

Dadansoddi ar gyfer Polisi



Analysis for Policy



Llywodraeth Cymru
Welsh Government

SOCIAL RESEARCH NUMBER: 77/2022

PUBLICATION DATE: 08/12/2022

Evaluability Assessment of the Healthy Wales: Healthy Weight Strategy

Mae'r ddogfen yma hefyd ar gael yn Gymraeg.

This document is also available in Welsh.

OGL © Crown Copyright Digital ISBN 978-1-80535-033-0

Evaluability Assessment of the Healthy Wales: Healthy Weight Strategy

Report by:



Views expressed in this report are those of the researcher and not necessarily those of the Welsh Government.

For further information please contact:

Victoria Seddon

Social Research and Information Division

Welsh Government

Cathays Park

Cardiff

CF10 3NQ

Tel: 0300 062 8348

Email: research.healthandsocialservices@gov.wales

Table of contents

Executive Summary	3
1. Background	7
2. Literature review: comparable evaluations	10
3. Stakeholder interviews	13
4. Logic model and theories of change	16
5. Evaluation framework.....	30
6. Evaluation approach.....	40
7. Evaluability assessment: toolkit.....	46
8. Evaluability assessment: examples.....	56
9. Practical evaluation issues	66
10. Conclusion and recommendations	73
References.....	76
Annex A – Theory of change evidence base	80
Annex B – Databank.....	83

Abbreviations

Abbreviation	Definition
AHWA	Amsterdam Healthy Weight Approach
AONB	Areas of Outstanding Natural Beauty
BMI	Body Mass Index
CBA	Cost-benefit analysis
GDPR	General Data Protection Regulation
CMP	Child Measurement Programme
HM Treasury	Her Majesty's Treasury
HWHW	Healthy Weight: Healthy Wales
IMTP	Integrated Medium Term Plan
NERS	National Exercise Referral Scheme
RDD	Regression discontinuity design
ToC	Theory of Change

Executive Summary

The Healthy Weight: Healthy Wales (HWHW) strategy is a 10-year plan developed by the Welsh Government to help make the healthier choice the easy choice for people in Wales. The strategy aims to shape the environments and settings people live in to better enable the population to make positive lifestyle choices to sustain or maintain a healthy weight.

Evaluation of the strategy and its component policies is expected to be key to ensuring the strategy is achieving its aims to the greatest extent possible and on a value-for-money basis, as well as informing ongoing decision-making. However, for this to happen, an evaluability assessment is required to assess the feasibility and extent of evaluation activities. Alma Economics has been commissioned by the Welsh Government to assess the evaluability of the HWHW strategy and its underlying policies. The assessment considers the breadth of evaluation types set out in HM Treasury's Magenta Book and seeks to determine how they can be best combined to provide a holistic evaluation approach.

Review of comparable evaluations

As part of our assessment of the evaluability of the Healthy Weight: Healthy Wales strategy, we conducted a literature review aiming to understand recent approaches to the evaluation of strategies comparable to HWHW either in terms of complexity or subject matter.

The key finding of this review is that evidence on evaluations of strategies and component parts is very scarce. Arguably none of the studies reviewed offered much insight into how to produce a comprehensive evaluability assessment for HWHW and this area represents an evidence gap in the literature.

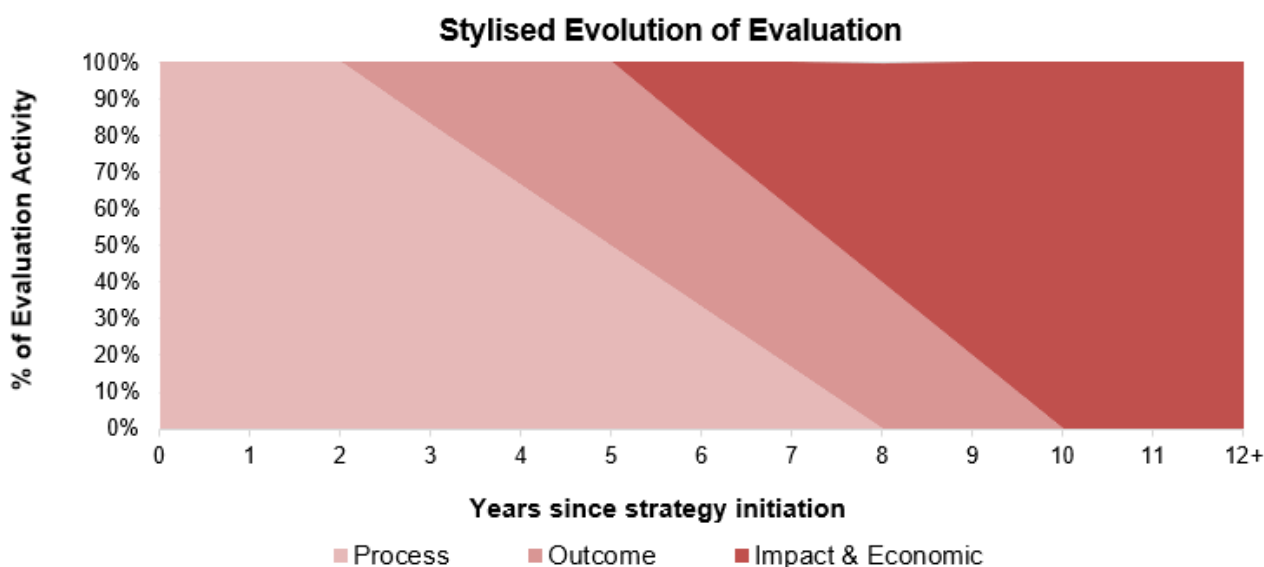
Logic model and theories of change

One of the objectives of this evaluability study was to develop a robust analytical framework which sets out the mechanisms through which the activities for the Healthy Weight: Healthy Wales strategy will impact obesity related outcomes in Wales. For this purpose, we have developed an integrated framework of Theories of Change (ToCs) and Logic Models which provides a structured way of thinking about if and how the strategy as a whole will realise its goal of preventing and reducing obesity. The framework explicitly lays out the causal chains to demonstrate the theoretical pathway through which activities will sequentially translate into outputs, short-term outcomes and long-term impacts.

The Healthy Weight: Healthy Wales strategy is highly complex in the sense that it looks to use a holistic and systems-based approach to tackle obesity. As such, it houses tens of policies and programmes which look to address the issue from a range of angles. To account for this complexity, while providing a structured and useful way to think about the causal mechanisms underlying the strategy’s work, we have developed a framework which combines a logic model with six Theories of Change. The approach combines the breadth which can be achieved using logic models with the depth added by applying ToC methodology.

Evaluation framework

According to HM Treasury’s Magenta Book (2020), there are three main types of evaluation activity: process evaluation, impact evaluation and value-for-money (economic) evaluation. The evaluability assessment also considers outcome evaluation as a fourth evaluation type, which can be seen as a sub-category of impact evaluation that focuses on short-run changes in outcomes.



These four categories seek to address the following questions:

- **Process** – What can be learned from how the intervention was delivered?
- **Outcome** – What difference has an intervention made in the short run?
- **Impact** – What difference has an intervention made in the long run?
- **Economic** – Was this intervention a good use of resources?

An important consideration for the overarching evaluation approach is around how resources should be allocated between the different types of evaluation at different points in the strategy's lifespan. This decision should be made based on two principal factors:

- The need to ensure that, as a whole, the evaluation provides a holistic picture of the strategy's performance.
- The availability of data at different points in the strategy's lifespan.

Across the strategy's lifespan, it is expected that the evaluation's emphasis will shift from process evaluation early on towards impact and economic evaluation as the requisite data becomes available. Outcome evaluation will form a bridge between these two phases by providing evidence of any emerging changes in outcomes at the earliest opportunity such that these findings can be used to inform the strategy's final delivery plans.

Evaluation approach and evaluability toolkit

In order to provide a structured way to think about evaluation of HWHW, four considerations are suggested to assess evaluability of different aspects of HWHW, as outlined below.

- **Data Availability** – What can be done with the data which is currently collected or that could feasibly be collected during the strategy's lifespan?
- **Cost** – Is the cost of any potential evaluation realistic in gross terms and in terms of its proportionality to the cost of implementing the programme?
- **Timeliness** – Is the evaluation sufficiently timely as to allow its results to be fed back into the strategy's planning decisions?
- **Robustness** – Is the evaluation both sufficiently robust in its methods and sufficiently holistic in its methods as to provide credible evidence of the strategy's success?

The toolkit provides a structured way for prioritising evaluation in a resource-constrained environment. In theory, there could be sufficient data and resources for every HWHW policy to have a comprehensive process, outcome, impact and economic evaluation, but in practice this isn't feasible due to limitations of data, resources and, in some cases, the availability of suitable robust evaluation techniques.

It also provides an approach on how to consider the feasibility of the evaluations during the lifetime of the strategy. However, it is acknowledged that evaluations could be determined on a case-by-case basis and there can be instances where an evaluation could go ahead even if the toolkit produces a different recommendation. The toolkit is applied in two steps:

- **Step 1** – The feasibility of the evaluation type being considered for the programme/policy in question is scored out of 4 for each of the four key considerations.
- **Step 2** – These four scores are combined to form a rounded average score which is specific to the programme/policy and the type of evaluation proposed. This average score will provide guidance as to whether that type of evaluation is feasible and desirable for the programme/policy in question. A rounded average of 3 or 4 indicates a recommendation for evaluation. The exception to this rule is the case where there is a score of 0 out of 4 for any of the four key considerations which then overrides the average and indicates that evaluation is not recommended.

A full scoring system is specified and worked examples is provided in section 7 and 8 to show how this could work in practice.

Conclusion and recommendations

A conclusion that can be drawn from the research conducted is that there is a lot of scope for process, impact and economic evaluation, and it is feasible to take some forward for the overall strategy and for several individual policies. There will also be some parts of HWHW that are not feasible to evaluate individually due to constraints of data, cost, timeliness or robustness. The evaluability assessment would need to be refined as and when further detail become available on the strategy and individual policies.

The cut-off point for where to draw the line between what is and is not evaluable depends on some decisions that have not been made, such as the total spend on HWHW evaluation over its lifetime and the nature of future policy changes made to implement the strategy. We have provided a toolkit for assessing where to draw the line as these decisions are made.

Some recommendations that will assist in increasing evaluability:

- Apply the toolkit early and review regularly.
- Consider monitoring plans at an early stage and decide where additional data collection is needed.
- Collect baseline data, especially for evaluation of systems change driven by the overall HWHW strategy.

1. Background

- 1.1 Alma Economics has been commissioned by the Welsh Government to assess the evaluability of their Healthy Weight: Healthy Wales strategy to prevent and reduce obesity. For this purpose, we have assessed the available options for evaluating a strategy of this kind and provided recommendations regarding which of them is most appropriate and why. The assessment considers the breadth of evaluation types set out in HM Treasury's Magenta Book and seeks to determine how they can be best combined to provide a holistic evaluation approach.
- 1.2 Another key component of this work has been to provide an in-depth understanding of the 'pathways of change' associated with the HWHW strategy. For this purpose, we have developed a logic framework which sets out the mechanisms through which the strategy's activities are expected to lead to improvements in people's health and wellbeing.
- 1.3 The remainder of this report is structured as follows. The remainder of Section 1 provides some background information regarding the context of obesity in Wales and the Healthy Weight: Healthy Wales strategy itself. Section 2 presents the findings of the literature review into the approaches to strategic evaluation. Section 3 displays the logic framework we have developed and outlines the underlying assumptions. Section 4 summarises outputs from stakeholder interviews used to inform the evaluability assessment. Sections 5 and 6 set out theory and approaches for evaluation, including explanation of relevant evaluation techniques and data sources. Sections 7 and 8 explain the toolkit we have developed for assessing the evaluability of programmes and policies which make up the HWHW strategy and provide examples of applying the toolkit. Section 9 sets out some of the practical concerns which will be relevant for future evaluators to bear in mind. Finally, Section 10 outlines our conclusions and recommendations.

Healthy Weight: Healthy Wales

- 1.4 The Healthy Weight: Healthy Wales strategy is a 10-year plan developed by the Welsh Government to help make the healthier choice the easy choice for people in Wales. The strategy aims to shape the environments and settings people live in to better enable the population to make positive lifestyle choices to sustain or

maintain a healthy weight. By empowering the people of Wales in this way, the strategy aims to provide current and future generations with the best possible start in life and the chance to live longer, better and happier lives.

- 1.5 The Well-being of Future Generations (Wales) Act can be seen as setting the foundations upon which the strategy has been designed (Welsh Government, 2015). The strategy uses the Act's five ways of working as a set of guiding principles for its systems-based approach to encouraging healthy weight in Wales.¹ The ways of working are designed to enable leadership at every level and are intended to maximise the strategy's impact upon current and future generations' lived experiences. The strategy is also intended to directly contribute to the Act's goals of creating a healthier, more equal and resilient Wales.
- 1.6 The strategy began in earnest in 2019 when the government set out their first delivery plan for 2020-22. This was the first in a series of five 2-year delivery plans which will set out the strategy's activities as they evolve over its lifespan.² By breaking the 10-year period down into 5 delivery plans, the strategy can be regularly updated to meet any emerging or changing needs.
- 1.7 The strategy's work will look to engender progress along four broad themes:
- **Healthy environments** – aims to support people to be able to make healthy choices. This will focus on changing the way people shop, the way they eat out, the way they travel, or how they use outdoor spaces.
 - **Healthy settings** – will develop supportive environments to promote healthier choices. This includes childcare settings, schools and higher and further education, workplaces and community settings.
 - **Healthy people** – providing advice, information and support. This includes providing people with the opportunity to regularly discuss their lifestyle choices with health and care professionals.
 - **Leadership and enabling change** – to drive improved leadership and accountability to deliver Healthy Weight: Healthy Wales across all sectors.

¹ The Well-being of Future Generations (Wales) Act's five ways of working are long-term thinking, integration, involvement, collaboration and prevention.

² Unfortunately, the Covid-19 pandemic derailed aspects of the 2020-22 delivery plan and this was reflected in a revised delivery plan for 2021-22.

Evaluation in the context of Healthy Weight: Healthy Wales

- 1.8 As part of the strategy, the Welsh Government has committed to continuously monitor and evaluate progress. Evaluating the strategy will provide benefits from both a learning and an accountability perspective and will be crucial to ensuring that resources are allocated to policies and programmes which are shown to deliver results.
- 1.9 The evaluation component of the strategy will play a major role in ensuring alignment between the strategy and the Future Generations Act's five ways of working (Welsh Government, 2015). Evaluation is the only way to ensure that decisions regarding which aspects of the strategy should be modified, expanded, continued or discontinued can be made with the best possible evidence. As a dual benefit, the evaluation could drive public and stakeholder collaboration and involvement, if designed correctly. Including qualitative evaluation methods which seek the input of these groups would allow them to directly influence evolution of the strategy going forward.
- 1.10 How this commitment to evaluation could and should look in practice is a non-trivial consideration given the complexity of the strategy, the constraints around data availability and budget, and the need for any evaluation to be sufficiently robust and timely. Any potential approach should come as close to satisfying all of these demands as possible, so a structured way to think about how the trade-offs between these competing requirements should be judged is required.
- 1.11 A strategy of this scope and complexity could be evaluated in a variety of ways. The purpose of this report is to develop an evaluability assessment to determine which option, or combination of options, is appropriate for the Healthy Weight: Healthy Wales strategy.

2. Literature review: comparable evaluations

- 1.12 As part of our assessment of the evaluability of the Healthy Weight: Healthy Wales strategy, we conducted a literature review which aimed to understand recent approaches to the evaluation of strategies comparable to HWHW in terms of complexity and subject matter. This review considered literature published in the past 5 years and found a small number of relevant approaches for evaluation using qualitative and quantitative techniques.
- 1.13 The key finding of this review was that evidence of holistic and comprehensive evaluations of relevant strategies and their component parts is scarce. Arguably none of the studies reviewed offered significant insight into how a comprehensive evaluation of HWHW should look and this area represents a clear evidence gap in the recent literature. That said, a number of evaluations provided some insights which, when combined, go some way to indicating what an evaluation of the Healthy Weight: Healthy Wales strategy might look like.
- 1.14 Perhaps the most relevant example in the recent literature is that of the Amsterdam Healthy Weight Approach (AHWA) given its high degree of similarity to HWHW in terms of its scope and goals. The AHWA is strategy which uses a whole-systems approach to improve children's physical activity, diet and sleep through action in the home, neighbourhood, school and city (Sawyer et al., 2021). A review of AHWA's success is based on sets of output, outcome and impact goals which are updated every two to three years. The idea is for the initial delivery plans to focus on outputs, with the focus then shifting towards outcomes and impacts as the strategy matures. The goals include metrics such as:
- Activities: '50 additional water fountains installed in public spaces, 1,734 healthy eating consultations, 2 published neighbourhood recipe books, 300 health ambassadors to promote healthy lifestyles, 14 healthy sports canteens, 11 neighbourhoods committed to a joint Healthy Weight Pact'.
 - Outputs: '1,200 severely obese children identified and being treated, two-thirds of 14 and 16-year-olds walk or cycle to school, more children drink a maximum of two glasses of sugary drinks a day' (UNICEF, 2020).

- 1.15 Although AHWA has not been formally evaluated, this approach provides an indication of how the focus of a strategic evaluation might shift over time with regards to outputs, outcomes and impacts, and provides indicators which could be important for assessing HWHW's success.
- 1.16 Another relevant study relates to the evaluation of Leeds's Childhood Obesity Programme. The evaluation used quantitative methods to assess the programme's impact on obesity rates both at the city level and specifically within areas of heightened deprivation (Rudolf et al., 2018). They conducted a simple regression which compares childhood obesity rates in Leeds to those of surrounding neighbourhoods and England (as a semi-counterfactual) over time. They also stratified the results by area deprivation as this is highly correlated with obesity. Therefore, the strategy exploits similarities between areas with and without the intervention to create a type of natural experiment.
- 1.17 While the subject matter is comparable, this is an evaluation of a much narrower programme from both a thematic and geographic perspective when compared to HWHW, focussing purely on analysis of end outcomes (i.e. obesity levels). As a result, the narrow focus of this evaluation is unlikely to be fit for purpose for something as broad as HWHW. However, this example could be useful to understand how the quantitative aspect of an evaluation of the strategy as a whole may look, for example at the national or regional level.
- 1.18 Given the limited recent evidence on evaluation of obesity prevention and reduction strategies, we have also considered two systematic reviews. The reviews consider evaluations of complex obesity prevention interventions and add perspective on how slightly older literature attempted to deal with the unique challenge of evaluating whole-systems approaches to obesity. Bagnall et al., (2019) considered 33 previous studies and found that process and economic evaluation techniques were particularly underrepresented in recent evaluation approaches and suggested that they should be made more prominent going forward. This is because they can enable better outcomes for both the current intervention and future interventions by enabling decision makers to find ways to overcome barriers to effectiveness and learn from past experiences as well as in determining how value for money can be achieved.
- 1.19 The Medical Research Council (2021) also suggests that process evaluation can help determine why an intervention fails or has unanticipated consequences, or

why it works and how it can be optimised. Such findings can facilitate further development of the intervention programme theory. They also highlight the importance of stakeholder engagement as being essential for prioritising research questions, the co-development of programme theory, choosing the most useful research perspective, and overcoming practical obstacles to evaluation and implementation. However, it also recommends that researchers should be mindful of conflicts of interest among stakeholders and use transparent methods to record potential conflicts of interest.

- 1.20 Similarly, Karacebeyli et al., (2018) conducted a review of quantitative outcome evaluations and concluded that a multi-setting, multi-component childhood obesity prevention intervention that relies on active participation of community stakeholders must first invest time and effort into community engagement before embarking on outcome evaluation. The importance of stakeholder engagement is discussed further in Section 9. The review points to the fact that a wholesystems approach to reducing and preventing obesity is still a relatively new methodology and, correspondingly, high-quality examples of its evaluation are sparse.
- 1.21 In summary, the literature has shown that there have been very few recent attempts to evaluate a strategy which is similar to HWHW in terms of scope, complexity and focus on obesity. An exception to this general finding is the evaluation approach for the AHWA. UNICEF's 2020 report provides an example of how the evaluation of an extremely similar strategy was approached, which has direct implications for the evaluation assessment. Additionally, the evaluation of the Leeds Childhood Obesity Programme provides an example of how the strategy's high-level impact on obesity outcomes might be quantitatively evaluated.

3. Stakeholder interviews

- 1.22 To inform the development of this report, 18 stakeholders who had a detailed understanding of HWHW were interviewed. This included representation from the Welsh Government, Obesity Alliance Cymru, Powys Teaching Health Board, Cardiff and Vale University Health Board, Cwm Taf Morgannwg University Health Board, Swansea Bay University Health Board and Aneurin Bevan University Health Board.
- 1.23 These interviews sought to better understand the HWHW strategy and underlying policies, as well as discuss the role of monitoring and evaluation to inform the evaluability assessment toolkit. The interviews covered relevant policies and interventions, monitoring and data collection and considerations for evaluability.
- 1.24 The table below summarises key discussion points from the interviews, which were used as a starting point to shape the evaluability assessment.

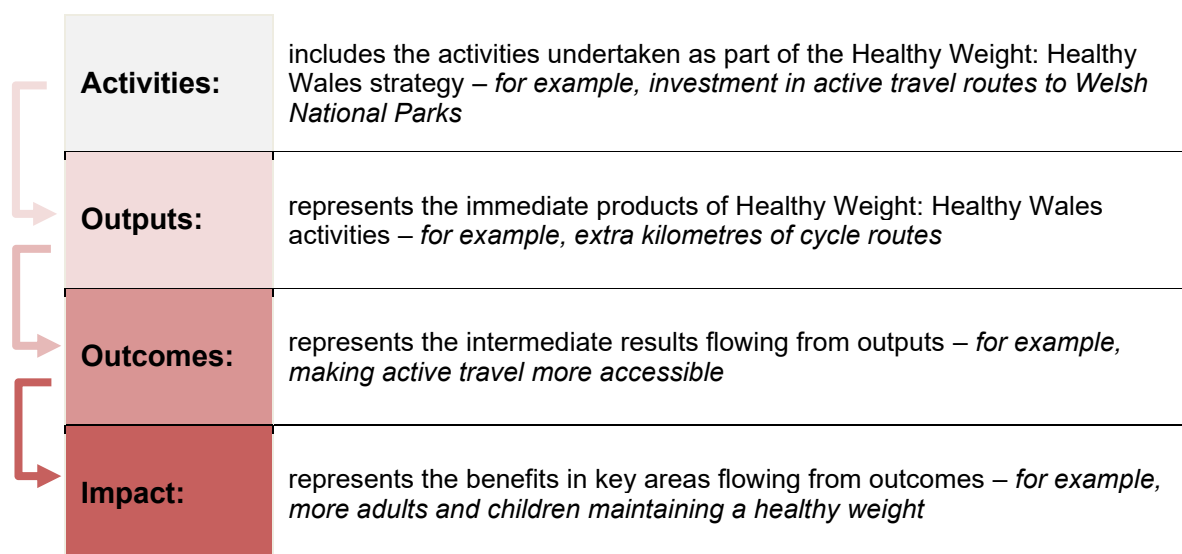
Theme	Key discussion points
Policies and interventions	<ul style="list-style-type: none">• An example of a relevant intervention could include changes in advertising for unhealthy food and drink.• There could also be restrictions on the location of takeaways near schools.• Partially banning high-fat, sugar and salt foods (HFSS) could help force changes in behaviour. These restrictions could apply differently to different providers who are currently all under the same restrictions, but have different levels of interaction with childhood obesity.• Targeting children in their early years is crucial to achieving impact.• Having affordable and enough choices is important.
Theory of Change	<ul style="list-style-type: none">• Covid-19 has exacerbated health inequalities and made people more aware of the need to be healthy.• There needs to be a focus on making healthy choice easier rather than informing and increasing offerings. It is about things being accessible, visible, affordable and viable.• A key challenge to impact is that obesity has become normalised.

Theme	Key discussion points
	<ul style="list-style-type: none"> • Brexit, Covid-19 and the cost-of-living crisis will worsen deprivation and make already more expensive healthy foods relatively even more expensive.
Suggested sources for data	<ul style="list-style-type: none"> • National Survey for Wales. • Child Measurement Programme for Wales. • Sports Wales. • Public health outcomes framework. • National implementation board. • School sports survey. • School Health Research Network.
Data limitations	<ul style="list-style-type: none"> • Collection of data needs to be significantly improved. The process needs to be more thoughtful and it needs to be communicated to those collecting the data in terms of why it should be a priority to avoid current issues of missing data. • Currently, only good data on process indicators exists, with relatively little on outputs and outcomes. • National measurements need to be taken more frequently. • Data collection across regions are fragmented and therefore needs to be aligned. • Population level data is limited, especially for those who do not interact with monitoring services and are missing from the data. • There is a data gap for the last two years due to school closures as a result of the Covid-19 pandemic. • More data should be collected by GPs (e.g., height and weight), but this will require giving GPs an incentive to collect the data. • Breakdowns on certain groups are not collected, such as for ethnicity groups and disability groups.
Considerations for evaluations	<ul style="list-style-type: none"> • It is not always clear from evaluations if the strategy meets the objectives, which should be clearly defined. • We need to determine if there is variance of impact across health boards. • The impact on inequality and across different regions should be considered. • HWHW should be judged on whether it makes activity more accessible and affordable as they are major barriers.

Theme	Key discussion points
	<ul style="list-style-type: none"> • There should be a focus on children's outcomes, as changing their behaviours is both easier and has longer lasting impact. • Stakeholder engagement and buy-in is crucial. • Qualitative tools should also be used to capture peoples' behaviours and decisions. • It should be evaluated whether the interventions have led to a shift in behaviour and decisions. • Evaluations need to be long term as outcomes will not be seen for a while, but there should also be short-term outputs/outcomes.
Indicators for evaluation	<ul style="list-style-type: none"> • Body Mass Index. • Obesity-specific quality of life (QSQL) using a contextualised questionnaire. • Impact on health services (cost & demand). • Rates of chronic disease. • Process indicators. • Percentage of adults/children with two or more healthy lifestyle behaviours (from the Wellbeing of Future Generations Act). • People's overall physical and mental health. • NHS waiting lists. • Referral rates by GPs to exercise centres/facilities. • Reductions in the prevalence of obesity-related illnesses (i.e., diabetes). • Percentage of children entering reception/primary school/secondary school who are obese or overweight. • Changes in the distribution of weight, not just those over the overweight or obese thresholds. • Measure participation levels in active travel and exercise. • Use surveys to test if people's understanding of how weight gain happens and how weight loss can happen has improved since HWHW. • An indicator of success would be the degree to which the NHS and other health organisations have integrated HWHW into their plans. • A measure of how much of the food consumed in Wales is 'healthy'.

4. Logic model and theories of change

- 1.25 One of the objectives of this evaluability study is to develop a robust analytical framework which sets out the mechanisms through which the activities for the Healthy Weight: Healthy Wales strategy will impact obesity-related outcomes in Wales. For this purpose, we have developed an integrated framework of Theories of Change (ToCs) and Logic Models which provides a structured way of thinking about if and how the strategy as a whole will realise its goal of preventing and reducing obesity. The framework explicitly lays out the causal chains to demonstrate the theoretical pathway through which activities will sequentially translate into outputs, short-term outcomes and long-term impacts.
- 1.26 The development of this framework was supported by an extensive stakeholder engagement process which included stakeholders from Welsh Health Boards, Welsh Government, academia, Public Health Wales and Obesity Alliance Cymru.
- 1.27 The distinct elements in this framework are defined as follows:



Activities:	includes the activities undertaken as part of the Healthy Weight: Healthy Wales strategy – <i>for example, investment in active travel routes to Welsh National Parks</i>
Outputs:	represents the immediate products of Healthy Weight: Healthy Wales activities – <i>for example, extra kilometres of cycle routes</i>
Outcomes:	represents the intermediate results flowing from outputs – <i>for example, making active travel more accessible</i>
Impact:	represents the benefits in key areas flowing from outcomes – <i>for example, more adults and children maintaining a healthy weight</i>

- 1.28 It is important to note that, while the logical sequence is visually presented as linear, this is a clear simplification of reality. The ways in which the outputs of each activity will interact with each other to generate the changes in expected outcomes and impacts are highly numerate and complex. As such, attempting to include them all in the graphics would defeat the point of the exercise which is to provide a useful and practical framework for thinking about the mechanisms underlying the strategy.

1.29 A key distinction to make is between a theory of change and a logic model. While the terms are often used interchangeably, they actually refer to slightly different methods. Our approach harnesses the comparative advantages of each to address the particular challenge of modelling a strategy with the complexity of HWHW. We use the following working definitions:

- **Theory of Change:** A flow chart, diagram or description which displays **how** and **why** the activities that will take place will manifest themselves through changes in outcomes and/or impacts. Theories of Change for the six HWHW sub-themes are detailed in the next section.
- **Logic Model:** A graphic which represents the theory of how an intervention produces its outcomes. It represents, in a **simplified** way, a hypothesis or 'theory of change' about how an intervention works. A logic model has been produced for the overarching HWHW strategy given the complexity of the programme.

1.30 The distinction between these two methods is the depth at which they describe the logical sequence instigated by a policy, programme, event or strategy. A logic model only sets out the logical sequence, the **how**, whereas a ToC includes a consideration of the causal mechanisms which drive the sequence, or the **why**.

The overarching approach

1.31 The Healthy Weight: Healthy Wales strategy is highly complex in that it looks to use a holistic and systems-based approach to tackle obesity. As such, it houses tens of policies and programmes which look to address the issue from a range of angles. To account for this complexity, while providing a structured and useful way to think about the causal mechanisms underlying the strategy's work, we have developed a framework which combines a logic model with six theories of change. The approach combines the breadth which can be achieved using logic models with the depth added by applying theory of change methodology. The logic model provides a broad overview of the strategy, whereas the theories of change go in detail into the policy and programme level causal chains.

Logic Model	Healthy Weight: Healthy Wales Strategy					
Theories of Change	Healthy Food Environments	Healthy Active Environments	Healthy Learning Settings	Healthy Work and Community Settings	Healthy People: Prevention & Early Intervention	Healthy People: Targeted & Specialist Services

- 1.32 The strategy-level logic model provides a simplified representation of the logical sequences underlying the strategy as a whole. The logic model also brings together all the sub-components of the strategy and provides an overview of how they will complement each other in achieving the strategy's goals. In this case, a logic model is the preferable tool as it provides an accessible way to consider the breadth of the strategy and the ways in which different components will work together.
- 1.33 This overarching logic model is complemented by six theories of change which provide a theoretic grounding for the causal mechanisms underlying six of the strategy's principal sub-themes. The ToCs are divided by sub-theme as it represents a level of granularity at which each ToC can comprehensively set out the relevant causal mechanisms at play while remaining sufficiently accessible as to be practically useful. That said, none of the sub-components of the strategy operate in isolation, and so it is implicit that many of the anticipated outcomes and impacts are only achievable as a result of the systems-based approach upon which the strategy is based. This may also have implications for evaluation activities, as the outcomes expected in each project evaluation may differ compared to the overall strategy evaluation which combines all the projects.
- 1.34 The fourth theme of the strategy, leadership and enabling change, is embedded throughout the seven graphics as an enabler of the other themes and the strategy as a whole. This theme is distinct to the others in that it is explicitly tailored towards creating a framework and architecture which enhances the work done under the other themes rather than having standalone policies or programmes itself. While it is equally vital to the strategy's success, its nature is sufficiently distinct as to merit an alternative approach to its inclusion in this framework.

Logic model

- 1.35 The model is principally broken down into three columns and four rows. The rows reflect the activities, outputs, outcomes and impacts structure outlined in paragraph 1.20. Each row houses seven sections covering the six sub-themes as well as the leadership and enabling change component, which, as is described previously, is seen as an enabler that enhances the other three themes. Visually, this interaction with the other three themes is represented by its spanning across the other theme's columns.
- 1.36 Each section lays out the high-level activities, outputs or outcomes for the respective (sub-)themes, with the activities in one sub-theme directly leading to the outputs associated with that sub-theme and so on. Long-term impacts are an exception in that these impacts will only be achieved as a result of the combined outcomes generated by all four of the themes. As a result, they are not attributable to any theme or sub-theme in isolation, but rather to the strategy as a whole. Visually, this is reflected by the absence of any theme or sub-theme related divisions in this row.

	Healthy Environments	Healthy Settings	Healthy People
Activities	Food Environment <ul style="list-style-type: none"> Incentivise responsible advertising. Promote reformulation of food and drink. Create barriers to accessing unhealthy foods. Encourage informative packaging. Support provision of healthy food and drink. 	Healthy Learning <ul style="list-style-type: none"> Promote a healthy diet and early movement in childcare settings. Integrate healthy nutrition physical activity into school curriculums. Provide healthy food and drink and active travel routes connecting schools and communities. 	Prevention and Early Intervention <ul style="list-style-type: none"> Provide new and expectant parents with additional practical and financial support, particularly to those in deprived areas. Encourage communication and engagement between healthcare professionals and the public around the benefits of maintaining a healthy weight.
	Active Environment <ul style="list-style-type: none"> Expand active travel network. Invest in safe cycle routes and pedestrian zones. Create new green spaces and expand access to National Parks and AONB. Improve sports and play facilities. 	Healthy Work & Community Settings <ul style="list-style-type: none"> Embed healthy lifestyle programmes within the workplace. Increase healthy food and active travel options at workplaces. Increase provision of healthy food and drink in community settings. 	Targeted and Specialist Services <ul style="list-style-type: none"> Expand use of the NERS in Wales. Develop dietetic led programmes. Provide specialist weight management services to those in need. Enable access to bariatric surgery for those in need.
	Integrated Leadership and Enabling Change <ul style="list-style-type: none"> Develop a systems-based approach based on the future. 		

Outputs	Food Environment <ul style="list-style-type: none"> Shift towards advertisement of healthier products. More availability of healthy food and drink. Unhealthy products facing extra taxes. More information for consumers. Less access to unhealthy foods in communities. 	Healthy Learning <ul style="list-style-type: none"> More training programmes for early year settings workers. A curriculum which integrates nutrition and physical activity throughout. A larger active travel network around campuses. More provision of healthy food and drink at schools. 	Prevention and Early Intervention <ul style="list-style-type: none"> Widened support for new and expectant parents to encourage positive parenting practices. More channels of communication between healthcare professionals and the public, including the use of informative campaigns.
	Active Environment <ul style="list-style-type: none"> A larger active travel network. More pedestrian and bicycle friendly communities and routes. More and better used green spaces. Additional and improved sports and play facilities. 	Healthy Work & Community Settings <ul style="list-style-type: none"> New healthy lifestyle programmes and policies for workplaces and communities. An expanded active travel network and healthier meal options at workplaces. More healthy food options available in community settings. 	Targeted and Specialist Services <ul style="list-style-type: none"> More referrals via NERS. Expanded provision of weight management services, dietetic led programmes and bariatric surgery, particularly to those in deprived areas.
	Integrated Leadership and Enabling Change <ul style="list-style-type: none"> A systems-based approach to obesity and healthy weight in Wales. Publication of regular technical reports which illustrate progress. 		

Short-term Outcomes	Food Environment <ul style="list-style-type: none"> Increased access to healthy food and drink. Less consumption of saturated fat, salt, and sugar. More informed consumers. 	Healthy Learning <ul style="list-style-type: none"> Active recreation and travel options are an increasingly appealing choice. Children and young people have a better appreciation of what constitutes a healthy diet and a healthy active lifestyle. 	Prevention and Early Intervention <ul style="list-style-type: none"> Expanded access for new and expectant parents to the support they need to make healthy decisions for their children. Healthy parenting choices are increasingly the easy choice.
	Active Environment <ul style="list-style-type: none"> More access areas for active leisure like green spaces, National Parks and AONB Active travel options are increasingly safe, practical, and appealing choices for travel. 	Healthy Work & Community Settings <ul style="list-style-type: none"> Increased access to healthy active travel options for commuting More accessible and affordable healthy food and drink options in workplaces and communities. 	Targeted and Specialist Services <ul style="list-style-type: none"> More access to a range of specialised services for local communities, particularly for communities in areas of deprivation.
	Integrated Leadership and Enabling Change <ul style="list-style-type: none"> An environment in which strong leadership, accountability and governance are evident at every level of the strategy and where they act as a basis for progress along all dimensions 		

Long-term Impacts	Individuals <ul style="list-style-type: none"> Healthier nutrition habits for children and adults. Improved physical and mental wellbeing. More children and adults maintaining a healthy weight. 	<ul style="list-style-type: none"> Reduced child and adult obesity. Reduced incidence of obesity induced chronic disease. More children starting school at a healthy weight.
	Society <ul style="list-style-type: none"> Reduced health and diet inequalities. A more sustainable environment. Less sickness absences in the workplace. Reduced expenditure on obesity-related healthcare costs 	

Theories of change

1.37 The theories of change are all structured in the same way:

- The **problems** highlight the key issues with the status quo in relation to the theme in question.
- There is a **hypothesis** which sets out in words the logic by which the theme's activities will address the problems.
- **Activities and outputs** are as defined in paragraph 1.20.
- **Short-term outcomes** are separated into expanding choices and shaping choices. Instances of **expanding choices** refer to increases in access to certain healthy opportunities or choices, such as an increase in access to high-quality play and sports facilities in areas of deprivation. **Shaping choices** refers to outcomes where an existing healthy choice has been made relatively easier or more appealing when compared with an existing unhealthy choice, such as consumers being more aware of the calorie content of their purchases.
- **Long-term impacts** are separated into **improved wellbeing**, which refers to benefits for individuals, and **wider benefits for society**, which refers to benefits for society as a whole. While each of the activities undertaken as part of the strategy can be at least tangentially linked to each of the long-term benefits identified in the logic model, only those which can be directly traced back to a particular theme's activities are included in the ToCs.
- The **integrated leadership & enabling change** section is consistent across the six ToCs and sets out the ways in which this area of the strategy will underlie and enhance the work done in each of the other sub-themes.

1.38 **Wider influences** represent the national and global trends which are out of control of policy makers or delivery partners, and which will act as either enablers or constraints for the sub-theme to achieve its goals. The hypotheses as well as the links between the strategy's activities and the eventual impacts are supported by evidence from the academic and grey literature. A summary of key findings are listed below, with the full list of sources available in Annex A.

Sub theme	Summary of findings
Food environment	<ul style="list-style-type: none"> Adults exposed to food advertising increased their consumption of unhealthy snacks by 28%.³ Taxation on unhealthy food and drink led to reduced consumption of these products and led to firms reformulating their products.⁴ Calorie labelling led to a decrease in calorie consumption.⁵
Active environment	<ul style="list-style-type: none"> Population-level health benefits of active travel include reduced obesity and increased rates of physical activity.⁶ The Safe Routes to School program in the US led to up to a 25% increase in walking and cycling.⁷ Evidence suggests that local-area proportions of green space may be associated with happiness and life satisfaction.⁸ Attending a school with sub-standard infrastructure and facilities decreases the probability of participating in physical activity in adulthood. Physical inactivity increases the risk of cancer, heart disease, stroke and diabetes by 20–30%.⁹
Healthy learning settings	<ul style="list-style-type: none"> Healthy childcare settings: providing children with healthy meals led to higher diet quality scores.¹⁰ Healthy schools: schools with bad infrastructure and facilities decreased the probability of doing physical activities in adulthood.¹¹
Healthy work and community settings	<ul style="list-style-type: none"> Provision of healthy food services on worksites led to more consumption of fruits and vegetables and less consumption of fast food.¹²
Prevention and early intervention	<ul style="list-style-type: none"> Substandard diet quantity and/or quality contributes to the loss of developmental potential and life-long health and economic disparities.¹³
Targeted & specialised services	<ul style="list-style-type: none"> Substandard diet quantity and/or quality contributes to the loss of developmental potential and life-long health and economic disparities.¹⁴

³ Zimmerman, F.J. and Shimoga, S.V., 2014. The effects of food advertising and cognitive load on food choices. *BMC Public Health*, 14(1), pp.1-10.

⁴ Economics observatory, 2021. Would taxes on unhealthy foods reduce obesity?

⁵ Crockett, R., King, S., Marteau, T., Prevost, A., Bignardi, G., Roberts, N., Stubbs, B., Hollands, G., Jebb, S., 2018. Nutritional labelling for healthier food or non-alcoholic drink purchasing and consumption. *Cochrane Database of Systematic Reviews* 2018, Issue 2.

⁶ Pucher, J., Buehler, R., Bassett, D.R. and Dannenberg, A.L., 2010. Walking and cycling to health: a comparative analysis of city, state, and international data. *American journal of public health*, 100(10), pp.1986-1992.

⁷ McDonald, N.C., Steiner, R.L., Lee, C., Rhoulac Smith, T., Zhu, X. and Yang, Y., 2014. Impact of the safe routes to school program on walking and bicycling. *Journal of the American Planning Association*, 80(2), pp.153-167.

⁸ Houlden, V., Weich, S. and Jarvis, S., 2017. A cross-sectional analysis of green space prevalence and mental wellbeing in England. *BMC public health*, 17(1), pp.1-9.

⁹ (a) Black, N., Johnston, D.W., Propper, C. and Shields, M.A., 2019. The effect of school sports facilities on physical activity, health and socioeconomic status in adulthood. *Social Science & Medicine*, 220, pp.120-128.

(b) WHO, 2022. Physical activity.

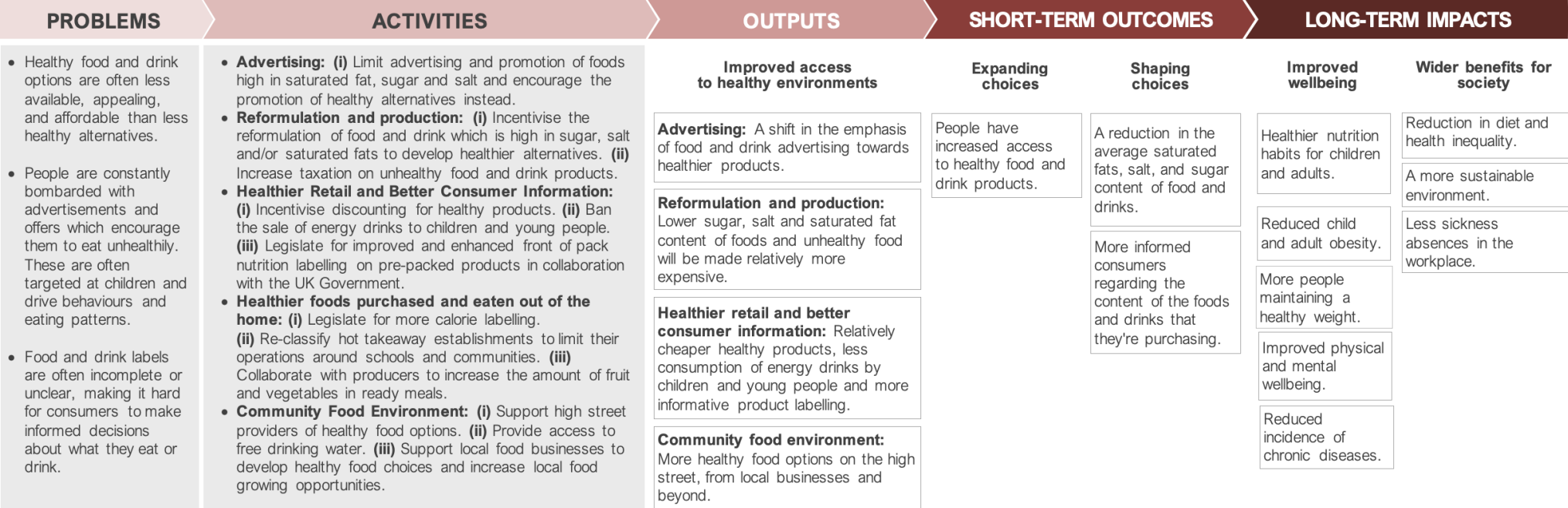
The Theories of Change for each of these sub-themes are as follows.

-
- ¹⁰ Yoong, S.L., Grady, A., Seward, K., Finch, M., Wiggers, J., Lecathelinais, C., Wedesweiler, T. and Wolfenden, L., 2019. The impact of a childcare food service intervention on child dietary intake in care: an exploratory cluster randomized controlled trial. *American Journal of Health Promotion*, 33(7), pp.991-1001.
- ¹¹ Black, N., Johnston, D.W., Propper, C. and Shields, M.A., 2019. The effect of school sports facilities on physical activity, health and socioeconomic status in adulthood. *Social Science & Medicine*, 220, pp.120-128.
- ¹² Dodson, E.A., Hipp, J.A., Gao, M., Tabak, R.G., Yang, L. and Brownson, R.C., 2016. The impact of worksite supports for healthy eating on dietary behaviors. *Journal of occupational and environmental medicine/American College of Occupational and Environmental Medicine*, 58(8), p.e287.
- ¹³ Hurley, K., Yousafzai, A. and Lopez-Boo, F., 2016. Early Child Development and Nutrition: A Review of the Benefits and Challenges of Implementing Integrated Interventions. *Advances in Nutrition*, Volume 7, Issue 2.
- ¹⁴ Douglas, I.J., Bhaskaran, K., Batterham, R.L. and Smeeth, L., 2015. Bariatric surgery in the United Kingdom: a cohort study of weight loss and clinical outcomes in routine clinical care. *PLoS medicine*, 12(12), p.e1001925.

Healthy Food Environments

Hypothesis

If food environments are redesigned such that healthy foods can be made more available, affordable, and appealing, making the healthy choice will become the obvious choice. Furthermore, by restricting the advertising of unhealthy products and by strengthening the labelling requirements for food and drink, consumers will be empowered to make more informed choices at the shops and supermarkets.



INTEGRATED LEADERSHIP & ENABLING CHANGE

- The Wellbeing of Future Generations (Wales) Act:** The Act's five ways of working (long-term, prevention, integration, collaboration, and involvement) provides the right tools and frameworks to the organisations and partnerships supporting the strategy to demonstrate how they are delivering and fulfilling obligations.
- Systems Based Approach:** Delivery partners will be encouraged to sign up to a charter which aims to empower healthier lifestyles and promote physical activity. This charter, through the National Implementation Board, will hold to account each partner who can support and enable change to happen across Wales.
- Communications and Engagement:** A communications plan will help bring healthy diets and active lifestyles to the forefront of the national discourse. This will be combined with a strong engagement approach that listens and conveys key messages to partners and communities.

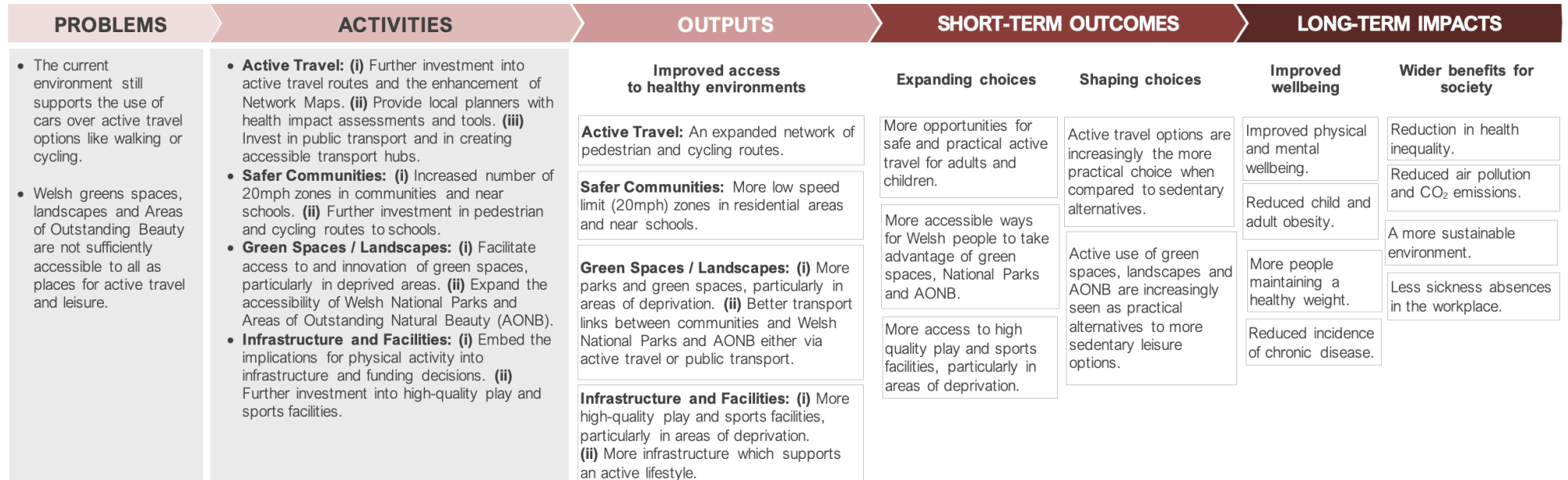
WIDER INFLUENCES

- The cost-of-living crisis will make any healthy food or drink options which are more expensive than unhealthy alternatives less accessible and appealing to consumers.
- The perception of what constitutes a healthy weight has been changing over time to the point where unhealthy weights are now often perceived as 'normal'.
- (Covid-19) The increased vulnerability of obese individuals to Covid-19 has increased awareness of the dangers of obesity.
- (Covid-19) Lockdown restrictions may have created persistent shocks in eating habits, both positive and negative.

Healthy Active Environments

Hypothesis

By continuing to expand the active travel infrastructure in Wales and by making active travel options safer, more accessible, and more affordable, active travel will become the obvious choice over more sedentary alternatives. Furthermore, if the accessibility of Welsh green spaces, National parks and Areas of Outstanding Natural Beauty is improved, many more Welsh people will be able to take advantage of these places for active outdoor activities and leisure.



INTEGRATED LEADERSHIP & ENABLING CHANGE

- The Wellbeing of Future Generations (Wales) Act:** The Act's five ways of working (long-term, prevention, integration, collaboration, and involvement) provides the right tools and frameworks to the organisations and partnerships supporting the strategy to demonstrate how they are delivering and fulfilling obligations.
- Systems Based Approach:** Delivery partners will be encouraged to sign up to a charter which aims to empower healthier lifestyles and promote physical activity. This charter, through the National Implementation Board, will hold to account each partner who can support and enable change to happen across Wales.
- Communications and Engagement:** A communications plan will help bring healthy diets and active lifestyles to the forefront of the national discourse. This will be combined with a strong engagement approach that listens and conveys key messages to partners and communities.

WIDER INFLUENCES

- There is a trend towards children and young adults spending more time playing videogames and using screens instead of participating in outdoor activities and this often persists into adulthood.
- The perception of what constitutes a healthy weight has been changing over time to the point where unhealthy weights are now often perceived as 'normal'.
- The cost-of-living crisis will make paying for the travel and equipment associated with participation in certain sports and activities more difficult for everyone but particularly for lower income families.
- (Covid 19) Lockdown restrictions may have created persistent shocks in participation in outdoor activities, both positive and negative.
- (Covid-19) The increased vulnerability of obese individuals to Covid-19 has increased awareness of the dangers of obesity.

Healthy Learning Settings

Hypothesis

If the benefits of a healthy diet and an active lifestyle are given increased prominence in the Welsh curriculum and learning settings are reimagined to make healthy choices easier, then children and young people will be better prepared to make healthy choices in their youth and as they progress into adulthood.

PROBLEMS	ACTIVITIES	OUTPUTS	SHORT-TERM OUTCOMES		LONG-TERM IMPACTS	
<ul style="list-style-type: none"> Currently, many learning settings are not set up with physical activity and healthy eating in mind and this leads to unnecessary obstacles between students and making healthy choices. Topics related to healthy nutrition, food preparation and physical activity are not always given sufficient priority in the curriculum to support children and young people in making healthy choices both in their youth and later in life. 	<ul style="list-style-type: none"> Healthy Childcare Settings: (i) Support childcare settings to align themselves with Welsh Government's food, nutrition and physical activity guidance. (ii) Support interventions that promote early movement and play. Healthy Schools: (i) Modify curriculum to include more topics related to healthy nutrition and physical activity. (ii) Train staff so that they can support pupil learning about food, physical activity, nutrition and hydration. (iii) Develop a range of integrated programmes that promote healthy eating, physical activity, recreation and school sport. (iv) Integrate considerations of food and physical activity into the assessment criteria for schools. Healthy Higher, Further and Other Educational Settings: (i) Incentivise provision of healthy food and drink in campuses and sites. (ii) Carry out targeted campaigns aimed at promoting a healthy lifestyle choices. (iii) Improve active travel routes in campuses and encourage physical activity among students. 	<p>Improved access to healthy settings</p> <p>Healthy Childcare Settings: More training programmes on food, nutrition and physical activity for early year settings workers. (ii) More interventions which promote early movement and play.</p> <p>Healthy Schools: (i) A curriculum which informs children and young people about the benefits of a healthy diet and an active lifestyle and teachers who are trained to deliver the curriculum effectively. (ii) A range of integrated programmes that promote healthy eating, physical activity, recreation and school sport. (iii) A school assessment criteria which rewards schools for providing effective education on the benefits of a healthy diet and an active lifestyle.</p> <p>Healthy Higher, Further and Other Educational Settings: (i) More healthy food and drink options on campuses including more fruits and vegetables on canteen menus. (ii) More active travel routes to campuses.</p>	<p>Expanding choices</p> <p>More accessible and appealing opportunities for children and young people to participate in physical activity and sports.</p> <p>More accessible and affordable healthy food and drink options at schools.</p> <p>More opportunities for safe active travel to schools.</p>	<p>Shaping choices</p> <p>Active recreation and travel options are increasingly the more appealing choice when compared to sedentary alternatives.</p> <p>Children and young people will grow up with a better appreciation of what constitutes a healthy diet and a healthy active lifestyle.</p>	<p>Improved wellbeing</p> <p>Healthier nutrition habits for children.</p> <p>Improved physical and mental wellbeing for children.</p> <p>More children starting school at a healthy weight.</p> <p>More children maintaining a healthy weight.</p> <p>Reduced child obesity.</p> <p>Reduced incidence of chronic diseases in children.</p>	<p>Wider benefits for society</p> <p>Reduction in diet and health inequality.</p> <p>A more sustainable environment.</p> <p>Less sickness absence at school.</p>

INTEGRATED LEADERSHIP & ENABLING CHANGE

- The Wellbeing of Future Generations (Wales) Act:** The Act's five ways of working (long-term, prevention, integration, collaboration, and involvement) provides the right tools and frameworks to the organisations and partnerships supporting the strategy to demonstrate how they are delivering and fulfilling obligations.
- Systems Based Approach:** Delivery partners will be encouraged to sign up to a charter which aims to empower healthier lifestyles and promote physical activity. This charter, through the National Implementation Board, will hold to account each partner who can support and enable change to happen across Wales.
- Communications and Engagement:** A communications plan will help bring healthy diets and active lifestyles to the forefront of the national discourse. This will be combined with a strong engagement approach that listens and conveys key messages to partners and communities.

WIDER INFLUENCES

- There is a trend towards children and young adults spending more time playing videogames and using screens instead of participating in outdoor activities and it often persists into adulthood.
- The perception of what constitutes a healthy weight has been changing over time to the point where unhealthy weights are now often perceived as 'normal'.
- The cost-of-living crisis will make any healthy food or drink options which are more expensive than their unhealthy alternatives less accessible and appealing to consumers.
- The cost-of-living crisis will make paying for the travel and equipment associated with participation in certain sports and activities more difficult for everyone but particularly for lower income families.
- (Covid-19) The increased vulnerability of obese individuals to Covid-19 has increased awareness of the dangers of obesity.
- (Covid-19) Lockdown restrictions may have created persistent shocks in participation in outdoor activities, both positive and negative.

Healthy Work & Community Settings

Hypothesis

If Welsh workplaces are incentivised and supported to create a healthier setting for their employees then this will help support the adult population in Wales to remain active, to promote community engagement and to maintain a healthy weight.

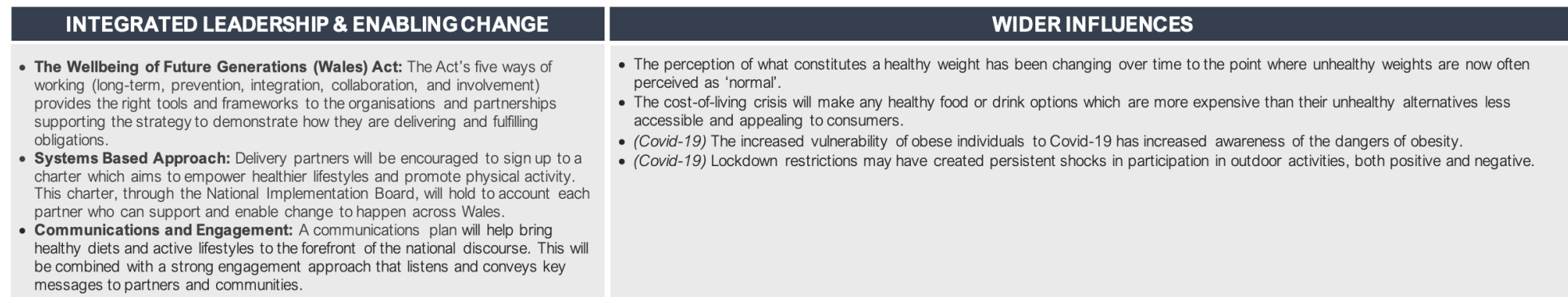
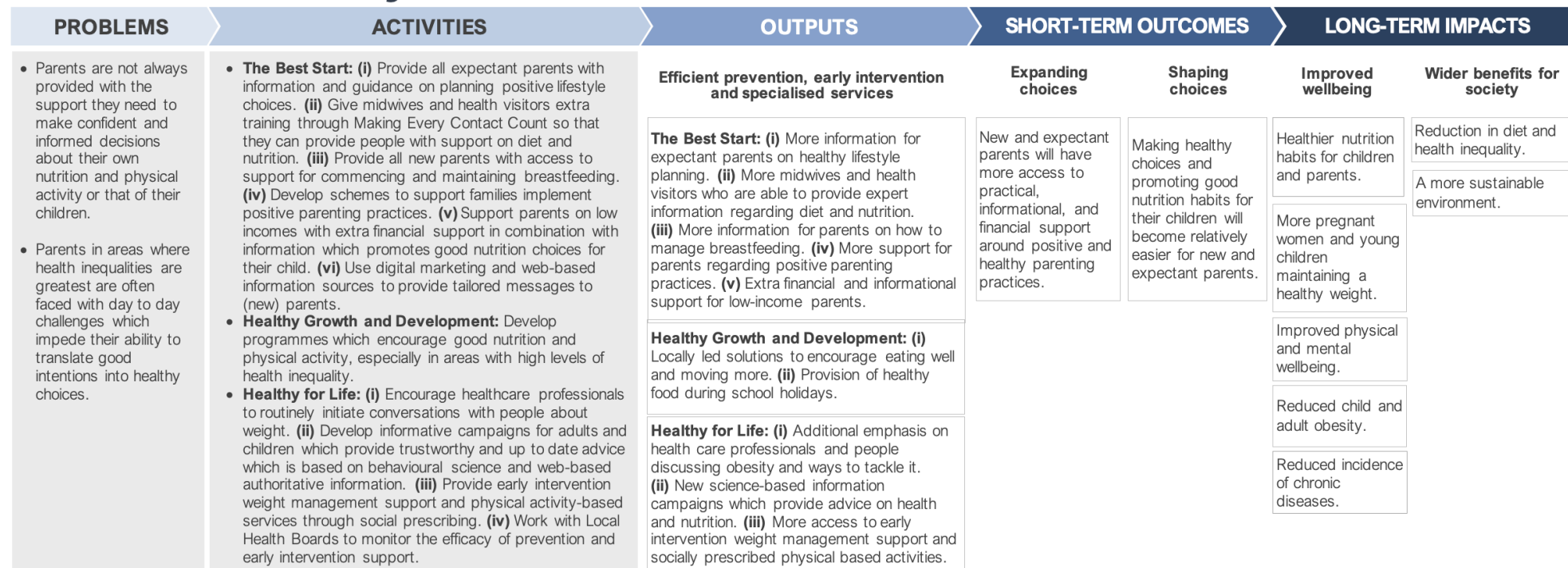
PROBLEMS	ACTIVITIES	OUTPUTS	SHORT-TERM OUTCOMES		LONG-TERM IMPACTS	
<ul style="list-style-type: none"> Work and community settings are not always focused on promoting a healthy lifestyle. The food and drink offered are generally unhealthy and workplaces are often geared towards having a sedentary lifestyle. 	<ul style="list-style-type: none"> Healthy Workplaces: (i) Support the enhancement of workplace healthy lifestyle programmes and policies. (ii) Pilot and embed programmes with public sector employers to test food and physical activity initiatives. (iii) Use the health and wellbeing national award status to motivate workplaces to meet the food and drink provision criteria. Healthy NHS: (i) Develop and implement strict criteria which ensure healthy food and drink provision at NHS sites. (ii) Develop clear active travel plans for all NHS sites. (iii) Stipulate that Local Health Boards and Trusts have to demonstrate progress in health and activity related support as well as weight management services to their employees. Healthy Communities: (i) Increase healthy food and drink provision across settings such as leisure, cinemas and community centres. (ii) Explore how a variety of levers could change the way in which food selling practices operate. 	<p>Improved access to healthy settings</p> <p>Healthy Workplaces: (i) More and wider reaching healthy lifestyle programmes and policies for workplaces and communities. (ii) Innovative food and physical activity initiatives in public sector workplaces. (iii) A wider use of the wellbeing national award status to motivate better workplace adherence to food and activity standards.</p> <p>Healthy NHS: (i) More healthy food options at NHS sites (canteens, vending machines etc.) with more servings of fruits and vegetables. (ii) More active travel options to NHS sites. (iii) A monitoring system to assess Local Health Boards and Trusts level of support to their staff around health and activity.</p> <p>Healthy Communities: (i) More healthy food and drink options in settings such as leisure, cinema and community centres. (ii) New tools to influence food selling practices.</p>	<p>Expanding choices</p> <p>More accessible and affordable healthy food and drink options in workplaces and communities.</p> <p>More accessible healthy active travel options for commuting.</p>	<p>Shaping choices</p> <p>Healthy food and drink options are increasingly the more appealing choice.</p> <p>Active travel options are increasingly the more appealing choice when compared to more sedentary commute modalities.</p>	<p>Improved wellbeing</p> <p>Healthier nutrition habits.</p> <p>More people maintaining a healthy weight.</p> <p>Reduced child and adult obesity.</p> <p>Improved physical and mental wellbeing.</p> <p>Reduced incidence of chronic diseases.</p> <p>An increase in job satisfaction.</p>	<p>Wider benefits for society</p> <p>Reduction in diet and health inequality.</p> <p>A more sustainable environment.</p> <p>Less sickness absences in the workplace.</p>

INTEGRATED LEADERSHIP & ENABLING CHANGE	WIDER INFLUENCES
<ul style="list-style-type: none"> The Wellbeing of Future Generations (Wales) Act: The Act's five ways of working (long-term, prevention, integration, collaboration, and involvement) provides the right tools and frameworks to the organisations and partnerships supporting the strategy to demonstrate how they are delivering and fulfilling obligations. Systems Based Approach: Delivery partners will be encouraged to sign up to a charter which aims to empower healthier lifestyles and promote physical activity. This charter, through the National Implementation Board, will hold to account each partner who can support and enable change to happen across Wales. Communications and Engagement: A communications plan will help bring healthy diets and active lifestyles to the forefront of the national discourse. This will be combined with a strong engagement approach that listens and conveys key messages to partners and communities. 	<ul style="list-style-type: none"> The perception of what constitutes a healthy weight has been changing over time to the point where unhealthy weights are now often perceived as 'normal'. The cost-of-living crisis will make any healthy food or drink options which are more expensive than their unhealthy alternatives less accessible and appealing to consumers. The cost-of-living crisis will make paying for the travel and equipment associated with participation in certain active leisure and active travel activities more difficult for everyone but particularly for lower income families. (Covid-19) The increased vulnerability of obese individuals to Covid-19 has increased awareness of the dangers of obesity. (Covid-19) Lockdown restrictions may have created persistent shocks in participation in outdoor activities, both positive and negative.

Healthy People: Prevention & Early Intervention

Hypothesis

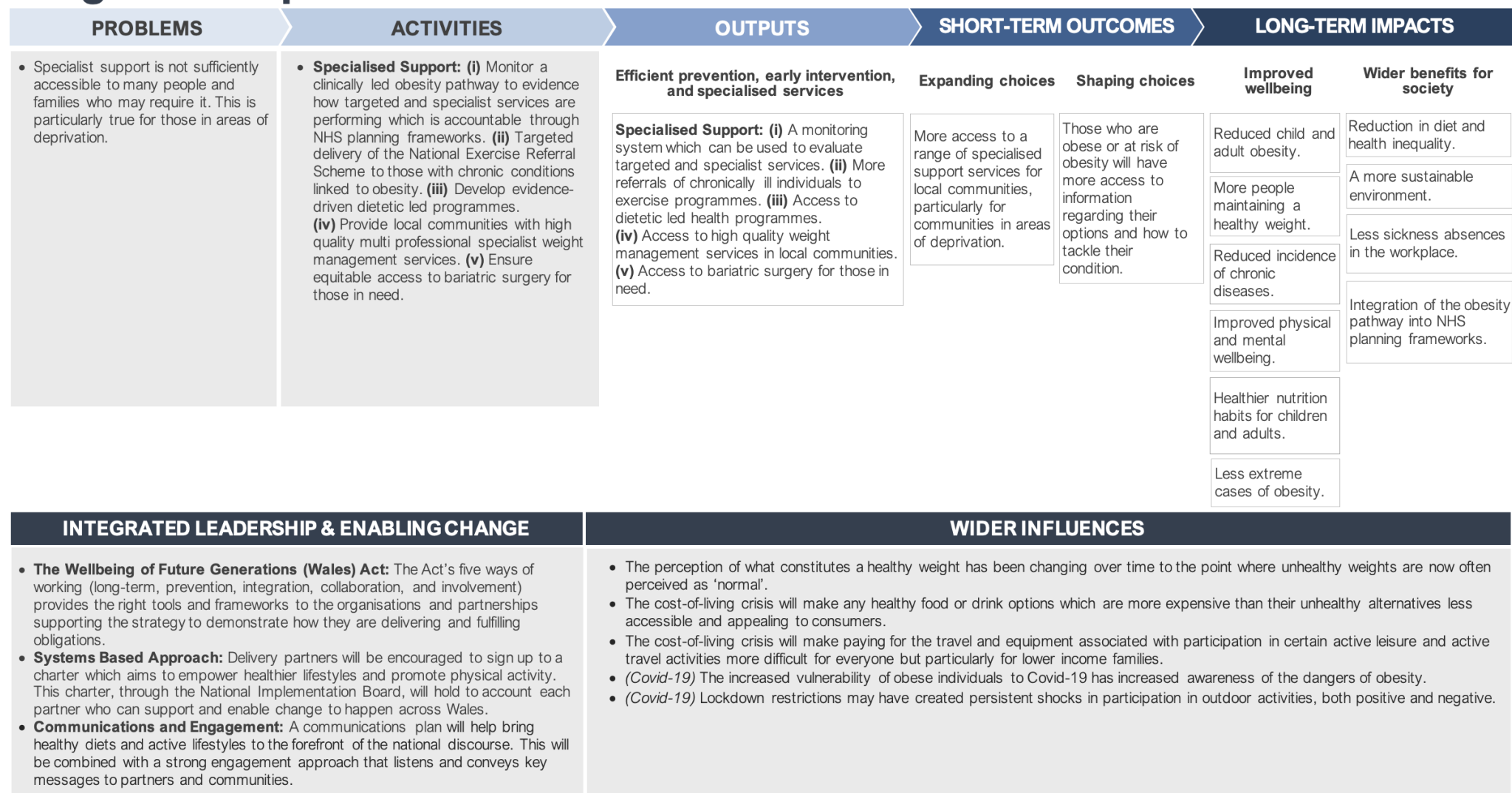
If new and expectant parents, and particularly those in areas of deprivation, are provided with more support regarding the healthy choices they can make for their children, they will be more able to establish healthy habits at this critical part of childhood. Moreover, by instilling these behaviours at an early age they are more likely to persist into adulthood.



Healthy People: Targeted & Specialist Services

Hypothesis

Providing targeted and specialist support to local communities will help prevent cases of obesity and, where obesity is already present, these services can reduce the chances of their becoming extreme cases.



5. Evaluation framework

Evaluation types

- 1.39 According to HM Treasury's Magenta Book (2020), there are three main types of evaluation activity: process evaluation, impact evaluation and value-for-money (economic) evaluation. Our framework also considers outcome evaluation as a fourth evaluation type which can be seen as a sub-category of impact evaluation that focuses on short-run changes in outcomes (Public Health England, 2018). These four categories seek to address the following questions:

Process – What can be learned from how the intervention was delivered?

Outcome – What difference has an intervention made in the short run?

Impact – What difference has an intervention made in the long run?

Economic – Was this intervention a good use of resources?

- 1.40 The reason we include outcome evaluation separately here is due to the mismatch between the standard time horizon of an impact evaluation and the strategy's scope for incorporating evaluation results into strategic design decisions in the short to medium run. Considering some degree of short-term impact (outcome) evaluation adds explicit scope for this kind of evaluation to feed into the design of later delivery plans.
- 1.41 HM Treasury's Magenta Book (2020) explains process, impact and economic evaluation in great detail, and so this section will only focus on their most important aspects and techniques for the evaluation of HWHW.

Process evaluation

- 1.42 Process evaluation, which is sometimes referred to as implementation evaluation, assesses the degree to which an intervention is being implemented as intended, whether the design is working and which aspects of it are working less well and why. The table below outlines suggested areas that could be covered in a process evaluation.

Dimension	Example questions for process evaluation
Criteria for success	<ul style="list-style-type: none"> • Is there a clear description of the implemented programme and what it is trying to achieve? • How is success defined? • Is there a shared understanding of key objectives and desired outcomes across stakeholders? • How are the objectives and intended outcomes reflected in how the success of the intervention will be determined?
Learning lessons on implementation	<ul style="list-style-type: none"> • What appeared to work well? What did not? • What lessons have you learned? What could be done differently next time? • What barriers still exist? What ought to be changed or improved? • Has there been any inconsistencies in roll out across different groups?
Information dissemination and education	<ul style="list-style-type: none"> • What method has been used for passing on funding to providers? • What appeared to work well? What did not? • What lessons were learned? What could be done differently next time? • What barriers still exist? What would you want to see changed or improved? • Has there been any inconsistencies in roll out across different groups?
Monitoring and continuous learning	<ul style="list-style-type: none"> • How is success defined? What desired outcomes would this involve? • Is there a coherent plan to collect data for the purpose of ongoing monitoring? • Would the collected data enable success to be measured? • To what extent can data be consistently and periodically collected reliably at a sufficient level of detail and stratification? • What processes are in place to enable data reporting and escalation to the right governance bodies if needed?

1.43 Process evaluations typically use a mixed-methods approach and will cover both subjective and objective measures of how well an intervention has been implemented. This holistic approach to evaluation will be a key characteristic of

any evaluation of HWHW given the human-centric nature of the changes which the strategy aims to achieve. That said, certain qualitative evaluation techniques can be particularly resource intensive, and so resource constraints may dictate that techniques which involve direct stakeholder engagement (i.e., interviews, focus groups and case studies) might need to be prioritised for certain components of the strategy over others.

- 1.44 Another key consideration for process evaluation in the context of HWHW is going beyond assessing which aspects of an intervention are working more or less well and to consider how this differs between demographic and regional groups. Ensuring that an assessment of these differences is included in any analysis will be a crucial part of assessing the strategy's progress towards two of its core goals, reducing health inequality and reducing diet inequality. This is discussed further in Section 9.

Outcome and impact evaluation

- 1.45 Impact evaluation can be done using a broad range of methods, with the balance between quantitative and qualitative methods largely dependent on the availability and quality of data. With that in mind, it is likely that any impact evaluation of the HWHW strategy or its components will involve some degree of a mixed-methods approach.
- 1.46 From a quantitative perspective, methods like regression discontinuity design (RDD),¹⁵ difference-in-differences,¹⁶ propensity score matching¹⁷ and synthetic control methods¹⁸ provide feasible ways of causally estimating the impact of certain policies/programmes as well as that of the wider strategy on a range of outcomes.
- 1.47 Qualitatively, the use of tools such as case studies, observational studies, interviews and focus groups will add an additional layer of context to the quantitative work undertaken. Many measures of progress are challenging to

¹⁵ RDD is a quasi-experimental method that aims to determine the causal effects of interventions by assigning a cut off or threshold above or below which an intervention is assigned.

¹⁶ This is a quasi-experimental approach that compares the changes in outcomes over time between a population enrolled in a programme (the treatment group) and a population that is not (the comparison group).

¹⁷ A method in which the researcher uses statistical techniques to construct an artificial control group by matching each treated unit with a non-treated unit of similar characteristics. Using these matches, the researcher can estimate the impact of an intervention.

¹⁸ This evaluates a treatment effect in comparative case studies to determine the impact of an intervention. It creates a synthetic version of treated units by weighting variables and observations in the control group.

- estimate empirically and so the experiences and opinions of stakeholders often provide insight which would be missed when considering quantitative data alone.
- 1.48 Theory based evaluation could be considered where there are significant limitations to conducting quasi-experimental techniques, particularly where there are data limitations, budget limitations or where there is a need to explore why an intervention would not work, which tends to not be considered in experimental studies. Theory based evaluations are conducted by first developing a Theory of Change or Logic models, as done in Section 4 and then gathering existing evidence and conducting analysis (where possible) to assess the cause-and-effect mechanism in question.
- 1.49 Both quantitative and qualitative impact evaluation methods can provide great scope for considering the inequality implications of an intervention. By considering how impacts differ between demographic and regional groups, they can provide insight into who is benefiting most from the strategy and, consequently, provide evidence of the strategy's impacts on health and diet inequality. Practical considerations of determining these groups and how they should be engaged are discussed in Section 9.
- 1.50 Short-term outcome evaluation will involve the same methods set out in this section, but will likely lean more heavily on qualitative techniques than is the case for long-term impact evaluation. This is due to the lack of appropriate data available at the point at which an outcome evaluation would be conducted in the context of HWHW.

Economic evaluation

- 1.51 Economic evaluation, otherwise known as value-for-money evaluation, compares the costs and benefits of interventions to assess the degree to which they provide a return on investment. The scope of these evaluations can vary substantially, with a key determinant of scope being the way in which benefits are defined. A wider, more holistic evaluation may look to monetise social welfare benefits, whereas a more limited analysis may only concern benefits which are strictly economic.
- 1.52 This wider evaluation methodology is often referred to as cost-benefit analysis (CBA), which is likely to be the most appropriate tool for economic evaluation in the context of HWHW. With cost-benefit analysis, it is possible to capture the

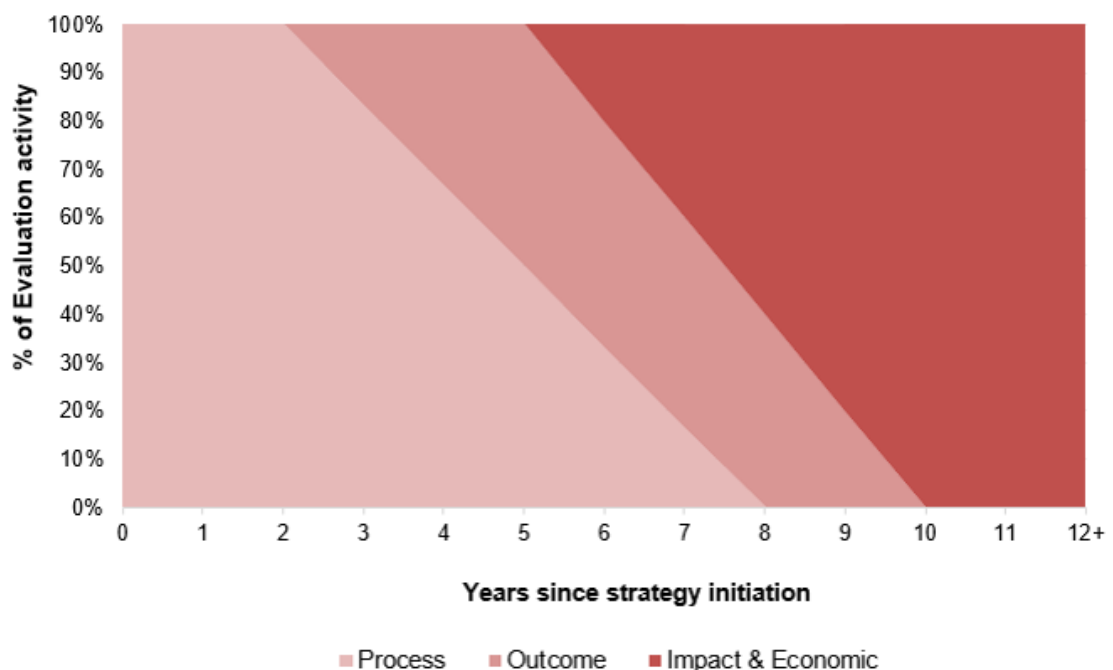
wider benefits to society related to reductions in inequality, air pollution and CO₂ emissions. HM Treasury's Green Book provides comprehensive guidance on how best to estimate such benefits as well as on how to design and execute a robust CBA.

- 1.53 The quality of a Green Book compliant CBA is largely dependent on the quality of the data inputs it is based on. In the preferred scenario, quantified estimates of the studied intervention's impacts are derived from experimental, quasi-experimental or synthesis impact evaluation methods. This highlights the need for economic evaluation to be carried out in tandem with rigorous impact evaluation. In the absence of this evidence, the grey and academic literatures can provide an evidence base of relevant studies carried out in comparable contexts which can be used to estimate benefits, however this approach relies on many additional assumptions.
- 1.54 Another facet of CBA which will likely be useful for any economic evaluation of HWHW is distribution analysis. This technique can be used to consider costs and benefits as they relate to different population groups and geographies, thus it can contribute to the understanding of the distributional effects of an intervention. In practice, this includes simply appraising and comparing the benefits and costs for different groups that are deemed relevant for the objectives of the intervention or strategy. [HM Treasury's Green Book](#) suggests weightings could be applied to the costs/benefits of each group to enable easy aggregation into a single quantified estimate. For example, assigning higher weight to low income households to reflect a desire for income redistribution would mean that any £1 benefit for lower income households would mean more than £1 of benefits for higher income households.
- 1.55 Updated guidance from the Medical Research Council (2021) recommends that 'broad ranging approaches such as cost benefit analysis or cost consequence analysis, which seek to capture the full range of health and non-health costs and benefits across different sectors, will often be more suitable for an economic evaluation of a complex intervention than narrower approaches such as cost effectiveness or cost utility analysis'.

Evolution of the evaluation

- 1.56 An important consideration for the overarching evaluation approach is around how resources should be allocated between the different types of evaluation at different points in the strategy's lifespan. This decision should be made based on two principal factors:
- The need to ensure that, as a whole, the evaluation provides a holistic picture of the strategy's performance.
 - The availability of data at different points in the strategy's lifespan.
- 1.57 It is not feasible to provide exact percentages for how much of the budget should be spent on each evaluation type ex-ante as doing so would require knowledge on the exact makeup of the strategy in terms of policies and programmes, something which will evolve over time. That said, it is possible to provide general guidance regarding the rough distribution of spending between evaluation types and how this will develop over the strategy's lifespan. Figure 1 provides stylised guidance on how the focus of the evaluation should shift over time.¹⁹

Figure 1 - Stylised evolution of the evaluation



¹⁹ Impact and economic evaluation are grouped together here due to the overlapping data requirements of the two evaluation types. See Evaluation Types for further explanation.

- 1.58 Across the strategy's lifespan, it is expected that the evaluation's emphasis will shift from process evaluation early on towards impact and economic evaluation as the requisite data becomes available. Outcome evaluation will form a bridge between these two phases by providing evidence of any emerging changes in outcomes at the earliest opportunity, such that these findings can be used to inform the strategy's final delivery plans.
- 1.59 **Years 0 to 2.** Initially, it will make sense to prioritise process evaluation for two reasons. Firstly, doing so will allow practitioners to determine which policies and programmes are working well and why, and adjust future delivery plans accordingly. Secondly, it is possible to collect process evaluation data at a much earlier stage than is the case for outcome data, which allows for process evaluation to be completed much sooner than is the case for other evaluation types.
- 1.60 **Years 2 to 5.** After this initial phase, the changes in outcomes driven by the strategy's early projects and programmes will be sufficiently mature for the first instances of outcome evaluation to take place. These early outcome evaluations will provide practitioners with an early indication of how effective the strategy's initial policies have been at engendering change. During this interim period, spending will be split between process and outcome evaluation, with outcome evaluation forming a larger share of spending over time.
- 1.61 **Years 5 to 8.** From year 5 onwards, the groundwork for impact and economic evaluation can begin to be put into place. It likely will not be possible to begin impact and economic evaluation until this point in the strategy's lifespan due to the delay between project implementation and the point at which impacts might be visible in the data. Ambitious impact evaluations may begin as early as year 6 with preparation beginning a year earlier. During this period, the use of process evaluations will be phased out as the scope for feeding their results into future delivery plans is diminished. It may also be possible to conduct small-scale piloting of approaches to explore what is practically feasible with the data and identify any issues. Outcome evaluation will continue to provide early indications of the effectiveness of certain projects and programmes.

- 1.62 **Years 8 to 10.** Between years 8 and 10, the spending will be split between outcome, impact and economic evaluation, with the emphasis shifting towards impact and economic evaluation as the requisite data becomes increasingly available.
- 1.63 **Year 10 onwards.** For the years after the strategy's conclusion, it will make sense to conduct impact and economic evaluation of the strategy as a whole. Follow-up evaluations will allow for an evaluation of the degree to which the impacts identified in the first round of impact evaluation were persistent.
- 1.64 The stylised guidance is, of course, just that, and how the strategy develops over time will have a determining influence on the way the evaluation evolves. As an example, in the case where the strategy vastly expands its activities in its final years, it may make sense to postpone the majority of the impact and economic evaluation spending until even later than is prescribed to ensure that the impact of these activities is captured.
- 1.65 The MRC (2021) suggests that: 'economic modelling could be undertaken at the feasibility stage to assess the likelihood that the expected benefits of the intervention justify the costs (including the cost of further research), and to help decision makers decide whether proceeding to a full scale evaluation is worthwhile. Depending on the results of the feasibility study, further work might be required to progressively refine the intervention before embarking on a full scale evaluation'.

Databank

- 1.66 As part of the evaluation framework, we have constructed a databank which seeks to map the data universe in Wales as it relates to the HWHW strategy. The idea is for the databank to act as a first point of reference for future evaluators as they look to design their evaluation. The databank also serves as a tool for highlighting data gaps and areas in which the Welsh Government might feasibly be able to collect additional primary data. The databank can be found in Annex B.
- 1.67 The databank is designed to be as user friendly as possible and is broken down by evaluation type such that an evaluator can focus on the data sources which are most likely to be relevant for their particular evaluation. Furthermore, the first column highlights the sub-theme and activity (as defined in the theories of change) for which the dataset is most likely to be applicable; this again narrows

the evaluator's search. The tables also set out the granularity of data, the indicators available from the data source which are most likely to be relevant and where data for a counterfactual scenario might be sourced.

- 1.68 The data sources collated here include national statistics, survey data, administrative data, private sector data and non-traditional sources of big data (e.g., google searches or satellite data). There are also suggestions for new areas of data collection which are marked as "Suggested" in the source column.
- 1.69 The approach to data collation varies by evaluation type given their different data demands. For process evaluation, it is standard practice to collect primary monitoring data for certain key performance indicators (KPIs), so in a lot of cases it will not be necessary to rely on existing data sources to support these kinds of evaluations. With that in mind, we focused on adding a few measures which could be used in addition to any monitoring data and on certain aspects of the strategy where a different approach might be required (i.e., the travel network and the prevalence of green spaces). Outcome and impact evaluation indicators have been combined into one table given the overlap in the outcomes considered in each of these evaluation types. In this section, we have included as broad a range as possible of indicators and measurements that could be used to estimate the changes in outcomes and behaviours which have been induced by the strategy and its sub-components. In the case of economic evaluation, we sourced data which would be most useful for monetising the long-term impacts of the strategy set out in the logic model and theories of change.
- 1.70 During our scoping exercise, we identified a number of data gaps and specific areas where additional data collection could be of great value for the evaluation of HWHW. In some of these cases, such as for process evaluation in general, this gap was expected as, by definition, process evaluations heavily rely on the collection of primary data during the programme's/policy's operational period.
- 1.71 In other cases, such as in the case of representative measurements of BMI in Wales, this gap is more troublesome given the measurement's importance to the strategy. While relevant data sources exist, such as the National Survey for Wales and the Child Measurement Programme (CMP), both have their limitations. The National Survey for Wales is self-reported data and the sample is not large enough to be representative at the local authority level. The CMP only

considers children aged 4 to 5, so the picture of what is happening with the rest of the child population is left blank.

1.72 In summary, the areas where our data scoping exercise has highlighted two primary data gaps which could feasibly be filled:

- *Measurements of height and weight for a representative sample of adults in Wales.* GPs and other primary care services provide a feasible avenue for generating an additional dataset of height and weight in Wales. Height and weight measurements may also need some standardisation of measurement units.
- *Measurements of outcomes related to the healthy work and community settings sub-theme.* Data on workplace and hybrid working behaviours are sparse and so setting in place monitoring systems of food provision and active travel usage in government departments, local authorities, related organisations and the NHS might provide an avenue for assessing behaviour change in this area. Monitoring here might include the use of surveys, use of purchasing data from workplace canteens and data on carpark usage.

1.73 While the databank is designed to be a helpful and broad source of data for future evaluators, it does have certain limitations which should be noted. Firstly, it is not feasible to capture every potentially relevant data source given the fact that the exact nature of the strategy's policies and programmes going forward is currently unclear. As such, a brief additional data search to check if the databank can be supplemented to cover any new areas of policy/programming or to include novel data sets should be carried out prior to evaluation. Another potential concern is the fact that, while the data sources listed in the Databank are current at the moment of writing this report, it is not possible to guarantee that these data streams will still be collected at the point of evaluation or across the whole period relevant to the evaluation. This risk is particularly pertinent for data collected by private sector entities.

6. Evaluation approach

- 1.74 Given the literature review' finds that there exist very few examples of holistic and comprehensive evaluations of strategies similar to Healthy Weight: Healthy Wales, we have largely developed a novel evaluation approach. That said, the literature does provide two useful considerations for our work which we have looked to integrate into our approach. Firstly, that the focus of the evaluation should evolve over time. Secondly, that any evaluation of an obesity reduction and prevention strategy should include an impact on obesity indicators as a key metric of the strategy's success.
- 1.75 The development of our approach was also supported by an extensive stakeholder engagement process which included stakeholders from Welsh Health Boards, the Welsh Government, academia, Public Health Wales and Obesity Alliance Cymru.
- 1.76 In order to provide a structured way to think about the evaluation of HWHW, we set out four key considerations for a toolkit to assess the evaluability of different aspects of HWHW:

Data Availability – What can be done with the data which is currently collected or that could feasibly be collected during the strategy's lifespan?

Cost – Is the cost of any potential evaluation realistic in gross terms and in terms of its proportionality to the cost of implementing the programme?

Timeliness – Is the evaluation sufficiently timely as to allow its results to be fed back into the strategy's planning decisions?

Robustness – Is the evaluation both sufficiently robust in its methods and sufficiently holistic in its methods as to provide credible evidence of the strategy's success?

- 1.77 These four considerations, combined with the implicit need to evaluate the strategy at both the strategic and programme/policy levels, have led to our suggestion for a three-tiered approach to the evaluation of the Healthy Weight: Healthy Wales Strategy.
- A process evaluation of the strategy's implementation to deliver process improvements which could ultimately help to improve outcomes.

- An impact evaluation of the strategy's ability to deliver systems change and a change in trends of obesity-related outcomes and behaviours to ultimately produce better outcomes
- An economic evaluation to determine if the intervention achieves value for money whereby the benefits outweigh the costs of intervention. There can also be other economic considerations such as equality impacts.

A tailor-made toolkit which can be applied to each programme/policy to separately assess whether process, outcome, impact or economic evaluation is appropriate for that particular programme/policy.

Strategic level

- 1.78 The Healthy Weight: Healthy Wales strategy has ambitious goals for creating systems change in the way that prevention and reduction of obesity is approached. As such, its ability to achieve this ambitious goal should logically be a core part of the strategy's evaluation. Furthermore, the strategy represents Wales's long-term strategy to prevent and reduce obesity, so any evaluation which did not assess the degree to which it has been successful in impacting the outcomes and behaviours related to obesity would be incomplete.
- 1.79 Assessing both systems change and changes in outcomes and behaviours necessitate the use of impact evaluation, so it is through this lens which we will view the strategic level component of the evaluation.

Systems change

- 1.80 To assess the strategy's ability to induce systems change, we propose that the evaluation focuses on three key areas:
- NHS and Health Board planning frameworks²⁰
 - School curriculums
 - Public behaviours and decision making
- 1.81 With respect to NHS and Health Board planning frameworks, the evaluation will need to assess whether the strategy has led to a change in the way that these frameworks operate. Are healthy weight and obesity integrated into planning decisions in a way which supports Welsh people by making healthy choices the

²⁰ Health Boards refer to these plans as Integrated Medium-Term Plans (IMTPs).

easy choices, and in a way which is consistent with the sentiment of the Future Generations Act. The evaluation of this dimension of systems change will necessitate an exclusively qualitative approach. The qualitative impact evaluation toolkit (i.e., case studies, observational studies, interviews and focus groups) should be applied to as broad a sample of planners, practitioners and members of the public as is feasible to assess the degree to which they have observed a tangible shift in priorities.

- 1.82 The assessment of systems change in school curriculums will require a similar approach. The strategy sets out explicit means by which it will look to bring healthy diets and physical activity to the heart of the school curriculum via initiatives like The Health and Well-being Area of Learning and Experience. If the strategy is going to have its desired impact on the youth of today and the adults of tomorrow, then assessing progress on this dimension is of critical importance. Surveys, interviews and focus groups with administrators, teachers, parents and students will be the optimal methods for assessing if the desired shift in learning settings has been achieved.
- 1.83 An assessment of the strategy's ability to influence behaviour and decision making will require a distinct approach. An assessment of the degree to which the strategy has been successful in changing reported behaviours and decision making will provide a strong indication of the level of systems changes which has been achieved. A mixture of surveys (baseline and post-strategy) and analysis of non-traditional data sources like google search data will provide a means to answer this question.
- 1.84 Additionally, the strategy places strong emphasis on empowering people to make healthy choices and so assessing the ways in which people feel more able to make these choices regarding diet and exercise will provide insight into the strategy's success in this dimension. Survey methods will again be the most appropriate tool for assessing this progress.
- 1.85 Although surveys for assessing health behaviour change are subjective and biased-prone by their very nature, Karacabeyli et al. (2018) conclude that surveys are still useful as they pragmatically allow for widespread administration with relatively low cost and staffing needs.

Outcomes

- 1.86 While the systems change that the strategy is looking to engender is about shaping people's environments and settings, the end goal of achieving this systems change is to help prevent and reduce obesity. With this in mind, we propose two key outcomes which should be evaluated as part of the strategic component of the evaluation.
- Obesity rates
 - Prevalence of healthy weight
- 1.87 Obesity and healthy weight are at the heart of the strategy and so considering how the strategy has impacted these metrics at the national, health board and local authority levels will be a major factor in assessing the strategy's success. This evaluation should consider separately the impact on these outcomes for adults and children as well as consider any inequalities in impacts between different demographic groups and geographies. The choice of these groups and respective engagement strategies are discussed in Section 9.
- 1.88 There are a number of ways that this aspect of the strategy level evaluation could be approached, but the two preferred approaches are difference-in-differences (with matching) and regression discontinuity design (RDD). In both cases, England and the other UK countries provide the best basis for creating a counterfactual against which Welsh performance can be compared. As a result, RDD and difference-in-differences have an advantage over other counterfactual approaches because they leverage existing data, natural counterfactual comparisons between neighbouring geographies and apply a less complex and subjective analytical technique. An evaluation of Leeds's Childhood Obesity Programme (Rudolf et al., 2018) provides a useful example of where these geographical similarities have been exploited to develop a counterfactual.
- 1.89 In the case of a difference-in-differences approach, people in Wales would be compared to those in England who did not benefit from the HWHW strategy. The difference between the Welsh and English outcomes before the strategy would be compared with the difference after the strategy's implementation, with the difference between these differences representing the strategy's impact. Matching could be used to improve the robustness of this method by directly

comparing people who are similar with regards to observables like age, gender and level of deprivation.

- 1.90 A regression discontinuity approach would use the Welsh border with England as a dividing line between those who benefited from HWHW and those who did not. By comparing those who live very close to the border on the Welsh side with those who live very close to the border on the English side, you strengthen the assumption that these communities will be extremely similar such that any difference in outcomes can be linked to the HWHW strategy alone. This makes assumptions about limited cross-border movement and similarities of any relevant behaviours and characteristics between those living in Wales and England, which will need to be tested for this approach to be valid.
- 1.91 Options for data sources which could underlie both approaches are the CMP (for children), the National Survey for Wales (for adults) and the additional GP measurements (for adults), which are suggested as a novel data stream in the Databank section. In the case of the CMP and the National Survey for Wales, the English CMP and the British Household Panel Survey (e.g., for health and usage of health services) can be used as data for the counterfactual.
- 1.92 Each of these datasets have limitations. The CMP currently only has one measure of 4 to 5 years and so it does not provide any insight on how children's weight progresses throughout the rest of their childhood.²¹ The National Survey for Wales is self-reported data and the sample is not sufficiently large to provide representative samples at the local authority level. The proposed data collection by GPs will take years to implement, even if the recommendation is undertaken, and the sample would only be representative of the sub-section of the population who visit the GP.²² Despite these limitations, these datasets represent the best data which could feasibly be operationalised to measure the strategy's impacts on these key outcomes.
- 1.93 Theory based evaluation could be considered where there are significant limitations for conducting quasi-experimental techniques, particularly where there are data limitations, budget limitations or where there is a need to explore why an

²¹ The addition of a second measure is underway and its implementation will go some way to mitigating this concern.

²² An additional concern for this proposed dataset is the degree to which it is viable in a world in which more and more consultations are being conducted online.

intervention would not work, which tends to not be considered in experimental studies. Theory based evaluations are conducted by developing a Theory of Change or Logic models, as done in Section 4, and then gathering existing evidence and conducting analysis (where possible) to assess the cause-and-effect mechanism in question.

- 1.94 Where there are concerns over the feasibility of evaluations, it is recommended to conduct smaller-scale pilots to explore evaluability and provide early indications of the effectiveness. The findings of these pilots could also provide valuable lessons earlier than a full evaluation.

7. Evaluability assessment: toolkit

Programme/policy level toolkit

- 1.95 The toolkit provides a structured way for prioritising evaluation in a resource-constrained environment. In theory, there could be sufficient data and resources for every HWHW policy to have a comprehensive process, outcome, impact and economic evaluation, but in practice this is not feasible due to limitations of data, resources and, in some cases, the availability of suitably robust evaluation techniques.
- 1.96 Certain programmes and policies are more suited to evaluation, and this will also vary between evaluation types. For example, it may be that it would be possible to conduct a holistic and timely process evaluation of a certain policy, but that the cost of this evaluation would be a disproportionately large share of the strategy's entire evaluation budget. Which of these competing factors is judged to be most decisive is unclear and so the toolkit provides a methodical way to consider the trade-offs between the competing constraints and requirements. The key constraints and requirements are those of data availability, cost, timeliness and robustness.
- 1.97 The toolkit is intended to be used in tandem with the guidance on the evolution of the evaluation. As such, it will make sense to initially focus on a programme's/policy's potential for process evaluation, then shift the focus towards impact and economic evaluation over the strategy's lifespan.
- 1.98 Certain aspects of the toolkit will need to be applied differently depending on the evaluation type in question, and this is made clear below on a factor-by-factor basis.
- 1.99 A final point to consider for the application of the toolkit is the case where a programme is not under the scope or control of the Welsh Government but rather scope or control of one of the many delivery partners who are contributing to the strategy's activities. In this scenario, it might be preferable from an operational standpoint to provide the toolkit to the partner organisation such that they can consider the feasibility of evaluating their programme internally.

Scoring

- 1.100 **Step 1** - The feasibility of the evaluation type being considered for the programme/policy in question is scored out of 4 for each of the four key considerations.
- 1.101 **Step 2** - These four scores are combined to form a rounded average score which is specific to the programme/policy and the type of evaluation proposed. This average score will provide guidance as to whether that type of evaluation is feasible and desirable for the programme/policy in question. A rounded average of 3 or 4 indicates a recommendation for evaluation. The exception to this rule is the case where there is a score of 0 out of 4 for any of the four key considerations, which then overrides the average and indicates that evaluation is not recommended.
- 1.102 A slightly different approach is necessary for process and economic evaluations given that data availability is not a potential issue in the same way as is the case for the other evaluation types. In the case of process evaluation, it is standard practice to collect primary data as a part of the evaluation (both qualitative and quantitative) and so concerns over the quality, robustness and accessibility of existing data streams is not relevant in this case. As detailed in the Evaluation type section, good economic evaluations use the results of an impact evaluation as a basis for their calculations and so tend to involve little primary data collection.
- 1.103 As such, for process and economic evaluations, the data availability aspect of the toolkit is not applicable and so the rounded average and the recommendation is based on the remaining three key considerations.
- 1.104 The scores for each of the key considerations have been designed such that a 3 or 4 out of 4 suggests that the proposed evaluation would be feasible and practical. They imply that there is sufficient data (or that new data could be feasibly collected), the cost is proportional, the results would be available in an appropriate time frame, or there is potential to use evaluation methods which are sufficiently robust and holistic as to be credible. Conversely, a 0 out of 4 implies that the data availability, cost, timeliness or robustness concerns are such that the evaluation proposed would be practically impossible.

Data availability

- 1.105 Data availability is not linear, and there are often cases where some data exists, but it may not be of sufficient quality or granularity to be useful for evaluation. To address this non-linearity and to account for the subjectivity of assessing data availability and quality, we have developed a set of four questions to assess whether a proposed evaluation would be viable from a data availability perspective. Each question will be scored as 0, ½ or 1, with the sum of the four question scores representing the score for data availability as a whole. If the answers to any of the four questions is 0, the overall score for data availability is also 0.
- 1.106 The databank provides a first point of reference for answering the four questions on data availability, but a brief supplementary data search might be necessary due to the limitations of the databank (see the Databank section for more detail on limitations).

Question 1 – Does appropriate data exist?

- 1.107 The first question is designed to assess whether or not an appropriate measure of the output, outcome or impact is available or feasible to collect. For example, if one is looking to measure how many people received a certain weight management service, is the data collected on attendance or registration? The distinction is that people might register and not attend, so registration data would be an imperfect way of measuring the output in question. This kind of indirect measurement is referred to as a proxy.
- 1.108 Where appropriate data is not already in existence, initiating a new stream of data collection may be an option. As such, the score for this question will also consider where existing data could be supplemented with additional data collection.

Table 1 – Data availability question 1 answer matrix

Answer	Score
Yes, an appropriate measure of the exact outcome(s) exists	1
Yes, an appropriate proxy measure of the outcome(s) exists	1/2
No appropriate (proxy) measures exist, but it would be feasible to collect the required data	1/2
No appropriate measures or proxies exist and it would not be feasible to collect additional data	0

Question 2 – What is the level of granularity of the data?

- 1.109 The level of granularity of the data is a key determinant of the rigour of evaluations of this kind. More granular data can be used to directly compare those ‘treated’ by an intervention to those who have not, which allows for the identification of effects that would not be visible in aggregated data. Furthermore, granular data can be used to provide insights into the ways in which a programme’s/policy’s impacts differed between different demographic groups and geographies.
- 1.110 In certain cases, it might be workable to execute a credible evaluation with more aggregate level data (i.e., health board or national level), but this would only be the case for programmes/policies which are sufficiently large scale to induce impacts at these aggregate levels. A national policy to provide free and healthy school meals would be an example of such a policy. Providing targeted weight management services to 100s or even 1,000s of people would be an example of a policy for which aggregated data would not be appropriate.

Table 2 - Data availability question 2 answer matrix

Answer	Score
Household or individual level	1
Health board or national level data (large-scale programme/policy)	1/2
Health board or national level data (small-scale programme/policy)	0

Question 3 – Does data for a robust counterfactual exist at the same level of granularity?

- 1.111 For evaluations to provide credible results, they require the construction of a credible counterfactual (see Section 5 for more explanation). As a result, not only is it necessary to collect data on the programme's/policy's impact on those 'treated', but also on a comparable group of people who were not 'treated'. This question relates to the availability of data on those not 'treated' or the control group.

Table 3 - Data availability question 3 answer matrix

Answer	Score
Yes	1
No, but they exist at a more aggregated level which would still be useable for the evaluation	½
No, they either do not exist or only exist at a level of aggregation which would not be compatible with the proposed evaluation	0

Question 4 - Is the data available and easy to access?

- 1.112 The degree to which the data required for an evaluation is freely and easily accessible is an important aspect of the reliability of this data stream as a basis upon which to build an evaluation. If the data comes from a private sector provider, there is a risk that the data stops being collected or that access is restricted while the evaluation is being carried out. With this in mind, this question looks to assess the reliability of an evaluator's access to the requisite data.
- 1.113 However, it is acknowledged that some privately held data can be important for evaluations compared to what is publicly available. In such instances, it is important that this data can be reliably and easily accessed and verified on an ongoing basis in the future.

Table 4 - Data availability question 4 answer matrix

Answer	Score
All critical data is (or will be) can easily be collected and managed by the Welsh Government or other public sector partners going forward	1
Critical data cannot easily be collected and verified by evaluators on an ongoing basis in the future	0

Cost






There are two main cost considerations for programme/policy evaluation in the context of HWHW: (i) whether the cost is proportional to the cost of implementation, and (ii) how the cost fits in with the strategy's wider evaluation budget. To that end, we have developed an index which combines both considerations into one measure. This measure, the Cost Index, is calculated by dividing the ratio of the evaluation cost to the evaluation budget by the ratio of the programme's/policy's cost of implementation to the total budget for the HWHW strategy. The evaluation budget is the annual evaluation budget, and the strategy budget is the annual strategy budget.

1.114

$$\text{Cost Index} = \left(\frac{\text{Evaluation Cost (£)}}{\text{Evaluation Budget (£)}} \right) \div \left(\frac{\text{Implementation Cost (£)}}{\text{Strategy Budget (£)}} \right)$$

- 1.115 The index is designed to incorporate both potential concerns into one, easily interpretable metric. If the evaluation cost represents a share of the evaluation budget which is disproportionate to the scale of the programme/policy in relation to the wider HWHW strategy, the Cost Index will be greater than 1. The larger the value for the index, the more disproportionate the cost.
- 1.116 The index can be applied to each of the evaluation types in essentially the same fashion, with the only difference being the evaluation budget. The evaluation budget should be the evaluation budget for that specific type of evaluation, with the evaluation specific budget determined on the basis of the guidance set out in the Evolution of the Evaluation section.

Table 5 - Cost answer matrix

Cost Index	Score
$x > 1.45$	
$1.30 < x \leq 1.45$	
$1.15 < x \leq 1.30$	
$1.00 < x \leq 1.15$	
$x \leq 1.00$	

Timeliness

- 1.117 The timeliness of an evaluation is crucial for ensuring that the evaluation can achieve one of its main purposes: informing future policy decisions. Different types of evaluation have different natural timelines due to the time it takes for impacts to be observable in the data compared to process indicators, which can be measured right away, for example. As a result, we have separate answer matrices for process, outcome, and impact and economic evaluations. We group impact and economic evaluation together for the reasons discussed in Section 5. ‘Time to results’ refers to the time between the programme’s/project’s initiation to the point where the evaluation’s results are available.

Table 6 – Timeliness answer matrix (Process)






Time to Results	Score
$x > 3$ years	
$2.5 \text{ years} < x \leq 3 \text{ years}$	
$2 \text{ years} < x \leq 2.5 \text{ years}$	
$1.5 \text{ years} < x \leq 2 \text{ years}$	
$x \leq 1.5 \text{ years}$	

Table 7 - Timeliness answer matrix (Outcome)











Time to Results	Score
$x > 6$ years	
$5 \text{ years} < x \leq 6$ years	
$4 \text{ years} < x \leq 5$ years	
$3 \text{ years} < x \leq 4$ years	
$x \leq 3$ years	

Table 8 - Timeliness answer matrix (Impact/Economic)

Time to results	Score
$x > 12$ years	
$10 \text{ years} < x \leq 12$ years	
$8 \text{ years} < x \leq 10$ years	
$6 \text{ years} < x \leq 8$ years	
$x \leq 6$ years	

Robustness

- 1.118 Similar to data availability, assessing robustness is not as linear as is the case for timeliness and cost. There are several concerns related to robustness, including the degree to which it is possible to create a robust counterfactual in the case of quantitative methods, as well as the degree to which the evaluation is holistic, appropriately combining qualitative and quantitative methods.
- 1.119 To address this non-linearity and subjectivity, we have developed a set of four questions to assess whether a proposed evaluation would be sufficiently robust. Each question is scored as 0 or 1, with the sum of the four question scores representing the score for data availability as a whole.

Question 1 – Is there scope to include both quantitative and qualitative analysis?

- 1.120 For an evaluation to be sufficiently holistic, it is desirable to have some degree of a mixed-methods approach. Quantitative analysis generally offers a significant advantage in the breadth which can be covered, whereas qualitative methods can often add a level of depth which is not achievable with quantitative methods alone. As such, this question addresses this very basic issue with a view at incentivising the use of a combination of methods in the programme/policy level evaluations.
- 1.121 This question is not relevant for economic evaluation and should be assigned as 1 by default when the toolkit is being applied to potential economic evaluations. This is because economic evaluations rely on pre-existing data and evidence to produce analysis and therefore would not include any significant collection of qualitative or quantitative information.

Table 9 – Robustness question 1 answer matrix

Answer	Score
Yes	1
No	0

Question 2 – Is it feasible to use quasi-experimental techniques?

- 1.122 In the absence of randomised controlled trials, quasi-experimental methods are the gold standard with regards to impact evaluation and, accordingly, their use leads to substantially more credible results than is the case for alternative methods.
- 1.123 This question is not relevant for process or economic evaluation and should be assigned as 1 by default when the toolkit is being applied to potential process or economic evaluations. This is because quasi-experimental techniques are only used for impact evaluation. Process evaluations tend to focus on less technical analysis, while economic evaluations use pre-existing data and would take any existing quasi-experimental analyses as given, if available.

Table 10 – Robustness question 2 answer matrix

Answer	Score
Yes	1
No	0

Question 3 – Are there studies of similar policies in similar contexts which can be used for robustness checks?

- 1.124 Past studies of similar policies can also provide a basis for sense checking the findings of an evaluation. If the evaluation’s findings are within the reasonable bounds of estimates of similar policies in similar contexts, that provides further credibility to the results.

Table 11 - Robustness question 3 answer matrix

Answer	Score
Yes	1
No	0

Question 4 – Is there an additional data source which can be used for robustness checks?

- 1.125 There are a plethora of issues which can arise with individual datasets, such as unrepresentativeness, sample size and accuracy. As such, there is great benefit in having more than one available data set with which the evaluation could be conducted. The analysis of additional data sets as robustness checks can help mitigate the risks of issues which can arise when using individual datasets.

Table 12 - Robustness question 4 answer matrix

Answer	Score
Yes	1
No	0

8. Evaluability assessment: examples

- 1.126 In order to demonstrate how the toolkit should be applied in practice, we have developed three worked examples which illustrate how to apply the toolkit to selected policies and programmes in the strategy's 2022-24 delivery plan. As not all the relevant information is available at the time of writing, some assumptions have been made regarding the evaluation's relative cost and the programme's/policy's timeline.
- 1.127 Throughout these examples, it will be made clear where the assessments are based on assumptions as opposed to existing data or information.

Example 1 (Process evaluation)

- 1.128 This worked example considers the potential process evaluation of the programme, which invests in a behavioural change intervention to encourage parents and carers, and those planning a pregnancy, to make a lifestyle adjustment.

Data availability

- 1.129 Unlike outcome and impact evaluation, process evaluations do not rely on the availability of output or outcome data, especially as the lack of monitoring data could be a finding of the process evaluation itself. Therefore, data availability is not deemed as a relevant criterion in an evaluability assessment of process evaluations. Score = N/A.

Cost

- 1.130 In order to assess the feasibility of this impact evaluation from a cost perspective, we use the Cost Index formula set out in the toolkit. We again have to make certain educated assumptions regarding the cost of the proposed evaluation, the strategy's evaluation budget and the implementation cost of the policy in question. We have used the strategy's overall budget of £52m as the strategy budget. To estimate the process evaluation budget, we begin by using the rule of thumb that roughly 10% of an intervention's budget should be set aside for evaluation, which implies a total evaluation budget of £5.2m. From there, we propose that 15% of that evaluation budget should be set aside for process evaluation, which implies a process evaluation budget of £0.78m. We then

assume that the cost of the programme will be £3m and the cost for the process evaluation is £50,000. Inputting these numbers into the Cost Index formula, as is done below, leads to a Cost Index of 1.11.

$$\text{Cost Index} = \left(\frac{£0.05\text{m}}{£0.78\text{m}} \right) \div \left(\frac{£3.0\text{m}}{£52.0\text{m}} \right) = 1.11$$

- 1.131 This Cost Index is between the 1.00 and 1.15 thresholds, between which the cost score is 3 and so the cost is deemed to be proportionate to the evaluation budget.

Timeliness

- 1.132 In the case of the evaluation's time horizon, the time requirements of each evaluation are unique and will depend on resources, capacity and data availability. That said, we have no reason to expect that the process evaluation of this programme will be longer than is standard and so, if we assume a 6-month design period, a 1-year monitoring period and a 3-month evaluation period, then results would be expected within 21 months of the programme's initiation. 21 months would imply a timeliness score of 3 as it is under the 2-year threshold (see Table 6).

Robustness

- 1.133 In order to assess the robustness for the proposed evaluation, we answer the four questions set out in the toolkit.

Q1 – Is there scope to include both quantitative and qualitative analysis?

- 1.134 The proposed process evaluation has scope to combine both qualitative and quantitative analysis. Monitoring data on how many parents were reached by the programme could provide a basis for quantitative evaluation, whereas qualitative monitoring data and interviews could be conducted with both practitioners and expectant parents to gauge their opinions on what did and did not work for the programme. Score = 1.

Q2. Is it feasible to use quasi-experimental techniques?

- 1.135 The toolkit explains that this question is not relevant to process evaluation and should be scored as 1 by default. Score = 1.

Q3. Are there studies of similar policies in similar contexts which can be used for robustness checks?

1.136 No, there does not seem to be any sufficiently similar process evaluations of similar programmes in a comparable context. Score = 0.

Q4. Is there an additional data source which can be used for robustness checks?





1.137 No, there does not seem to be any additional datasets which could be used for the process evaluation of this programme. Score = 0.

1.138 The answers to the four questions lead to a combined score for robustness of 2 out of 4.

Conclusion

1.139 Overall, the application of the toolkit to this policy has led to a rounded average score of 3, indicating that a process evaluation would likely be feasible in this case (see Table 13). The case in favour of evaluation is strong from a cost and timeliness perspective, with the only source of concern being that there are limited options for conducting a robustness check on the evaluation's results.

Table 13 - Worked example 1 score summary

Key consideration	Score
Data availability	N/A
Cost	
Timeliness	
Robustness	
Rounded average	
Recommendation	<i>Do evaluate</i>

Example 2 (Impact evaluation)

- 1.140 This worked example considers the potential impact evaluation of the programme which invests £1m into electric bike pilot schemes on the rates of active travel among the recipient communities.

Data availability

- 1.141 In order to assess the data availability for the proposed evaluation, we answer the four questions set out in the toolkit.

Q1. Does appropriate data exist?

- 1.142 Yes, an appropriate measure of the exact outcome exists. The National Survey for Wales provides self-reported data on people's physical activity, including cycling. Score = 1.

Q2. What is the granularity of the data?

- 1.143 The data is available at the individual and household level. Score = 1.

Q3. Does data for a robust counterfactual exist at the same level of granularity?

- 1.144 Yes, the British Household Panel Survey and the Health Survey for England provide comparable data on physical activity at the individual and household level. Score = 1.

Q4. Is the data available and easy to access?

- 1.145 Yes, the National Survey for Wales is under the purview of the Welsh Government. Score = 1.
- 1.146 The answers to the four questions lead to combined score for data availability of 4 out of 4.

Cost

- 1.147 In order to assess the feasibility of this impact evaluation from a cost perspective, we use the Cost Index formula set out in the toolkit. At this point, we again have to make certain educated assumptions regarding the cost of the proposed evaluation, the strategy's evaluation budget and the implementation cost of the policy in question. We have used the strategy's overall budget of £52m as the strategy budget. To estimate the impact evaluation budget, we begin by using the rule of thumb that roughly 10% of an intervention's budget should be set aside for

evaluation, which implies a total evaluation budget of £5.2m. From there, we propose that 30% of that evaluation budget should be set aside for impact evaluation, which implies an impact evaluation budget of £1.56m. The programme cost is defined as £1m. For the estimate of the impact evaluation cost, we have estimated this at £50,000 on the basis of our past work on impact evaluation. Inputting these numbers into the Cost Index formula, as is done below, leads to a Cost Index of 1.67.

$$\text{Cost Index} = \left(\frac{£0.05\text{m}}{£1.56\text{m}} \right) \div \left(\frac{£1.0\text{m}}{£52.0\text{m}} \right) = 1.67$$

- 1.148 This Cost Index is marginally above the 1.45 threshold, after which the cost score is set to 0, so this disproportionate cost will lead to a recommendation that this policy should not be impact evaluated. The score of 0 out of 4 for cost overrides the rounded average score.
- 1.149 To sense check this result, we can compare the impact evaluation cost of £50,000 to the total impact evaluation budget of £195,000. Intuitively, spending almost 1/4 of the budget on the evaluation of one programme would in itself be enough to support our recommendation that the cost of impact evaluation would be disproportionate in this case. As discussed in Section 6, theory-based approaches can be useful for assessing impact where empirical evaluations are not feasible.

Timeliness

- 1.150 In the case of the evaluation's time horizon, the time requirements of each evaluation are unique and will depend on resources, capacity and data availability. As we do not have information on these aspects at this point, we can only base this assessment on assumptions. We might reasonably expect to observe a significant impact on the use of active travel instead of sedentary alternatives within 5 years or less. Our estimate is shorter than might be expected for a country-wide policy, as targeted pilots are often shown to have larger and quicker impacts than is the case for policies or programmes rolled out at the national level. 5 years would imply a timeliness score of 4 as it is under the 6-year threshold (see Table 8).

Robustness

1.151 In order to assess the robustness for the proposed evaluation, we answer the four questions set out in the toolkit.

Q1. Is there scope to include both quantitative and qualitative analysis?

1.152 Yes, there is scope for a rigorous quantitative impact evaluation using the data outlined in the data availability questions, and there is also scope for a qualitative aspect to the evaluation. Surveys and interviews with administrators and beneficiaries of the programme could add depth to the evaluation and shed light on the mechanisms by which the programme impacted their physical activity habits. Score = 1.

Q2. Is it feasible to use quasi-experimental techniques?

1.153 Yes, methods such as difference-in-differences and time series econometric analysis are two examples of ways in which this programme's impact could be evaluated using quasi-experimental techniques. In both cases, the evaluation would compare the level of physical activity done by those who stand to benefit from the programme before and after its implementation. In the case of difference-in-differences, this difference would also be compared to the difference observed in sample of people who did not benefit from the programme. Score = 1.

Q3. Are there studies of similar policies in similar contexts which can be used for robustness checks?

1.154 Yes, one such example could be the 2016 Institute for Employment Studies (IES) report on the impact of the Cycle to Work Scheme. Score = 1.

Q4. Is there an additional data source which can be used for robustness checks?

1.155 No, we are not aware of any additional data sources in existence which could be used for robustness checks. Score = 0.






1.156 The answers to the four questions lead to a combined score for robustness of 3 out of 4.

Conclusion

1.157 Table 14 displays the final results for an illustrative application of the toolkit to this programme. The proposed impact evaluation of the programme was shown to be promising from a data availability, timeliness and robustness perspective, but its

disproportionate cost led to a recommendation of not proceeding with the evaluation. Note that there could still be a decision to proceed with an evaluation depending on the prioritisation of this evaluation versus others, as well as understanding what type of evaluation could be feasible with a more proportional budget.

Table 14 - Worked example 2 score summary

Key consideration	Score
Data availability	
Cost	
Timeliness	
Robustness	
Rounded average	
Recommendation	<i>Do not evaluate (as Cost score = 0)</i>

Example 3 (Economic evaluation)

1.158 This worked example considers the potential for an economic evaluation of the collection of policies and programmes which focus on good nutrition throughout the school day. They include:

- Extending the provision of free school meals to all primary school pupils so children have access to two healthy meals a day, alongside the free school breakfast initiative, instilling healthy eating habits early on in life. The free school breakfast initiative has been robustly [evaluated by Cardiff University](#), so there could be the potential to further develop the approach for the roll-out and evaluation of free school meals.
- Reviewing the regulations on school food nutrition in line with the latest nutritional standards and guidelines, and updating current standards. This will include exploring food labelling, local procurement, climate change and transport, school provision and timing of the school day.
- Building upon the work with Veg Power to positively promote fruit and vegetables through primary schools across Wales.

- Promoting wider programmes, such as Designed to Smile, to encourage positive behavioural messages around oral and dental health and the impact of sugary food on teeth and gums.

1.159 This worked example is designed to show the potential for applying the toolkit to a group of policies rather than just a single one in isolation. A good economic evaluation should be based on the results of a rigorous impact evaluation, so for this example, we will assume that this impact evaluation has already taken place and that the impact of this suite of policies on nutrition in schools has been estimated.

Data availability

1.160 This criterion is deemed as not relevant for economic evaluations. This is because economic evaluations use pre-existing evidence and analysis, such as from impact evaluations that would have already been conducted. Economic evaluations may use some data collection and analysis, but this would typically be limited. Therefore, this aspect of the toolkit is not applied for this example. Score = N/A.

Cost

1.161 In order to assess the feasibility of this economic evaluation from a cost perspective, we use the Cost Index formula set out in the toolkit. We again have to make certain educated assumptions regarding the cost of the proposed evaluation, the strategy's evaluation budget and the implementation cost of the policy in question. We have used the strategy's estimated overall budget of £52m as the strategy budget. To estimate the impact evaluation budget, we begin by using the rule of thumb that roughly 10% of an intervention's budget should be set aside for evaluation, which implies a total evaluation budget of £5.2m. From there, we propose an economic evaluation budget of £500,000. We also assume that the combined cost of the programmes and policies is £4m and that the estimated cost of the economic evaluation is £70,000. Inputting these numbers into the Cost Index formula, as is done below, leads to a Cost Index of 1.33.

$$\text{Cost Index} = \left(\frac{£0.07\text{m}}{£0.50\text{m}} \right) \div \left(\frac{£4.0\text{m}}{£52.0\text{m}} \right) = 1.82$$

1.162 This Cost Index is above the threshold of 1.82, which suggests that the evaluation cost is high.

Timeliness

- 1.163 In the case of the evaluation's time horizon, given that the economic evaluation will be based on a previous impact evaluation, the time necessary to conduct this preliminary evaluation will need to be factored in. We might, therefore, reasonably expect the impact evaluation to conclude within 8 years, with economic evaluation conducted in the following year. 9 years would imply a timeliness score of 2 as it is between the 8 and 10 year thresholds (see Table 8).

Robustness

- 1.164 In order to assess the robustness for the proposed evaluation, we answer the four questions set out in the toolkit.

Q1. Is there scope to include both quantitative and qualitative analysis?

- 1.165 The toolkit explains that this question is not relevant to economic evaluation and should be scored as 1 by default. Score = 1.

Q2. Is it feasible to use quasi-experimental techniques?

- 1.166 The toolkit explains that this question is not relevant to economic evaluation and should be scored as 1 by default. Score = 1.

Q3. Are there studies of similar policies in similar contexts which can be used for robustness checks?

- 1.167 Yes, one such example would be Vale et al. (2012) which focused on school food policy in Newcastle. Score = 1.

Q4. Is there an additional data source which can be used for robustness checks?





- 1.168 It is not feasible to assess this without knowing the exact benefits which the economic evaluation will look to estimate. In our experience conducting CBAs, more often than not there is a way to use an alternative estimate or data series to sense check the analysis. With that in mind, we will assume that for this scenario this option exists. Score = 1.

- 1.169 The answers to the four questions lead to a combined score for robustness of 3 out of 4.

Conclusion

1.170 Table 15 displays the final results for the illustrative application of the toolkit to this suite of programmes. The final rounded score of 2 leads to a recommendation not to evaluate, which is reflective of the cost being deemed disproportionately high and the time horizon after which results would be available, leaving little scope for findings to be fed back into the strategy's planning processes. Despite this, the score was on the margins and the proposed evaluation had the potential to be methodologically robust and so, in a different cost environment it might become viable.

Table 15 - Worked example 3 score summary

Key consideration	Score
Data availability	N/A
Cost	
Timeliness	
Robustness	
Rounded average	
Recommendation	<i>Do not evaluate</i>

9. Practical evaluation issues

1.171 There are a number of practical issues common to all evaluation methodologies which need to be kept in mind by an evaluator:

- data collection
- data protocols
- setting the baseline

Data collection

1.172 While our recommendation does discuss the need for additional data collection in certain areas (see Section 5), specific aspects about the practicalities of data collection will necessarily need to be determined by the evaluators during the design of each particular piece of evaluation work.

1.173 In the cases where the collection of new data streams is being considered, evaluators should bear in mind the following practical concerns:

- The costs associated with the data collection and management.
- Whether the data which is feasible to collect represents a large enough sample to be used for robust and meaningful evaluation.
- The sampling design: the degree to which the sample which could feasibly be collected would be representative of the population they are studying.
- Whether there are relevant examples of this data being collected in similar contexts which can be used as guidance.
- Whether the data is being collected by a public or private sector partner and the level of access to and control of the data that this entails.
- The granularity of data that can be feasibly collected, such as for breakdowns at the level of household/firm, health board or national level data.
- The availability of comparable data which could be used as a counterfactual for this new data stream in the case where the data is to be used for quantitative outcome or impact evaluation.

Data protocols

- 1.174 Future evaluators must adhere to required data confidentiality and security standards.²³ Confidentiality refers to the evaluators' agreement to ensure any data gathered from the evaluation's research participants is not improperly divulged. Security refers to the evaluators' ability to keep data gathered from research participants secure and in line with legal regulations regarding data protection.
- 1.175 To comply with the new United Kingdom General Data Protection Regulation (UK GDPR), the evaluators should:
- Assign a data privacy role to one project member to explicitly address UK GDPR regulation, responsible for storing, transferring and disposing of data correctly.
 - Undertake Privacy Impact Assessments of the evaluators' own facilities or any third-party data processing parties and, if necessary, revise any existing standard contracts or security policies to meet UK GDPR regulations.
 - Create an official protocol for participants wanting to access or delete their data. This could be communicated through consent forms.
 - Create an official protocol for data breaches.
- 1.176 Though data collected for this evaluation is all likely to be anonymous, it is still likely to be disclosive (i.e., someone could identify an individual based on the anonymous information about them). As well as ensuring that data is held securely, the data privacy lead should ensure that any data reported in interim or final reports is not disclosive.

Setting the baseline

- 1.177 All of the additional avenues of data collection discussed in this evaluability assessment, aside from those related to process evaluation, will need to have a baseline data collection round against which progress can be measured. The precise dates of the baseline will depend on the particularities of the evaluation in question. For example, data streams for strategic level evaluation need to be

²³ The term "data" is used in the widest sense. It includes quantitative files, such as statistics and numerical files, and qualitative materials, such as interview transcripts and field notes. Research data may include audio and video formats, geographical information and websites.

baselined as soon as feasibly possible in order for the future evaluation to capture the impact of the strategy in its entirety. On the other hand, for outcome or impact evaluations of individual policies or programmes, a baseline round of data collection should take place as soon as possible, before the point of project implementation.

Stakeholder engagement

- 1.178 Stakeholder engagement in evaluations can be important to enable the findings of evaluations to stimulate practical outcomes that can inform the strategy and improve evidence-led decision making in the future. For example, stakeholder engagement can ensure that key findings are expressed clearly and meaningfully in a way that can be easily implemented by those less familiar with the technical detail of the evaluation. Further, involvement of stakeholders can ensure the recommendations capture a diverse range of perspectives which are relevant to delivering improvements to the strategy or future programmes, to ultimately deliver better value for money. Lastly, the involvement of stakeholders can generate support for the evaluation, which in turn could support the credibility and implementation of its findings.
- 1.179 Updated guidance from the Medical Research Council (2021) highlights the importance of stakeholder engagement as being essential for prioritising research questions, the co-development of programme theory, choosing the most useful research perspective, and overcoming practical obstacles to evaluation and implementation. However, it also recommends that researchers should be mindful of conflicts of interest among stakeholders and use transparent methods to record potential conflicts of interest.

The table below summarises how stakeholders can be engaged throughout the evaluation process.

Evaluation phase	Detail
Planning	Stakeholders are people or parties that could be affected by the programme evaluation. Stakeholders and their role in the evaluation should be identified at the start of the evaluation. Stakeholders can be used to identify short term and long term goals of the evaluation, which could ultimately be used

Evaluation phase	Detail
	<p>to determine the success of the programme. Stakeholders can also help to improve the uptake of the evaluation's findings by helping to determine the target audience. Their input can also be helpful to identify and mitigate key risks of the programme.</p>
Implementation	<p>Stakeholder engagement can help to overcome any unforeseen issues, such as with recruitment and retention, and the ability of the evaluation to follow the project plan. Process evaluations should involve a diverse range of stakeholders to produce a representative view on how implementation can be improved, which in turn could improve the ability of the programme to achieve its objectives.</p>
Completion	<p>The evaluation of long term outcomes and value for money assessments can be supported by bringing together the diverse perspectives of stakeholders who could support the development of logic models and Theory of Change. Stakeholders may also suggest targeting specific outputs or outcomes that can yield the greatest traction with decision makers.</p>
Dissemination and reporting	<p>The successful dissemination and implementation of the evaluation findings is influenced by the ability of the evaluators to develop a comprehensive dissemination plan which targets the right audience in the right way at the right time. Given these audiences could have varying interests and perspectives, it is critical that stakeholders representing such views are engaged in the development of the dissemination plan and the reporting content to ensure the presentation/language is tailored in the right way. Successfully engaged stakeholders can also help champion</p>

Evaluation phase	Detail
	key findings and ensure they are adopted in future decision making.

Capturing different groups in evaluation

- 1.180 Different groups and their experiences of an intervention should be captured in an evaluation, especially if there is reason to believe that the effectiveness of an intervention could differ across groups. For example, obesity and deprivation are highly correlated, so interventions tackling obesity may have differing impacts across those with varying deprivation levels. Similarly, the implementation of a national programme may lead to very different levels of success at a local level, perhaps because implementation is fragmented locally or there are varying levels of local support. In such instances, ignoring these differences across these groups could result in an evaluation missing key opportunities to improve outcomes for those groups, therefore limiting the overall success of the programme.
- 1.181 Although there is not a set list of groups that should be considered, it is often best practice to engage with stakeholders, as well as review existing evidence and precedent, to identify which groups could have differing experiences. Evaluators may also consider groups relating to the nine specific protected characteristics in the Equality Act 2010 which requires consideration by public bodies on whether people with these characteristics will be disproportionately affected by the impacts of a decision (i.e., age, disability, gender reassignment, marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex, and sexual orientation).
- 1.182 Some groups may be more readily engaged in the evaluation than others, which could increase the risk of biases and unrepresentative findings. Once relevant groups have been identified, evaluators should tailor their research and evaluation methods to acknowledge any differences in the ability of some groups to engage or be captured in the data. The table below provides some guidance on how information collection methods may differ across some example groups.

Groups	Appropriateness of information collection methods
Demographics	<ul style="list-style-type: none"> • Interviews and focus groups: this could be appropriate for certain demographic groups who may find it easier to share information face to face rather than through digital or written mediums (e.g., the elderly). • Surveys and polls: if conducted digitally, this could suit groups with a high level of digital literacy. Conversely, elderly groups which are typically associated with lower levels of digital engagement may find it harder to complete online surveys. In such instances, enabling both digital and non-digital methods could be helpful. • Output/performance metrics and qualitative observational studies could be used as a method to systemically collect data cross all groups, providing that this data is representative.
Regions/localities	<ul style="list-style-type: none"> • Interviews and focus groups: it may not be feasible to conduct interviews and focus groups in every region, depending on the scale of the evaluation. • Surveys and polls and output/performance metrics: these data collection techniques could favour evaluations which seek to collect detailed information systematically over a wide geography. However, digital literacy may also need to be considered if this is something that varies by region.
Deprivation	<ul style="list-style-type: none"> • Deprived groups may find it financially difficult to participate in in-person focus groups and interviews due to the cost of travel.

Groups	Appropriateness of information collection methods
	<ul style="list-style-type: none">• Similarly, online communication may also be impacted as deprived groups may be unable to afford internet access.• Collecting any performance metrics or engaging stakeholders could clarify the extent to which the experience of different deprivation groups may differ.

10. Conclusion and recommendations

- 1.183 A conclusion that can be drawn from the research conducted is that there is a lot of scope for process, impact and economic evaluation, and it is definitely feasible to take some of this forward at the level of both the overall strategy and a number of individual programmes/policies. There will also be some parts of HWHW that are not feasible to evaluate individually due to constraints of data availability, cost, timeliness or robustness.
- 1.184 The cut-off point for where to draw the line between what is and is not evaluable depends on some decisions that have not yet been made (e.g., the total spend on HWHW evaluation over its lifetime and the way in which the policies and programmes which make up the strategy will evolve over time). We have provided a toolkit for assessing where to draw the line as these decisions are made.
- 1.185 There could still be a decision to proceed with an evaluation depending on the prioritisation of this evaluation versus others, as well as understanding what type of evaluation could be feasible with a more proportional budget.
- 1.186 The remainder of this section provides practical next steps following this report to refine and inform future assessments.

Step 1: Collect information

- 1.187 Evaluability assessments should be refined as the understanding behind the strategy is developed, such as in terms of the interventions, respective budgets and the availability of information. The below box presents a stylised checklist of the types of information that should be collected to refine the evaluability assessment.

Checklist of information requirements for the evaluability assessment:

- The objectives of the strategy, as well as the criteria for success, are clearly defined. Key outcomes are measurable with data readily available.
- The strategy and interventions are defined, with detailed understanding about the associated activities, potential impacts and cohorts.
- There is a coherent logic connecting the inputs, activities, outputs and outcomes at both the programme and project level.
- The budgets at the strategy and intervention level are properly defined, including the amounts dedicated to evaluation.
- There is a clear organisational process for planning, executing and learning from evaluation findings.
- There is a dedicated workstream or structure (e.g., staff) for managing evaluation work.
- There is a comprehensive plan for data collection and monitoring of interventions and the strategy.
- Timeframes for evaluation work are clear.

Step 2: Plan the evaluability assessment

1.188 The toolkit described in Section 7 provides a transparent and repeatable methodology to conduct the evaluability assessment. However, it is acknowledged that evaluations could be considered on a case-by-case basis such that an evaluation could still be conducted despite the toolkit deeming it to not be feasible.

1.189 Some recommendations that will assist in increasing evaluability could include:

- **Apply the toolkit early and review regularly** – decisions on evaluability of HWHW policies will vary over time as decisions are made on the amount of resources available for evaluation and as new policies are implemented. To maximise the value from evaluation budgets, it is recommended that the toolkit is applied and updated regularly as new decisions are made with respect to evaluation budgets and policies. Early prioritisation will assist in deciding where to focus data collection efforts and provide longer lead times for planning evaluation activities.

- **Consider monitoring plans at an early stage and decide where additional data collection is needed** – best practice involves considering monitoring and evaluation at an early stage of policymaking, with monitoring often key to evaluability. A number of data sources already exist and some data can be collected retrospectively, but improving ongoing data collection is key to maximising evaluability.
- **Collect baseline data, especially for evaluation of systems change driven by the overall HWHW strategy** – for evaluating the ability of the HWHW strategy to achieve systems change, it is recommended that baseline data is collected as soon as possible to create an evidence base on current practices and attitudes. This will enable evaluators in subsequent evaluations to collect similar data to compare with the baseline to assess whether systems change has been achieved. As discussed in the chapter on practical evaluation issues, it is important to consider whether other policies within HWHW also require baseline data collection to increase evaluability.

Step 3: Implement the findings of the evaluability assessment

- 1.190 The evaluability assessment should ultimately inform decision making on what should be evaluated and how. These recommendations should be clearly articulated as an outcome of the assessment, with the appropriate decision-making structures to act on the findings.
- 1.191 It is possible that the evaluability assessment could be an iterative process, as it may stimulate further discussions on some of the interventions and evaluation planning, particularly if specific barriers are identified to evaluability that can be overcome relatively quickly. It may also be that the perceived evaluability may change once the programme has been implemented and there is a better understanding of data collection and likely evaluability budget.

References

- Bagnall, A., Radley, D., Jones, R., Gately, P., Nobles, J., Van Dijk, M., Blackshaw, J., Montel, S. and Sahota, P., 2019. Whole systems approaches to obesity and other complex public health challenges: a systematic review. *BMC Public Health*, 19(1).
- Brown, A.D., Whelan, J., Bolton, K.A., Nagorcka-Smith, P., Hayward, J., Fraser, P., Strugnell, C., Felmingham, T., Nichols, M., Bell, C. and Le, H.N., 2022. A Theory of Change for Community-Based Systems Interventions to Prevent Obesity. *American Journal of Preventive Medicine*, 62(5), pp.786-794.
- Department for Business, Energy & Industrial Strategy, 2021. Valuation of greenhouse gas emissions: for policy appraisal and evaluation [Data Set], [Valuation of greenhouse gas emissions: for policy appraisal and evaluation - GOV.UK](#)
- DFID, FCO and MoD, 2017. External evaluation of the National Action Plan on women, peace and security: endline report. *The UK Government*.
- Digital Health and Care Wales, 2022. Diagnostic and therapy services waiting times (2005-2022) [Data Set], [AssetName - Asset Manager](#)
- Frontier Economics, 2022. Estimating the full costs of obesity.
- Gadsby, E.W., Hotham, S., Eida, T., Lawrence, C. and Merritt, R., 2020. Impact of a community-based pilot intervention to tackle childhood obesity: a 'whole-system approach' case study. *BMC public health*, 20(1), pp.1-12.
- Greater Manchester Combined Authority (GMCA), 2014. Greater Manchester CBA model [Data Set], [Research: Cost Benefit Analysis - Greater Manchester Combined Authority](#)
- Housing Associations' Charitable Trust (2014), 2014. Social Value Bank [Data Set], [UK Social Value Bank | HACT](#)
- HM Revenues and Custom, 2022. Payrolled employments in the UK by region, industry and nationality, *UK Government* [Data Set], [Payrolled employments in the UK by region, industry and nationality, from July 2014 to June 2021 - GOV.UK](#)
- HM Treasury, 2020. Magenta Book: Central Government guidance on evaluation, *UK Government*.
- HM Treasury, 2022. Green Book: Appraisal and evaluation in central government, *UK Government*.
- Institute for Employment Studies, 2016. Impact of the Cycle to Work Scheme: Evidence Report.
- Internal Oversight Change, 2021. Evaluation of the UNESCO Strategy for Action on Climate Change (2018-2021), *UNESCO*.

- Karacabeyli, D., Allender, S., Pinkney, S. and Amed, S., 2018. Evaluation of complex community-based childhood obesity prevention interventions. *Obesity Reviews*, 19(8), pp.1080-1092.
- Kumanyika, S.K., 2019. A framework for increasing equity impact in obesity prevention. *American Journal of Public Health*, 109(10), pp.1350-1357.
- Kwon, J., Reeve, E., Mann, D., Swinburn, B. and Sacks, G., 2022. Benchmarking for accountability on obesity prevention: evaluation of the Healthy Food Environment Policy Index (Food-EPI) in Australia (2016–2020). *Public health nutrition*, 25(2), pp.488-497.
- Medical Research Council, Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, Boyd KA, Craig N, French DP, McIntosh E, Petticrew M, Rycroft-Malone J, White M, Moore L. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*. 2021 Sep 30;374:n2061. doi: 10.1136/bmj.n2061. PMID: 34593508; PMCID: PMC8482308.
- Mackenbach JP, Meerding WJ, Kunst AE, 2011. Economic costs of health inequalities in the European Union. *Journal of Epidemiology & Community Health*;65:412-419.
- Moon, H., Yoo, S.H. and Huh, S.Y., 2021. Monetary valuation of air quality improvement with the stated preference technique: A multi-pollutant perspective. *Science of The Total Environment*, 793, p.148604.
- NatCen Social Research. 2016. Welsh Health Survey, 2015. [data collection]. *UK Data Service*. SN: 8090 [Data Set], DOI: [10.5255/UKDA-SN-8090-1](https://doi.org/10.5255/UKDA-SN-8090-1)
- NatCen Social Research, Food Standards Agency, 2020. Food and You Survey, 2018. [data collection]. 2nd Edition. *UK Data Service*. SN: 8574 [Data Set], DOI: [10.5255/UKDA-SN-8574-2](https://doi.org/10.5255/UKDA-SN-8574-2)
- National Survey for Wales, [National Survey for Wales | GOV.WALES](https://gov.wales/national-survey-for-wales)
- NHIPPCC (National Health Information and Performance Principal Committee), 2017. The Australian Health Performance Framework. *Adelaide: COAG Health Council*.
- NHS, 2022. Electronic Staff Record (ESR) [Data Set], [Electronic Staff Record \(datadictionary.nhs.uk\)](https://datadictionary.nhs.uk)
- NHS Digital, National Diabetes Audit Programme [Data Set], [National Diabetes Audit dashboards](https://dashboards.nhs.uk)
- Office for National Statistics, 2022. Labour productivity [Data Set], [Labour productivity - Office for National Statistics \(ons.gov.uk\)](https://ons.gov.uk/labour-productivity)
- Pérez-Escamilla, R., Lutter, C.K., Rabadan-Diehl, C., Rubinstein, A., Calvillo, A., Corvalán, C., Batis, C., Jacoby, E., Vorkoper, S., Kline, L. and Ewart-Pierce, E., 2017. Prevention of childhood obesity and food policies in Latin America: from research to practice. *Obesity Reviews*, 18, pp.28-38.

- Pringle, A., McKenna, J. and Zwolinsky, S., 2017. Linking physical activity and health evaluation to policy: Lessons from UK evaluations. *In Routledge handbook of physical activity policy and practice* (pp. 425-440). Routledge.
- Public Health England, 2018. Outcome evaluation: evaluation in health and wellbeing, *UK Government*.
- Public Health Wales, 2021. Child Measurement Programme for Wales Report 2020-21 [Data Set], [CMP 2020-21 - Public Health Wales \(nhs.wales\)](#)
- Public Health Wales, 2016. Measuring Inequalities [Data Set], [Measuring inequalities 2016 - Public Health Wales \(nhs.wales\)](#)
- Ramos, M.F., Weber, J., Thompson, J.A., Werk, L., Shuell, J., Woods, K. and Bamdad, T., 2021. Evaluation of “Healthy Way to Grow”: An obesity prevention program in early care and education centers. *Early Childhood Education Journal*, 49(4), pp.553-566.
- Rudolf, M., Perera, R., Swanston, D., Burberry, J., Roberts, K. and Jebb, S., 2019. Observational analysis of disparities in obesity in children in the UK: Has Leeds bucked the trend?. *Pediatric obesity*, 14(9), p.e12529.
- Sawyer, A., den Hertog, K., Verhoeff, A.P., Busch, V. and Stronks, K., 2021. Developing the logic framework underpinning a whole-systems approach to childhood overweight and obesity prevention: Amsterdam Healthy Weight Approach. *Obesity science & practice*, 7(5), pp.591-605.
- StatsWales: Adult General Health and Illness [Data Set], [Adult general health and illness \(gov.wales\)](#)
- StatsWales: Adult Lifestyles [Data Set], [Adult lifestyles \(gov.wales\)](#)
- StatsWales: Air Quality [Data Set], [Air quality \(gov.wales\)](#)
- StatsWales: Child Lifestyles [Data Set], [Child lifestyles \(gov.wales\)](#)
- StatsWales: Council Tax Collection [Data Set], [Collection \(gov.wales\)](#)
- StatsWales: Life Expectancy [Data Set], [Life expectancy \(archive\) \(gov.wales\)](#)
- StatsWales: Mental Health and Wellbeing [Data Set], [Mental health and wellbeing \(gov.wales\)](#)
- StatsWales: NHS Programme Budget [Data Set], [NHS programme budget \(gov.wales\)](#)
- The Climate Change Committee, 2021. Industrial decarbonisation policies for a UK net zero target.
- Theis, D.R. and White, M., 2021. Is obesity policy in England fit for purpose? Analysis of government strategies and policies, 1992–2020. *The Milbank Quarterly*, 99(1), pp.126-170.
- UNICEF, 2020. The Amsterdam Healthy Weight Approach: Investing in healthy urban childhoods: A case study on healthy diets for children.

- Vale, L., Adamson, A., Critchley, J., Rushton, S., Armstrong, N., Donaldson, C., 2012. Economic evaluation: evaluating the short-term impacts of the school food policy and experimental modelling of longer term impacts. *Public Health Research Consortium*.
- Welsh Government, 2015. Well-being of Future Generations (Wales) Act 2015.
- Welsh Government, 2019. Healthy Weight: Healthy Wales. Our long term strategy to prevent and reduce obesity in Wales.
- Welsh Government, 2020. Healthy Weight: Healthy Wales 2020 to 2022 delivery plan.
- Welsh Government, 2021a. Healthy Weight: Healthy Wales 2021 to 2022 delivery plan.
- Welsh Government, 2021b. National Survey for Wales, 2020-2021. [data collection]. UK Data Service. SN: 8870 [*Data Set*], DOI: [10.5255/UKDA-SN-8870-1](https://doi.org/10.5255/UKDA-SN-8870-1)
- Welsh Government, 2022a. Healthy Weight: Healthy Wales 2022 to 2024 delivery plan.
- Welsh Government, 2022b. Active Travel Approved Routes [*data set*], [Active Travel Approved Routes | DataMapWales \(gov.wales\)](#)

Annex A – Theory of change evidence base

Evidence base for the Food Environment theory of change

Description	Source
Advertising: Adults exposed to food advertising increased their consumption of unhealthy snacks by 28%.	Zimmerman, F.J. and Shimoga, S.V., 2014. The effects of food advertising and cognitive load on food choices. <i>BMC Public Health</i> , 14(1), pp.1-10.
Reformulation and production: Taxation on unhealthy food and drink led to reduced consumption of these products and led to firms reformulating their products.	Economics observatory, 2021. Would taxes on unhealthy foods reduce obesity?
Healthier retail and better consumer information: Calorie labelling led to a decrease in calorie consumption.	Crockett, R., King, S., Marteau, T., Prevost, A., Bignardi, G., Roberts, N., Stubbs, B., Hollands, G., Jebb, S., 2018. Nutritional labelling for healthier food or non-alcoholic drink purchasing and consumption. <i>Cochrane Database of Systematic Reviews 2018</i> , Issue 2.

Evidence base for the Active Environment theory of change

Description	Source
Active Travel: Population-level health benefits of active travel include reduced obesity and increased rates of physical activity.	Pucher, J., Buehler, R., Bassett, D.R. and Dannenberg, A.L., 2010. Walking and cycling to health: a comparative analysis of city, state, and international data. <i>American journal of public health</i> , 100(10), pp.1986-1992.
Safe Communities: The Safe Routes to School program (SRTS) in the US led to up to a 25% increase in walking and cycling.	McDonald, N.C., Steiner, R.L., Lee, C., Rhoulac Smith, T., Zhu, X. and Yang, Y., 2014. Impact of the safe routes to school program on walking and bicycling. <i>Journal of the American Planning Association</i> , 80(2), pp.153-167.
Green Space/Landscapes: Evidence suggests that local-area proportions of green space may be associated with happiness and life satisfaction.	Houlden, V., Weich, S. and Jarvis, S., 2017. <i>A cross-sectional analysis of green space prevalence and mental wellbeing in England</i> . <i>BMC public health</i> , 17(1), pp.1-9.

Infrastructure and Facilities: Attending a school with sub-standard infrastructure and facilities decreases the probability of participating in physical activity in adulthood (Black et al., 2019). Physical inactivity increases the risk of cancer, heart disease, stroke and diabetes by 20–30% (WHO).

(a) Black, N., Johnston, D.W., Propper, C. and Shields, M.A., 2019. The effect of school sports facilities on physical activity, health and socioeconomic status in adulthood. *Social Science & Medicine*, 220, pp.120-128.
 (b) WHO, 2022. Physical activity.

Evidence base for the Healthy Learning Settings theory of change

Description	Source
Healthy Childcare Settings: Providing children with a food service of healthy meals led to higher diet quality scores.	Yoong, S.L., Grady, A., Seward, K., Finch, M., Wiggers, J., Lecathelinais, C., Wedesweiler, T. and Wolfenden, L., 2019. The impact of a childcare food service intervention on child dietary intake in care: an exploratory cluster randomized controlled trial. <i>American Journal of Health Promotion</i> , 33(7), pp.991-1001.
Healthy Schools: Schools with bad infrastructure and facilities decreased the probability of doing physical activities in adulthood.	Black, N., Johnston, D.W., Propper, C. and Shields, M.A., 2019. The effect of school sports facilities on physical activity, health and socioeconomic status in adulthood. <i>Social Science & Medicine</i> , 220, pp.120-128.

Evidence base for the Healthy Work and Community Settings theory of change

Description	Source
Healthy Workplaces: Provision of healthy food services on worksites led to more consumption of fruits and vegetables and less consumption of fast food.	Dodson, E.A., Hipp, J.A., Gao, M., Tabak, R.G., Yang, L. and Brownson, R.C., 2016. The impact of worksite supports for healthy eating on dietary behaviors. <i>Journal of occupational and environmental medicine/American College of Occupational and Environmental Medicine</i> , 58(8), p.e287.

Evidence base for the Prevention and Early Intervention theory of change

Description	Source
The Best Start: Substandard diet quantity and/or quality contributes to the loss of developmental potential and life-long health and economic disparities.	Hurley, K., Yousafzai, A. and Lopez-Boo, F., 2016. Early Child Development and Nutrition: A Review of the Benefits and Challenges of Implementing Integrated Interventions. <i>Advances in Nutrition</i> , Volume 7, Issue 2.

Evidence base for the Targeted & Specialised Services theory of change

Description	Source
Specialised Support: Bariatric surgery is a proven method for achieving weight reduction.	Douglas, I.J., Bhaskaran, K., Batterham, R.L. and Smeeth, L., 2015. Bariatric surgery in the United Kingdom: a cohort study of weight loss and clinical outcomes in routine clinical care. <i>PLoS medicine</i> , 12(12), p.e1001925.

Annex B – Databank

Data for process evaluation

Sub-theme (activity)	Indicator	Description	Source	Data Granularity
Active environment (Active travel)	Length of active travel network (km)	The length of the active travel network. This can be at the national, health board and local levels.	Welsh Government (2022b)	N/A
Active environment (Green spaces/ landscapes)	% of urban area which is green space	Satellite data can be used to show the evolution of green spaces versus grey spaces in urban areas.	<i>Suggested</i>	N/A
Targeted and specialist services (Specialised support)	Waiting times for dietetics	Waiting times for specified diagnostic and therapy services.	Digital Health and Care Wales (2022)	Local Health Boards

Data for outcome and impact evaluation

Sub-theme (activity)	Indicator	Description	Source	Granularity	Counterfactual
Food environment (Healthier nutrition habits)	Consumption of fruit and vegetables	Number of portions of fruit and vegetables eaten yesterday.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
Food environment (Healthier nutrition habits)	Consumption of unhealthy foods	Whether or not a respondent has started eating less food which is high in sugar, (saturated) fat or salt in the last 6 months.	Food and You Survey	Individual/household level	Comparable data is available for the other countries of the UK
Active Environment (Improved physical wellbeing)	Obesity-specific quality of life (QSQL)	A contextualised questionnaire which asks questions like "Do you feel your weight impedes your daily tasks/activities."	<i>Suggested</i>	Individual/household level	N/A
Active Environment (Active travel)	Rate of physical activity (children)	The number of days in a week that the child is active for 1 hour or more.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
All sub-themes & the strategy as a whole	Behaviour change and decision making relating to obesity-related activities	Self-reported survey data on changes to habits.	<i>Suggested</i>	N/A	Comparable data is available for the other countries of the UK
Healthy learning (All activities)	Consumption of soft drinks, fruits and vegetables (children)	Whether the child (a) drinks coke (b) drinks diet coke (c) eats fruit (d) eats vegetables every day.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
Food environment (All activities)	Purchasing habits	Purchasing data on over 100,000 food and drink products	Kantar	Individual/household level	Kantar has data for other

Sub-theme (activity)	Indicator	Description	Source	Granularity	Counterfactual
		which includes the caloric, fat, carb, sugar, protein, fibre and sodium contents of the products.			countries which could be used as comparators
Active environment (Active travel)	Use of active travel for commuting to school (children)	Modality of transport and distance of commute.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
Active environment (All activities)	Participation in physical activity	Minutes of moderate to vigorous physical activity per week.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
All sub-themes & the strategy as a whole	Children's BMI	Height, weight and BMI for children aged 4 to 5 years old.	Child Measurement Programme (CMP)	Individual/household level	Comparable data is available for the other countries of the UK
Targeted and specialist services (Specialised support)	Participation in physical activity by obese individuals	Combination of BMI and physical activity indicators.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
All sub-themes & the strategy as a whole	Obesity and healthy weight	BMI.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
Healthy work and community settings (All activities)	Availability of healthy food and drink at places of work	A survey targeted at workers in private and public sector asking questions such as "Does your workplace encourage healthy eating habits?" and "Are healthy	<i>Suggested</i>	Individual/household level	N/A

Sub-theme (activity)	Indicator	Description	Source	Granularity	Counterfactual
All sub-themes & the strategy as a whole	Rates of diabetes	Prevalence of type 2 diabetes. food options available either in your place of work or in the surrounding area?"	National Diabetes Audit Programme	Individual/household level	Comparable data is available for the other countries of the UK
All sub-themes & the strategy as a whole	Physical health and wellbeing	Self-reported general health and prevalence of illness.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
All sub-themes & the strategy as a whole	Mental wellbeing	Warwick-Edinburgh Mental Well-being Scale.	National Survey for Wales	Individual/household level	Comparable data is available for the other countries of the UK
All sub-themes & the strategy as a whole	NHS costs	Annual NHS expenditure by disease category.	StatsWales: NHS programme budget	National	Comparable data is available for the other countries of the UK
All sub-themes & the strategy as a whole	Air pollution	Air Quality Indicators for average NO2, PM2.5 and PM10 concentrations across local authority areas and health board areas, derived from modelled data for each square	Stats Wales: Air quality indicators	Local authority, health board	Comparable data is available for the other countries of the UK

Sub-theme (activity)	Indicator	Description	Source	Granularity	Counterfactual
All sub-themes & the strategy as a whole	Rate of absence due to sickness	kilometre in Wales, measured in $\mu\text{g}/\text{m}^3$. Sickness absence rates of workers in the UK labour market, including number of days lost and reasons for absence.	ONS	Health board level	Comparable data is available for the other countries of the UK

Data for economic evaluation

Indicator	Description	Source	Units
Improved mental wellbeing	NHS cost savings associated with improved mental health.	Greater Manchester Combined Authority (2014)	£ per person per year
Improved general health	The monetary value of improvements in self-reported health.	Housing Associations' Charitable Trust (2014)	£ per person per year
Reduction of obesity	Social annual cost of obesity.	Frontier Economics (2022)	% of GDP
Improved physical wellbeing	The monetary value of keeping physically fit.	Housing Associations' Charitable Trust (2014)	£ per person per year
Improved physical wellbeing	The monetary value of walking regularly.	Housing Associations' Charitable Trust (2014)	£ per person per year
Reduced air pollution	Environmental prices for air pollutants.	Moon et al. (2021)	Euro per kg
Reduced emissions	Carbon values.	Department for Business, Energy & Industrial Strategy (2021)	£ per tonne
Less sickness absences	Labour productivity.	Office for National Statistics (2021)	Output per worker per year
Reduced health and diet inequalities	Cost of health inequality.	Mackenbach et al. (2011)	% of GDP

