

How to Monitor for Population Health Outcomes: Guidelines for developing a monitoring framework

Public Health Intelligence
Occasional Bulletin No. 44

Author

Dr John Wren from the Public Health Intelligence (PHI) Unit
in the Ministry of Health wrote this guide.

Citation: Ministry of Health. 2007. *How to Monitor for Population Health Outcomes: Guidelines for developing a monitoring framework. Occasional Bulletin No. 44.*
Wellington: Ministry of Health.

Published in July 2007 by the
Ministry of Health
PO Box 5013, Wellington, New Zealand

ISBN 978-0-478-19150-9 (print)
ISBN 978-0-478-19153-0 (Internet)
HP 4414

This document is available on the Ministry of Health website:
<http://www.moh.govt.nz>



MANATŪ HAUORA

Foreword

Public Health Intelligence (PHI) is the epidemiology group of the Ministry of Health. PHI carries out the Ministry's statutory responsibility to monitor the health of the New Zealand population by analysing population health outcomes and risks and determinants, and examining inequalities across regional boundaries and between population groups. An important role for PHI is the delivery and dissemination of epidemiology evidence for the development of policy and decision-making in the health sector.

How to Monitor for Population Health Outcomes presents guidance to public and population health programme managers, and interested others, on how to develop indicators to monitor progress on achieving population health outcomes sought from their programmes. This guideline has been designed to complement the information and guidance presented in the Ministry's *Guide to Developing Public Health Programmes: A generic programme logic model* (Ministry of Health 2006). The need to develop a robust outcomes monitoring framework is driven by an increased expectation from central government that there will be a focus on results in the design and delivery of publicly funded services. This expectation is highlighted by the requirements of the Public Finance Act 2004 and the Crown Entities Act 2004¹ (The Treasury and State Services Commission 2007a; b).

We welcome your comments and suggestions about the contents, and any additions or clarifications you might have.



Barry Borman
Manager (Epidemiologist)
Public Health Intelligence

1 Crown entities includes District Health Boards.

Acknowledgements

The author wishes to particularly thank:

- Geoff Stone, Senior Advisor, Centre for Social Research and Evaluation, Ministry of Social Development, for his contribution to the development of definitions in the Glossary and insights into outcomes monitoring in social settings other than public health
- Craig Wright, Senior Advisor (Statistics), Public Health Intelligence (PHI), Ministry of Health, for his significant preparation in Part C of the discussion on the small numbers problem and the statistical ability to detect change in health outcomes of interest; and
- Dyfed Thomas and Sarah Gerritsen, (PHI), for their preparation of Part D.

The author also wishes to thank the following external peer reviewers for their insightful comments:

- Anne Dowden, Director – Evaluation, Research New Zealand
- Rob Smith and Tony Walzl, Allen and Clarke Policy and Regulatory Specialists
- Ray Prebble, Editor, Macmillan and Prebble.

The following people provided valuable internal peer review:

- Dr Chris Wong, Public Health Physician, Ministry of Health
- Sarah Gerritsen, Senior Advisor (Population Health Research), PHI
- Dr Kirstin Lindberg, Senior Advisor (Public Health Medicine), PHI.

Contents

Foreword	iii
Acknowledgements.....	iv
Part A: Introduction	1
How to use this guide	1
Why measure performance?	2
Part B: What Objectives and Associated Outcomes Need to be Monitored?	3
Introduction: outcomes monitoring programmes	4
What should you measure to demonstrate the success of the programme?	4
What makes good outcome statements that are measurable?	6
Have you used action words in active statements?	6
Are they SMART objectives and/or outcomes?	7
Monitoring double-barrelled objectives.....	8
Prioritising objectives/outcomes for measurement	8
Management selection criteria for prioritising what is essential to monitor in public health programmes	10
Instrument selection criteria for prioritising what is possible to monitor appropriately in public health programmes	12
Additional questions and guidance for selecting outcome measures and indicators	13
Do you need a new outcome indicator or instrument?	13
Match the measurement instrument to the objective	14
Cost-effectiveness measures.....	15
Decide on the type of instrument	15
Collecting and analysing information.....	16
Part C: Issues to Consider in Outcomes Monitoring	17
Outcomes monitoring as a tool for ‘continuous programme improvement’: the problem of accountability, attribution and performance management	17
Outcomes monitoring versus evaluation	19
The problem of small numbers	19
Detecting change in small numbers	20

Part D: Available Data and Information from Public Health Intelligence.....	24
Survey data.....	24
Administrative data.....	26
PHIOnline – www.phionline.moh.govt.nz	27
Glossary	30
References	44
Further reading.....	45

Figures:

Figure 1: A framework for prioritising what outcomes to monitor.....	9
Figure 2: Simulated effect of 15% reduction in violent offences.....	22
Figure 3: Power to detect change due to an intervention, for specified incidence, population size and 10% incidence reduction.....	23
Figure 4: The PHIOnline home page	27
Figure 5: PHIOnline single map interface	28
Figure 6: PHIOnline double map interface	29

Tables:

Table 1: Distinguishing objectives, outcomes and outputs	5
Table 2: List of action words for writing performance objectives.....	6
Table 3: Defining SMART objectives.....	7
Table 4: Checklist and score card for prioritising outcomes and indicators for monitoring	13
Table 5: Summary of New Zealand Health Monitor surveys, 2002–2012	25
Table 6: Summary of administrative data sources.....	26

Part A: Introduction

Key Points

- This *How to* aims to help public and population health programme managers to develop timely, appropriate and cost-effective outcome measures for programmes funded by government.
- Measuring programme performance is essential for maintaining and enhancing support for public health programmes.
- The process of developing a logic model and an outcomes monitoring framework is intended to facilitate a process of ‘continuous programme improvement’.
- This *How to* should be read in conjunction with *A Guide to Developing Public Health Programmes: A generic programme logic model* (Ministry of Health 2006).

How to use this guide

This *How to* aims to provide guidance to public and population health programme managers and planners – whether they be in District Health Boards, non-governmental organisations or government departments – on how to select timely, appropriate and cost-effective outcome measures for public health programmes.

Throughout this *How to* we aim to provide plain English advice, checklists and examples to guide you through the process of selecting which outcomes to measure and how they should be measured. We provide information about some issues that are likely to arise from adopting a more rigorous outcomes-focused framework to monitoring public health programmes. We also outline the types of data held by the Ministry of Health that may be freely used to help monitor the performance of your public health programmes.

The guide is structured into five parts.

- **Part A: Introduction** outlines the purpose of this document and where it fits in relation to other Public Health Intelligence (PHI) publications.
- **Part B: What Objectives and Associated Outcomes Need to be Monitored?** describes a process for translating policy goals and objectives into measurable outcome statements, and a process for selecting and prioritising the outcomes to be monitored.
- **Part C: Issues to Consider in Outcomes Monitoring** discusses a number of key issues arising from the application of an outcomes monitoring approach to publicly funded programmes and small populations.
- **Part D: Available Data and Information from Public Health Intelligence** introduces PHIONline, a free public access internet-based service to a wide range of government statistics that may be useful in monitoring a population health programme.
- A **Glossary** provides an alphabetical list of terms, definitions and explanatory text relating to outcomes monitoring language.

This *How to* builds on the material presented in *A Guide to Developing Public Health Programmes: A generic programme logic model* (Ministry of Health 2006). The *Guide* provides a general introduction on how to design and implement comprehensive and measurable public health programmes in New Zealand, and includes some material on designing outcome measures (Ministry of Health 2006). This *How to* significantly expands on that material, and material presented at a PHI analytical workshop on Monitoring Public Health Outcomes held in October 2006. Ideally the *Guide* should be read before this *How to*.

The Ministry of Health's Leading for Outcomes website <http://www.leadingforoutcomes.org.nz/> is another Ministry-led initiative that promotes and uses an outcomes framework to improve how we think about and improve health in New Zealand, starting with cardiovascular disease (CVD) and diabetes. The website sets out a model of risk factor and disease progression, an outcomes hierarchy and an indicators framework for measuring progress towards achieving better CVD and diabetes health outcomes in New Zealand, with a view to extension of the approach to other health and disability areas.

Why measure performance?

The main motivation for introducing management tools such as programme logic models, outcomes monitoring and managing for results is an increased expectation from central government that the design and delivery of publicly funded services will have a strong results focus. This expectation is highlighted by the requirements of the Public Finance Act 2004 and the Crown Entities Act 2004² and associated guidance documents (The Treasury and State Services Commission 2007a; b) for the development of statements of intent by departments and Crown agencies.

To sum up, measuring performance:

- facilitates change and improvement
- is a mechanism for accountability
- supports planning and decision-making relating to resources
- can highlight areas requiring further work.

Ideally, performance measures for public health programmes should provide information about:

- a change in health status and health determinants achieved in priority population groups, including changes in inequalities
- resource and service utilisation
- the programme's responsiveness to the target population.

² Crown entities includes District Health Boards.

Part B: What Objectives and Associated Outcomes Need to be Monitored?

Key Points

1. It is not always possible to monitor every objective and outcome, so you will need to establish a list of prioritised objectives/outcomes for monitoring.
2. The process of planning an outcomes monitoring framework begins with translating policy goals and objectives into SMART objectives:
 - Specific
 - Measurable
 - Achievable (sometimes 'Accurate' or 'Action-oriented' are used)
 - Relevant (sometimes 'Realistic' is used)
 - Time-based.
3. Management and instrument criteria can be used to assess whether it is possible and essential to monitor the objective or outcome of interest. Management selection criteria are:
 - attribution (accountability)
 - centrality
 - cost-benefit
 - robustness to withstand public scrutiny
 - timing.

Instrument selection criteria are:

 - availability
 - reliability
 - sensitivity
 - validity.
4. Planning for outcomes monitoring requires decisions about:
 - What information is required? What is the data going to be used for?
 - What data is already available? Is new data really needed?
 - Who is going to collect the data?
 - What type of measurement instrument could or should be used?
 - How many measurement instruments are to be used?
 - Who is going to do the analysis?
 - When is the data required by to inform timely decision-making?
5. Table 4 provides a checklist and scorecard to help prioritise which objectives and outcomes are essential to monitor.
6. See the Glossary for a full list of terms and associated definitions used throughout this *How to*, and extra explanatory comment and examples illustrating the use of the terms.

Introduction: outcomes monitoring programmes

Establishing which outcomes need to be – and can be – successfully monitored requires careful thought and planning. Preparing an outcomes monitoring plan will help you to establish a successful monitoring programme.

An outcomes monitoring plan is a process for the routine systematic collection and recording of timely information about aspects of a programme to assess whether progress is being made on achieving the programme's objectives, and how the programme could be improved. Information is timely when it contains data directly relevant to the programme and – importantly – is available in a timeframe where it can usefully inform decisions about whether the programme is performing as planned, and what things need to be changed to improve the performance of the programme if required.

Collecting baseline data at the start of a programme is vital to be able to assess what change has occurred over time, particularly when you are trying to attribute a change to an effect of the intervention.

In preparing the outcomes monitoring plan you will need to assess:

- what needs to be measured to demonstrate success
- how things should be measured and by whom (deciding what indicators are going to be used, after considering issues such as validity, reliability, sensitivity, attribution, availability of suitable instruments, cost-benefit)
- the timeframes for delivering information about the performance of the programme against the stated objectives to decision-makers and key stakeholders
- whether a formal evaluation process is required.

A good outcomes monitoring process:

- identifies and prioritises the outcomes that are essential to monitor and that can be monitored appropriately
- provides timely information to key decision-makers and stakeholders about the progress made on achieving the desired outcomes
- uses outcome measures robust enough to withstand public scrutiny
- is cost effective
- includes baseline data relevant to the intervention.

What should you measure to demonstrate the success of the programme?

Critical to any successful outcomes monitoring plan is to identify what could and should be measured in order to show that the programme is being implemented as planned, and that progress is being made to achieve the desired health outcome(s). The steps involved are:

1. identify what *could* be measured
2. prioritise these into those that *should* be measured if resources allow
3. identify the essential few that *must* be monitored.

Underpinning this approach is the assumption that it is not always feasible – or necessary – to measure everything.

The first place to look to determine what could and should be monitored is to identify what the key programme objectives, outcomes and outputs are. You should be able to identify these from the ‘logic model’ developed for the programme. A logic model is a planning tool that describes the purpose, what, where, when, and how of the programme being implemented.

Ideally, programme objectives, outcomes and outputs should be clearly distinguished from each other (see Table 1 for definitions of these terms) and expressed in action words. Expressing outcomes statements clearly makes it easier to identify what actions need to be undertaken, what changes need to take place, what the desired end result is, and consequently what needs to be monitored.

Table 1: Distinguishing objectives, outcomes and outputs³

<p>Objectives are statements about the results a programme seeks to achieve. Any programme must have at least one objective.</p> <p>Objectives may form a hierarchy that moves from a limited set of high-order objectives that are synonymous with aims or goals to be achieved in the long term (five to seven years). Underneath high-level objectives are more intermediate-level objectives that are to be achieved in a three- to five-year time frame, and which must be achieved in order to attain the high-level objective. The lowest levels of objectives are immediate or operational objectives that must be achieved first – typically in one or two years.</p> <p>Objectives may be translated directly into ‘outcomes’ if they deal with only one issue. However, double-barrelled objectives will require multiple outcome measures to be developed.</p>
<p>Outcomes are specific statements about the intended change in public health-related attitudes, knowledge, behaviours, or physical health status in the target population(s) sought by undertaking the planned public health activity. In some situations ‘process’ outcomes may be desirable.</p> <p>Process outcomes typically measure the amount of effort put into a programme and the quality of the service provided. They can also be appropriate where it is important to monitor community support for a programme. Process outcomes that measure effort can be expressed as ‘outputs’.</p>
<p>Outputs are things (such as goods) produced, services delivered, events held, or participation generated resulting from the activities undertaken.</p>

3 See the Glossary for a full list of other terms and associated definitions used throughout this *How to*, and extra explanatory comment and examples illustrating the use of the terms.

What makes good outcome statements that are measurable?

Have you used action words in active statements?

In some cases, even though a programme's objectives and outcomes can be identified, they may be stated in a way that makes them difficult to measure or unsuitable for programme planning or contracting purposes. This can happen because the language used is too passive in tone, or is too complex or abstract. Consequently, in some situations it may be necessary to translate the identified objectives into language that is more suitable for monitoring. Ensuring the language used is active rather than passive will help to do this.

Active statements involve 'action' words, and a list of action words identified as suitable for public health is provided in Table 2.

Table 2: List of action words for writing performance objectives

Accept	Adopt	Advocate	Analyse	Arrange
Approve	Appraise	Bargain	Calculate	Care
Change	Choose	Classify	Categorise	Challenge
Chart	Compare	Conduct	Construct	Contrast
Co-operate	Check	Defend	Define	Demonstrate
Describe	Develop	Differentiate	Discriminate	Draw
Evaluate	Execute	Explain	Express	Fill out
Forecast	Formulate	Generate	Identify	Inform
Instal	Interview	Judge	Justify	Label
List	Locate	Manipulate	Modify	Name
Operate	Organise	Outline	Persuade	Plan
Prepare	Prescribe	Produce	Purchase	Question
Rank	Recall	Recognise	Reflect	Remove
Research	Resolve	Review	Select	Sort
Specify	State	Study	Take	Tell
Translate	Use	Write		

Source: Bartholomew et al 2006

Take, for example, the following objective:

Communication plans *will be developed* to ensure stakeholders are kept informed.

This sentence contains the idea of what needs to be done, but responsibility for doing it and therefore an emphasis on the action needed is lost because of the passive form. Compare this with the alternative wording, using an active form and action words:

The committee will develop communication plans to ensure stakeholders are kept informed.

Using a bullet list for objectives lets you start each objective with an action word, which can be a forceful way of conveying the objectives; for example:

The committee will:

- *develop* communication plans to ensure stakeholders are kept informed
- *prepare* a report on the nutrition in schools open day
- *inform* parents about the progress their children have made.

Are they SMART objectives and/or outcomes?

Another good approach to identifying which objectives and/or outcomes are able to be measured, or to reconfigure them so that they are measurable, is to ask yourself, are they SMART objectives (Iverson 2003)? SMART objectives/outcomes are defined in Table 3.

Table 3: Defining SMART objectives

Item	Definition
Specific	An objective should address a specific target or accomplishment. Specific implies that an observable action, behaviour or achievement is described, which is also typically linked to an identifiable change in rate, number, percentage or frequency.
Measurable	A method should be established to indicate that an objective has been met. That is, there should be a system, method or procedure for tracking and recording the change in behaviour or action towards which the objective is directed.
Achievable*	Though not necessarily easy or simple, the objective should be feasible – that is, capable of being achieved. Objectives should be limited to what can realistically be done with available resources, and ideally the resulting change should be ‘attributable’ to the action undertaken.
Relevant*	An objective should be significant to the people involved in the programme (from beneficiaries to the programme’s sponsoring organisation), and the objectives should be capable of having an impact or making a change.
Time-based	An objective should be achievable within a specific timeframe. Generally this takes the form of a start and end date. The time may be short (two or three months, up to two years), medium (three to five years), or long term (five years to seven years).

Source: Iverson A. 2003. *Preparing Program Objectives: Theory and Practice*. The International Development Research Centre: Evaluation Unit, Toronto

* What A and R stand for is inconsistent in the literature. A is sometimes given as ‘accurate’, ‘action-oriented’, ‘accountable’ or ‘attributable’. R is sometimes given as ‘relevant’ or ‘realistic’. In some situations, these alternative phrases may be more useful, in which case use the term that is most appropriate or that most clearly helps to describe the intent of the objective.

Monitoring double-barrelled objectives

Double-barrelled objectives contain multiple components, and if well written have the advantage of being able to convey multiple objectives in one single sentence rather than using several sentences. However, they are also more complex to monitor. Using the above criteria, we need to ask the question: Are double-barrelled objectives SMART? Here are two examples of double-barrelled objectives.

- Objective 1: Enable people with chronic conditions to improve their health, slow progress of their condition(s), and maintain independence wherever possible by aligning community and hospital services across [] DHB.
- Objective 2: Reduce the incidence of cancer, diabetes, and cardiovascular disease by 20% over the next five years.

The first objective has multiple components in that it refers to chronic conditions, and talks about slowing the progress of the conditions, maintaining independence and aligning services. It could be argued that this objective, as written, is a good high-level objective in that it uses some action words that provide an overall general direction for the DHB, and it complies adequately with the SMART criteria in Table 3 in terms of a high level objective statement.

However, for operational and monitoring purposes it does not adequately comply because it does not specify the chronic conditions, time frame and services. All of these issues would have to be addressed when developing the logic model and outcomes monitoring plan to achieve this high-order objective. Multiple outcome measures would have to be used to assess whether this objective was achieved, because different measures would need to be used for each component.

The second objective also has multiple components, but it clearly specifies the chronic conditions of interest, the size of change sought, and the timeframe for achieving the changes. The objective complies with the criteria in Table 3, although the objective would also require the use of multiple outcome measures – a different measure for each chronic condition – for progress on its achievement to be monitored appropriately.

In answer to the question, are double-barrelled objectives SMART? the best advice is generally to avoid double-barrelled objectives when writing low-order (ie, operational) objectives. This will make planning clearer and the objectives easier to monitor. Where double-barrelled objectives may be desirable (eg, when writing high-level objectives), make sure they comply as closely as possible with the criteria in Table 3. Also, remember that where double-barrelled objectives are used, multiple outcome measures will generally have to be developed to monitor each component of the objective statement.

Prioritising objectives/outcomes for measurement

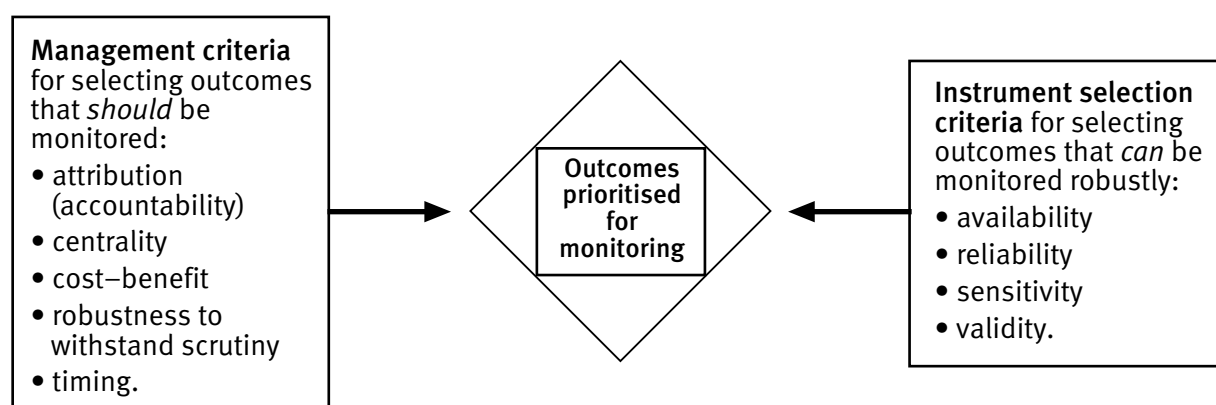
After identifying and clarifying the range of objectives and/or outcomes that could be monitored in an ideal world, it is highly likely that it will not be feasible – or necessary – to monitor all the objectives and outcomes planned for. Consequently, a process of prioritisation will have to take place to select the objectives and outcomes that are both essential and possible to monitor.

We suggest two types of criterion to use as aids in making decisions about which outcomes should be and can be monitored:

- management considerations
- instrument selection considerations.

Management considerations are concerned with promoting the development of an appropriate and robust monitoring system. Instrument selection considerations are concerned with establishing scientific credibility for the monitoring system. Figure 1 illustrates how these two types of criterion relate to each other. They are also further discussed in the following pages. Although the criteria can be defined separately, in practice they are interrelated and a judgement will have to be made as to where the balance lies between the merits of each.

Figure 1: A framework for prioritising what outcomes to monitor



By examining each objective and indicator using the criteria set out below, you will be able to develop a prioritised selection of the outcomes and associated indicators that *can* and *should* be measured appropriately. To help with this, it may be useful to use a scorecard such as that provided in Table 4. The scorecard provides a method for rating each outcome and indicator against the criteria. Note that more than one indicator can be used to monitor an outcome, but it may not be necessary or cost effective to do so.

The scorecard is simple to use: just assign a ‘1’ score to each criterion that has been successfully met. All the scores are added up, and the outcomes and associated indicators with the highest score should receive the highest priority for inclusion in the outcomes monitoring plan. A slightly more complex approach, which would provide better differentiation between possible indicators, would be to use a scoring approach that allows for decimal points; for example, instead of scores of 1 or 2, a score of 1.5 could be allocated. In some situations a more sophisticated approach may be desirable, in which case a ‘weighted’ system could be applied to the scorecard. In a ‘weighted’ approach, some criteria would be judged as more or less important than others, and consequently a higher or lower range of possible scores could be allocated to the selected criteria. For example, because of the size and nature of the programme it may be decided that the instrument selection criteria of validity and reliability are so important that they are worth double points compared to the others (ie, the results are going to be weighted (biased) towards indicators that score well on those factors).

Management selection criteria for prioritising what is essential to monitor in public health programmes

Attribution (accountability)

Attribution is the extent to which change in the outcome of interest is associated with the type of activity undertaken. Attribution is an important issue to consider because it has implications for how well a programme's activities can be said to have resulted in any changes observed, and for assigning the level of accountability for the success or failure of a programme or its components.

Strong attribution requires being able to establish a clear and unambiguous causal link between what you do and what happens – something that is often problematic for many public health programmes. When assigning attribution, you need to beware of any 'attributional bias' that results from over-attributing a change to any particular activity.

Centrality

How important is the outcome of interest to establishing the success of the programme? The more central the outcome, the more important it is to measure it appropriately. Where data is not available, think carefully about developing a new measure, or using an appropriate 'proxy' measurement (see 'Availability' below).

Whether a programme is deemed successful or not is not always related to whether the central health outcome of interest has been achieved. A programme may fail in one aspect but still be successful if other outcomes of importance have been achieved.

Cost–benefit

'Cost–benefit' refers to the balance between the cost of using and/or developing a measure and the benefit that will be gained from implementing it. For many programmes, you will need to make a judgement about whether the cost of developing a monitoring regime or measure is worth the benefit to be gained from the information provided. Where the cost outweighs the benefit, then consider using an appropriate proxy indicator, or information from a pre-existing monitoring regime.

In general, the proportion of the total budget spent on monitoring should be in the region of 5% to 20%, depending on the type and size of the programme. If the programme is a pilot initiative, the results of which could be used to significantly influence whether the programme is expanded into a major effort, then 20% of the total budget could be appropriate. If the programme has a large budget and is well established, applied in standard ways and supported by evidence for its effectiveness, then a budget allocation of 5% may be more appropriate for monitoring purposes.

Cost per output is probably the best method to use when calculating the cost of a monitoring programme or a component of it. For example, the cost of purchasing information from an existing data set may be cheaper than doing it yourself. Or, it may be cheaper to use a proxy measure than to use a direct measure (assuming the proxy measure is robust enough to withstand public scrutiny and the information trade-off is acceptable). Note that cost effectiveness is just one of a number of criteria that should be used in planning an outcomes monitoring regime.

Robustness to withstand scrutiny

Robustness will be provided by ensuring that the instrument selection criteria (presented below) are observed. Be very clear about why a particular set of outcomes has been selected for monitoring and not others, and why particular measurement instruments have been selected over others. Ask yourself the following.

- What is being monitored?
- Why and how is it going to help to assess the effectiveness or efficiency of the programme as a whole?
- Why is it essential?
- When must the information be available?
- How is the data to be collected?
- Who is going to do the collection and analysis?
- Is the proposed approach cost effective?
- Will the information be provided in a timely manner?
- Will the monitoring system withstand scrutiny by stakeholders?

For public health programmes funded by government money, it is important that assessments of the effectiveness of the programme be able to withstand public scrutiny. In this context, we suggest it is usually advisable not to attribute a programme's success to a single criterion: robustness is provided by the strength of logic of the total outcomes monitoring regime.

Timing

Timing has two aspects to it. The first concerns establishing 'attribution' (see above). In this case, timing refers to establishing the time sequence between when an activity took place and when a change in the desired outcome was observed. The second aspect relates to being able to report to stakeholders on the performance of a programme in a timely manner, which also means the proposed indicator must be able to be produced within an appropriate timeframe.

As a general rule, the timeframe for short-term objectives can be as short as two to three months or up to two years. The medium term is defined as three to five years, and a timeframe of five to seven years is usual for achieving long-term objectives (The Treasury and State Services Commission 2007a; b).

In practical terms, it may not be feasible to monitor/measure progress on attaining short-term, or even medium-term, objectives using national data. For example, information from national survey and administrative data sets is typically not available anywhere from 6 to 36 months after the data was originally collected due to data checking and quality control processes. This may mean that it is not feasible to use information from these data sources as indicators because the data is not available soon enough.

Instrument selection criteria for prioritising what is possible to monitor appropriately in public health programmes

Availability

This refers to whether data, or a measuring instrument, already exists at the local or national level that could be used to help monitor the outcome of interest. Where possible, use existing data that may be sourced from within your organisation or a range of government agencies or organisations such as public health organisations in your area. A good place to start seeing what data is already available to help monitor your programme is at PHIONline, www.phionline.moh.govt.nz, which includes data at national, District Health Board and Territorial Local Authority level (see Part D for an introduction to the range of information available to you on PHIONline). Using existing information sources can save a considerable amount of time and effort.

Where such data exists, efforts should be made to use that data or instrument rather than invest in developing new data sources or instruments. If issues such as cost, complexity and/or timeliness prevent direct measurement of the outcome of interest, consider using a proxy indicator. If achieving the central outcome is vital, and data is not readily available, careful thought needs to be given to spending extra effort on developing a new measure, or on identifying a suitable proxy measure.

Reliability

Reliability is the extent to which a measure, when used repeatedly in the same way, will produce the same or a similar result. For example, if you were interested in monitoring a person's weight and used a weighing machine that produced a different reading each time it was used (when all other factors were the same), then the machine would not be a reliable instrument to use.

Sensitivity

Sensitivity refers to how well a measure is able to detect when a change has occurred in the outcome being monitored. For example, a weighing machine that was able to differentiate a change in weight by 100 gram graduations is more sensitive than a machine that can only detect changes in 500 gram amounts.

Specificity is another criterion that is particularly appropriate in clinical settings, where it is used to assess how accurate a clinical test is in identifying people at risk.

Validity

Validity refers to how well a measurement, index or indicator reflects the outcome it is intended to measure. For an outcome measure to be valid, it should be both sensitive and specific.

Often there is more than one way of measuring an outcome, and the most valid measure may not be practical in all situations. For example, the best way to measure body composition (particularly lean and fat mass) is under-water densitometry (weighing), but this is not practical outside small clinical studies. Therefore, measures of weight adjusted for height (eg, body mass index, BMI) are often used to reflect body composition, particularly body fat mass. Although BMI is correlated with body fat mass, this relationship varies according to body build, age and ethnicity. Therefore, BMI is not recommended for assessing excess body fat mass (obesity) at an individual level.

However, BMI can be useful for assessing obesity at a population level if used and interpreted appropriately. BMI based on direct measurement of height and weight is more valid than BMI based on self-reported height and weight, particularly if measurements are taken by trained observers using appropriate equipment and standardised procedures.

Ultimately, the choice of measure for a particular programme will depend on a number of factors, including validity and practicality.

Table 4: Checklist and score card for prioritising outcomes and indicators for monitoring

Outcome of interest	Indicator name	Criteria									Total score	Priority
		At	Av	Ce	Cb	Re	Rb	S	T	V		
Outcome 1												
Output 1												
Objective 1												
Criteria abbreviations At Attribution (Accountability) Av Availability Ce Centrality Cb Cost–benefit Re Reliability Rb Robustness to withstand public scrutiny S Sensitivity T Timing V Validity												

Source: Adapted from Table 1 ‘How to select indicators’ (page 68) in United Nations Development Programme Evaluation Office 2002

Additional questions and guidance for selecting outcome measures and indicators

Do you need a new outcome indicator or instrument?

Think carefully before developing a new outcome indicator or instrument, ask yourself the following.

- Why is a new indicator/instrument needed?
- Is it because there is a problem with an existing indicator/instrument, or is it because it really is a new outcome?
- Can we use existing data creatively to answer the question?

Every instrument has its own strengths, weaknesses and peculiarities – the perfect instrument does not exist. If you are unfamiliar with the instrument that best suits your plans, seek help from someone who knows about it already. Public Health Intelligence at the Ministry of Health have extensive expertise in measuring public health outcomes. If in doubt, seek their advice.

Match the measurement instrument to the objective

Choose the level of observation

What is the focus of interest? Is it an individual or a group of people? If it is an individual, perhaps the measurement is of an individual's change in knowledge and behaviour due to the intervention, or maybe it involves observing a change in their health status over time. Is the focus on groups of patients (eg, patients from a particular age group, with a specific disease, or submitted to a certain intervention), or their ability to access services? If the intervention's utility or the general quality and cost-effectiveness of different care systems is the main interest, compare the quality of care between different systems; say, between primary and secondary care or between geographic locations.

Formulate and describe the measure's aims

What is your aim for the measure? Do you want to describe, compare or evaluate health outcomes? The selection of your instruments is highly related to the endpoints of your project. What do you want to use the instrument for?

The principal uses for a health measure are as follows.

- **A health status measure** can be used as an indicator, measuring the current health condition/ state of a person or a population group at a point in time. In addition to validity, both reproducibility and specificity to the chosen health condition are important. Reproducibility is particularly important when undertaking a robust pre- and post-assessment of change following the intervention.
- **A health outcomes measure** can be used as a comparison, relating differences at different points (eg, before and after intervention). For this type of action, sensitivity and responsiveness to change are important. Put simply, the measure must be able to register small changes in people's health.
- **A health outcomes assessment** implies that, apart from being an outcome measure, it is an attempt to use the information through feedback to the users of the information, including government funders of the programme. Apart from achieving the outcomes sought, government funders may also be interested in cost-effectiveness measures.
- **Process outcome monitoring** can use both qualitative and quantitative measures to report on process evaluation issues. This type of information helps to provide the 'story behind the statistics', and is useful for reporting on the following aspects that may be critical to the monitoring and success of the programme:
 - community capacity
 - service/agency capacity
 - compliance with good practice
 - programme integrity.

Examples of these types of measure are goal attainment scaling, global assessment scales and other rubric-based scoring.

Cost-effectiveness measures

Cost effectiveness refers to the balance between the cost of implementing the intervention and the benefit or effectiveness gained from programme. There are many ways of measuring cost effectiveness. The three approaches suggested by (The Treasury and State Services Commission 2007a; b) as being appropriate for Crown entities are:

- **cost-benefit analysis**, which is seen as the ‘gold standard’ for establishing cost effectiveness
- **cost per unit of impact/outcome**, which is a simpler approach (assuming it is feasible to quantify the impact/outcome gained)
- **cost efficiency**, which is a method by which the *cost per output* is measured. This approach is likely to be particularly appropriate where the service provided (eg, a public health information service) is remote from any health outcome.

Cost per output is probably the most useful method when calculating the cost of a monitoring programme or a component of it. For example, the cost of purchasing information from an existing data set may be cheaper than doing it yourself. Note that cost-effectiveness is just one of a number of criteria that should be used when planning an outcomes monitoring regime.

Decide on the type of instrument

It is important to note that the ‘psychometric’ qualities of the instrument you choose must be able to support your goals and objectives. This means it is essential that the instruments are valid and reliable (see above).

- In general, a *condition-specific measure* will have a narrow focus but will contain considerable detail in the area of interest. If you are interested, say, in one disease condition, and the assessment is mainly of symptoms and function, then use a condition-specific measure.
- If a specific domain, such as daily functioning or mental wellbeing in different populations, is your interest, use a *dimension-specific instrument*.
- If you are interested in general health or in the interaction between different conditions, or if you are interested in populations that may include healthy people, then you should use *generic instruments*.
- If you think the influence of other diseases or conditions that you have not measured may influence the results of the problem or the disease of interest, *combine disease-specific and generic instruments*.

No one instrument will prove satisfactory for all purposes. You may need to combine instruments because a reasonable instrument does not exist. But beware: when possible, use the instruments in their original form. Do not change them or use only parts of them. Validation covers only the complete instruments that were tested (see above on reliability and validity).

Be careful with instruments that are taken from one country to be used in another. Cross-cultural validation needs to follow strict rules. Even a survey that has been validated in Canada, the UK or the US may not be valid in New Zealand because of different cultural dynamics and language uses. The formal validation of an instrument is a costly and time-consuming process. How much of this work you do depends on your resources. And don’t forget the practicalities: the necessary time to fill in questionnaires and the costs of mailing and analysis. Think about your target group: for example, not every instrument suits children or older people.

The measures selected must align with the programme's goals and objectives. It is therefore important to understand the programme's goals and objectives so that appropriate selections can be made from the many measurement instruments available. In particular, you will need to determine whether an intervention's effect is to be measured or descriptively assessed.

You will need broadly validated instruments that have been used in other studies if you want the intervention to describe the health status of a defined population or a specific disease category. Short, feasible and reliable instruments are recommended if care providers are to use them in their clinical work.

Collecting and analysing information

Selecting the right measures is only part of the process of establishing a monitoring system. You will also need to think about how the information will be collected and processed. There is no point choosing the perfect set of measures if there is no feasible way to collect and use the information.

Ask the following questions.

- What is the cost of collecting the required information?
- How easy is it to access the information?
- Are there administrative, privacy and ethical issues?
- Has the cost of data analysis been allowed for?
- Who is going to do the analysis?
- Who is the audience for or user of the information?

It takes time and skill to correctly analyse and interpret data, and this needs to be allowed for when planning a monitoring regime. The amount of time to allow depends on the complexity of the analysis and the type of information collected. Ask the following types of questions.

- What type of analysis is required? That is, what type of questions have to be answered: descriptive or explanatory? (Generally, explanatory analysis will require more work than descriptive analysis to prepare.)
- Who is going to use the information?
- What decisions will be made with the findings?
- When does it need to be done by?
- Who is going to do it? Do they have the skills?

In general, where the users of the information are senior decision-makers and/or the funding for the programme is significant, the more robust the analysis needs to be in order to withstand scrutiny, and consequently the more time should be allocated to this task. For major regional and national programmes involving the analysis of multiple indicators and the preparation of complex tables and report writing, a time period of several weeks or months is not unreasonable.

Part C: Issues to Consider in Outcomes Monitoring

Key points

1. Outcomes monitoring is intended to help:
 - a. facilitate a process of ‘continuous programme improvement’
 - b. build confidence in, and support for, public health programmes.
2. Outcomes monitoring is not a cheap replacement for a traditional evaluation process.
3. Outcomes monitoring may, or may not, include a traditional formative process and impact evaluation exercise.
4. Beware of the problem of small numbers.

Outcomes monitoring as a tool for ‘continuous programme improvement’: the problem of accountability, attribution and performance management

The main motivation for introducing management tools such as programme logic models, outcomes monitoring and managing for results within the state sector is an increased expectation from central government that there will be a focus on results in the design and delivery of publicly funded services.

This expectation is highlighted by the requirements of the Public Finance Act 2004 and the Crown Entities Act 2004⁴ (The Treasury and State Services Commission 2007a; b). This legislation means that there is now an increased emphasis on ‘maximising results for the available resources’, and managers must report on the impacts, outcomes, or objectives that a programme seeks to achieve or contribute to, and how it will contribute to implementing government policy directions (The Treasury and State Services Commission 2007a; b). Increasingly, these expectations are now being transmitted into the services contracted for by government agencies.

The reporting requirements in the acts include stating the:

- rationale for the main types of interventions planned and the results expected
- risks identified and how they will be managed
- main measures that will be used to monitor progress and performance – including the cost effectiveness of the activities
- how the programme will link to other government agency programmes to support them, including those activities undertaken by non-government organisations (The Treasury and State Services Commission 2007a; b).

4 Crown entities includes District Health Boards.

Implicit in this approach is the idea of ‘accountability’ for achieving the objectives set and outcomes sought. As a general principle, the degree to which accountability can be maintained depends on the ability of the person or organisation to meaningfully influence the changes sought. Strong accountability can only be maintained where there is a clear and unambiguous causal link (ie, attribution) between cause and effect.

For many government activities, including public health programmes, establishing clear causal links between an intervention and change in the outcome sought – in this case population health – is difficult. A number of practical difficulties must be overcome (Nutbeam 1998). Consequently, when assigning attribution in public health programmes, you should beware of any ‘attributional bias’ that would result in over-attributing a change to a single programme activity.

It is likely that success will not be achieved through a single activity, but by undertaking a range of activities and working with other programmes and across-government agencies to achieve the objectives. This means that using outcomes monitoring as a strict contract performance management tool is problematic. This does not mean that non-government organisations and providers of government services should not be accountable for the choice of methodology and quality of implementation undertaken to achieve the health objectives.

The use of logic models and outcomes monitoring should be seen as a way of instituting a ‘continuous improvement cycle’ in the design and delivery of publicly funded programmes (The Treasury et al 2005), rather than as a contract performance management tool to be feared. With the continuous improvement approach, the performance of organisations is focused on their ‘understanding, reviewing, and learning from the efficiency and effectiveness of their operations’ (The Treasury et al 2005). The approach provides a tool for helping programmes to evolve over time in response to changes in the environment and as programme objectives are achieved.

This approach is a recognition that being able to attribute outcome changes to particular interventions or outputs ‘won’t always be feasible’ (The Treasury and State Services Commission 2007b). This does not absolve service providers from any responsibility and accountability for designing, implementing, monitoring and reporting on the effectiveness and efficiency of their programmes. It does mean that where establishing attribution is difficult, it is even more important that a clear logic (rationale) for the intervention be stated, and that where a number of similar but small programmes are being implemented, all the programmes use the same reporting framework and attribution takes place at the group level rather than at the individual programme provider level.

In terms of contract performance management, providers should be reassured that the approach recommended here includes considering the full range of reasons for meeting or not meeting a target. The important thing is that there is a robust programme planning and monitoring plan that will enable us to answer questions such as:

- What impact have our interventions had?
- What else was going on at the same time that also had an impact?
- Can we learn from our performance monitoring, and how can we improve in response to the lessons?

Outcomes monitoring versus evaluation

Along with the new emphasis on using logic models for planning and monitoring for health outcomes there has been an implicit shift towards a monitoring process that relies on quantitative methods using epidemiological and biostatistical methods, rather than traditional formative and process evaluations. This shift in emphasis should not be seen as something particularly new, but rather as a reflecting the ongoing evolution in public management thinking and practice. In this case, the use of logic models can be seen as an evolving from ‘Managing by Objectives’ to ‘Managing for Outcomes’. Similarly, ‘outcomes monitoring’ reflects an evolution in the practice of ‘evaluation’.

- Traditional evaluation of public health programmes tended to focus on qualitative methods and formative and process questions, although current practice is to use multiple methods.
- Monitoring for outcomes tends to use quantitative methodologies and focuses on endpoints.

However, depending on what the programme is supposed to achieve, formative and process evaluation may still form a legitimate part of an outcome monitoring framework. For example, where community acceptance of a new type of intervention is critical to achieving the ultimate health outcome sought, then a formative and process evaluation would be appropriate.

A full programme evaluation that includes formative, process and outcomes evaluation is likely to be appropriate for major national-level campaigns. Such evaluation examines long-term changes in health status and the determinants of health. These include changes in knowledge, awareness and behaviour; shifts in social, economic and environmental conditions; as well as changes to public policy and health infrastructure.

Outcome or impact evaluation also seeks to measure the reduction in health status inequities between population subgroups. In this approach, it is important to identify and measure short-, medium- and long-term outcomes to ensure the ongoing support and relevance of the activity for those whose agendas are shorter term. Outcomes evaluation also uses indicators as benchmarks, or proxy measures, to assess the extent to which objectives have been met. Matching objectives to associated indicators in a logic model helps to ensure the availability of relevant data sources for programme evaluation.

Note that a monitoring regime using quantitative methodologies should *not* be seen as a cheap replacement for evaluation. Both approaches require skilled staff to advise on the most robust methods and measures to use and to undertake the data analysis, and both rely on robust administrative systems.

The problem of small numbers

Monitoring the outcome of an intervention requires a health indicator that can, among other things:

- detect the outcome sought from the intervention
- be used to provide a baseline measure against which future measurements, which commence after the intervention has started, are compared
- consistently be used over the time period of interest.

The first is usually the hardest to determine, but statistical power analysis calculations and the explanations that form the logic behind the use of the indicator can help decide whether you should even begin a monitoring process using the indicator. The second is usually easy to determine, because it simply requires data before and after the intervention. The third is more difficult than it appears because it can be easy to introduce changes in the way the data is collected over time. Assessing whether the indicator has been consistently used requires someone with a good knowledge of the process.

A good health status indicator measures either:

- **prevalence** (the number people with the condition in the population, divided by the total number of people in the population), which can be expressed as a ratio; or
- **incidence** (the number of new cases of the condition in the population divided by the total number of people in the population) of a disease, condition or risk factor, which can be expressed as a ratio. Generally, incidence is for a specified period of follow-up of the population of concern (eg, one year).

In both measures, the numerator or population from which one identifies the cases must match the denominator population.

The problem with these types of indicators in New Zealand is that either or both the numerator (the number of people with the health status of interest) or the denominator (the number of people in the population as a whole) may be very small. When the numbers are small, then statistically it can be difficult to measure the prevalence or incidence, and it also becomes very difficult to detect whether a change has occurred, let alone to decide whether any detectable change is related to any effect of the intervention.

Example 1: Illustrating mortality incidence: New Zealand's crude mortality rate in 2006

In 2006, 28,390 deaths were registered in New Zealand and, as of 30 June 2006, 4,127,000 people were estimated to be living in New Zealand. So the crude mortality rate is 28,390 divided by 4,127,000, which equals 0.0069. This translates to 1 death per 145 people, or 688 deaths per 100,000 people in the denominator population.

This is an incidence rate. Mortality rates are always incidence rates due to the fact that there is no prevalent pool, because the cases are all deaths in New Zealand and so they match the same population as the denominator.

Detecting change in small numbers

The ability of an indicator to detect the effect of an intervention depends on a combination of factors, including:

1. **sensitivity** (the ability to pick up accurately a change in the outcome measured)
2. **specificity** (how accurate a test is at identifying people at risk – this is particularly important in clinical testing situations)
3. **power** (the probability of detecting an effect, or degree of change, of a specified size).

Determining the sensitivity and specificity of an indicator is often complicated and requires specialist knowledge beyond the scope of this *How to*. The power of your indicator to detect changes in the outcome of interest depends on the actual size of the effect (degree of change) taking place, or that is expected to take place based on prior experience or the literature. The larger the change, the easier it is to detect.

The power to detect also depends on the background prevalence or incidence of the indicator of interest. Finally, the size of the population of interest and the variability in the indicator go hand in hand to determine how likely it is your indicator will detect the intervention.

Following are two examples showing how the problem of small numbers in the New Zealand population inhibits our ability to statistically identify change in the health status of the population due to the effects of an intervention.

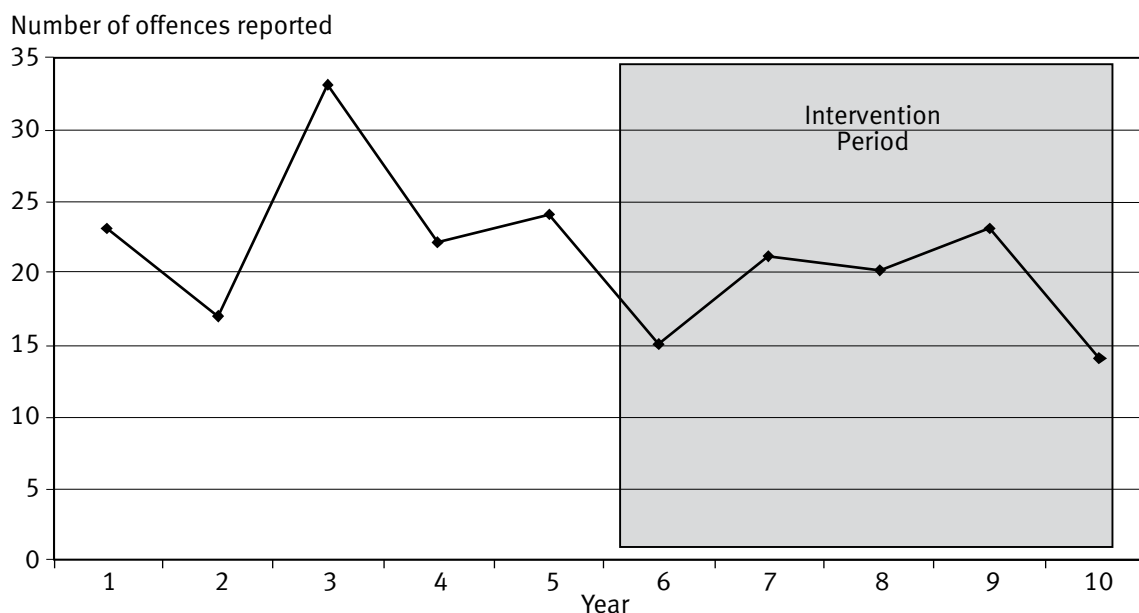
Example 2: Power to detect a change in the number of violent offences in a small community

A small community has implemented a programme to reduce interpersonal violence. This community collects data on violent offences reported to the local police from one year before the programme was initiated. The community has a population of 1000 people and there are 25 violent offences reported to the police in the year before the programme. Overseas evidence from similar programmes indicates a reduction in reported offences of 15%. This translates to an annual incidence of $25/1000$ or 0.025 offences per person per year. A reduction of 15% would be just under four fewer cases a year in this community.

Under this scenario, and if the reduction is actually 15%, the power of this indicator to statistically detect the reduction in the following year would be 2.7%. That means if we were to run the programme for another 37 years we would on average detect the difference once in that time. We would conclude that an indicator with higher incidence would be needed to monitor the change made due to the programme.

Figure 2 illustrates this scenario. It simulates a 10-year period for our small community: five years before and five years after the violence prevention intervention was initiated at the beginning of year six. The pattern highlights the usual statistical variation that one could expect to see. Note that in this example, because of the small numbers, it would be very hard to argue statistically that the programme has been successful – even though it might have been.

Figure 2: Simulated effect of 15% reduction in violent offences



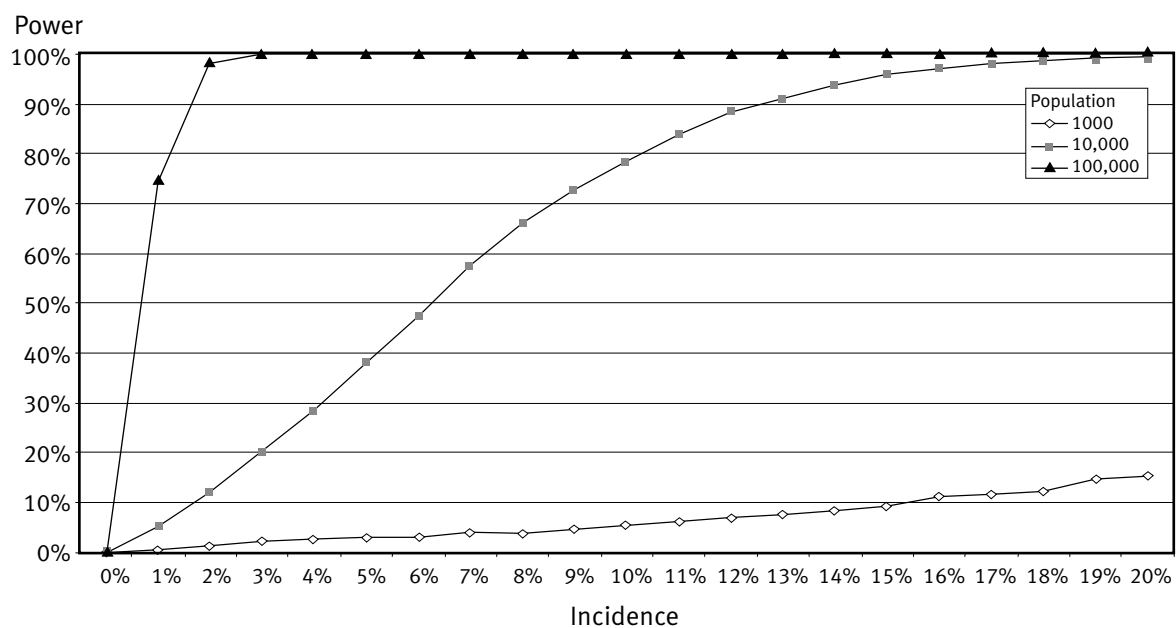
Source: Craig Wright, Public Health Intelligence

Example 3: Power to detect the effects of a health intervention

The larger the population, or the larger the effect change that we want to detect, the easier it is to detect a change. The interaction between the size of the underlying incidence of the health issue, a desire to detect a 10% change, and population size is illustrated in Figure 3. Here, the three lines indicate the statistical power to detect a 10% change in three populations comprising 1000 people, 10,000 people, and 100,000 people, for a health problem that has an incidence of 1 to 20% in the population.

It can be seen that the population of size 1000 never has a power over 20% for the whole incidence range of 0–20%, making it unlikely to detect interventions in a population of this size. Population 10,000 reaches a reasonable level of power (ie, 80%) at about 9% incidence, so it would be unlikely that one would want to use indicators with under a 10% incidence in communities of size 10,000 when expecting a reduction of 10%. Finally, indicators in the large 100,000 size population are likely to detect a 10% decrease in the indicator above 2% incidence (with 80% confidence).

Figure 3: Power to detect change due to an intervention, for specified incidence, population size and 10% incidence reduction



Source: Craig Wright, Public Health Intelligence

Part D: Available Data and Information from Public Health Intelligence

Key points

1. Public Health Intelligence has data you can use:
 - a. New Zealand Health Monitor Survey programme population survey data
 - b. Administrative data, which includes hospital records, disease registrations, and mortality data.
2. Go to PHIOne (www.phionline.moh.govt.nz) to access data.

Public Health Intelligence (PHI) has two main types of data you can utilise as indicators for outcomes monitoring: survey data collected from population surveys, and administrative data, which includes hospital records, disease registrations, and mortality data.

Survey data

PHI is responsible for the New Zealand Health Monitor (NZHM) survey programme, which is a population-based, integrated, ongoing survey programme. Data is collected for the following three major health information domains:

- health outcomes (health status, disease states)
- causes of these outcomes (social and environmental determinants, risk and protective factors)
- health services (access, utilisation, need, coverage, quality, responsiveness, cost).

There are two main types of vehicles for collecting information in the NZHM programme: serial cross-sectional surveys and serial cohort studies. Table 5 presents a summary of the NZHM surveys.

Table 5: Summary of New Zealand Health Monitor surveys, 2002–2012

Survey	Topic / data areas	Frame (target population)	Sample	Mode	Frequency
New Zealand Health Survey	Chronic diseases, biological and behavioural risk factors, reported health status, health service utilisation, sociodemographics.	All New Zealanders	12,500 adults (15+ years) and 5000 children (birth to 14 years)	Face-to-face, computer-assisted (CAPI) questionnaire plus anthropometric measurements in respondent's home.	Previously 1992/93, 1996/97, and 2002/03. Now every three years (2006/07, 2009/10 etc).
Te Rau Hinengaro New Zealand Mental Health Survey	Prevalence, severity, impairment and treatment of major mental health disorders.	New Zealand adults (16+ years)	Approx. 13,000	Face-to-face CAPI questionnaire in respondent's home.	Approximately every 10 years. Previously 2004, next planned for 2014.
New Zealand Nutrition Surveys	Food and nutrient intake, factors influencing dietary intake, nutritional status and nutrition-related status.	New Zealand adults (15+ years) or New Zealand children (5–14 years)	Approx. 4000–5000	24-hour dietary recall and food frequency questionnaire, self-administered questionnaire, plus examination, in respondent's home or at school for children.	Every five years, alternating between adult and child (adult 1997 next 2007/08; child 2002, next 2012).
New Zealand Oral Health Survey	Oral health status, oral health beliefs, attitudes, knowledge and practices.	All New Zealanders	Approx. 6000 to 8000	Face-to-face CAPI questionnaire and oral examination.	Every 10 years from 2008.
New Zealand Tobacco Use Survey	Tobacco use and the psychosocial correlates of smoking behaviours. Prevalence and consumption data available from the NZ Health Survey in 3rd year.	New Zealand adults (15 to 64 years)	Approx. 4000 to 6000	Face-to-face CAPI questionnaire in respondent's home.	Two out of every three years (2005, 2006, 2008, 2009, etc).
New Zealand Alcohol and Drug Use Survey	Alcohol and illicit drug use, and the behaviours associated with alcohol and drug use.	New Zealand adults (16 to 64 years)	Approx. 8000	Face-to-face CAPI questionnaire with audio-assisted self-complete section (A-CASI) in respondent's home.	Every two years from 2007.

Source: Ministry of Health 2005.

In addition to the cross-sectional surveys listed above, PHI is also involved in serial cohort studies (often referred to as record linkage studies). These include the New Zealand Census – Mortality Study and Cancer Trends.

Descriptive reports on the results of the above surveys can be found on PHI's website. Access to the unit record survey data sets is available on application; see: www.moh.govt.nz/phi/surveys.

Administrative data

PHI often utilises administrative data collected by the New Zealand Health Information Service (NZHIS). These data sources include hospital discharges, cancer registrations and mortality data. Table 6 describes these data sources in more detail.

Table 6: Summary of administrative data sources

Short name of data source	Description/keywords	Source of data	Period covered	Delay
Mortality data	Mortality data from the Mortality Data Collection	NZHIS	1970–2003	3 years
Foetal mortality	Foetal and infant mortality data from the Mortality Data Collection	NZHIS	1988–2003	3 years
Cancer registrations	All cancers	NZHIS	1950–2004	3 years
	Priority sites (lung, female breast, cervix, prostate and colorectal)	NZHIS	1950–2004	2 years
Notifiable diseases	Notifiable diseases from Environmental Science and Research's (ESR's) schedule	ESR	1997–2006	18 months
Sexually transmitted infections	Sexually transmitted infections from sexual health clinics	ESR	1997–2003	18 months
Public hospital discharges	Public hospital discharge from the National Minimum Data Set (NMDS)	NZHIS via DHBSPF	1971–2006	6 months
Filtered public hospital discharges	Filtered public hospital discharge from the NMDS	NZHIS via DHBSPF	1988–2006	6 months
Private hospital discharges	Private hospital discharge from the NMDS	NZHIS	1980–1995, 2001–2003	Irregularly
Birth registrations	Live and still birth registrations from the BDM registrations	BDM via SNZ	1980–2006	18 months
Mental health services	Mental health services from the Mental Health Information National Collection (MHINC)	NZHIS	2001–2006	6 months

Notes: DHBSPF = District Health Board Service Provision Framework; BDM = Births, Deaths and Marriages; SNZ = Statistics New Zealand.

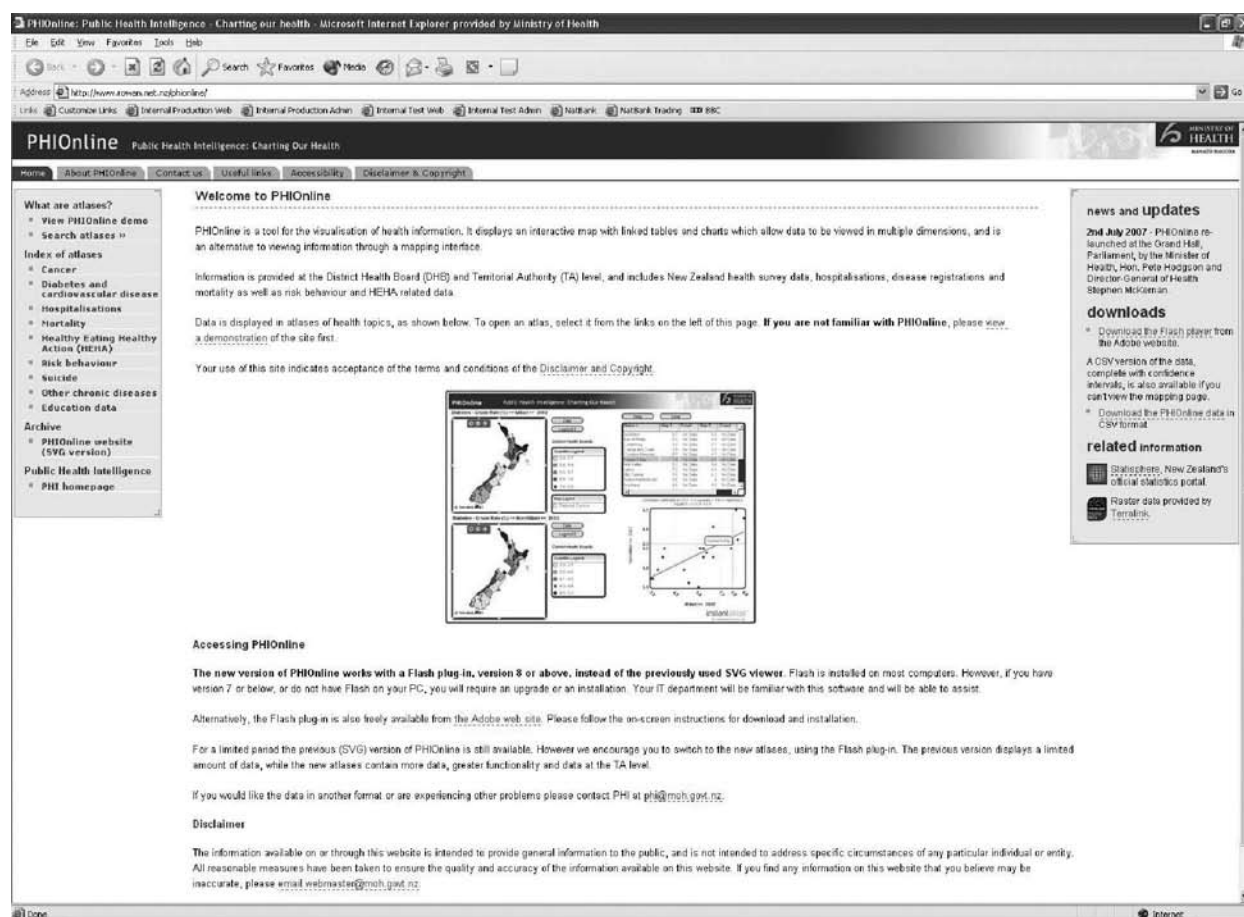
All of the above data sources are:

- in SAS version 8 file format
- at unit record level
- licensed to Public Health Intelligence, Ministry of Health.

PHIOnline – www.phionline.moh.govt.nz

National administrative data and survey data can often be disaggregated to the regional level, which may be more useful for your monitoring purposes. The main access portal for this information is PHIOnline. It is a powerful visualisation tool and provides an alternative way to access health information through a mapping interface rather than traditional text-based documents. Data is displayed for different DHB as well as at the Territorial Local Authority level, allowing you to visualise variability within your DHB.

Figure 4: The PHIOnline home page



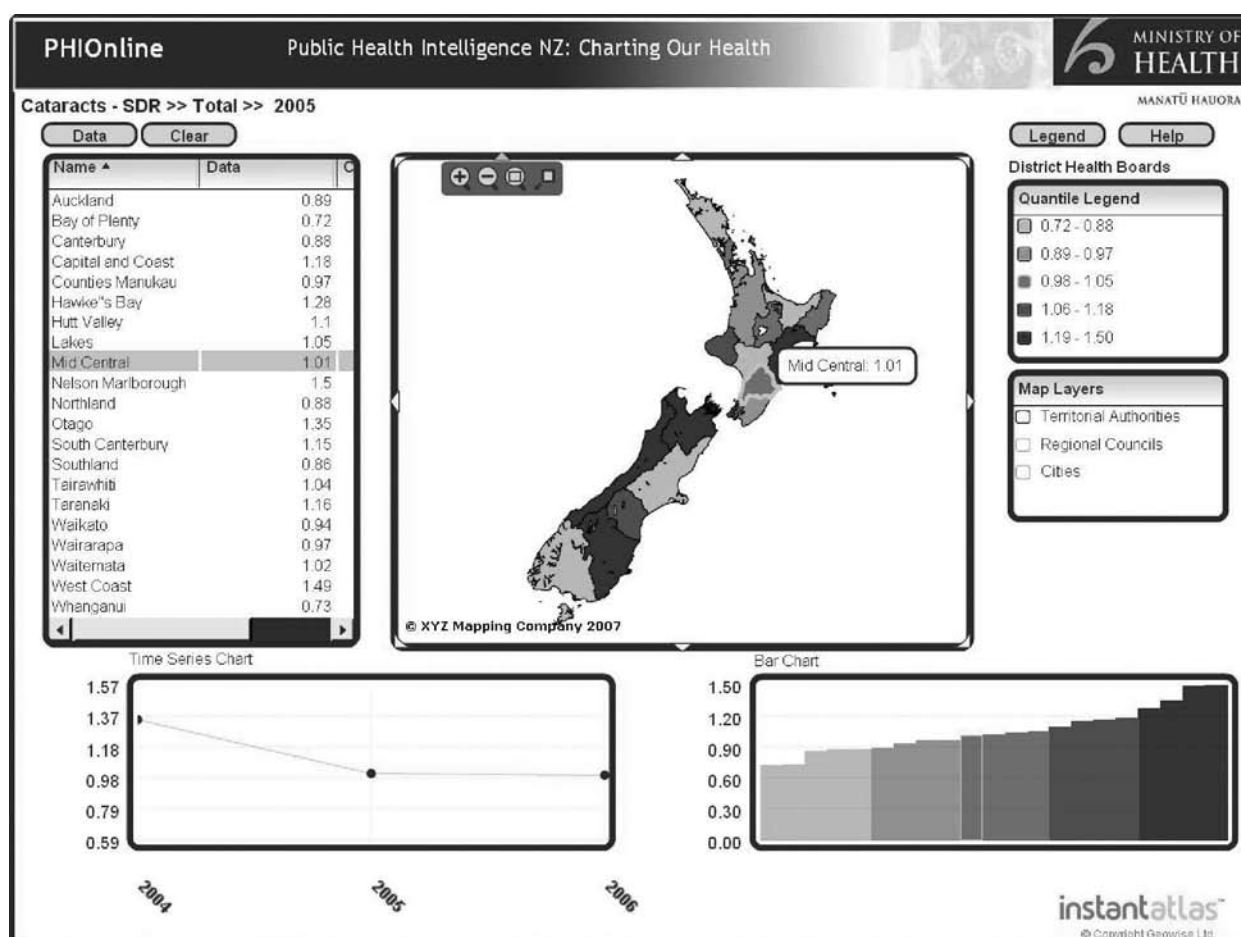
The home page provides information for users on accessing the visualisation tool. Each visualisation is thought of as an 'atlas', and the information has been grouped in accordance to the health priorities found in the New Zealand Health Strategy. To view the atlases, just click one of the links listed on the left-hand side of the home page shown in Figure 4.

PHIOnline incorporates a number of formats to view health information on the one web page. The interface consists of an interactive map (or maps) with linked tables, charts and graphs. The standard view of the website shows the one map as the centrepiece of the interface (Figure 5).

The map is thematically drawn in relation to the range and classification of the particular data set.

PHIOnline offers a number of classifications, which the user can experiment with. To complement the map, a table of the data is also displayed, as well as a time series of the data set. You can rank the data in the table by value, highest-to-lowest or lowest-to-highest, as well as by alphabetical order. Data is displayed by region as well as providing a national figure. You can download all data displayed on the site. The site also contains extensive metadata for each data set.

Figure 5: PHIOnline single map interface

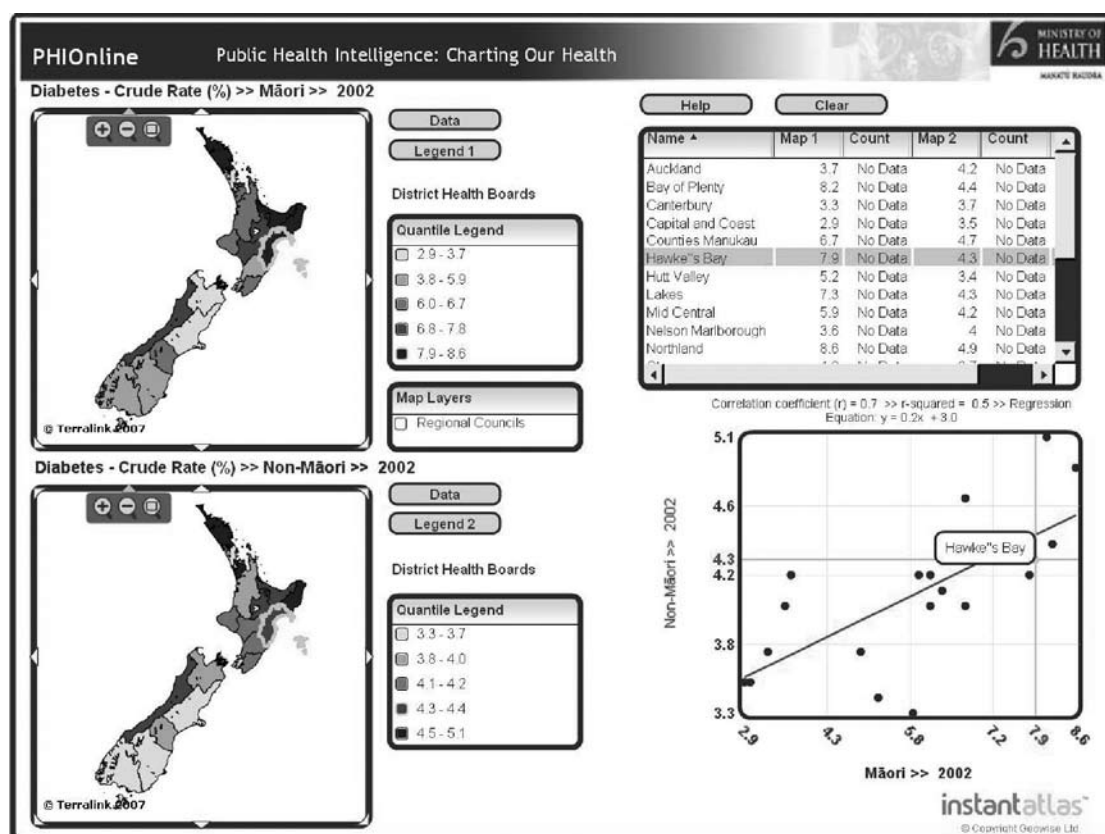


Another interface allows you to compare two different health data sets, displaying two maps on the one web page (see Figure 6). A scatter plot for both data sets is also displayed.

In addition to survey and administrative data, PHIOnline contains data on:

- oral health
- Get Checked diabetes
- water quality
- elected services
- problem gambling.

Figure 6: PHOnline double map interface



Glossary

Term	Definition	Explanatory comment and example
Attribution	<p>Attribution refers to the extent to which change in the outcome of interest is associated with the type of activity undertaken. Strong attribution requires being able to establish a clear and unambiguous causal link between cause and effect. This is often difficult for many public health programmes.</p> <p>Testing for attribution involves identifying a causal theory and then confirming the causal mechanism, using at least three of the following ways. Check:</p> <ul style="list-style-type: none"> • with other observers – did they notice the same things for the same hypothesised reasons? • that programme results relate to the programme content • that client impact (results) occurred close in time to delivery of the activity • that there are no other sufficiently plausible explanations • for a pattern of telltale outcomes suggesting one cause or another • that increased dose (more activity) leads to increased effect (or greater results) • by using a comparison group. <p>(For more, see Davidson 2005: 70–84.)</p>	<p>Attribution is an important issue to consider because it has implications for how well a programme's activities can be said to have resulted in any changes observed, and for assigning the level of accountability for the success or failure of a programme or its components.</p> <p>When assigning attribution, beware of any 'attributional bias' that would result in over-attributing a change to a single programme activity.</p> <p>For many public health programmes, establishing clear causal links between an intervention and change in a health outcome is difficult, and there are a number of practical difficulties that must be overcome. Where establishing attribution is problematic, it is important that a clear logic (rationale) for the intervention be stated, and that where a number of similar programmes are being implemented all the programmes use the same reporting framework, and attribution takes place at the group level rather than at the individual programme provider level.</p> <p><i>Example</i></p> <ol style="list-style-type: none"> 1. If a service provider is contracted to provide an information service, then information on outputs (eg, number of materials distributed, number of enquiries responded to in a time frame, length of time to respond to requests) is appropriate. This sort of information will help to provide cost efficiency information that will give information about the level of performance provided by the service. However, the justification for this type of intervention relies on the strength of the logic for the supply of the service.

Term	Definition	Explanatory comment and example
		<p>2. If a service provider is contracted to undertake workforce development, a training evaluation may include follow-up surveys to find out the extent to which training content has been used to respond to and/or resolve a client's needs.</p> <p>3. Where change in the health outcome of interest is hard to identify (eg, suicide prevention in a small community), then the effectiveness of the intervention might be assessed by grouping standardised data from all the programme providers into one collective result for the type of intervention as a whole.</p>
Accountability	Accountability refers to holding a person or an organisation responsible for obtaining the objective set through implementation of the intervention.	
Balanced scorecard	The balanced score card model explicitly links organisational strategy, organisational capacity, service delivery, service development and client outcomes. It is characterised by feedback loops using lead (measuring effort) and lag (measuring results) indicators.	
Causal or explanatory theory	Good programme implementation relies on a few critical principles, foremost of which is well-founded and clearly stated causal (or explanatory) theory. This involves defining a core problem, identifying factors that are proven to (or potentially) contribute to the problem, and identifying the downstream effects this problem creates. Clear and relevant programme objectives (see Objectives) seek to change key contributing factors. Suitable indicators (see Indicators) can then be developed that check whether any changes a programme makes to contributing factors actually produce a reduction in problem effects.	<p>In the family violence area one problem might be defined as children's exposure to parental family violence. Contributing factors can range across:</p> <ul style="list-style-type: none"> • limited care-giver awareness about the negative short- and long-term effects of family violence on children • general volatility in one or both parents • insufficient relationship and problem-solving skills with one or both caregivers

Term	Definition	Explanatory comment and example
		<ul style="list-style-type: none"> • substance abuse and/or mental illness • environmental stressors such as unemployment/low income, overcrowded housing, social isolation, a larger family, children's behaviour problems, lack of social support services <p>Problem effects may include:</p> <ul style="list-style-type: none"> • children's safety (physical injury, mental health) • children's behaviour (fretting, acting out, aggression) • children's development (engagement with school, positive socialisation). <p>Using a 'problem tree' approach allows appropriate solutions to be identified and a theory of change or action to be formulated (see Theory of change/action).</p> <p>For more, see: <i>Problem Tree Analysis</i> (ODI) http://tinyurl.com/2eg9gj <i>Problem, Alternatives, Objectives Tree</i> (Participation) http://tinyurl.com/2f551s</p>
Cost-effectiveness Cost-benefit Cost per unit of impact/outcome/output Cost-efficiency	<p>Cost effectiveness is the balance between the cost of developing a measure or introducing a new monitoring regime, and the benefit that will be gained from implementing it. There are many ways of measuring cost effectiveness. The three approaches suggested by (The Treasury and State Services Commission 2007a; b) as appropriate for Crown agencies and entities are:</p> <ul style="list-style-type: none"> • <i>cost-benefit analysis</i>, which is seen as the 'gold standard' for establishing cost effectiveness • <i>cost per unit of impact/outcome</i>, which is a simpler approach to use (assuming it is feasible to quantify the impact / outcome gained) 	<p>For many programmes, a judgement will have to be made as to whether the cost of developing a monitoring regime or measure is worth the benefit to be gained from the use of the information provided. Where the cost outweighs the benefit, consider using an appropriate proxy indicator, or whether information from an already existing monitoring regime could be used.</p> <p>In general, the proportion of the total budget spent on monitoring should be in the region of 5% to 20%, depending on the type and size of the programme. Where the programme is a pilot initiative, the results of which could be used to significantly influence whether the programme is expanded into a major effort, then 20% of the total budget could be appropriate.</p>

Term	Definition	Explanatory comment and example
	<ul style="list-style-type: none"> <i>cost-efficiency</i>, which is a method by which the <i>cost per output</i> is measured. This approach is likely to be particularly appropriate where the service provided (eg, a public health information service) is remote from any health outcome. 	<p>However, if the programme has a large budget and the programme is well established, applied in standard ways and supported by evidence for its effectiveness, then a budget allocation of 5% may be appropriate for monitoring purposes.</p> <p>Cost per output is probably the best method to use when calculating the cost of a monitoring programme or a component of it. For example, the cost of purchasing information from an existing data set may be cheaper than doing it yourself. Or, it may be cheaper to use a proxy measure than to use a direct measure – assuming the proxy measure is robust enough to withstand public scrutiny and the information trade-off is acceptable. Note that cost effectiveness is just one of a number of criteria that should be used in planning an outcomes monitoring regime.</p>
Data Data sets	<p>Data is factual information that after analysis is used to reason or make decisions. The word is commonly used to refer to either single or multiple pieces of information, and in technical reports is often used in the plural form (as in ‘data are collected’).</p> <p>Data sets refers to collections of pieces of information that are collected for particular purposes. ‘Data sets’ is synonymous with databases.</p>	<p>For example, the Pharmhouse data set collects pieces of information about the prescribing of drugs subsidised in New Zealand by Pharmac.</p> <p>The New Zealand Health Information Service is responsible for the collection and maintenance of many of the large national health-related data sets, such as mortality and morbidity data. Other government agencies are responsible for other data sets that can be useful for monitoring public health programmes. A large number of these data sets are hosted on the <i>PHIONline</i> website (www.phionline.moh.govt.nz).</p>

Term	Definition	Explanatory comment and example
Evaluation Formative Process Impact	<p>Evaluation is a research process that aims to assess systematically and objectively the relevance, performance and success of ongoing or completed programmes. Evaluation aims to provide information about whether underlying assumptions or theories used in the programme development were valid, what worked and what did not work, and why. Typically evaluation aims to determine the relevance, efficiency, effectiveness, impact and sustainability of a programme. Evaluation should allow for lessons learnt to be transferred into other programmes.</p> <p>A comprehensive evaluation programme usually involves three parts:</p> <ul style="list-style-type: none"> • formative evaluation • process evaluation • impact/outcome evaluation. <p><i>Formative evaluation</i> focuses on the establishment of the programme. It is particularly useful when the programme is new, untested and requires community acceptance if it is to be successful. It seeks to understand what the conditions for programme success are supposed to be.</p> <p><i>Process evaluation</i> focuses on how well the programme functions. Are the conditions for success present in the way the programme is being delivered?</p> <p><i>Impact/outcome evaluation</i> assesses how successful the programme has been in delivering the desired outcomes. Impact evaluation should always try to identify unintended benefits and any negative programme consequences.</p>	

Term	Definition	Explanatory comment and example
	<p>Other terms that may be associated with evaluation include:</p> <ul style="list-style-type: none"> • output • outcome • impact. 	
Evidenced-based public health (EBPH)	<p>Definitions of what ‘evidenced based public health’ means are evolving (Kohatsu et al 2004). In the context of the <i>Guide to Developing Public Health Programmes: A generic programme logic model</i> (Ministry of Health 2006) and this <i>How to</i>, the definition by Brownson (2003) is preferred:</p> <p>EBPH is the development, implementation, and evaluation of effective programs and policies in public health through application of principles of scientific reasoning, including systematic uses of data and information systems and appropriate use of program planning models.</p>	<p>See also Kohatsu (2004) and Jenicek (1997) for alternative formulations defining EBPH. Kohatsu’s (2004) formulation is applicable to the policy development process, while Jenicek’s (1997) definition is an application of evidence-based medicine concepts to public health practice.</p> <p>(See also Public health.)</p>
Impacts	<p>Impacts are the contribution made to the achievement of an outcome by a specified set of actions (The Treasury and State Services Commission 2007b).</p>	<p>Impacts should be positive, although unintended consequences from an action may occur that result in either positive or negative contributions to the achievement of the desired outcome.</p> <p>Note: this definition and use of ‘impact’ are different to how it is defined and used in evaluation methodology (see also Evaluation).</p>

Term	Definition	Explanatory comment and example
Indicators Quantitative Qualitative Proxy	Indicators are either quantitative or qualitative measures that assess the direction and size of change in the thing being measured. <i>Quantitative indicators</i> are numerical measurements that lend themselves to statistical analysis. (This <i>How to</i> focuses on this type of indicator). <i>Qualitative indicators</i> are primarily measures taken in a textual (narrative) form (ie, non-numerical) and may or may not lend themselves to quantification. Analysis of qualitative indicators requires the use of qualitative analytical techniques. <i>Proxy indicators</i> are measures that provide an approximate estimate of change in the outcome of interest. Proxy measures may be appropriate to use when direct measures cannot be used because of concerns about ethics, cost, complexity of measurement or timeliness. For example, hospitalisations for intentional self-harm are used as an acceptable proxy measure of suicide attempt, and BMI is usually accepted as a proxy for obesity. (See also Instrument .)	Time is an important factor in assessing what change has occurred. Indicators may measure change over the short, medium or long term. What constitutes short, medium or long term depends on the context and the issue(s) the programme is tackling. Health indicators directly or indirectly (ie, proxy) measure a health-related characteristic of an individual, population or the environment. The indicator may measure one or more health aspects (quality, quantity and time) (Nutbeam 1998).
Inputs	Resources put into a programme to carry out an activity. Another term for input is 'effort'.	Inputs can be human, material, financial or expressed as time. Effort describes everything that 'we' use and do for 'them' – the programme recipients.
Instrument	An instrument is a numerical or text-based measuring method that shows the extent or amount or quantity or degree of something of interest.	The instrument is the process used to produce the desired public health indicator. (See Indicators .)
Measure(s)	See Indicators and Instrument .	
Metric(s)	See Indicators and Instrument .	
Monitor	See Outcomes monitoring .	

Term	Definition	Explanatory comment and example
Objectives High-order/general Specific/intermediate Low-order/operational SMART	<p>Objectives are statements about the results a programme seeks to achieve. Any programme must have at least one objective. Objectives may be translated directly into 'outcomes' where they deal with only one issue. However, double-barrelled objectives will require multiple outcome measures to be developed.</p> <p>Objectives may form a hierarchy that moves from a limited set of high-order objectives that are synonymous with aims or goals, to lower-order or more specific or intermediate objectives, as follows.</p> <p><i>High-order or general objectives</i> are statements about the over-all or long-term effects or ultimate outcomes that are expected to be obtained as a result of the programme, or to which the programme contributes to as part of a whole-of-government initiative.</p> <p>Attainment of high-order objectives may or may not be directly attributable to any one programme. Long-term objectives are generally those to be achieved in 5 to 10 years.</p> <p><i>Specific or intermediate-level objectives</i> are statements about the intermediate outcomes that are expected to be achieved by the programme, and that are more directly attributable to the programme's activities. These objectives are generally those to be achieved in 3 to 5 years.</p> <p><i>Low-order or operational objectives</i> are statements about the immediate outcomes (which may be expressed as outputs in some circumstances) expected to be achieved or accomplished from the intervention that are directly attributable to the activities undertaken. These objectives are generally those to be achieved in 1 to 2 years.</p>	<p>In terms of the Public Finance Act 2004, objectives are not expressly defined, and consequently can be seen as meaning a goal or aim. However, because some core departmental activities that service the operation of government do not fit within the definition of an outcome given in the Act, objectives are also seen as being similar to impacts and outcomes. (See Outcomes.) (The Treasury and State Services Commission 2007a.)</p> <p>In health terms, health goals (objectives/aims) are general statements of intent and direction set for a health programme, which may include a health target to achieve (Nutbeam 1998). In this document, the term 'objective' is used to identify particular results the programme seeks to achieve. For monitoring purposes, the result may be expressed in terms of an outcome or an output. (See Outcomes and Outputs.)</p> <p>Following is an example of a hierarchy of objectives.</p> <p><i>High-order objective</i> (synonymous with vision, aim or goal):</p> <ul style="list-style-type: none"> • Enable people with chronic conditions to improve their health, slow progress of their condition(s), and maintain independence wherever possible by aligning community and hospital services across [] DHB. <p><i>Intermediate-level objective associated with achieving the high-level objective:</i></p> <ul style="list-style-type: none"> • Reduce the incidence of cancer, diabetes, and cardiovascular disease by 20% respectively over the next 5 years.

Term	Definition	Explanatory comment and example
	<p>Objectives should be written as <i>SMART</i> objectives:</p> <ul style="list-style-type: none"> • Specific • Measurable • Achievable (sometimes referred to as 'accurate' or 'action-oriented') • Relevant (sometimes referred to as 'realistic') • Time-based. 	<p>Low-order or operational objectives associated with achieving the intermediate level:</p> <ul style="list-style-type: none"> • Increase the level of knowledge by 50% in the target population groups about the causes and prevention of breast and prostate cancer, diabetes and cardiovascular disease. • Increase the uptake of cancer screening services by 50% in the high-risk populations in the next 3 years. • Increase access by 50% to early diagnosis services in rural communities in the next 2 years.
Outcomes Intermediate Cross-agency Process	<p>Outcomes are specific statements about the intended change in public health-related attitudes, knowledge, behaviours, or physical (including mental) health status in the target population(s) sought by undertaking the planned public health activity.</p> <p><i>Intermediate outcomes</i> (also called interim outcomes) are steps along the way to the desired end outcome. They are often smaller changes that need to happen before the final desired outcome can be reached.</p> <p><i>Cross-agency outcomes</i> are outcomes where the collective effort of more than one government agency or programme is required if the outcome is to be achieved.</p> <p><i>Process outcomes</i> typically measure the amount of effort put into a programme and the quality of the service provided. They can also be appropriate where it is important to monitor community support for a programme. Process outcomes that measure effort can be expressed as outputs.</p> <p>(See Process outcome monitoring.)</p>	<p>Health outcomes should be expressed in such a way that indicates the type, direction and extent of change in health knowledge, behaviour or status sought in an individual or population group.</p> <p><i>Example</i></p> <ul style="list-style-type: none"> • Rates of smoking cessation in young women under 20 years of age increase by 20% at the end of the programme. <p>Attainment of the outcomes is measured through indicators (see Indicators). Ideally, the health outcome sought should be directly attributable to the intervention, but this is usually problematic in health promotion activities (Nutbeam 1998).</p> <p>This definition can be seen as a specific application to public health of the definition of an 'outcome' provided in the Public Finance Act 2004 (s2(1)): 'A state or condition of society, the economy or the environment and includes a change in that state or condition.'</p>

Term	Definition	Explanatory comment and example
		<p><i>Cross-agency outcomes</i> should probably be the rule rather than the exception in comprehensive public health programmes. In such situations, it is important that the roles and responsibilities of the participating agencies are agreed – including who is responsible for monitoring outcomes – before the programme begins.</p>
Outcomes monitoring	A plan for the routine, systematic collection and recording of information about aspects of a programme over time. The purpose is to assess whether progress is being made on achieving the programme objectives.	<p>Preparing the outcomes monitoring plan requires assessing:</p> <ul style="list-style-type: none"> • what needs to be measured to demonstrate success • how things are going to be measured and by whom (deciding what indicators are going to be used, after considering issues such as validity, reliability, sensitivity, attribution and feasibility) (see Indicators.) • the timeframes for delivering information about the performance of the programme against the stated objectives to decision-makers and key stakeholders in a timely manner. <p>Outcomes monitoring may or may not include a formal evaluation process depending on the objectives of the programme. (See Evaluation.)</p>
Outputs	Outputs are things (eg, goods) produced, services delivered, events held, or participation generated resulting from the activities undertaken. (See Outcomes and Process outcome monitoring).	<p>This definition is consistent with that provided in the Public Finance Act 2004. Common public health outputs are things such as information services provided, meetings held or attended, and educational/training or social marketing products or services delivered. Such outputs are intended to have an impact that assists or results in achieving the desired public health outcome, such as a change in health knowledge and behaviour.</p> <p>(See Outcomes and Process outcome monitoring.)</p>

Term	Definition	Explanatory comment and example
<p>Process outcome monitoring</p>	<p>This is like outcome monitoring but includes keeping track of how well a programme is being implemented in its different aspects. Process outcome monitoring provides the basis for programme staff to assure themselves that good outcomes are associated with good programme delivery. If the desired outcomes are not achieved, then the agency will have sufficient information to assess which aspects of programme delivery need to be developed or changed.</p> <p>(See also Balanced scorecard.)</p>	<p>A programme theory of change will indicate what kinds of inputs and processes are likely to produce a prevention effect. For example, a home-visiting programme might detail the following as vital to good client outcomes:</p> <ul style="list-style-type: none"> • appropriately experienced, skilled and supervised staff • staff who are empathetic and warm in their interactions • a family/whānau assessment that identifies strengths and challenges likely to help or hinder child-rearing (eg, includes a screen for family violence, substance abuse, depression) • the assessment process encourages co-operation and reduces defensiveness • families are visited at sufficiently regular intervals • ideally the ethnicity of workers and primary caregivers are matched to maximise engagement • there is a developmental model to compare children's progress against • the developmental model is easy to communicate to caregivers and they 'buy' it • developmental outcomes for children and conditions that hinder normal development are systematically tracked • guidance is provided in a responsive way (ie, is timely, solution focused, realistic and checked for understanding) • goal-setting is used to enhance caregiver motivation and to prioritise actions. <p>Regular supervision meetings and/or programme reviews provide space for clients' progress to be checked against the quality of programme delivery. Any changes in client plans, programme design or professional development can be considered a process outcome.</p>

Term	Definition	Explanatory comment and example
<p>Public health</p> <p>Population-based public health</p> <p>Public health action</p>	<p>There is a range of terms and definitions relating to public health, including public health action and population-based public health. The World Health Organization (1998) defines public health as:</p> <p>The science and art of promoting health, preventing disease, and prolonging life through the organized efforts of society.</p>	<p>It can be argued that this definition omits the important point that public health primarily focuses on the health of the population or population sub-groups and not the health of specific individuals. This focus on populations in public health is identified more explicitly in more recent definitions.</p> <ul style="list-style-type: none"> • ‘Public health is concerned with improving the health of the population, rather than treating the diseases of individual patients’ (Department of Health 2004 cited in Public Health electronic Library). • ‘[Public health is] concerned with the health of the community as a whole’ (<i>American Journal of Public Health</i> 2006, website (http://www.ajph.org/)). • Public health action is ‘Collective action for sustained population-wide health improvement ... and to reduce health inequalities. Responsibility for such action is not confined to the health sector but should include all sectors whose actions affect the health of populations. Populations can be geographic and/or defined by factors such as ethnicity, gender, age, sexual orientation, income etc.’ (Beaglehole et al 2004; Public Health Advisory Committee 2006). • ‘Public health is the organised response by society to protect and promote health, and to prevent illness, injury and disability. The starting point for identifying public health issues, problems and priorities, and for designing and implementing interventions, is the population as a whole, or population sub-groups.’ (National Public Health Partnership 2006). <p>(See also Evidence-based public health.)</p>

Term	Definition	Explanatory comment and example
Reliability	Reliability is the extent to which a measure, when used repeatedly in the same way, will produce the same or a similar result.	For example, if we were interested in monitoring a person's weight and used a cheap weighing machine that produced a different reading each time it was used (when all other factors were the same), then the machine would not be a reliable instrument to use.
Robustness to withstand scrutiny	It is important that assessments about the effectiveness of public health programmes funded by government money be able to withstand scrutiny by all the stakeholders and other interested parties.	We recommend that for most public health programmes it is advisable that a programme's success not be attributed to a single criterion: robustness is provided by the totality of the outcome monitoring regime.
Sensitivity	Sensitivity refers to how well a measure is able to accurately detect when a change has occurred in the outcome being monitored.	For example, a weighing machine that was able to differentiate a change in weight by 100 gram graduations is more sensitive than a machine that could only detect changes in 500 gram graduations.
Target	A target is the amount of change desired, or expected, in a specific time period through the implementation of the health programme.	Targets should be based on specific and measurable changes in health outcomes. Setting targets helps define the process towards achieving the health objective set, although targets require the existence of an appropriate indicator (Nutbeam 1998). (See Indicators .)
Theory of change/ action	A theory of change/action is a statement about how practitioners believe individual, inter-group and social/ systemic change happens, and how the proposed actions will produce positive results. A good theory of change explicitly describes the assumptions and causal links thought to be important for achieving the health outcome sought.	Where initial evidence for the effectiveness of an intervention is lacking, a robust theory of change should be articulated to justify the choice of intervention. Used in this way, a theory of change provides the 'logic model' framework to plan and justify the intervention selected. A theory of change approach may be particularly appropriate when interventions are aimed at minority population groups.

Term	Definition	Explanatory comment and example
Timely data/ information	Information is timely when it contains data directly relevant to the programme and – importantly – available in a time frame where it can usefully inform decisions about whether the programme is performing as planned and what things need to be changed to improve the performance of the programme, if required.	Some national data sets are scheduled to be collected on 5-yearly data. Such data is unlikely to be timely for performance monitoring purposes if the programme will only run for 3 years, say. Similarly, national mortality and morbidity data is typically only available 18 to 24 months after the event. Such information may not be timely for some programmes, although local data may be available within a more useful time frame.
Timing	Timing refers to the time period between when an activity took place and when a change in the desired outcome was observed.	Timing is an important aspect in establishing ‘attribution’, and being able to monitor whether a programme is working.
Validity	<p>Validity refers to how well a proposed instrument measures the outcome of interest. Some outcomes may be measured directly, while others will require a proxy method that results in a proxy measure. Generally, the more direct a measure is, the more valid it is likely to be.</p> <p>There is often more than one way of measuring an outcome of interest, and there can be considerable debate about how ‘valid’ a measure actually is.</p>	<p>For example, the best way to measure body composition (particularly lean and fat mass) is under-water densitometry (weighing), but this is not practical outside small clinical studies. Therefore, measures of weight adjusted for height (eg, body mass index, BMI) are often used to reflect body composition, particularly body fat mass. Although BMI is correlated with body fat mass, this relationship varies according to body build, age and ethnicity. Therefore, BMI is not recommended for assessing excess body fat mass (obesity) at an individual level.</p> <p>However, BMI can be useful for assessing obesity at a population level if used and interpreted appropriately. BMI based on direct measurement of height and weight is more valid than BMI based on self-reported height and weight, particularly if measurements are taken by trained observers using appropriate equipment and standardised procedures.</p> <p>Ultimately, the choice of measure for a particular programme will depend on a number of factors, including validity and practicality.</p>

References

- Bartholomew L, Parcel G, Kok G, et al. 2006. *Planning Health Promotion Programmes: An intervention mapping approach*. San Francisco. Jossey-Bass.
- Beaglehole R, Bonita R, Horton R, et al. 2004. Public health in the new era: improving health through collective action. *Lancet* 363: 2084–6.
- Brownson RC. 2003. *Evidence-based Public Health*. Oxford: Oxford University Press.
- Davidson J. 2005. *Evaluation Methodology Basics: Nuts and bolts of sound evaluation*. Thousand Oaks. Sage.
- Department of Health. 2004. *Public Health Electronic Library Standards for Better Health*. London: Department of Health. London.
- Iverson A. 2003. *Preparing Program Objectives: Theory and practice*. Toronto: The International Development Research Centre Evaluation Unit.
- Jenicek M. 1997. Epidemiology, evidenced-based medicine, and evidence-based public health. *Journal of Epidemiology*. Dec; 7(4): 187–97.
- Kohatsu ND, Robinson JG, Torner JC. 2004. Evidence-based public health: an evolving concept. *American Journal of Preventive Medicine* 27: 417–21.
- Ministry of Health. 2005. *The New Zealand Health Monitor: Updated strategic plan*. Wellington: Ministry of Health.
- Ministry of Health. 2006. *A Guide to Developing Public Health Programmes: A generic programme logic model*. Wellington: Ministry of Health.
- Ministry of Health. 2007. *Leading for Outcomes*. <http://www.leadingforoutcomes.org.nz> Wellington: Ministry of Health.
- National Public Health Partnership. 2006. *Public Health Classifications Project Phase One: Final report*. Melbourne: National Public Health Partnership.
- Nutbeam D. 1998. Health promotion glossary. *Health Promotion International* 13: 349–64.
- Public Health Advisory Committee. 2006. *Health is Everyone's Business: Working together for health and wellbeing: A report to the Minister of Health on the implications of a changing context for public health in New Zealand*. Wellington: Public Health Advisory Committee.
- The Treasury, State Services Commission. 2007a. *Guidance and Requirements for Crown Entities: Preparing the Statement of Intent*. Wellington: The Treasury, State Services Commission.
- The Treasury, State Services Commission. 2007b. *Guidance and Requirements for Departments: Preparing the Statement of Intent*. Wellington: The Treasury, State Services Commission.

The Treasury, State Services Commission, Department of the Prime Minister and Cabinet. 2005. *Guidance for Crown Entities: Planning and managing for results*. Wellington: The Treasury, State Services Commission, Department of the Prime Minister and Cabinet.

United Nations Development Programme Evaluation Office. 2002. *Handbook on Monitoring and Evaluating for Results*. New York: United Nations Development Programme Evaluation Office.

World Health Organization. 1998. *Health Promotion Glossary*. Geneva: World Health Organization.

Further reading

Ernst K. 2004. *The Canadian Outcomes Research Institute's Program Logic Model*. Calgary: The Canadian Outcomes Research Institute.

European Research Group Health Outcomes. 1996. Choosing a health outcomes measurement instrument: general advice proposed by the European Research Group on Health Outcomes Measures (ERGHO). *Quality of Life Newsletter* 15: 7–8. URL: [http://www.mapi-research.fr/pdf/newsletter/qol15 0.PDF](http://www.mapi-research.fr/pdf/newsletter/qol15%20.pdf). Accessed 3 November 2005.

