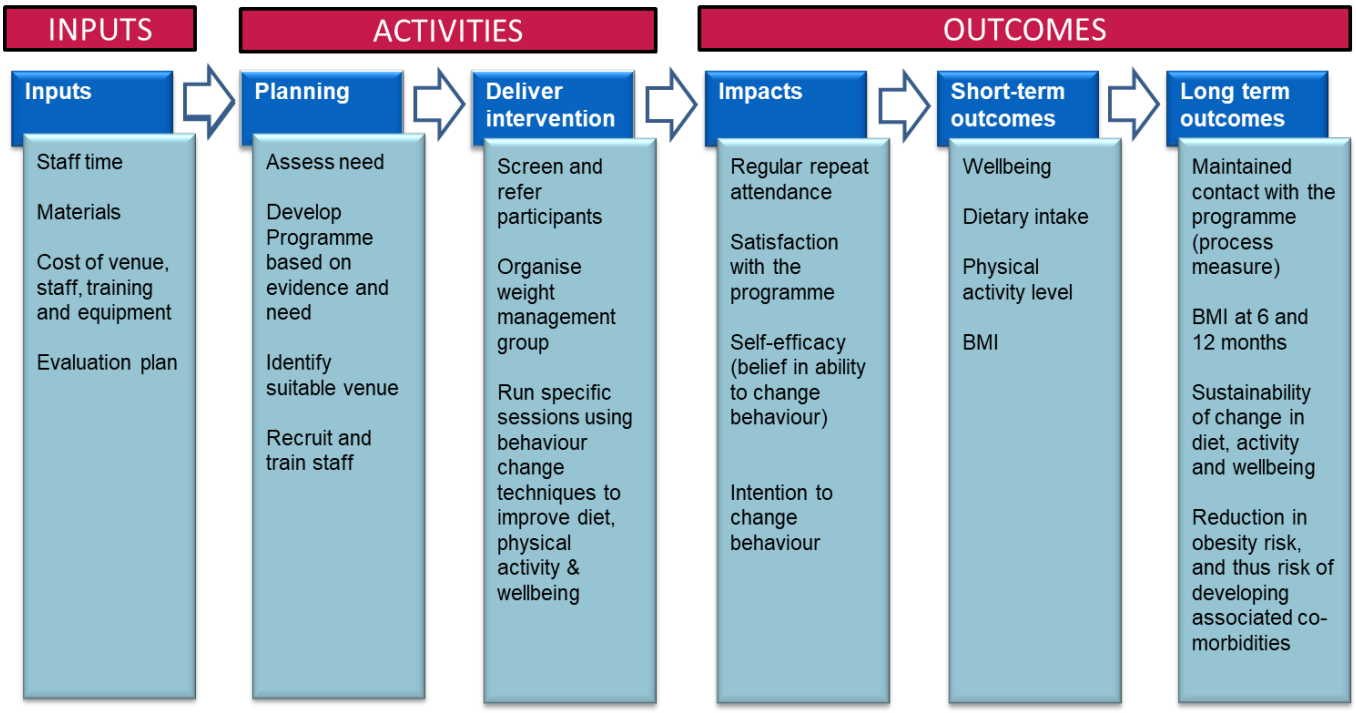
Other outcomes, such as increased participation in physical activity, reduced caloric intake or increased consumption of fruit and vegetables, would be interesting outcomes that might help explain why the intervention has or hasn't worked. These would be secondary outcomes, measured by secondary outcome indicators.

It is important to be clear about the priorities attached to primary and secondary outcomes. For example, an intervention may have the objective of reducing the participants' weight as a result of reducing the consumption of takeaway and processed foods; in this instance weight reduction is the primary outcome and reduction in consumption of takeaway and processed foods is the secondary outcome.

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 anticipated 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 that the project is likely to achieve its objectives.

Figure 1: Example logic model for a simplified adult weight management project

NB: detailed guidance on developing and delivering a weight management intervention is provided in the PHE commissioners and providers guide.[2]



Standard Evaluation Framework checklist

This section presents the core elements of the standard evaluation framework. Essential criteria are presented as the minimum recommended dataset for evaluating a weight management intervention. Desirable criteria are additional data that would enhance the evaluation. The supporting guidance, in section 4, describes why particular criteria have been categorised as essential and desirable, and gives further information on collecting data.

A data capture tool has been created by PHE to support the collection of individual participant data for the monitoring and evaluation of child[17] and adult[18] weight management interventions. This data capture tool aligns with this framework: the minimum data collected within the tool is shaded below.

Criterion	Essential	Desirable
Part 1: intervention details		
1. Title/name of intervention	X	
2. Aims and objectives (including primary and secondary outcomes)	X	
3. Intervention timescale (duration, frequency and number of sessions provided)	X	
4. Intervention delivery dates	X	
5. Duration of funding (including dates)		X
6. Location and setting	X	
7. Description of intervention • intervention content (including details of any individual tailoring or protocol modifications) • delivery method • details of quality assurance mechanisms (including adherence and fidelity assessments, if undertaken) • follow on support/maintenance (if provided)	X	
8. Rationale for intervention (including theoretical basis)	X	
9. Intervention staff and core competencies required	X	
10. Equipment and resources required	X	

11. Use of incentives		X
12. Method of recruitment and referral	X	
13. Participant consent mechanism	X	
14. Participant eligibility criteria	X	
15. Cost of intervention per participant	X	
16. Cost to participant		X
17. Detailed breakdown of cost (including any in kind or developmental costs)		X
18. Type of evaluation and evaluation design	X	
19. Details of equality impact assessment	X	
20. Relevant policy and performance context		X
21. Details of health needs assessments that have been conducted		X
22. Contact details	X	
23. Commissioner(s) of the intervention and sources of funding		X
24. Declaration of interest		X
25. Details of type and extent of any clinical involvement		X
Part 2: demographics of individual participants		
26. Age	X	
27. Sex	X	
28. Ethnicity	X	
29. Disability	X	
30. Measure of socioeconomic status (eg IMD acquired from home postcode)	X	
31. Religion	X	
32. Employment status	X	
33. Sexual orientation	X	
34. Known co-morbidities	X	

35. Additional information: including medical history, alcohol use, medication use, referral to other services, housing tenure, marital status, smoking status, pregnancy, breast feeding and number of children, social support needs, and details of parental weight status (required for child weight management interventions).		X
Part 3: baseline data		
36. Height and weight (to calculate BMI or BMI-z/centile for interventions in children under 18 years)	X	
37. Additional proxy measures for adiposity		X
38. Measure(s) of wellbeing	X	
39. Measure(s) of dietary intake and behaviour		X
40. Measure(s) of physical activity levels and behaviour		X
41. Potential facilitators of, and barriers to, lifestyle change		X
Part 4: follow-up data		
Impact evaluation		
42. Follow-up data on height, weight (BMI) to be collected at: • the end of the active intervention, • 6 months after the end of the active intervention • 12 months after the end of the active intervention	X	
43. Follow-up data on height, weight (BMI) collected 24 months after the end of the active intervention		X
44. Follow-up data on wellbeing to be collected at: • the end of the active intervention, • 6 months after the end of the active intervention • 12 months after the end of the active intervention	X	
45. Follow-up data on wellbeing collected 24 months after the end of the active intervention		X
46. Follow-up data on additional measures of adiposity to be collected at: • the end of the active intervention, • 6 months after the end of the active intervention • 12 months after the end of the active intervention If collected at baseline		X

47. Follow-up data on dietary intake and behaviours to be collected at: • the end of the active intervention, • 6 months after the end of the active intervention • 12 months after the end of the active intervention If collected at baseline		X
48. Follow-up data on physical activity levels to be collected at: • the end of the active intervention, • 6 months after the end of the active intervention • 12 months after the end of the active intervention If collected at baseline		X
49. Follow-up measures on potential facilitators of and barriers to lifestyle change (if collected at baseline)		X
Process evaluation		
50. Number referred and referral source	X	
51. Number recruited	X	
52. Number attended each session or contact point	X	
53. Number completed (ie attended 75% of all sessions of the active intervention)	X	
54. Number of participants at each follow-up point	X	
55. Reasons for dropping out		X
56. Details of any unexpected outcomes and/or adverse events	X	
57. Participants' satisfaction	X	
58. Plans for sustainability	X	
Part 5: analysis and interpretation		
59. Process evaluation	X	
60. Impact evaluation: Summary of results compared to baseline (for primary and secondary outcomes)	X	
61. Impact evaluation: Detail of any further analyses or statistical methods used		X
62. Economic evaluation		X
63. Limitations and generalisability		X

Explanatory notes

The list of criteria above is separated into essential and desirable.

'Essential' criteria are the minimum data and information required to carry out a basic evaluation of an intervention that sets out to manage or reduce body weight including the primary prevention of weight gain. They also align with the minimum data required to populate the PHE weight management service data capture tools for adults and children.[2]

'Desirable' criteria are recommended to improve the quality of the evaluation so that practitioners, commissioners and evaluators can:

- understand more about what went well, and what did not go well when carrying out the intervention
- understand whether or not the delivery of the evaluation needs to be modified
- be confident about conclusions that are drawn from any analysis and interpretation of collected data

Part 1: intervention details

1. Title/name of intervention: essential

A record of the name or title of the intervention, for example, Lose Weight, Feel Great.

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

What does the intervention aim to do? Does it have a primary weight management outcome target such as reducing body weight, BMI or BMI-z or centile? Does it have a secondary outcome target such as increasing a measure of wellbeing such as self-esteem, reducing caloric intake and increasing activity levels? These are the key outcomes which need to be carefully measured, against which the intervention will be evaluated. It may be helpful to refer to the above section on logic models.

Aims and objectives need to be as clear as possible and, ideally, SMART, that is: Specific, Measureable, Achievable, Realistic and Time-bound. See the PHE introductory guide to evaluation[5] for more detail.

3. Intervention timescale: essential

How long will the intervention run for each group of participants? How many sessions, episodes or events will be delivered? How long is the active intervention intended to last? For example, ‘the intervention was delivered in 24 two-hour sessions, twice a week for twelve weeks’.

Please note, the duration of the active intervention may differ from the duration of the service.

4. Intervention delivery dates: essential

This includes dates for the initial referral or recruitment, first point of contact and any subsequent contacts and follow-ups.

5. Duration of funding (including dates): desirable

What are the start and finish dates for the service? The active intervention may be run a number of times throughout the duration of a commissioned service.

6. Location and setting: essential

Where is the intervention taking place? For example the location could be a GP surgery, school or community centre. It may be that it takes place in several settings and they should all be included here. It may be useful to mention if any transport is being provided as part of the intervention.

7. Description of intervention: essential

The headings below are only intended as a guide. It may be that these points are described differently for a particular intervention.

Active intervention content
It is important to provide a clear description of the intervention content, so it is obvious what the results of your evaluation are attributable to, and helpful to others who may wish to adopt the approach used. Clearly state what the active intervention is going to do, and how it is going to do it. List all of its major techniques and theoretical components (eg motivational interviewing), and activities (eg weekly one hour healthy cooking demonstrations followed by a one hour low impact aerobics classes). It is important here to state if the intervention is tailored to participants needs (eg activities may be tailored to fitness or ability levels, or advice tailored to specific dietary requirements). It is also important to note if modifications to the original intervention had to be made, what these were and why they were made. Where possible, provide links to, or append, intervention handbooks, protocols, participant information or delivery materials.

Delivery method
How will the intervention be delivered? For example, face-to–face meetings, by telephone or online, or a combination of these. Who is the intervention aimed at? For example, individuals, families or particular groups.

Details of quality assurance mechanisms
What mechanisms are in place to ensure the intervention is being delivered in the way in which it was planned? This is particularly important if the intervention sets out to use a particular delivery method or style such as motivational interviewing.

Examples of quality assurance mechanisms are spot-checks carried out by an external assessor, or self-assessment check-lists that can be used by the deliverer of the intervention. Include details of any relevant health and safety checks, risk assessments and Disclosure and Barring Service[19] checks if the intervention involves children or vulnerable adults.

Follow on support/maintenance
Some interventions also provide additional follow on or maintenance support after the active intervention has been delivered, and it is important to record details of this and where it occurs.

For example, ‘after the 12 week active intervention, all clients are invited to a free weekly weigh in session at the community centre and provided with the option to attend a 40 minute Zumba class at a cost of £2/client’, or ‘clients were provided with access to a supporting website, which included access to a user forum and behaviour change prompts to help support relapse prevention’.

8. Rationale for intervention (including theoretical basis): essential

It is essential to state the reasoning behind the design of the intervention and the methods that will be used. State the theories or scientific evidence on which the intervention is based. What is the theoretical or scientific basis that suggests the intervention will be successful in its aims and objectives? This could be peer reviewed research studies, NICE guidance on obesity interventions, theories about health promotion and behaviour change orthe PHE commissioning guidance.[2] If your intervention also includes an innovative component it is important to state what this is, and your rationale for including it eg insight work with the local target community informed adaptations to the dietary component of a previously used intervention, to make it more culturally appropriate.

9. Intervention staff and core competencies required: essential

Who is designing and delivering the intervention? For example, school nurse, community nutritionist, health trainer, health professional or teacher. What is their background, expertise and specific training and qualifications. What is their role in the intervention?

What are the core skills needed by everyone involved in delivering the intervention? For example, facilitation skills, cooking skills, experience of working with children, young people, or individuals with disabilities, communication skills, and basic knowledge of nutrition or physical activity.

What are the basic training requirements for the delivery of the intervention? Do they have to have a qualification in weight management, fitness training, psychology or nutrition? Do they have to be trained to a specific level, or be a member of a certain professional body? What intervention specific training is provided, and how is it delivered? Further information about training requirements can be found in the guidance for weight management intervention providers and commissioners.[2]

This information is important in helping others who may want to replicate the approach taken by this intervention.

10. Equipment and resources required: essential

Is a particular type of venue required? For example, one with a kitchen or gym. Are specific resources needed such as pictures of 'eat-well guides', specific physical activity equipment, cooking equipment, food and measurement tools such as pedometers and calibrated scales?.

11. Use of incentives: essential

Participant incentives: Have any participant incentives been provided to encourage individuals to take part in the intervention and, if so, what are they? Have incentives been provided for first attendance or completion of the intervention? If incentives have been used it is important to record their use and uptake as this may have an impact on the success of the intervention and the sustainability of any behaviour change.

Provider incentives: Some intervention commissioners may provide incentives to intervention providers such as payment based on attendance rates or results. It is also important to record these incentives which may provide insight into the completeness of findings.

12. Method of recruitment and referral: essential

How have participants been recruited to the intervention? What percentage of those that are eligible have been recruited? Has there been a referral process or was it self-referral? For example, have participants been referred by a GP or have leaflets and posters been used to advertise in GPs, social media or community centres so participants can sign up themselves? Please give brief details here of any sampling process that was undertaken, if applicable (ie if the intervention is delivered as part of a formal trial or pilot). Was there any targeting of particular groups by, for example, advertising the intervention in certain communities or at specific locations? The method

by which people have been recruited should be taken into account when carrying out the evaluation. For example, a self-referred group of participants may be more motivated (and ready to change) than participants referred by a professional, and thus may be more likely to start and complete the intervention.

13. Participant consent mechanism: essential

The appropriate mechanism for gaining participant consent must be considered. The nature of consent will vary for different groups of people. For example:

- those able to consent for themselves
- those with parental responsibility and consenting on behalf of a child or young person under the age of 16 years
- those who lack the capacity to consent.

Policy guidance on seeking consent from different groups and a wide range of consent forms are available from the Health Research Authority website.[20]

The Data Protection Act must be adhered to when collecting personal data from individuals. A data protection statement should be given to participants before any personally identifiable data is collected. It should explain exactly which personal data is being held, why, where it will be held, and who will have access to the data. This is particularly important when collecting sensitive data such as home postcode, ethnicity and socioeconomic status. More information about the requirements of the Data Protection Act can be found on the UK government website.[21]

14. Participant eligibility criteria: essential

Participants who have been referred or have self-referred should meet pre-defined criteria. For example, the target population may be adults aged 18 years and over from x local authority, with a BMI>/=30, but excluding women who are pregnant and participants with a BMI >40. It is helpful to provide an explanation for any exclusion criteria eg a local tier 3 intervention is available for populations with a BMI 40+, and the intervention has not been approved for use during pregnancy. Further information on eligibility criteria is provided in the guidance for weight management intervention providers and commissioners.[2]

15. Cost of intervention per participant: essential

This describes the cost of running the intervention and should be calculated by dividing the total cost of running the intervention by the number of participants who attended at least one session. This information is important for an economic analysis of whether or not the intervention is good value for money. It also enables commissioners to judge whether the resources required to run the intervention are available.

16. Cost to participant: desirable

It should be noted if participants are charged for any part of the intervention. Some interventions may also charge a small fee for any follow on interventions that are offered after the active intervention has reached completion. There may also be additional participant costs such as travel or child care which is important information to capture, as this could provide further insight into differences in attendance. Conducting interviews with participants may be helpful in exploring the impact of these broader participant costs.

17. Detailed breakdown of cost: desirable

A detailed breakdown of all intervention costs is important for a full economic analysis, in order to judge whether or not it is good value for money. Take into account costs incurred in planning stages as well as during the delivery stages. Some examples of input costs are staff time, transport, venue hire, equipment, publicity and incentives. It is especially important to factor in 'invisible' or 'in kind' costs. For example, a room in a local authority leisure centre may be hired free of charge as part of a partnership agreement with the local authority or Clinical Commissioning Group. However, these cost needs to be taken into account so that if the intervention is repeated, financial resources can be planned accurately.

18. Type of evaluation and evaluation design: essential

The way in which an evaluation is designed to collect data, and the method by which data may be analysed to measure impact, should be recorded here. For example, does the evaluation involve a pre and post design? Is there a control group or control population? Was formative research conducted to inform the development of the intervention? Does the evaluation use qualitative and/or quantitative data? This information is important as it will determine what inferences can be made about the evaluation findings.

See 'Evaluation of public health interventions: an introductory guide' for a more detailed explanation of evaluation designs.[5]

19. Details of equality impact assessment: essential

Public bodies have a duty to undertake equality impact assessments (EIA) under race, gender and disability equality legislation. It is important to provide an equality impact assessment as part of its overall evaluation. It can give valuable information if particular outcomes are seen in different groups, and can be useful in informing servie-redesign or commissioning activity. An equality impact assessment 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.'

The Equality Act 2010 made it unlawful[b] to discriminate against people with a 'protected characteristic'. These characteristics are:

- age
- disability
- gender reassignment
- marriage and civil partnership
- pregnancy and maternity
- race
- religion or belief
- sex
- sexual orientation

These characteristics therefore need to be taken into account in any equality impact assessment. Information and guidance on the Equality Act 2010, including age discrimination and public sector Equality Duty, can be found on the .gov website: https://www.gov.uk/guidance/equality-act-2010-guidance.

20. Relevant policy and performance context: desirable

It may be useful to show how an intervention fits into any strategic policies, or whether it is a priority intervention as outlined in, for example, a joint strategic needs assessment.[22]

For example, 3 national indicators directly relating to obesity and overweight prevalence are included in: the Pubic Heath Outcomes Framework:[23]

- excess weight in 4-5 year olds
- excess weight in 10-11 year olds
- excess weight in adults

It is also important to demonstrate alignment with the PHE guidance for weight management commissioners and providers.[2]

[b] Unlawful behaviour is: *"Under the Act people are not allowed to discriminate, harass or victimise another person because they have any of the protected characteristics. There is also protection against discrimination where someone is perceived to have one of the protected characteristics or where they are associated with someone who has a protected characteristic."* https://www.equalityhumanrights.com/en/equality-act/equality-act-faqs [bullet 3]

21. Details of health needs assessments that have been conducted: desirable

Whilst both commissioners and providers should consider undertaking Equality Impact Assessments to ensure that the needs of protected characteristics are considered. In addition, local areas should consider undertaking health equity audits of intervention provision to identify areas in which interventions may not be equitable.

Has a health needs assessment been conducted that identifies a gap in this intervention being provided for the target population? Information may come from a specific needs assessment conducted for the intervention, or it may be available from other sources. For example, data relating to health inequalities and gaps in intervention provision may already be available from policy documents such as the Joint Strategic Needs Assessments (JSNAs) or Children and Young People's Plans.

If information is not readily available from these documents, other existing datasets may be helpful:

- Public Health Outcomes Framework: https://www.gov.uk/government/collections/public-health-outcomes-framework
- National Child Measurement Programme: http://content.digital.nhs.uk/ncmp
- Health Survey for England: http://content.digital.nhs.uk/healthsurveyengland
- Active Lives Survey: http://www.sportengland.org/research/active-lives-survey/
- National Diet and Nutrition Survey: https://www.gov.uk/government/collections/national-diet-and-nutrition-survey
- Health Profiles: https://fingertips.phe.org.uk/profile/health-profiles
- Census: https://www.ons.gov.uk/census/2011census

When using data to identify gaps in intervention provision and to justify resource allocation, it is important to assess the quality of the data being used. For example, how robust is the data at the geographical level at which you wish to use it? How old is the data? How well validated is the tool used to collect the data? If it is estimated data, how has it been modelled and how accurate an estimate is it likely to be? A more robust approach could be to use findings from a number of different data sources and support these by carrying out local research. This could include the use of local health and well-being questionnaires, focus groups or face-to-face interviews with the target population or community.

22. Contact details: essential

Give a list of the key people involved in the intervention planning, delivery and evaluation. This should include all contact details, and details of staff positions, as staff may change jobs during the course of the intervention.

23. Commissioner(s) of the intervention and sources of funding: desirable

Where has the funding come from to commission the intervention, and who has commissioned it? For example, ‘the funding has come from the core public health grant, and the intervention has been commissioned by the public health team’.

24. Declaration of interest: desirable

This covers any potential conflicts of interest in carrying out the intervention and evaluation. This is particularly important if the evaluation is funded by an agency that could be perceived as wanting to influence 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 a non-pecuniary interest (where someone may have publicly expressed a clear opinion on the intervention in question, and this may influence their impartiality).[24]

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 or evaluation should not go ahead as planned; it may be acceptable to state how potential conflicts are going to be avoided.

25. Details of type and extent of any clinical involvement: desirable

Will any clinicians be involved at any stage of the intervention? This includes during development, delivery and carrying out quality assurance of the delivery. In some cases it may be appropriate for the provider to inform GPs that their patients are participating in an intervention.

Part 2: demographics of individual participants

A PHE data capture tool is provided to collect participant age, sex, disability, home postcode (which can be converted into IMD), ethnicity, religion, sexual orientation, employment status, marital status, housing tenure, number of children, service referrals, smoking status and known co-morbidities (eg Cardiovascular disease, Dyslipidaemia, Hypertension ,Osteoarthritis, Sleep apnoea, Type 2 diabetes, Generalised anxiety disorder, Depression, Chronic back pain).[2]

26. Age: essential

It is essential to record the age of all participants in the intervention.

Examples of age categories from national population surveillance studies that could be used for comparative purposes are:

- Census 2011 used: 0-4, 5-9, 10-14, 15-19, 16 or 19-24, 25-34, 35-44, 45-54, 55-64, 65-74 and 75+ years
- Health Survey for England uses: 2-10, 11-15 and 2-15 and adults are classed as 16 years and over
- National Child Measurement Programme uses: children in Reception year (aged 4-5 years) and in Year 6 (aged 10-11 years)

27. Sex: essential

Record the sex of all participants. This is useful for identifying whether or not the intervention tends to be more effective for males or females, and in assessing whether the intervention is appealing and accessible to both sexes.

28. Ethnicity: essential

It is standard practice in healthcare 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, it is still important information for raising understanding about the extent to which uptake and response to the intervention may vary between different ethnic groups. For example, if the intervention is aimed at women aged 45-55 years in a local community, of which 25% of the female population are of Bangladeshi ethnicity, and less than 2% of the participants are Bangladeshi, it can be concluded that there is something about either the intervention itself or the publicity for the intervention that is not engaging with the Bangladeshi women. In this case, further research and community development work may be needed to engage with these communities and the intervention or publicity amended accordingly. This is particularly important as there are specific ethnic groups with a higher than average prevalence of obesity.[25] The PHE key performance indicator guide for adult tier 2 services stipulates a locally defined target to ensure the needs of high risk groups are met.[2]

Under the Public Sector Equality duty set out in the Equalities Act 2010, public bodies are required to analyse the effect of their organisation's functions on all protected groups. It requires equality considerations to be reflected into the design of policies and the delivery of services, including internal policies, and for these issues to be kept under review. Public authorities will therefore not be able to meet the duty unless they have enough usable information. Ethnic monitoring demonstrates that policies for equality are

working in practice. It is a way of identifying potential discrimination and whether policies promoting equality of opportunity and good relations between different racial groups are being implemented.[26]

It is recommended that public authorities and their partners use the following Census 2011 categories for ethnic monitoring in England and Wales:[27]

White:

- English/Welsh/Scottish/Northern Irish/British
- Irish
- Gypsy or Irish Traveller
- any other White background, please describe

Mixed/Multiple ethnic groups:

- White and Black Caribbean
- White and Black African
- White and Asian
- any other Mixed/Multiple ethnic background, please describe

Asian/Asian British:

- Indian
- Pakistani
- Bangladeshi
- Chinese
- any other Asian background, please describe

Black/ African/Caribbean/Black British:

- African
- Caribbean
- any other Black/African/Caribbean background, please describe

Other ethnic group:

- Arab
- Any other ethnic group, please describe

29. Disability: essential

The Equality Act 2010 defines disability as a physical or mental impairment that has a 'substantial' and 'long-term' negative effect on your ability to do normal daily activities'.[28]

The Public Sector Equality Duty (PSED) was introduced in April 2011 as part of the Equality Act 2010.[26] At the heart of the PSED is the requirement that public bodies must have due regard to the need to:

- eliminate unlawful discrimination
- advance equality of opportunity
- foster good relations

In the context of obesity, monitoring of disability is important as physical or learning difficulties in adults and children can confound the outcome of a weight management intervention. There is increasing evidence to suggest that many disabling conditions such as arthritis, mental health disorders, learning difficulties and back ailments increase the risk of obesity.[29] Children and adults with learning or physical difficulties may be at a higher risk of obesity. This may be related to genetic or metabolic complications associated with a particular disorder, use of medications with a side-effect of weight gain or practical issues related to inappropriate eating or physical activity.[30]

It is important to address health inequalities by tailoring interventions and information for particular groups. This may be relevant for individuals with obesity and disabilities, whose needs are different from that of standard practice. NICE offers detailed guidance on the identification, assessment and management of obesity including for people with a condition associated with an increased risk of obesity.[31]

30. Measure of socioeconomic status: essential

An indicator of socioeconomic status should be recorded. There is evidence associating lower socioeconomic status with the likelihood of overweight and obesity in children and adults. Evidence from the Health Survey for England (2012) shows that mean BMI varied significantly by equivalised income quintile. Mean BMI was highest among those in the fourth lowest income quintile (27.6kg/m² in men, 28.3kg/m² in women) and lowest among those in the highest income quintile (27.2kg/m² for men and 26.0kg/m² for women).[32] However the relationship between obesity and income was found to be different for men and women in the 2012 survey. Among women the prevalence of obesity was highest in the bottom two income quintiles. Among men there was no significant difference in the prevalence of obesity or overweight.

Evidence from the National Child Measurement Programme also shows a clear relationship between childhood obesity and low socio-economic status with more than

twice as many children with obesity in the most deprived areas, when compared to the most affluent areas.[33]

There is currently only a limited understanding of the reasons for this association, making it important to build evidence of effectiveness among different socioeconomic groups. It is also useful to monitor uptake of intervention by different socioeconomic groups, to ensure they are not systematically excluding any groups through their design, delivery, recruitment or referral methods.

Another indicator of socio-economic status that is often used in public health is the Index of Multiple Deprivation, which can be assigned from a home postcode.

The Index of Multiple Deprivation (IMD) combines a number of indicators covering a range of economic, social and housing issues and creates a single deprivation score for each small area in England. This allows areas to be ranked according to their level of deprivation and can be derived from postcodes. These rankings have been produced at Lower Layer Super Output Area level (LSOA), of which there are 32,482 in England.[34] LSOAs can be mapped against postcode which allows an individual's address to be given a general IMD ranking. Any ranking given is 'modelled' against a number of criteria and relates to an overall ranking for an area which may not necessarily be indicative of the characteristics of an individual household. Further information about IMD is available on the Communities and Local Government website.[35]

Interventions for children should, where possible, collect this information about their parents or carers.
Analysis of socio-economic data can be complex so 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. Religion: essential

Religion can be recorded using the following census criteria:

- Christian
- Buddhist
- Hindu
- Jewish
- Muslim
- Sikh
- other religion
- Nno religion
- person prefers not to say

Collecting data on religion may help in terms of improving understanding about the extent to which uptake and response to the weight management interventions may vary between different religious groups, and whether adaptations to dietary or physical activity components are required to accommodate different religious practices.

32. Employment status: essential

Participant employment status (or parental/carer employment status in child interventions) is an important consideration as it will help plan intervention timetables to ensure they fit within participants working patterns. There is also evidence to show that some occupations that involve shift work, may present additional barriers to successful weight loss.[36]

Employment status is also an important standard indicator of socioeconomic status used in the Census 2011. The National Statistics Socio-economic Classification (NS-SEC) is a structured occupationally-based classification that also includes categories for the non-employed.

The NS-SEC categories are:

1. Higher managerial, administrative and professional occupations
   1.1 Large employers and higher managerial and administrative occupations
   1.2 Higher professional occupations
2. Lower managerial, administrative and professional occupations
3. Intermediate occupations
4. Small employers and own account workers
5. Lower supervisory and technical occupations
6. Semi-routine occupations
7. Routine occupations
8. Never worked and long-term unemployed
9. Not classified (this includes Students; Occupations not classified or inadequately described; and Not classifiable for other reasons)

Further information on these categories and how they have been derived is available on the Office for National Statistics (ONS) website, and an online tool is available to assist with correct coding.[37]

33. Sexual Orientation: essential

As a protected characteristic it is important to record sexual orientation, using the following categories:

- heterosexual or straight
- gay or lesbian

- bisexual
- other not listed
- person does not know or is not sure
- person prefers not to say

34. Known co-morbidities: essential

Co-morbidities known to be associated with obesity include:

- cardiovascular disease
- dyslipidaemia
- hypertension
- osteoarthritis
- sleep apnoea
- type 2 diabetes
- generalised anxiety disorder
- depression
- chronic back pain.

Recording known co-morbidities is important for 2 reasons:

1. There is evidence to suggest that weight management may be more difficult for individuals with certain co-morbidities such as diabetes.[38]
2. It is important to ensure that where appropriate, reasonable adjustments are made to accommodate the additional needs of participants with co-morbidities eg delivery staff are appropriately trained, and materials and resources adapted.

35. Additional Information: desirable

The suggested information below may be helpful in determining why an intervention may be more or less successful in some individuals than others. The PHE weight management data capture tools provide options to collect additional data on housing tenure, marital status, smoking status, number of children, parental weight status (which is deemed essential for child interventions), and referral to other services.[17,18] Any relevant supplementary fields can also be added. .

Medical history
This can indicate confounders such as participants with an existing clinical disorder that is could influence weight status such as thyroid dysfunction or disordered eating. It may also include: participants who have undergone bariatric surgery, taking bariatric medication (such as orlistat), or medications known to be associated with weight gain (such as steroids and some anti-psychotics); and noting those patients who are at high risk of developing diabetes.

Smoking status*:*
There is a relation between smoking and obesity. A recent large study found that while current smokers were less likely to have obesity than those who had never smoked, former smokers were more likely to have obesity than both current smokers and those who had never smoked. Among smokers, the risk of obesity increased with the amount smoked and former heavy smokers were more likely to have obesity than former light smokers.[39]

Marital Status
Marital status is a 'protected characteristic' so it is illegal to discriminate against anybody on the basis of their marital status, which can be recorded using the following census criteria:

- single
- married/civil partner
- divorced/person whose civil partnership has been dissolved
- widowed/surviving civil partner
- separated

This is a useful indicator as there is a significant amount of evidence indicating that the BMI of married people tend to be correlated.[40] There is also some evidence that people gain weight when married and lose weight when no longer married.[41] It is therefore useful not only to collect data on whether the subject is married (or cohabitating) but also the weight and height of the subject's partner.

Alcohol intake
Whilst the relationship between alcohol intake and obesity is complex, there is an association between alcohol and obesity, as often individuals are unware of the caloric content of alcohol, and alcohol can also increase food intake.[42] It may therefore be an important confounder in any weight loss intervention.

Pregnancy, breastfeeding and number of children:
There is evidence to indicate that the incidence of maternal obesity at the start of pregnancy is increasing, with the incidence of obesity rising with number of pregnancies.[43] Pregnant and breast feeding women also have additional nutritional and care requirements. It is therefore important to know whether participating women become pregnant, as NICE do not currently recommend weight loss interventions during pregnancy.[31] It may be sensible to exclude any women that fall pregnant during a weight management intervention to avoid skewing the data collected.

Parental weight status (essential for child interventions)
A number of studies have shown children are more likely to be overweight or have obesity if one or both of their parents are overweight or has obesity.[44] Therefore collecting parental weight status can be a very useful indicator when evaluating child weight management interventions, and is therefore a required field in the PHE child data collection template. However, it is acknowledged that whilst useful, may not always be possible to record if parents do not wish to be measured.

Referral to other services
Data on referral to other services may provide an indication of the number of participants with complex circumferences such as families in crisis, participants with social care requirements, and smoking, drinking or substance mis-use where this data is not otherwise collected. Referral to these services may suggest additional challenges in the lives of these participants that may present additional barriers to their weight loss attempts.

Housing tenure
Housing tenure can provide an alternative indicator of socio-economic status and can be collected using the Census categories are as follows:

- owner occupied: owned outright
- owner occupied: owned with mortgage or loan
- owner occupied: shared ownership
- social rented: rented from council
- social rented: other social rented
- private rented: private landlord or letting agency
- private rented: employer or a household member
- private rented: relative or friend of a household member
- private rented: other
- living rent-free

Part 3: baseline data

For all weight management interventions, the key long-term health outcome will be to reduce the risk of health-related conditions associated with obesity which can lead to premature mortality and obesity-related morbidity. Given that it is impractical to track very long-term health outcomes, it may be necessary to use intermediate or shorter-term health outcomes or markers such as changes in food choices, increased physical activity, decreased sedentary behaviour, and sustained improvements in weight status.

It is extremely important to collect baseline data before an intervention begins. It is quite common for health improvement/promotion projects to begin without collecting any

baseline data. This creates a significant challenge for any evaluation as it is extremely difficult to look back and collect data with any confidence.

36. Height and weight (to calculate BMI/BMI-z/centile): essential

It is essential to collect the height and weight of participants in order to calculate their BMI. Even if changing BMI is not the primary objective of the intervention, or BMI is not expected to change in the short term, height and weight should be measured throughout any weight management intervention to assess any changes in BMI in the long term.

BMI is defined as the person's weight in kilograms divided by the square of their height in metres. It is one of the most common ways of measuring a person's weight status. It is used because, for most people, it correlates with their proportion of body fat.[45] It is also a relatively easy, cheap and non-intrusive method for establishing weight status. However, BMI only provides an indication of body fatness and other factors such as fitness (muscle mass), ethnic origin and puberty can alter the relationship between BMI and body fatness, and must therefore be taken into consideration. There are also additional measurements such as waist circumference and skin thickness which can be collected. These help to provide an indication of body fat distribution, which is not provided by BMI.

In the UK weight status is classified using the following BMI categories:

- underweight: <18.5kg/m2
- healthy weight: 18.5–24.9kg/m2
- overweight: 25–29.9kg/m2
- obesity I: 30–34.9kg/m2
- obesity II: 35–39.9kg/m2
- obesity III: 40kg/m2 or more.

However, the use of lower BMI thresholds (23kg/m^2 to indicate increased risk and 27.5kg/m^2 to indicate high risk) have been recommended to trigger action to reduce the risk of conditions such as type 2 diabetes in Black African, African-Caribbean and Asian (South Asian and Chinese) groups.[31]

For all adults, 18 years and over, the BMI ranges for underweight, healthy weight, overweight and obesity do not change with age and are the same for both men and women. However, because children and teenagers are still growing, BMI as an indicator of adiposity is different for boys and girls, and changes with age. It is therefore not possible to use the adult BMI categories for young people under 18 years old. In the UK, the UK 1990 growth charts[46] are used to calculate the weight status of children and

young people between two and 18 years old, and UK-WHO growth charts for children 0-4 years.[c]

The UK 1990 growth charts were compiled from measurements on boys and girls collected during 11 British surveys carried out between 1978 and 1990. They show the growth patterns of these UK children. The UK-WHO growth charts are designed using data from the new WHO standards, and describe the optimal pattern of growth for all young children, setting breast feeding as the norm.

A child's BMI and age can be plotted on these charts to work out the BMI centile. The centile is the most commonly used method of interpreting an individual child's BMI and indicates the relative position of the child's BMI as compared with the reference population of children of the same age and sex. For example, a child with a BMI centile of 95 has a BMI higher than 95% of children of the same age and sex in the reference population, and only 5% of children of the same age and sex in the reference population have a higher BMI. If a child's BMI is lower than the second centile, it means that only 2% of children from the reference population have a lower BMI than they do.

However whilst centiles are generally used within clinical practice, for research often a BMI-z or standard deviation score is used instead of a centile (for more information on this see the final section below on data analysis).

In the UK weight status in children is classified as follows:

Weight status classification	**For population monitoring (eg national reports)**	**For clinical practice (eg when feeding back to parents)**
underweight	≤2nd centile	≤2nd centile
healthy weight	>2nd centile to <85th centile	>2nd centile to <91st centile
overweight	≥85th centile to <95th centile	≥91st centile to <98th centile
obesity	≥95th centile	≥98th centile
severe obesity	≥99.6th centile	≥99.6th centile

The UK90 BMI reference is available on printed growth charts, where the centiles are shown evenly spaced at 2 thirds of a standard deviation. This means the 0.4th, 2nd, 9th, 25th, 50th, 75th, 91st, 98th and 99.6th centiles are shown.

Information on how to purchase the full UK90 growth charts and a BMI calculator for individual height and weight measurements is available on the Harlow Healthcare website.[47] When calculating the BMI of a group of children it is possible to use a free

[c] https://www.rcpch.ac.uk/Research/UK-WHO-Growth-Charts

excel plug in: LMS growth, which can be downloaded here:
http://www.healthforallchildren.com/?product=lmsgrowth

More information about measuring BMI in children can be found in this simple guide:
http://webarchive.nationalarchives.gov.uk/20170110173352/http://www.noo.org.uk/uploads/doc/vid_11762_classifyingBMIinchildren.pdf

It is important to ensure that standardised, easily reproducible protocols are used for measuring height and weight so that inaccuracies and inconsistencies are minimised. Ethical considerations should also be taken into account, particularly when weighing and measuring children. The National Child Measurement Programme (NCMP) and the Health Survey for England (HSE) have established basic standards for these procedures including using standardised weighing and measuring equipment, and ensuring the consistent posture and head positioning of participants when measuring height.[48] The key points about weighing and measuring from these surveillance programmes are summarised below.

Chair-bound participants and those with difficulty standing
If a participant is too unsteady on their feet for these measurements, do not attempt to take them. If they find it painful to stand or stand straight, do not attempt to measure height.

Rollameters,[d] sling scales, infant scales or chair scales can be used to measure length and weight of very young infants or children who may have difficulty standing.

Site
It is strongly advised, if possible, to measure height and weight on a floor which is level and not carpeted.

Equipment
Portable stadiometers are advised. This is a collapsible device with a vertical rule and a sliding head plate or paddle.

General points about measuring adult and child height:

- take measurement without shoes on
- check the stadiometer is assembled correctly
- the participant should stand with their feet flat against the base with their back as straight as possible against the rule but not leaning on it. They should be facing forwards with their arms at their sides
- move the participant's head so that the Frankfurt Plane is in a horizontal position (The Frankfurt Plane is an imaginary line passing through the external ear canal and

[d] http://www.healthforallchildren.com/product-category/shop/heightlength-measuring-equipment/.

across the top of the lower bone of the eye socket, immediately under the eye) – use the diagram in the NCMP operational guidance to identify this.[48] This position is important for an accurate reading. An additional check is to ensure that the measuring arm rests on the crown of the head, that is the top back half.
- if the measurement has been done correctly, the participant should be able to step off the stadiometer without ducking
- height should be recorded in centimetres to the first decimal place eg 120.3cm (it is important not to round up or down)

General points for weighing adults and children:

- participant should remove their shoes and any heavy clothing, outer garments or heavy jewellery
- the participant should stand with their feet together in the centre of the scales
- ensure calibrated scales are used
- measure weight in kilograms, to the first decimal place eg 20.6kg (it is important not to round up or down)

Note: it is important that all weight and height data is measured not self-reported given there are known biases in the return of self-reported data, where height is often overestimated and weight often underestimated.[49]

37. Additional proxy measures for adiposity: desirable

There are a number of other proxy measures for adiposity or body fatness. Alternative estimates calculated from height and weight are the Ponderal or Rohrer's Index (which is weight/height^3) or Benn's Index (weight/height^p where p is determined by analysis of the population being investigated). Each have their own advantages and disadvantages, and they are sometimes used in specific instances, such as for newborns or in research studies, but by far the most common anthropometric measure for published prevalence figures is BMI.

Waist circumference measurement can be used to assess a patient's abdominal fat content or 'central' fat distribution. Central obesity is linked to a raised risk of type 2 diabetes and coronary heart disease. NICE recommends that waist circumference may be used, in addition to BMI, to assess risk in adults with a BMI of less than 35kg/m^2.[50] Although the first UK waist circumference charts for children can be found on the reverse of the growth charts provided by Harlow printing,[e] currently NICE do not recommend using waist circumference as a routine measure in children, although it can be used to give additional information on the risk of developing other long term health problems.[31]

[e] http://www.healthforallchildren.com/shop-base/shop/growth-charts/uk90-bmi-identification-charts/

Adult waist circumference cut points are defined as:

- high waist circumference: 94-102cm in men and 80-88cm in women
- very high waist circumference: >102cm in men and >88cm in women

Waist-hip ratio is another measure of body fat distribution. It is defined as waist circumference (in metres) divided by hip circumference (in metres). There is no consensus about appropriate waist–hip ratio thresholds. However, a raised waist–hip ratio is commonly taken to be 1.0 or more in men and 0.85 or more in women. NICE does not recommend the use of waist–hip ratio as a standard measure of overweight or obesity.[31]

Skin fold thickness-measures, using skin fold callipers on multiple sites, can be used to estimate the subcutaneous fat mass and distribution. This method requires some basic training and is more invasive than other measures, but the equipment is relatively inexpensive and portable. However, as with all anthropometric measurements, the possibility of measurement error is high without standardised equipment, measurement protocols and training. In addition, prediction models need to be used in order to derive percentage body fat value.[51]

Bio-impedance is an indirect measure of lean tissue mass, estimated by passing a low-level electrical current through the body and measuring voltage changes given lean tissue mass conducts and fat mass insulates.[52] There are a variety of equations for different instruments and for different age groups that can be used to estimate body composition. Like waist to hip ratio, it is not recommended that bio-impedance is used in isolation.

38. Measure of wellbeing: essential

Wellbeing is an important indicator for any weight management intervention, as it is critical in ensuring the intervention has a beneficial, not detrimental impact on participants' lives. Quality of life has been shown to be related to obesity, and reductions in obesity are related to improvements in health-related quality of life.[53] There are a number of validated tools that can be used to measure wellbeing. Measurement tools recommended in the PHE data capture tool are as follows.

Wellbeing assessment tools for adults:

- Warwick-Edinburgh Mental Well-being scale (WEMWBS): A 14 item scale with 5 response categories
- Short Warwick-Edinburgh Mental Well-being scale (SWEMWBS)

- SF-36: The Short Form-36 (SF-36): A 36 item questionnaire which measures Quality of Life (QoL) across eight domains
- EQ-5D: EQ-5D standardised health questionnaire: A simple descriptive profile and a single index value for health status

Guidance on measuring the wellbeing of children and young people is provided by Public Health England (2015).[54]

39. Measure(s) of dietary intake and behaviour: desirable

Obesity is caused by an imbalance between energy intake and energy expenditure. NICE guidance recommends that validated methods should be used to estimate dietary intake and physical activity. In principle, all instruments used to assess behaviour indirectly should be tested for reliability and validity. In this context, reliability refers to how likely it is that the same thing will be measured, each time it is used, either within the same or a different participant. Validity refers to how accurately the measurement reflects the actual behaviour. An instrument can therefore be reliable but not valid as it could measure the wrong thing, but do so consistently.[9] There are a number of ways of testing for reliability and validity.

Ideally, overall energy intake, broken down by food types would be measured during the active intervention and at follow-up. However, dietary intake is very complex and difficult to measure.[55] The 'gold standard' objective measure of energy intake is the 'doubly labelled water' method.[56] However these measurements are impractical and expensive for interventions outside a research setting, and are usually used as a validation method for other, indirect but more practical measures.

As a result, dietary intake measurement frequently relies on self-report instruments. Self-reporting can result in recall errors, over- and under-reporting of portion sizes and dishonest reporting. Through a phenomenon known as 'social desirability bias', participants are more aware of behaviours promoted by healthy lifestyle campaigns in the media, and report their behaviour to be more aligned with these messages, particularly for behaviours such as drinking alcohol.[57] This can be unconsciously or consciously done. There is also evidence from doubly labelled water validated studies to show that under-reporting is more common and most severe in subjects with obesity compared to those of normal weight.[58]

Self-report instruments to measure dietary intake fall into two categories; those that measure current diet such as 24-hour recall and food diaries; and those that assess food groups or dietary habits such as food frequency questionnaires (FFQs). The most reliable of these instruments is a 7-day weighed food diary. However these are time consuming to the participant and there is evidence to show that there tends to be under-reporting.[59] They also require analytical expertise and software to calculate the nutrient

and energy intake from diary information. Some studies have investigated the use of web-based or electronic tools to collect dietary intake data more accurately, particularly with children. One example a web program that uses a 24-hour recall methodology is intake24: https://intake24.co.uk/. Other studies have used mobile phone technology, online or 'app' based assessments to record dietary intake.[60]

Interventions should aim to measure the aspect of the diet that is the focus of the active intervention. For example, if the active intervention is focusing on reducing consumption of fatty foods, then it would be important to measure fat consumption at baseline and follow-up. However, most interventions are likely to focus on reducing overall energy consumption, and measurement of this is highly problematic. A review by the East Midlands Public Health Observatory[61] and further consultation during the development of this document have indicated that there are currently no validated, simple instruments available for measuring dietary intake that are not onerous or time consuming for the participant; or that do not require highly specialised analysis. This therefore remains a significant research gap: to ensure
that validated tools are available for collecting dietary intake data that are easy to use, relevant and indicative of improvements in weight status.

Some interventions may focus on increasing fruit and vegetable intake. This can act as a rudimentary indication of an individual's diet, though practitioners will wish to apply alongside other lifestyle measures to assess risk. There are some tools available that provide a measure of fruit and vegetable consumption, including the FACET [62] and CADET[63,64] tools.

However, when selecting any measurement assessment tool it is important to assess:

- does it measure the outcome you are interested in?
- how reliable is it?
- has it been validated in the population you are evaluating?
- has it been validated to measure change over time?

In some cases it might be useful and appropriate to collect information about dietary behaviour and eating patterns. Unhealthy meal patterns and snacking behaviour have been associated with obesity.[65] Questionnaires such as the Golan Family Eating and Physical Activity Habits questionnaire[66] attempts to understand whether eating patterns are, in fact as a result of hunger.

40. Measure(s) of physical activity levels and behaviour: desirable

If an intervention aims to increase physical activity as part of its approach to preventing or reducing obesity, then a measure of physical activity is an important aspect of the evaluation. Appropriate tools for measuring physical activity are outlined below.

Physical activity can be measured either by using objective methods or by using a questionnaire or other self-assessment tools. Objective measurements of activity include measures using motion sensors (which measure body movement in two or three planes) or pedometers (which measure number of steps taken). These can be used to calculate total energy expenditure over a day or week. Such instruments have the advantage of being well-validated and are generally accurate which overcomes some of the problems of recall bias found in surveys.

Issues to consider when using objective methods to measure physical activity include the storage of data, and the importance of using only tools that have been well-validated. Cheaper pedometers may be useful as motivational or promotional devices but they are unlikely to provide accurate measures of physical activity.[67]

More expensive motion sensors allow data to be uploaded to a computer, sometimes remotely. 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.[68] However, objective measures can be costly to use at a population level, and do not always provide data at the appropriate level of detail. For example, motion sensors cannot provide data on the mode of activity. For this reason, the focus here is on using self-assessment methods.

The former National Obesity Observatory (now PHE) conducted a rapid review on the most appropriate tools for measuring physical activity (and diet) in local level weight management interventions. Details of the recommended tools are available in a supplement to this review.[69]

41. Potential facilitators of, and barriers to, lifestyle change: desirable

Examples of potential facilitators of, and barriers to, lifestyle change are:

- health literacy: health-related knowledge, attitude, motivation, behavioural intentions, personal skills and self-efficacy[9]
- social action and influence such as social norms and public opinion
- other environmental influence such as workplace and home
- other psychosocial outcomes such as self-esteem, confidence, body image, positivity and improved quality of life

A wide range of theories, concepts and accounts of behaviour and behaviour change may be considered when evaluating the intervention. Psychosocial models of behaviour change allow determinants such as socioeconomic and cultural context, physical environment, attitudes, knowledge, self-perception and social norms to be identified and taken into account at baseline and follow-up. This type of information may help to tailor

particular aspects of the intervention, or to explain why the intervention has worked in the way it has for certain individuals. This approach is supported by NICE regarding behaviour change at population, community and individual levels.[70] NICE suggest planning and prioritising interventions that are 'tailored to tackle the individual beliefs, attitudes, intentions, skills and knowledge associated with the target behaviours.' In addition, psychosocial measures can be useful indications of the likelihood of positive behaviour changes being sustained beyond follow-up. For example, there is evidence to show that low self-esteem in children with obesity is associated with a lack of desire to take part in physical activity. In addition, a child's perception of their competence to take part in physical activity will also have an impact on their motivation to do so.[71] A measure of psychological wellbeing can ensure there are no adverse effects from the intervention on their wellbeing, for example as a result of stigmatisation.

A comprehensive guide to behaviour change techniques for adult tier 2 weight management has been provided by PHE as part of the commission and provider guide.[2] The guide provides relevant, evidence-based behaviour change techniques recommended for inclusion in weight management interventions for weight loss and weight loss maintenance.

Where the resource and expertise are available, it can be helpful to explore barriers and facilitators to change using qualitative methodologies, such as focus groups or interviews. This can help to provide understanding as to why the intervention may work better for some participants than others, and can provide useful insights to help inform future intervention developments and delivery.

Part 4: follow-up data

Whilst it is acknowledged that collecting long term data from weight management participants can be challenging, particularly when engaging with participants who may have struggled to maintain their weight loss, from an evaluation perspective this data collection is critical. At the end of the active intervention, impact is likely to be greatest. To measure only at this point might give an artificially inflated indication of the intervention's impact. Evidence shows that many people regain weight following an apparently successful intervention.[72] This is related to the statistical issue known as regression to the mean.[73] Behaviour change needs to be sustained to have health impacts. It is important to see if behaviours that change at six and 12 months after the active intervention has finished are maintained in the longer term.

It is important that, if practicable, procedures are put in place to ensure that follow-up height and weight data is measured and not self-reported. This is to avoid potential bias and also important when considering the impact of recall bias, ie how many of the original sample returned for a follow-up measure and were these likely to be just those individuals who were able to maintain their weight loss or continue to lose weight?

42. Follow-up data on height, weight (BMI): essential

To be collected at:

- the end of the active intervention,
- 6 months (post active intervention completion)
- 12 months (post active intervention completion)

43. Follow-up data on height, weight (BMI) collected at 24 months: desirable

44. Follow-up data on wellbeing: essential

To be collected at:

- the end of the active intervention,
- 6 months (post active intervention completion)
- 12 months (post active intervention completion)

45. Follow-up data on wellbeing collected at 24 months: desirable

46. Follow-up data on additional measures of adiposity: desirable

To be collected at:

- the end of the active intervention,
- 6 months (post active intervention completion)
- 12 months (post active intervention completion)
- 24 months (post active intervention completion)

If collected at baseline.

47. Follow-up data on dietary intake and behaviours: desirable

To be collected at:

- the end of the active intervention,
- 6 months (post active intervention completion)
- 12 months (post active intervention completion)
- 24 months (post active intervention completion)

If collected at baseline.

48. Follow-up data on physical activity levels: desirable

To be collected at:

- the end of the active intervention,
- 6 months (post active intervention completion)
- 12 months (post active intervention completion)
- 24 months (post active intervention completion)

If collected at baseline.

49. Follow-up measures on potential facilitators of and barriers to lifestyle change (if collected at baseline): desirable

If collected at baseline these measures should also be collected at the end of the active intervention, and again at 6 and 12 months after completing the active intervention (ideally for longer if possible).

Process evaluation

50. Number referred and referral source: essential

A critical aspect of evaluating an intervention is to determine the flow of participants. How many individuals were referred to the intervention and where were they were referred from? How many were repeat referrals?

51. Number recruited: essential

It is important to know that of those who were referred to the intervention, how many of these enrolled (ie attended at least one session)?

52. Number attended each session or contact point: essential

How many participants attended each exposure, episode, session or contact point? For example, if the active intervention is run twice a week for 10 weeks, how many participants attended each of these sessions?

53. Number completed: essential

How many participants completed the active intervention? The definition of what is meant by ‘completed’, should be consistent to understand whether those who complete the active intervention are more likely to benefit compared to those with poor attendance or who do not complete. PHE advise using the definition of ‘completion’ based on 75% attendance, as used in previous trials.[2]

54. Number of participants at each follow-up point: essential

How many participants attended each of the follow-up points and had the required data collected?

All of the above data should ideally be analysed on an ‘intention to treat’ basis. This means that the attendance and number completed should be calculated as a percentage of all the people initially referred or invited to the intervention. This removes the bias inherent in evaluations that present only weight loss among those who completed the intervention.

55. Reasons for dropout: desirable

While this information is not always the easiest to gather, it is vitally important to understand why participants might drop out of a intervention. This is particularly useful if the active intervention is going to be run more than once as part of a rolling intervention. This sort of information can be collected in a number of ways and is similar to gathering information on participants’ overall satisfaction with the intervention (see below). The difference here is that participants may have to be contacted directly. This needs to be handled sensitively so they do not feel like they are being chased. It is advisable to let participants know when they enter the intervention that, if they choose to leave, they will be contacted for feedback which will be used to improve how the intervention is delivered in the future. It is especially helpful to collect information about the demographics of people who have left the intervention, to investigate whether it is contributing to health inequalities.

56. Details of any unexpected outcomes and/or adverse events: essential

Above all it is important to ensure that any intervention does no harm, it is therefore essential that any unexpected outcomes or adverse events are systematically recorded and reported in the intervention evaluation.

Were there any unexpected side effects or outcomes from the active intervention? For example, did participants gain weight or did they have lower self-esteem by the end of the active intervention?

Unexpected outcomes, however, do not necessarily have to be negative, and there can be unexpected positive health outcomes.

57. Participants' satisfaction with the intervention: essential

All participants should be provided with the opportunity to feedback their satisfaction with the intervention. Satisfaction questionnaires are frequently used as part of evaluations, they can be a bespoke locally defined questionnaire, or standard questionnaires such as the NHS friends and family test: https://www.england.nhs.uk/ourwork/pe/fft/ . If participants are dissatisfied with the way in which an intervention is being delivered, or unhappy with an element of the overall design of the intervention, it is unlikely they will continue to attend. Consequently, the intervention is less likely to achieve its defined outcomes.

When undertaking research into 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 of research such as focus groups or semi-structured interviews. It may also be advisable for the deliverer of the intervention not to carry out the research. Participants may feel more able to be honest with another person whom they have not previously encountered as part of the delivery team.

These issues also apply to people who have dropped out of the intervention (see above). In many cases, interviews with such people are more likely to provide useful information about intervention improvement than talking to people who have completed successfully.

58. Plans for sustainability: essential

Consider whether plans have been made to ensure participants are offered the opportunity to continue with the intervention in some way. This will help the intervention's effect to be sustained over time. There may be resource or logistical implications for this type of long-term planning, and these should be included in the evaluation.

Part 5: analysis and interpretation

No matter how good the quality of the data collected, it is important to ensure correct analysis and interpretation. Otherwise, the data will not provide useful or robust information about the intervention.

It is beyond the scope of this document to give detailed guidance on data analysis. Readers should either seek the advice of a statistician or analyst, or consult one of the many good textbooks on study design and statistical methods.

59. Process evaluation: essential

All relevant information from the process evaluation section (criteria 50 to 58) should be presented as clearly as possible. This has two main functions: to help improve the intervention in the future; and to help anyone who wants to replicate the intervention in another area or setting.

60. Impact evaluation: Summary of results compared to baseline: essential

The bare minimum 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.

'Evaluation of public health interventions: an introductory guide (2015)'[5] describes the main study designs used for evaluations, all of which have appropriate analysis methods. In experimental designs, such as randomised controlled trials (RCTs), results are presented as a change in the intervention group compared to change in the control group. If the difference between the intervention and control group is statistically significant (usually expressed as a p value of <0.05), there can be confidence this was caused by the intervention itself, and not by some external factor. So, for example, in an RCT of a self-help intervention to promote active commuting, Mutrie et al. found that the intervention group was almost twice as likely to increase walking to work as the control group at six months.[74] This was expressed as an odds ratio of 1.93 (meaning the odds of walking to work in the intervention group were 1.93 times the odds of walking to work in the control group). Thus, there is confidence that the intervention is effective in increasing walking.

The stepped wedge cluster randomised trial is a relatively new pragmatic study design, that is increasingly being used for intervention evaluations with political or logistical constraints. The design uses random and sequential crossovers of clusters from control to intervention until all clusters are exposed. As a result more clusters are exposed to the intervention towards the end of the study than in its early stages, thus sample size calculations and analysis must make allowance for both the clustered nature of the design and the confounding effect of time. This approach has been described in detail by Hemming et al.[75]

Quasi-experimental designs usually include a control group. Unlike the RCT, they do not randomly allocate individuals to intervention or control. Like the RCT, results are stated in terms of differences between intervention and control. The main limitation is the lack of certainty that the difference between intervention and control group was due to the intervention. For example, in a pilot study of obesity prevention counselling for children, Kubik et al. compared counselling in one primary care practice with the usual

care provided by a control practice.[76] Following the pilot it was found that more intervention than control parents reported they intended to give their child five or more servings of fruits and vegetables on most days during the next 30 days (25% of parents versus 9% of control parents; p=0.049). This indicates that the intervention was successful, with the limitation that there might have been something different about the 'usual care' practice that influenced the results.

As stated in the introductory guide,[5] pre-experimental designs provide the weakest evidence and should only be used when other possibilities have been explored.[9] Like experimental designs, data from pre and post studies is usually presented as difference between data before and after the intervention. The limitation here is that we cannot be sure that any change would not have happened anyway, as there was no control group. For example, a weight management intervention may have reported weight loss, however this result could have been prompted by other external factors such as another intervention or campaign, or simply be due to secular trends.

61. Impact evaluation: Detail of any further analyses or statistical methods: desirable

It is beyond the remit of this document to detail the statistical methods that could be used in the analysis of collected data and the nature of the statistical methods used will vary depending on the research and evaluation study design. As a minimum the PHE performance indicators should be assessed.[2] Below are some key points about data analysis which may assist the evaluation.

Change in BMI z-score (for children and adolescents)
When evaluating children's weight management interventions it is preferable to measure the change in a child's BMI using the BMI z-score rather than the BMI centile. A BMI z-score or Standard Deviation Score indicates how many units of the standard deviation a child is above or below the average BMI value for their age group and sex.

For instance, a z-score of 1.5 indicates a child is 1.5 standard deviations above the average value and a z-score of -1.5 indicates a child is 1.5 standard deviations less than the average value. Approximately 95% of children will lie within two standard deviations from the average BMI. This is due to the fact that the BMI centile is relative to the reference population and a reduction in a child's BMI will not be consistent across the entire scale. For example, a change from the 98th to the 95th centile will typically be a larger reduction in actual BMI than a change from the 68th to the 65th centile. However, using the BMI z-score, any reductions in a child's BMI z-score will be consistent within each reference population, that is their age group and sex, regardless of their original BMI value.

Statistical significance
This describes the extent to which we can be certain that a result did not occur by chance. Statistical significance is usually expressed as a p value, often shown to be $p<0.05$. This means that there is a 5% possibility that the result occurred due to chance, and was not as a result of the intervention. Statistical significance is related to the power of a study, which can be determined through sample size calculations. These should be conducted before the study begins to calculate how many people are needed in the study to enable measurements that will be statistically significant. It is important to seek the advice of a statistician before a study commences to ensure that it is large enough and that the sample is constructed correctly.

Confidence intervals
These describe the range of possible values around an observed outcome. For example, there may be a mean change in body weight of one kilogram following an intervention, with 95% confidence intervals stated as 0.2 to 1.8. This means there is a 95% likelihood that the true amount of weight loss lies between 0.2 and 1.8kg.

62. Economic evaluation: desirable

It is really important to demonstrate the economic value of a weight management intervention. It is therefore advised that as a minimum, the cost per participant is provided as part of the evaluation. However some evaluations may require more detailed cost benefit analysis, which is beyond the scope of this guidance. For evaluators wishing to undertake more advanced economic assessments it would be advisable to seek advice from a health economist.

63. Limitations and generalisability: desirable

The generalisability of the intervention is how likely it is that the results of the intervention would be reproduced if the intervention were carried out in another group or in the whole population. When assessing generalisability, it is important to take into account the following: the target population, nature of the active intervention, nature of sampling and recruitment methods, length of follow-up, settings and resources needed. The more tightly controlled the intervention is, that is the higher the internal validity, the less likely it is that the results would be reproduced unless another intervention was conducted and controlled in the same way.

Conclusions

We hope this revised document has been helpful to those considering an evaluation of a weight management intervention. This revision has taken into account feedback received on the first edition of the standard evaluation framework, and has helped to establish consensus on the minimum dataset that needs to be collected for an evaluation to be effective and to help raise standards of evaluation. This should contribute towards the long-term aim of advancing knowledge on what works in combating the rise in obesity in this country.

Following several years of delivering evaluation training and working with service providers and commissioners, here are our top 10 tips for a successful evaluation:

Top 10 tips:

1. understand the needs of your end users: be innovative and work with them to tailor the best available evidence to their needs
2. prioritise process evaluation so you can learn and improve as you go
3. where possible used a mixed method approach to your evaluation as qualitative data can provide important insight and context to the interpretation of the quantitative data
4. ensure evaluation is embedded at the very start of planning a weight management intervention
5. where possible seek an independent evaluation, as this will enhance the transparency of the findings
6. use the PHE data capture tool to ensure data collection practices are standardised and all minimum requirements are met, and use validated tools where possible.
7. involve all stakeholders in the evaluation process so all expectations are managed
8. publish your findings so they strengthen the evidence base, and provide shared learning
9. remember to report both positive and negative outcomes – sometimes more can be learnt from what hasn't worked, than what has worked, and can be useful in informing service improvements
10. be pragmatic – if for some reason an essential criteria could not be collected – explain why

Glossary

Active intervention
The pre-defined weight management intervention, that doesn't include follow up.

Aim
A broad statement of intent setting out the purpose of the project. For example, 'the project aims to reduce obesity through counselling of children with obesity, and their families.'

Evaluation
Exploring the extent to which a project has achieved what it set out to do.

Focus group
A group of people who discuss an issue, led by a researcher. This generates qualitative data, usually in the form of transcripts. Sometimes used in preference to individual interviews, as some researchers believe that the 'group processes' involved will reveal more about people's beliefs and attitudes.

Follow up
The period after the active intervention.

Indicator
A measure of something which demonstrates a change in a particular outcome. For example the number of people attending a weight control session.

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 on the phone, structured (with fixed questions) or semi-structured (where discussion can be more flexible).

Milestone
A marker of progress, usually used to monitor whether a course of action is on track. Like a milestone on the road it tells you whether you are on the right track, how far you have travelled and how far you still have to go.

Monitoring
Collection of routine data that helps you assess whether projects are proceeding to plan. A sub-set of evaluation.

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.’

Outputs
Things that the project produces or activities that occur through the use of the resources in the project. For example, distribution of 300 leaflets to the target population.

Process evaluation
Evaluation which focuses on the process used throughout a project: it aims to see why the project meets or does not meet its aims and objectives; what went right and what went wrong; what can be learnt for future projects.

Qualitative research
Information that is reported in narrative form or which is based on descriptive information, such as diaries, open ended responses to questions and field notes.

Quantitative research
Information that is reported in numerical form, such as number of people attending and dropout rates.

Reliability
How likely it is that a measurement instrument will measure the same thing each time it is used.

Target population
People who the project aims to reach. These may be segmented by a number of factors including age, gender, ethnicity and social class.

Validity
How well something measures the ‘truth’. For example, to what extent a self-report food frequency questionnaire reflects the actual dietary intake. In the context of a research study validity can be either ‘internal’ or ‘external’. Internal validity is the extent to which differences between a study and a control intervention are real rather than a product of bias. External validity is the extent to which the results of the study can be made general for the wider population.

References

1. Coulton V, Dodhia S, Ells L, et al. National mapping of weight management interventions. London: Public Health England, 2015.
2. Public Health England. Weight management: guidance for commissioners and providers: guidance for commissioners and providers 2017. www.gov.uk/government/collections/weight-management-guidance-for-commissioners-and-providers.
3. Cavill N, Roberts K, Rutter H. Standard Evaluation Framework for Physical Activity Interventions. London: National Obesity Observatory, 2012.
4. Roberts K, Cavill N, Rutter H. The Standard Evaluation Framework for dietary interventions London: National Obesity Observatory, 2012.
5. Cavill N, Roberts KE, Ells L. Evaluation of weight management, physical activity and dietary interventions: an introductory guide: Public Health England, , 2015.
6. Hartmann-Boyce J, Johns DJ, Jebb SA, et al. Behavioural weight management programmes for adults assessed by trials conducted in everyday contexts: systematic review and meta-analysis. Obes Rev 2014;15(11):920-32 doi: 10.1111/obr.12220[published Online First: Epub Date]|.
7. Hartmann-Boyce J, Johns DJ, Jebb SA, et al. Effect of behavioural techniques and delivery mode on effectiveness of weight management: systematic review, meta-analysis and meta-regression. Obes Rev 2014;15(7):598-609 doi: 10.1111/obr.12165[published Online First: Epub Date]|.
8. Hillier-Brown FC, Bambra CL, Cairns JM, et al. A systematic review of the effectiveness of individual, community and societal level interventions at reducing socioeconomic inequalities in obesity amongst children. BMC Public Health 2014;14:834 doi: 10.1186/1471-2458-14-834[published Online First: Epub Date]|.
9. Nutbeam D, Bauman A. Evaluation in a Nutshell: a Practical Guide to the Evaluation of Health Promotion Programs. Sydney: McGraw Hill Publishers, 2006.
10. CONSORT. The CONsolidated Standards of reporting Trials. Secondary The CONsolidated Standards of reporting Trials 2010. www.consort-statement.org/consort-2010.
11. CONSORT. Pragmatic Trials. Secondary Pragmatic Trials 2010. www.consort-statement.org/extensions?ContentWidgetId=556.
12. Glasziou PP, Boutron I, Milne R, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;7(348) doi: 10. 1136/bmj. g1687. [published Online First: Epub Date]|.
13. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013;50(5):587-92 doi: 10.1016/j.ijnurstu.2012.09.010[published Online First: Epub Date]|.
14. Caetano R. Standards for reporting non-randomized evaluations of behavioral and public health interventions: the TREND statement. Addiction 2004;99(9):1075-80 doi: 10.1111/j.1360-0443.2004.00785.x[published Online First: Epub Date]|.

15. The Magenta Book Guidance for evaluation. London: HM Treasury 2011.
16. Weight Management Economic Assessment Tool. Secondary Weight Management Economic Assessment Tool 2016. webarchive.nationalarchives.gov.uk/20170110171618/https://www.noo.org.uk/visualisation/economic_assessment_tool.
17. Child weight management interventions: collect and record data Secondary Child weight management interventions: collect and record data 2017. www.gov.uk/government/publications/child-weight-management-interventions-collect-and-record-data.
18. Adult weight management interventions: collect and record data Secondary Adult weight management interventions: collect and record data 2017. www.gov.uk/government/publications/adult-weight-management-interventions-collect-and-record-data
19. The Disclosure and Barring Intervention (DBS) 2017. www.gov.uk/government/organisations/disclosure-and-barring-intervention/about.
20. Health Research Authority. Consent and participant information 2016. www.hra.nhs.uk/resources/before-you-apply/consent-and-participation/consent-and-participant-information/.
21. UK Government. Data Protection. 2016. www.gov.uk/data-protection/the-data-protection-act.
22. Health and Social Care Information Centre. Joint Strategic Needs Assessments. 2016. www.hscic.gov.uk/jsna.
23. Public Health England. Public Health Outcomes Framework 2015. www.phoutcomes.info/.
24. National Institute for Health and Care Excellence. Policy on conflicts of interest 2007.
25. National Obesity Observatory. Obesity and ethnicity. 2011. https://www.noo.org.uk/uploads/doc/vid_9444_Obesity_and_ethnicity_270111.pdf.
26. Equalities and Human Rights Commission. Equality Act 2010. Secondary Equality Act 2010 2016. www.equalityhumanrights.com/legal-and-policy/legislation/equality-act-2010.
27. Office for National Statistics. 2011 Census analysis: Ethnicity and religion of the non-UK born population in England and Wales 2016.
28. UK Government. Definition of disability under the Equality Act 2010 Secondary Definition of disability under the Equality Act 2010 2016. www.gov.uk/definition-of-disability-under-equality-act-2010.
29. Public Health England. Obesity and Disability: adults. 2013. www.noo.org.uk/uploads/doc/vid_18474_obesity_dis.pdf.
30. Public Health England. Obesity and disability: children and young people 2013
31. National Institute for Health and Care Excellence. Obesity: identification, assessment and management Secondary Obesity: identification, assessment and management 2014. www.nice.org.uk/guidance/cg189.
32. NHS digital. Health Survey for England 2012, 2013.
33. National Child Measurement Programme - England, 2016-17 NHS digital, 2017.

34. Office for National Statistics. Lower Layer Super Output Areas Secondary Lower Layer Super Output Areas 2016. https://data.gov.uk/dataset/lower-layer-super-output-areas-2001-to-lower-layer-super-output-areas-2011-to-local-authority-d.
35. UK Government. English indices of deprivation 2010: technical report 2011. www.gov.uk/government/statistics/english-indices-of-deprivation-2010-technical-report.10.1111/obr.12569 [published Online First: Epub Date]|.
36. Atkinson G, Fullick S, Grindey C, et al. Exercise, Energy Balance and the Shift Worker. Sports medicine (Auckland, NZ) 2008;38(8):671-85
37. Office for national Statistics. NS-SEC coding tool Secondary NS-SEC coding tool 2016. www.neighbourhood.statistics.gov.uk/HTMLDocs/dev3/ONS_NSSEC_discovery_tool.html
38. Brown TJ, O'Malley C, Blackshaw J, et al. Exploring the evidence base for Tier 3 weight management interventions for adults: a systematic review. Clin Obes, 7 2017;7:260-72
39. Dare S, Mackay DF, Pell JP. Relationship between Smoking and Obesity: A Cross-Sectional Study of 499,504 Middle-Aged Adults in the UK General Population. PLoS One 2015;10(4)
40. Jeffery RW, Rick AM. Cross-sectional and longitudinal associations between body mass index and marriage-related factors. Obes Res 2002;10(8):809-15 doi: 10.1038/oby.2002.109 [published Online First: Epub Date]|.
41. French SA, Jeffery RW, Forster JL, et al. Predictors of weight change over two years among a population of working adults: the Healthy Worker Project. Int J Obes Relat Metab Disord 1994;18(3):145-54
42. S M, M G. Obesity and alcohol: an overview Oxford, 2012.
43. Heslehurst N, Ells LJ, Simpson H, et al. Trends in maternal obesity incidence rates, demographic predictors, and health inequalities in 36,821 women over a 15-year period. Bjog 2007;114(2):187-94 doi: 10.1111/j.1471-0528.2006.01199.x [published Online First: Epub Date]|.
44. Parsons TJ, Power C, Logan S, et al. Childhood predictors of adult obesity: a systematic review. Int J Obes Relat Metab Disord 1999;23 Suppl 8:S1-107
45. National Obesity Observatory. Body Mass Index as a measure of obesity. 2009. www.noo.org.uk/securefiles/160329_1633//noo_BMI.pdf.
46. Cole TJ. Growth monitoring with the British 1990 growth reference. Archives of Disease in Childhood 1997;76(1):47-49 doi: 10.1136/adc.76.1.47 [published Online First: Epub Date]|.
47. Harlow Healthcare. Growth Charts Secondary Growth Charts 2016. www.healthforallchildren.com/hfac-support/.
48. Public Health England. National Child Measurement Programme operational guidance. 2017. www.gov.uk/government/publications/national-child-measurement-programme-operational-guidance.
49. NHS digital. Health Survey for England 2011 - Health, social care and lifestyles - summary of key findings. Leeds: NHS digital, 2012.

50. National Institute for Health and Clinical Excellence. Obesity: Guidance on the prevention, identification, assessment and management of overweight and obesity in adults and children. Quick reference guide 2 for the NHS. London, 2006:19.
51. Goran MI. Measurement issues related to studies of childhood obesity: assessment of body composition, body fat distribution, physical activity, and food intake. Pediatrics 1998;101(3 Pt 2):505-18
52. Hills AP, Kagawa M. Body composition assessment in children and adolescents Children Obesity and Exercise Oxford Routledge, 2007.
53. Kroes M, Osei-Assibey G, Baker-Searle R, et al. Impact of weight change on quality of life in adults with overweight/obesity in the United States: a systematic review. Curr Med Res Opin 2016;32(3):485-508 doi: 10.1185/03007995.2015.1128403[published Online First: Epub Date]|.
54. Public Health England. Measuring the mental wellbeing of children and young people London: Public Health England, 2015. www.gov.uk/government/publications/measuring-the-mental-wellbeing-of-children-and-young-people
55. Bressan J, Hills AP, Hermsdorff H. Eating behaviour in children and the measurement of food intake. In: Hills A, King N, NM N, eds. Children obesity and exercise: prevention treatment and management of childhood and adolescent obesity. Oxford Routledge, 2007.
56. Johnson RK. Dietary intake--how do we measure what people are really eating? Obes Res 2002;10 Suppl 1:63S-68S doi: 10.1038/oby.2002.192[published Online First: Epub Date]|.
57. Stockwell T, Zhao J, Greenfield T, et al. Estimating under- and over-reporting of drinking in national surveys of alcohol consumption: Identification of consistent biases across four English-speaking countries. Addiction 2016 doi: 10.1111/add.13373[published Online First: Epub Date]|.
58. Lichtman SW, Pisarska K, Berman ER, et al. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. N Engl J Med 1992;327(27):1893-8 doi: 10.1056/NEJM199212313272701[published Online First: Epub Date]|.
59. Little P, Barnett J, Margetts B, et al. The validity of dietary assessment in general practice. J Epidemiol Community Health 1999;53(3):165-72
60. Storey KE. A changing landscape: web-based methods for dietary assessment in adolescents. Curr Opin Clin Nutr Metab Care 2015;18(5):437-45 doi: 10.1097/mco.0000000000000198[published Online First: Epub Date]|.
61. Roberts K, Flaherty S, Jenner D, et al. A review and needs assessment of local level diet and nutrition data 2008.
62. Ashfield-Watt PA, Welch AA, Godward S, et al. Effect of a pilot community intervention on fruit and vegetable intakes: use of FACET (Five-a-day Community Evaluation Tool). Public Health Nutr 2007;10(7):671-80 doi: 10.1017/s1368980007382517[published Online First: Epub Date]|.
63. Cade JE, Frear L, Greenwood DC. Assessment of diet in young children with an emphasis on fruit and vegetable intake: using CADET--Child and Diet Evaluation Tool. Public Health Nutr 2006;9(4):501-8

64. Christian MS, Evans CEL Nykjaer C et al. Measuring diet in primary school children aged 8-11 years: validation of the Child and Diet Evaluation Tool (CADET) with an emphasis on fruit and vegetable intake. Eur J Clin Nutr 2015:69: 234-241
65. AP H, Byrne N. Children obesity and exercise Oxford Routledge 2007.
66. Golan M, Weizman A. Reliability and validity of the Family Eating and Activity Habits Questionnaire. Eur J Clin Nutr 1998;52(10):771-7
67. 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 doi: 10.1136/bjsm.2005.025296 [published Online First: Epub Date]|.
68. Crouter S, Schneider P, Bassett DJ. Spring-levered versus piezo-electric pedometer accuracy in overweight and obese adults. Med Sci Sports Exerc 2005;37(10):1673-9
69. Richardson D, Cavill N, Ells L, et al. Measuring Diet and Physical Activity in Weight Management Interventions: a Briefing Paper. Oxford National Obesity Observatory, 2011.
70. National Institute for Health and Care Excellence. Behaviour Change Secondary Behaviour Change 2016. http://pathways.nice.org.uk/pathways/behaviour-change.
71. Nicholls J, AT M. Development and its discontents: the differentiation of the concept of ability. In: Nicholls J, ed. Advances in motivation and achievement Vol 3. Greenwich JAI press, 1984.
72. National Institute for Health and Care Excellence. Weight management: lifestyle interventions for overweight or obese adults. 2014. www.nice.org.uk/guidance/ph53.
73. Bowling A. Research methods in Health, 2nd edition Buckingham Oxford University press, 2002.
74. Mutrie N, Carney C, Blamey A, et al. "Walk in to Work Out": a randomised controlled trial of a self help intervention to promote active commuting. J Epidemiol Community Health 2002;56(6):407-12
75. Hemming K, Haines TP, Chilton PJ, et al. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ : British Medical Journal 2015;350 doi: 10.1136/bmj.h391[published Online First: Epub Date]|.
76. Kubik MY, Story M, Davey C, et al. Providing obesity prevention counselling to children during a primary care clinic visit: results from a pilot study. J Am Diet Assoc 2008;108(11):1902-6 doi: 10.1016/j.jada.2008