



PATHFINDER

Guidance on Outcomes Focused Management

Supporting Paper : Benchmarking Using Outcome Information

Version 2.1 July 2003

This document provides guidance for agencies integrating outcome information into their decision-making processes. The document was produced by New Zealand's Pathfinder Project. More Pathfinder guidance documents are available on <http://io.ssc.govt.nz/pathfinder>.

We hope other outcome-based initiatives continue to develop the material presented in this suite.

Benchmarking Using Outcome Information

Benchmarking studies should include benchmarking outcomes as well as costs. Benchmarking requires inclusion of analysis of outcomes in terms of effectiveness – i.e. quality, appropriateness, and accessibility and equity.

NATSEM, University of Canberra, April 2001, p.4.¹

Purpose Statement

1 This paper discusses the use of outcome measures in benchmarking studies. The paper is written for agencies that have developed a precise definition of their “vital few” outcomes, and have robust outcome indicators or impact measures for some of these outcomes.

2 The target audience for this chapter is senior managers and business analysts who require an overview of benchmarking activities and the broader benefits it can offer your organisation.

Management Overview

3 The focus of this paper is on benchmarking results, that is, comparing outcomes to valid benchmarks. Benchmarking helps identify effective practices, so they can be analysed, documented and spread throughout the organisation. Systematic comparisons with similar organisations provide information about performance and help you identify areas for potential improvement. Benchmarking can help set performance goals, show improvement over time, and provide a basis for reporting results.

4 There are seven important considerations in using benchmarking effectively:²

- i focus on outcomes from the start. Precise definitions and good measures are needed;
- ii start with the ‘vital few’ outcomes that are expected from core services. Improving these services should generate the largest results;
- iii use multiple measures of outcomes, and include quality measures. Focusing on one dimension of performance may focus you too narrowly on a limited range of interventions. Multiple measures can help where information is patchy or uncertain.³ Information on costs, coverage and near-term results can help you interpret differences in outcomes;
- iv check that the benefits of benchmarking outweigh the costs.⁴;
- v benchmarking is an integral part of your strategic and business planning. Measurement, management and performance improvement are inherent in good planning, not a side issue;
- vi don’t reinvent the wheel. Sector knowledge can help to identify outcomes, methods, data and business units. Often the information is already held and just needs to be processed; and
- vii benchmarking is not an end in itself, but is a tool for performance improvement. Make sure that the lessons learned through benchmarking are transferred.

¹ Report to the Board of the ACT Health and Community Care Service on ACT Public Hospital Costs

² After, Improvement and Development Agency (UK), *Join the Club? Benchmarking for best value*, October 2000, p.2.

³ See recommendation #5 for enhancing the usefulness of performance data in Harry Hatry, Craig Gerhart and Martha Marshall, ‘Eleven Ways to Make Performance Measurement More Useful to Public Managers’, *Public Management*, September 1994. (accessed at: <http://icma.org/go.cfm?cid=1&gid=3&sid=101&did=114>)

⁴ For example, agencies should determine if there is room for improvement and if there are political, regulatory, and organisational or resource constraints that, in practice, make it very difficult to get improvement.

Making Benchmarking Work

5 Benchmarking offers management powerful insights into what works. Benchmarking projects can inform decision making and provide information for:

- a *Strategic planning.* There is a two-way relationship between strategy and benchmarking. Strategy can inform what to benchmark, who to benchmark with and what the results may mean. Benchmarking helps identify areas that could improve, and good practices that can be used.
- b *Practice improvement.* Benchmarking helps identify effective practices, so they can be analysed, documented, acknowledged as standards, and spread throughout the organisation.
- c *Organisational goals and targets.* Benchmarking can be used to set challenging goals for learning and to encourage improvement. Benchmarking can also ensure that the goals and targets are feasible and credible. Benchmarking can help you promote innovation and celebrate success.
- d *Performance monitoring and reporting.* Benchmarking can be used to provide information on the achievement of outcomes, and to report on performance in comparative terms.

6 Benchmarking is not a stand-alone activity. It is part of a determined approach to assess performance, spread best practice, reallocate scarce resources and improve outcomes. Benchmarking raises questions about historical performance. Management should therefore be prepared to respond to defensive or adverse reaction within the organisation⁵. To obtain best results⁶:

- a *Senior management provides leadership and support.* Visible support is required to fund and conduct benchmarking, get agreement on best practice and implement change. You need to set clear objectives to focus the work and avoid making performance comparison an end in itself.
- b *Focus on important outcomes and groups.* Benchmarking can be costly so focus on what matters most, usually the ‘vital few’ outcomes. Focus effort on major business units or areas where performance variation suggests there is scope for performance improvement. Nationally or internationally accepted outcome definitions help you make comparisons across institutions.
- c *Be willing to change.* Your organisation must be willing and able to change when the potential for performance improvement is identified. A culture that encourages best practice and comparison to peers underpins success.
- d *Strong links to best practice and business improvement.* Benchmarking is not a stand-alone activity – it is part of a systematic effort to test the underlying intervention logic and improve performance. Management should recognise the links between benchmarking, best practice, continuous improvement and strategic planning.
- e *Decisions follow analysis.* Your goal is to improve outcomes. Simply identifying differences in outcomes will not, in itself, lead to improvement. You must analyse how and why performance varies, and produce specific, actionable recommendations to management.

⁵ Common responses reported by Keehley et al (Benchmarking for best practices in the public sector, 1996) include:

- Scepticism and distrust - particularly if the results are unexpected or unflattering.
- “Shoot the messenger” – staff delivering unwanted results may be heavily criticised.
- Not invented here – methods identified as effective are discredited because they were not developed in-house or take account of unique circumstances. Often critics emphasise the lack of “like with like” comparisons.
- We’ll look into it – where results seem to be accepted, but are not used.
- We cannot afford it – benchmarking costs money better spent elsewhere.
- Do not have time – our staff are spread too thinly on tasks of greater value.

⁶ After the UK Audit Commission (*‘Getting better all the time: making benchmarking work, (2000)*

7 The most persuasive benchmarking will identify, analyse, document and promote practical ways of improving outcomes. This will be tempered with careful analysis to identify cultural, temporal, geographic or other characteristics that limit the identification and replicability of ‘best’ practice, and that can lead to blind imitation instead of encouraging innovative thinking.⁷

8 Benchmarking is not an exact science – it requires judgement and insight. Achieve the level of accuracy required to make good decisions. It is better to be roughly right than precisely ignorant.⁸

Benchmarking in Practice

9 Managing for Outcomes is about focusing management effort on achieving core results (outcomes). Benchmarking helps you identify management practices that achieve results.

10 There are two main types of benchmarking. **Internal benchmarking** compares processes and results across different branches or units of the same organisation. **External benchmarking** compares processes and results across different organisations trying to achieve the same outcomes.

11 Benchmarking activities need not be complex. Good benchmarking uses straightforward methods to determine practical answers and worthwhile improvements.⁹ Hatry outlines six commonly used methods that can be used when benchmarking outcomes or results:¹⁰

- i performance in previous periods;
- ii recognised standards and established targets;
- iii outcomes for different groups and service categories;
- iv different practices or interventions in pursuit of the same outcome;
- v performance of similar organisational units or geographical areas; and
- vi Performance of similar organisations elsewhere

12 Agencies should identify a small number of outcomes they wish to achieve (preferably their “vital few” outcomes) and robust measures for these outcomes (see Building Block 1). Benchmarking studies on interventions can also include comparisons of intermediate outcomes and near-term results.

13 While methods vary, benchmarking generally follows discrete steps (Exhibit 1).

Exhibit 1: A generic benchmarking cycle

Individual agencies will have their own approach to identifying best practice, pursuing continuous improvement and to strategic planning. Likewise there is no single approach to benchmarking although benchmarking generally involves the following generic steps.¹¹

- i identify the outcome(s) to be improved and the indicators to be measured;
- ii measure your own performance;
- iii identify, screen and recruit benchmarking partners (internal or external as appropriate);
- iv collect and analyse data, and determine if there is a performance gap;
- v communicate results;
- vi develop and implement action plan; and
- vii monitor results.
- viii Recalibrate benchmarks (*Repeat process...*)

⁷ Steven Cohen and William Eimicke (Columbia University), *Understanding and Applying Innovation Strategies in the Public Sector*, Presentation to the 57th Annual National Conference of the American Society for Public Administration, 1996, Atlanta Georgia.

⁸ H. Hatry, *Performance Measurement: Getting Results*, The Urban Institute, Washington DC, 1999, Chapter 9, p.99.

⁹ Australian Public Service Commission, *Raising the Standard*, para. 6.

¹⁰ H. Hatry, *Performance Measurement: Getting Results*, The Urban Institute, Washington DC, 1999, Chapter 9, p.99.

¹¹ After, Treasury Board of Canada Secretariat, *Guide X: Benchmarking and Best Practices: An Update to Guide VII*, March 1996, p. 3.

i Performance in previous periods

14 Comparing results with past periods (including longitudinal and time series analysis) is a simple and common type of benchmarking activity.

15 Data on past results for business units or interventions may be available if you have been actively monitoring important results. For new measures, past data may not initially be available but that should not stop new measures being collected. As more data is collected, comparisons will be possible.

16 Comparing results between current and previous reporting periods (e.g. monthly, quarterly, annually) helps managers to assess whether performance has improved or deteriorated over time. But there are two important considerations when making this sort of comparison. They are:

- i *Frequency of reporting.* Agencies should consider value and the cost of collecting more data when deciding how often they will report against a benchmark. If the outcomes change slowly, frequent reporting may have low value. Conversely, if data is already held by the organisation, more frequent reporting can often be obtained at minimal cost.
- ii *Comparing similar periods of cycles.* For programs that have seasonal variation, data from one period should be compared with data from same period in previous years. Similarly, labour market comparisons should be made at similar phases of the economic cycle.

17 Some indicators will not change rapidly, but are affected by cyclic variations. Both the expected rate and cycle of change will then guide decisions on the frequency of reviewing a benchmark.

Exhibit 2: Comparison with past periods: Road Fatalities in New Zealand 1996 to 2001

Casualty Types	1996	1997	1998	1999	2000	2001
Drivers	238	253	221	254	244	235
Passengers	151	164	139	141	132	123
Motorcycle Riders	42	52	47	39	29	33
Motorcycle Pillions	6	4	7	3	2	1
Pedestrians	63	54	71	63	35	52
Pedal cyclists	13	12	16	8	19	11
Other	1	-	-	-	1	-
Total	514	539	501	508	462	453

Source: <http://www.ltsa.govt.nz/research/fatal5yr.html>

18 Time series data may show complex patterns or be influenced by factors that cannot be clearly identified. Further analysis will then be required to identify the variables influencing these patterns. In some cases these variables may not be easily predictable. So take care when interpreting results.

ii Recognised standards and established targets

19 A recognised general standard for an outcome – e.g. one set by government, professional bodies or international convention - can be used to assess performance. It is important that these are outcome standards, not process or input standards. The latter will not be useful in benchmarking outcomes. Examples of such standards include environmental air quality and water quality.

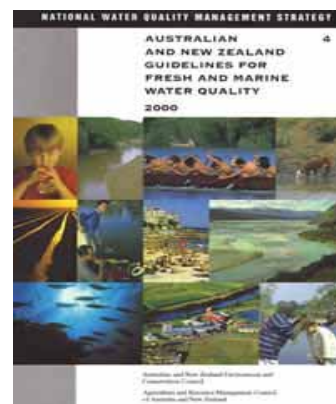
Exhibit 3: Using Standards: Australian and New Zealand Guidelines for Water Quality¹²

Since the early 1990s the Agriculture and Resource Management Council of Australia and New Zealand and the Australian and New Zealand Environment and Conservation Council have developed a joint framework for managing water quality. Known as the *National Water Quality Management Strategy*, a series of water quality policy, benchmark and industry guidelines sets out the agreed processes for identifying and protecting the 'environmental values' of surface, marine and groundwaters.

Two key strategy papers have been released:

- the Australian and New Zealand Guidelines for Fresh and Marine Water Quality (2000); and
- the Australian Guidelines for Water Quality Monitoring and Reporting (2000).

These documents provide a comprehensive approach to water quality issues, such as salinity and eutrophication.



20 Voluntary standards also provide useful metrics for identifying sites warranting management attention. But it is difficult to set standards without baseline data on past results. When data for a specific measure have never been collected or analysed before, management is unlikely to know what would be a reasonable level of performance. One approach in this situation is to rank units' or agencies' performance relative to one another, rather than against a fixed standard.¹³

21 Standards can also be established using a national average. Care should be taken when comparing results solely against an average as changes in a unit's or agency's performance may be due to changes in other units' or agencies' performance.

22 It is sometimes necessary to align national standards for each measure with procedures to correct for differences between regions or groups. An unintended side effect of modifying standards to reflect differences in demographic conditions is that it can reduce the incentive to provide appropriate services to populations identified as "hard-to-serve."¹⁴ Where this is a risk, correction procedures should take into account the characteristics of the groups receiving services. Alternatively, outcome indicators could be monitored for the whole population of interest as well as those receiving services.

¹² In New Zealand, *Drinking-Water Standards for New Zealand 2000* list the maximum concentrations of chemical, radiological and microbiological contaminants acceptable for public health in drinking-water. For community drinking-water supplies, the Standards also specify the sampling protocols that must be observed to demonstrate that the drinking-water complies with the Standards.

¹³ The US Department of Health and Human Services (HHS) dealt with this issue by rewarding the top states in each category of the Temporary Assistance for Needy Families (TANF) programme, rather than by establishing a fixed standard. See, US Department of Health and Human Services (HHS), Administration for Children and Families (ACF), and, Assistant Secretary for Planning and Evaluation (ASPE), *Report on Alternative Outcome Measures: Temporary Assistance for Needy Families (TANF) Block Grant*, December 2000. See Chapter 2 'Principles of Outcome-based Performance Measurement'.

This report is available at: <http://aspe.hhs.gov/hsp/alt-outcomes00>

¹⁴ HHS et al., *Report on Alternative Outcome Measures*.

23 Another method is to monitor improvements over time as well as absolute levels of performance. The US Department of Health and Human Services (HHS) took this approach in developing the TANF High Performance Bonus measures. HHS believes this gives states that have performed poorly a strong incentive to improve, even if they are unlikely to achieve top rankings. Since demographic conditions do not change much from year to year, improvements are likely to be caused by changes in operations rather than by underlying conditions.¹⁵ This approach is equally applicable to benchmarking the performance of similar organisational units or geographical areas.

24 Goal or target setting is an interesting variant of benchmarking against standards. Targets are typically set where agencies know disparities in outcomes exist between groups or business units, or to promote general improvement across a population. As an example, the *New Zealand Road Safety Strategy 2010* sets out the government's plan for road safety and includes explicit targets, such as:

- a A general target of no more than 300 fatalities and 4500 hospitalisations per year by 2010.
- b Specific regional and road user group targets for 2004.

Government agencies that contribute to road safety (Police, Land Transport Safety Authority, Transit New Zealand, ACC, Ministry of Transport) will report against these road safety targets.

One criticism of using standards and targets in benchmarking is the potential for unintended consequences. For example, agencies or individual business units may change policy or pervert practice in unintended ways.¹⁶ They may create incentives to alter policy or practice to improve measures rather than outcomes. Management, analysts and auditors should be alert to possible negative incentives and the unintended consequences that may arise. Using multiple measures will help mitigate this risk.

After, Courtney, et al., 'National Standards in Child and Family Services Reviews: Time to Improve a Good Idea', Paper prepared for the 2002 APPAM Fall Conference, Dallas, Texas, Nov.7 2002.

25 There are a number of factors to bear in mind when setting and benchmarking against targets:¹⁷

- a targets must be challenging, but achievable. Previous performance is relevant;
- b a target does not have to be an absolute value; ranges and rates are acceptable;
- c targets must be cost-effective to achieve, and discussed with interested parties;
- d targets should be backed by measurement systems and ongoing monitoring;
- e allow for statistical variation, extraneous factors, legislative change, and possible effects of cyclic factors and directional trends, which may affect achievement of the target; and
- f partial success must be treated as success, not failure.

Ensure that any targets you set create the behaviours you want - targets can produce odd behaviours. Target setting and monitoring systems must assess the impact of different factors that affect outcomes, and whether a target could be reached or not. This creates stronger incentives to strive for success.

¹⁵ Ibid. Care must be taken in relying on standards that have incremental increases each year. Management should be alert to the possibility that a maximum realistic level has been reached and that seeking any further improvement would have a negative benefit/cost. Mark Courtney, Barbara Needell and Fred Wulczyn, 'National Standards in Child and Family Services Reviews: Time to Improve a Good Idea', Paper prepared for the 2002 APPAM Fall Conference, Dallas, Texas, Nov.7 2002, provide a critique of HHS' methodologies for setting national standards. [Thanks to Rose Ryan & CYFS for this reference].

¹⁶ Conversely, not conducting benchmarking may allow the persistence of negative incentives or unintended consequences.

¹⁷ After, Hatry, *Performance Measurement*, pp.129-130; and National Road Safety Committee, *Road Safety Strategy 2010*, pp.27-41.

iii Outcomes for different groups and service categories

26 Outcomes for different groups and service categories can be used to identify whether outcomes are better or worse for particular subgroups within the population of interest. Where agencies have robust outcome indicators or impact measures, and can allow for variations in practice or throughput (e.g. client group composition) that may cause differences in outcome, results can be used to contrast the effectiveness of different business units in delivering what are supposed to be identical services.

An example of this approach to identifying priority populations for targeting is the Ministry of Health's publication *Our Health, Our Future: The Health of New Zealanders 1999*.¹⁸ The report provides information on population health outcomes and their distribution across subgroups within the population, and the causes or determinants of these outcomes. *Our Health, Our Future* includes a description of population health status in terms of the two dimensions of health: quantity and quality of life. The report then integrates these two dimensions using both health expectancy and health gap measures. Finally, the report analyses the scope for health gain. Variations in health outcomes between generations, genders, ethnic groups, and socio-economic groups are explored.

iv Different practices or interventions in pursuit of the same outcome

27 This type of benchmarking is used to compare alternative policies, processes or procedures to those currently in use. This approach is generally applied in three circumstances:

- i When new practices are introduced across the board to replace the old practice. Here results from the period before the change are compared with results after the change.
- ii When new practices are introduced into part of an operation and the old and new practices run concurrently. This allows alternatives to be tested without irreversible commitment.
- iii When deliberate or 'natural' variation in practice occurs.

28 When outcomes and the costs of different practices are known, this type of benchmarking allows management to consider the cost of alternative practices and factor in resource implications.

Variations in practice can cause programmes to fail. Cognitive Skills training is part of the Correctional Service of Canada's (CSC) Living Skills Programmes. The Cognitive Skills programme targets "faulty" thinking patterns that help maintain patterns of criminal offending. Cognitive deficits addressed include impulsive decision-making, narrow thinking, absence of goal-setting, and maladaptive inter-personal skills.¹⁹ The programme was piloted in two regions in 1988 and 1989 and resulted in a 10% reduction in re-offending. CSC began national implementation of Cognitive Skills Training in 1990, yet evaluation results indicated only a 1% to 2% effect on re-offending. Knowing this, CSC went back and focused on improving service delivery. Building on the Canadian experience, The United Kingdom recently achieved 11% to 15% reductions in re-offending with minimum risk offenders.

v Performance of similar organisational units or geographical areas (internal benchmarking)

29 This type of benchmarking is used to compare outcomes between business units that provide similar services to a comparable customer base. Results indicate which business units or geographical areas are performing well and which are performing poorly, relative to one another. They are used to judge best practice on results, as an alternative to making subjective judgements of service quality.

¹⁸ Ministry of Health, *Our Health, Our Future: The Health of New Zealanders 1999*, Wellington, December 1999.

¹⁹ Research Report No. R-41: The Impact of Cognitive Skills Training on Post-Release Recidivism among Canadian Federal Offenders, David Robinson, August 1995 Correctional Service Canada - Service correctionnel Canada.

30 It is important that either the services and recipients are reasonably similar across business units or geographical areas, or that allowance is made for variation in services or recipients. To produce the benchmarking results needed, the data must be able to be disaggregated to business unit or area.

31 An example of this approach is the NZ Department of Labour's six-monthly reports on regional economic activity and labour market conditions that compare regional data with national data.²⁰

Exhibit 4: The importance of consistent measures — comparing hospital costs

Three major studies compare hospital costs across Australia. Results were wrongly used to suggest that the Australian Capital Territory (ACT) had an expensive and inefficient public hospital system. A review found that the methods used were appropriate only if the purpose was to measure average cost. But the studies were not comparing "like with like". Variations in demographic composition, scope of clinical services, referral patterns, discharge options, patient risk, accounting practices, and scale all had a profound impact on relative service delivery costs.

It is more appropriate to compare the cost of providing particular treatments within different hospitals, rather than comparing hospitals delivering a different service mix. Comparing average costs for entire hospitals could not identify where services were provided at excessive cost.

From 'Executive Summary', *Report to the Board of the ACT Health and Community Care Service on ACT Public Hospital Costs*, NATSEM, University of Canberra, April 2001.²¹

32 The underlying assumption in this type of benchmarking is that business processes and client base are similar. Where this is not true, correction must be made for differences.

33 This type of benchmarking may be attractive because access to sensitive data and information is easier; standardised data is more readily available; and benchmarking can be done quickly and cheaply. There may be fewer barriers to transferring practices across the same organisation. But management should also consider whether the practices identified within the organisation represent the best practice that the organisation can achieve. External benchmarking may help here.

vi Performance of similar organisations elsewhere (external benchmarking)

34 Another common benchmarking approach is to compare your outcomes with outcomes achieved by similar organisations in other jurisdictions (or the private sector). Done well, external benchmarking can show how well you are doing in relative terms, and how much outcomes could be improved. It also helps identify best practice, which you can emulate (or which others could emulate).

Exhibit 5: Making meaningful international comparisons – a good practice example

The New Zealand Department of Corrections benchmarks performance indicators for Public Prisons against selected international jurisdictions. The Department has an 'International Benchmarking Catalogue' that describes definitions and counting rules adopted each jurisdiction and discusses the level of comparability of each performance indicator with New Zealand.²² This helps the Department to detect statistical anomalies and determine where performance variation is real.

35 Organisations and business environments are diverse and complex. Be careful when you are comparing outcomes and practices that you:

- a can access comparative information in a timely manner;

²⁰ Labour Market Policy Group, *Regional Labour Market Report; Auckland*, June 2002.

<http://www.lmpg.govt.nz/Labour-Market-Reports/Auckland%20report%20June%202002.pdf>

²¹ Available from: http://www.natsem.canberra.edu.au/pubs/otherpubs/canb_hospital/costs.html

²² [Thanks to the NZ Department of Corrections for providing a copy of its 'International Benchmarking Catalogue'].

- b understand differences in policies and measures that can distort outcomes (and correct for these where possible);
- c know why 'best practice' worked well in the other jurisdiction; and
- d confirm that the same success factors can be replicated in New Zealand.