

## Navigating from the heights of technical rigour to the swampy reality: Lessons from New Zealand in population health outcomes monitoring

John Wren

*Ministry of Health, Health and Disability Intelligence Group, Wellington, New Zealand*

*Correspondence to: John Wren, Research Team, ACC, Research & Development Group, PO Box 242, Wellington 6140, New Zealand.*

*Email: john.wren@acc.co.nz*

### Abstract

**Background:** The objective of this paper is to describe and discuss two documents produced by the New Zealand Ministry of Health concerning the monitoring of outcomes of public health programs.

The New Zealand Government is increasingly expecting planners and managers of publicly funded services to shift their focus from the delivery of 'outputs' towards achievement of 'outcomes'. Intervention logic models and outcomes monitoring are promoted by central government agencies as suitable management methods for implementing the change. [1-3]

**Methods:** To help managers design and implement comprehensive, effective and measurable population health programmes the Ministry of Health recently published two guidance documents. The first document provided guidance about how to plan programmes using a generic logic model approach.[4] The second set out in detail a process on how to monitor population health programmes.[5] The intent of the documents was to help managers navigate between the heights of technical rigour and the swamps of reality in the delivery of population health programmes.[6]

**Results:** A number of issues and implications for how population health programmes are planned monitored and performance assessed have been identified by the guidance documents. Issues include the problem of small numbers, understanding the difference between outcomes monitoring and traditional forms of evaluation, and outcomes monitoring being seen as a tool for punitive performance management rather than 'continuous programme improvement'. Implications include more time spent on the design of programmes. Planners will need to focus upon better sequencing of activities, setting more specific and time limited goals, and to be more informed about how to use research to inform the selection of interventions.

**Conclusions:** The guidance documents promoted by the Ministry of Health have provided useful advice about how to develop and use logic models and outcomes monitoring in the planning of population health programmes. Practitioners particularly welcomed the inclusion of a glossary that included explanatory comments and examples. Two key issues have been identified with the application of outcomes monitoring that need to be managed. The first is the issue of statistically small numbers associated with relatively rare health events. The second is to promote outcomes monitoring as a tool for continuous programme improvement, rather than as a potentially punitive.

*Key words: outcomes monitoring, logic model, planning, management, public health programmes*

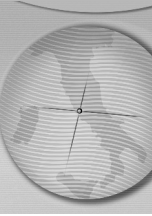
### Introduction

Intervention logic and outcomes monitoring are key components of the New Zealand public expenditure system.[3] Consequently, the New Zealand Government is increasingly expecting planners and managers of publicly funded services to shift their focus from the delivery of 'outputs' towards achievement of 'outcomes'. [1-3] Furthermore, in the health context 'achieving measurable progress on public health outcomes' is an objective under the Government's Achieving

Health for All People strategy.[7]

To help design and implement comprehensive, effective and measurable population health programmes two guidance documents have been published by the Ministry of Health, and over 15 workshops delivered to programme planners around the country in the last two years.[4, 5]

The first document described a generic logic model for the development of public health programmes that linked to key New Zealand health strategy documents and the Ottawa



Charter for Health Promotion framework.[8] The second publication provided advice on how to select appropriate, timely and cost-effective outcome measures, and discussed a number of practical issues with adopting an outcomes monitoring approach.[5] The approach and advice promoted in the second publication, and feedback from practitioners attending training workshops, are presented in the following sections.

### **A New Zealand model for monitoring population health outcomes at the regional and local level of delivery**

Under the rubric of 'Managing for Outcomes' the use of intervention logic models and outcomes monitoring plans has been promoted by the New Zealand Treasury and the State Services Commission as suitable management methods for changing from an outputs focus to an outcomes one.[1-3] In general policy terms outcomes are defined as the results desired from policy settings - the intended effects of a policy, while outputs are defined as the goods or services that provide the means to achieving the desired outcomes.[9]

Internationally, many governmental and non-governmental organisations that deliver community social services, including health care and support, have adopted the logic model concept. [10-12] A comprehensive logic model includes an outcomes monitoring plan. While good logic models appear simple, they require work to develop, and a number of practical difficulties must be overcome to measure outcomes and to demonstrate the success of a public health programme.[5, 13]

A key component of outcomes monitoring is to link outcomes to the relevant policy settings. The first task in doing this is to identify and define what the key programme objectives, outcomes and outputs are. To facilitate this process an extensive glossary was developed that set out definitions for commonly used terms, and provided some explanatory comment and examples.[5]

For population health, an outcome has been defined as 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 activity. Outcomes should be expressed in such a way that indicates the type, direction and extent of the change sought. Different types of outcomes may be set, including intermediate (interim) ones,

cross-agency outcomes, and process outcomes. Intermediate outcomes are the steps along the way to the desired end. Cross-agency outcomes are those where the collective effort of more than one agency is involved if the desired outcome is to be achieved. Process outcomes typically measure the effort put into a programme and the quality of the service provided. They can be appropriate where it is important to monitor community support for a programme. Process outcomes that measure effort can be expressed as outputs. Outputs are things such as goods produced, services delivered, events held, or participation generated resulting from the activities undertaken.

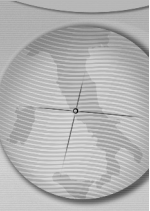
Objectives are defined as statements about the results a programme seeks to achieve. Objectives may form a hierarchy that begins with a limited set of high-order objectives that are synonymous with aims or goals to be achieved in the long term (five to seven years). Underneath high-level objectives are 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. 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.

A key to developing measurable outcomes is to use 'action' words. Action words help to make it clear what activities need to be undertaken, what changes need to take place, what the desired result is, and consequently what should be monitored. Bartholomew et al 2006, has identified a list of action words suitable for population health programmes.[14]

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?[15] SMART objectives are specific, measurable, achievable, relevant and time-based.

### ***Criterion for prioritising 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 identified. This means a process of prioritisation has to take place to select the outcomes to monitor. Two types of

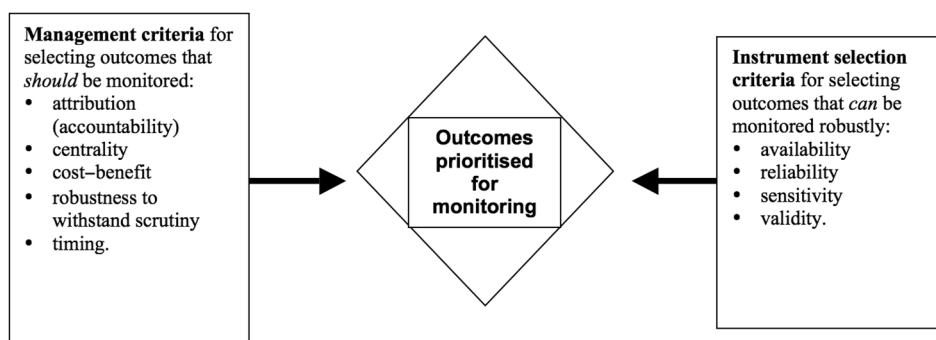


criterion have proven to be useful in practice to aid selection and prioritisation of outcomes for monitoring:

- 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 to inform decision-making.

Figure 1. A framework for prioritising what outcomes to monitor.



**Management selection criterion**

**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 to consider because it has implications for the degree to which accountability can be assigned 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, beware of any ‘attributional bias’ that results from over-attributing a change to a 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. In situations where data is unavailable to measure the outcome, give thought to either using an appropriate ‘proxy’ measure or to developing a new measure.

**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, a judgement will have to be made about whether the cost of developing a new measure is worth the benefit to be gained from the information provided. Where the cost outweighs the benefit, consider using an appropriate proxy indicator, or information from a pre-existing monitoring regime.

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,

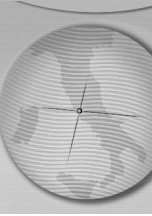
the cost of purchasing information from an existing data set may be cheaper than doing it yourself. Or, it may be cheaper to use a proxy measure than to use a direct measure (assuming the proxy measure is robust enough to withstand public scrutiny and the information trade-off is acceptable).

**Robustness to withstand scrutiny**

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. Because establishing clear attribution between cause and effect is often problematic, it is usually advisable not to attribute a programme’s success to a single criterion, rather, robustness is provided by the strength of the intervention logic and the robustness of the total outcomes monitoring plan. In situations where research evidence for the effectiveness of an intervention is lacking, it is particularly important that a robust outcomes monitoring plan is implemented.

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

- 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?

### Timing

Timing has two aspects to it. The first concerns establishing 'attribution'. 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. In New Zealand, 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 through publicly funded programmes [1, 2].

In practical terms it may not be feasible to measure progress in attaining short-term or even medium-term objectives using national data. For example, information from national survey and administrative data sets is typically 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

#### 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, from a range of government agencies, or public health organisations in your area. 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 investing 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

measuring the 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 of 100 grams is more sensitive than a machine that can only detect changes in 500 gram amounts.

#### Validity

Validity refers to how well a proposed instrument measures the outcome of interest. Some outcomes may be measured directly, while others will require a method that results in a proxy measure. Generally, the more direct a measure is, the more valid it is likely to be. Often there is more than one-way of measuring an outcome, and there can be considerable debate about how 'valid' a measure actually is.

### ***Balancing the criteria – prioritising outcomes to monitor***

Although the criteria outlined above 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. To help assist with this process a scorecard such as that provided in Figure 2 has been developed to aid decision-making. 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.

To use the scorecard at the most simple level: just assign a '1' score to each criterion that has been successfully met. Then add all the scores 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

Figure 2. Checklist and score card for prioritising outcomes and indicators for monitoring.

Checklist and scorecard for prioritising outcomes for monitoring												
Outcome of interest	Indicator name	Criteria									Total Score	Priority
		At	Av	Ce	Cb	Re	Rb	S	T	V		
Outcome 1	Indicator 1 Indicator 2...											
Output 1	If any..	Rate on a scale of 0 to 1 per criteria for each Outcome and Indicator										
Objective 1	Indicator 1 Indicator 2...											
List Objectives / Outcomes of interest for potential monitoring	List potential indicators	Criteria Abbreviation									Rank outcomes by score - check score for essential outcomes; set monitoring priority	
		Management criterion Attribution (Accountability) AAt Centrality Ce Cost – benefit Cb Robustness to withstand public scrutiny Rb Timing T Instrument criterion Availability Av Reliability Re Sensitivity S Validity V										

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 (i.e., the results are going to be weighted (biased) towards indicators that score well on those factors).

**Issues arising from the use of the approach**

Using the approach outlined, a number of national, regional and local level logic models and outcomes monitoring frameworks have been developed in New Zealand. The experience has highlighted a number of practical issues and implications with using the approach.

Two key issues that have continually arisen are the:

- fear of outcomes monitoring being used as a punitive contract / performance management tool, rather than as a process for ‘continuous programme improvement’
- problem of small numbers with some relatively rare health events, which affects the ability to statistically detect change due to the intervention.

Implicit in the results focus approach is the idea of ‘accountability’ for achieving the outcomes sought. As a matter of principle though, the degree to which accountability can be maintained depends on

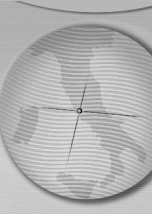
establishing a meaningful causal link (attribution) between cause and effect. However, for many traditional population health programmes establishing attribution involves overcoming a number of practical difficulties.[13] This means that using outcomes monitoring as a strict contract performance management tool is problematic.

The central government agencies in New Zealand have suggested that rather than seeing logic models and outcomes monitoring as a tool for performance management, they should be seen as a way of instituting a continuous improvement cycle in the design and delivery of publicly funded programmes. 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”. [16] This approach has been welcomed by practitioners.

A good health status indicator typically measures either prevalence or incidence. The problem with these types of indicators in New Zealand is that either or both the numerator or the denominator may be very small. When the numbers are small, then statistically it is difficult to measure the prevalence or incidence of the health event of interest. It also becomes very difficult to detect whether a change has occurred, let alone to decide whether any change is related to the intervention.

For example, a small New Zealand rural community may wish to implement a programme to reduce family violence. The community has a population of 1000 people and there are 25 violent offences reported to the police in the year before the programme commenced. This equates to an annual incidence of 25/1000 or 0.025 offences per person per year. A review of overseas evidence of programmes similar to that planned indicates it is realistic to expect a reduction in reported offences of 15%. This level of reduction would be approximately four fewer cases a year in the community.

In this scenario, if a reduction of 15% is actually achieved, the power to statistically detect the reduction in the following year would be 2.7%.



This means if the programme were to operate for another 37 years, on average, we would only be able to statistically detect the difference once in that time. Consequently, with this indicator we could not assess whether the programme had achieved the health outcome of interest. An indicator with higher incidence would be needed to monitor the change made due to the programme in this community. Alternatively, programme planners could decide to implement and monitor the programme across a number of small local sites. Assessment of the success of the programme would be made not upon a single site, but upon the grouped results. The effect of this is to increase the ability to detect the change sought.

The introduction of the approach described has also highlighted a need for planners and managers to:

- up skill in the use of programme planning methods, and to spend more time on planning programmes
- learn how to interpret and use research evidence to inform their planning and decision-making, rather than assuming traditional forms of intervention and methods of delivery remain relevant
- design programmes that are more specific, time limited, and better sequenced
- place more emphasis upon counting health 'outcomes' rather than health 'outputs'
- focus less upon traditional forms of formative and process evaluation that tend to use qualitative methods, compared to outcomes monitoring that tends to rely upon quantitative methods
- realise that outcomes monitoring is not a cheap replacement for evaluation.

### Conclusions

An outcomes monitoring plan is a process for the routine systematic collection and recording of timely information about aspects of a programme over time. The purpose is to assess whether progress is being made in achieving the programme's objectives. The adoption of such an approach is increasingly being required by government agencies as a means of promoting efficient and effective resource allocation. Preparing an outcomes monitoring plan requires assessing:

- what must be measured to demonstrate success;
- how things are going to be measured and by whom and
- the timeframes for delivering information about the performance of the programme to key stakeholders.

An outcomes monitoring plan is the final component of an intervention logic model that aims to systematically link programme activities to policy objectives and the measurement of progress to achieving desired outcomes.

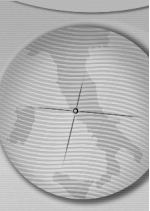
The inclusion of a glossary of common terms within the Ministry's guidance document was welcomed by attendees at the training workshops.[5] The glossary was a useful mechanism for providing a common language around 'how to' design and implement population health programmes using the logic model and outcomes monitoring approach.

When developing outcomes monitoring plans at the local and regional level of delivery two key issues have been identified that threaten the usefulness of the approach. The first is the need to manage the statistical problem of small numbers associated with the relative rarity of some health events in small populations. The second is that outcomes monitoring should be seen as a tool for 'continuous improvement' rather than solely as a method for performance management, which can be seen as a threatening or punitive process.

Other issues arising include the need for population health planners and managers to acquire new knowledge and skills around programme planning, to learn how to transfer research knowledge into operational delivery decisions, to think hard about how to best to sequence intervention activities, and to move beyond traditional ways of measuring the success or otherwise of a programme.

### References

- 1) The Treasury, State Services Commission. Guidance and Requirements for Departments: Preparing the Statement of Intent. Wellington, New Zealand: The Treasury, State Services Commission, 2007:29.
- 2) The Treasury, State Services Commission. Guidance and Requirements for Crown Entities: Preparing the Statement of Intent. Wellington, New Zealand: The Treasury, State Services Commission, 2007:23.
- 3) Webber D. Wrestling with Outcomes: The New Zealand Experience. *Agenda*, Vol 11, No 4 Australian National University, Canberra Reprinted in *Public Sector, Journal of The New Zealand Institute of Public Administration*, Autumn 2005:2004.
- 4) Ministry of Health. A Guide to Developing Public Health Programmes: A generic programme logic model. Wellington: Ministry of Health, 2006:51.
- 5) Ministry of Health. How to Monitor for Population Health Outcomes: Guidelines for developing a monitoring framework. Occasional Bulletin No. 44. Wellington: Public Health Intelligence, Ministry of Health, 2007:45.
- 6) Schon D. *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books, 1983.
- 7) Ministry of Health. *Achieving Health For All People: A framework for public health action for the NZ Health Strategy*. Wellington: Ministry of Health, 2003.
- 8) International Conference on Health Promotion. *The Ottawa*



Charter for Health Promotion. Ottawa: World Health Organisation, 1986.

9) Brumby J, Robinson M. Performance budgeting, an Overview. In International Seminar on Performance Budgeting. Brasilia: International Monetary Fund and Brazilian Ministry of Planning, Budget and Management, 2004.

10) Kellogg Foundation WK. W.K. Kellogg Foundation Evaluation Handbook: W.K. Kellogg Foundation, 1998.

11) United Nations Development Programme Evaluation Office. Handbook on Monitoring and Evaluating for Results. New York: United Nations Development Programme Evaluation Office, 2002:140.

12) United Nations Population Fund Division for Oversight Services. Programme Manager's Planning Monitoring & Evaluation Toolkit: Tool Number 1: Glossary of Planning Monitoring & Evaluation Terms. New York: United Nations Population Fund Division for Oversight Services, 2004:14.

13) Nutbeam D. Health promotion glossary. Health Promot Int 1998;13(4):349-64.

14) Bartholomew L, Parcel G, Kok G, Gottlieb N. Planning Health Promotion Programmes: An Intervention Mapping Approach. Jossey-Bass, 2006.

15) Iverson A. Preparing Program Objectives: Theory and Practice. Toronto: The International Development Research Centre: Evaluation Unit, 2003:51.

16) The Treasury, State Services Commission, Department of the Prime Minister and Cabinet. Guidance for Crown Entities: Planning and Managing for Results. Wellington, New Zealand: The Treasury, State Services Commission, Department of the Prime Minister and Cabinet, 2005:15.