

The role and value of benchmarking in organisational change within a New Zealand District Health Board

Phillipa Gaines

Lattice Consulting Ltd, New Zealand

Abstract

Benchmarking has assumed an important role in the performance management toolkit. However, the evidence for its effectiveness, particularly in healthcare settings, remains inconclusive. This presentation outlines the findings of an evaluation project that employed a three-stage process to explore the conditions that helped and hindered the use of benchmarking as a service improvement tool in a New Zealand District Health Board adult mental health service.

The first stage of the evaluation used a traditional mixed methods approach that involved interviewing a sample of participants, a review of the supporting documentation and an appraisal of some of the quantitative evidence using the local benchmarked reports. Q-methodology was used for the second stage of the evaluation to help address the lack of significantly new information that emerged from stage one. The approach and the findings from this second stage of the evaluation are the main focus of this presentation.

The four perspectives that emerged from the Q-sort emphasised the importance of sound systemic features (eg, leadership), a good information infrastructure (processes) and a strong values-driven approach (people) as well as a process for socialising 'new entrants' into the 'benchmarking club'. These findings present some challenges to those health service providers that are considering implementing benchmarking activities into their organisation. The inference is that the implementation of a benchmarking process without attention to each one of these key areas, as well as the interplay between them, is unlikely to succeed.

The implications for the evaluation field are two-fold. Firstly, complex subjects of enquiry may require a more adaptive response from evaluators at all stages of the evaluation. Secondly, evaluators may need to consider utilising a wider range of methods which could help to reveal information previously obscured by the 'noise' that is generated within complex systems.

1. Background to the National Mental Health KPI project

In 2004 the Ministry of Health initiated the development of a national key performance indicator (KPI) framework for New Zealand mental health and addiction services with the objective of supporting quality improvement efforts in the sector (Counties Manukau District Health Board, 2007). The development of the KPI Framework was undertaken as a partnership between the funder (Ministry of Health) and service providers with the intention that the final product would form the basis of national benchmarking activity in the mental health and addiction sector over the medium to long term.

The development of the National KPI Framework was informed by the model that was undertaken in Australia during 2005 (NMHWG Information Strategy Committee Performance Indicator Drafting Group, 2005), which in turn was informed by the model developed by the Canadian Institute of Health Information in 1999. The New Zealand KPI Framework was subsequently published in 2007 by Counties Manukau District Health Board and the first national benchmarking forums commenced in 2009. This phase of the National KPI Project involved only nine selected DHB adult mental health services and their NGO service provider partners.

The benchmarking initiative has subsequently evolved to include two national benchmarking streams as follows:

- (a) all twenty DHB adult mental health services and their nominated NGO partners,
- (b) all twenty DHB child & youth services and their nominated NGO partners.

Since the first benchmarking forum in 2009, much has been learnt about those organisational factors and environmental conditions that are more likely to generate service improvements with the use of a benchmarking process. Although benchmarking is now common practice in a number of different industries, including public health services, it has evolved to include different types of benchmarking (Codling, 1996), different models of benchmarking (Watson, 1993) and different approaches to the implementation of benchmarking (Longbottom, 2000), all of which further complicate the issue for those organisations that are attempting to introduce benchmarking into practice, as well as those evaluators that are wanting to assess its value as a service improvement tool.

There is also a paucity of information about benchmarking activity which is being conducted in the New Zealand environment, apart from research published by Mann and Grigg (2004) and a theory of benchmarking that was postulated by Moriarty (2011).

Given the increased use of benchmarking as a mechanism to help drive improvements in the mental health and addiction system now and into the future, it is timely to conduct an in-depth examination of the role and value of benchmarking in organisational change from the perspective of one of the District Health Boards that has participated in each phase of the National Mental Health KPI project (Northern DHB Support Agency Ltd, 2012). This evaluation was undertaken with the following two questions in mind:

1. To what extent is it possible to attribute service improvements within District Health Boards to the introduction of benchmarking?
2. What key factors seem to contribute to improvements in system performance and health gains for service users?

2. Methods used in the DHB evaluation

This evaluation was completed as part of a Post Graduate Diploma in Social Sector Research and Evaluation (Massey University). The DHB was selected on the basis that it had an established track record of internal benchmarking that dated back to July 2010 and it was considered by other DHBs to be an exemplary case.

The first phase of the evaluation involved a brief literature search on the theory of benchmarking, the different models of benchmarking and the models of change in health care services.

The second phase of the evaluation involved some traditional methods for the collection and analysis of data about the benchmarking process in this particular DHB, including:

- The views of a group of participants who were selected because of their current or previous leadership roles in either the service or one of the partner NGOs (via a process of semi-structured interviews).
- Supplementary documentation to provide additional contextual information about the DHB's operating environment as well as the stated strategic objectives for the service (eg, DHB's Strategic Plan, the Service Improvement Plan for the Mental Health Service).
- Benchmarked reports for the local DHB adult mental health service spanning the period June 2009 to July 2011. These reports were produced using administrative data and were routinely generated by the DHB for all teams within the adult mental health service. All of these reports were reviewed by the evaluator with a view to identifying those

service areas where a clear process shift was indicated via the use of two very basic statistical rules – ‘trends’ and ‘the clump of eight’ as described by Balestracci (2009).

The main themes that emerged from this phase of the evaluation were similar to those reported by Amaral and Sousa (2009) who identified three different types of factors that could act as barriers to internal benchmarking, as follows:

- (a) organisational issues (people, culture, context)
- (b) problems with project management (planning, leadership, business pressures) and
- (c) difficulties in accessing and comparing data.

The similarities between this evaluation and the Amaral and Sousa (2009) study was reassuring, albeit somewhat uninspiring, but what was of interest was the influence of some additional accountability drivers which were an integral part of the DHB benchmarking process (eg, the voices of service users, families/whānau and of Non Government service providers). This difference proved to be important in the final analysis.

The third phase of the evaluation evolved in response to the lack of significantly new information that emerged from phases one and two. The main proposition for phase three emerged as a result of me reflecting on my own personal experience of the benchmarking process and then comparing this experience with the results from phase two as well as the main themes from the benchmarking literature. This led to the hypothesis that the adoption of a ‘recipe approach’ to benchmarking was not sufficient in itself to deliver service improvement and that there was an intersubjective component which was critical to its success.

Phase three was never envisaged in the original evaluation design but it became necessary in order to deliver value to the key stakeholders, in accordance with the guiding principle of utilisation-focused evaluation (Patton, 2008).

The primary method used in stage two was Q-methodology (Brown, 1996; McKeown and Thomas, 1988). The Q-sort was selected as the best method for continuing to explore the complexity of the benchmarking issue because it enabled the evaluator to look at the pattern of individual responses with the aim of revealing ‘distinct points of view’ (Donner, 2001) that were possibly inaccessible using more traditional methods of inquiry, whilst at the same time continuing to maintain the integrity of each individual’s responses as a whole (Webler and Tuler, 2001).

In many ways the process for deciding on the methodological approaches that were used in this evaluation mirrors the improvement cycle that has been adopted by the benchmarking initiative ie, ‘Sherwert Cycle for Learning and Improvement’ in Deming (1994). This cycle emphasises the importance of reflecting on what you have done before moving onto the next action step in the cycle. The reflective step enables people to study the impact of their actions and to modify their approach accordingly before starting the cycle again.

The limitations of Q-methodology include concerns about its validity and reliability. The statements that were used for the Q-sort passed the content validity test by assessing them against the main themes that emerged from the literature on benchmarking. The actual Q-sort represents people’s subjective views on a topic so it is automatically accepted as being a valid expression of their opinion. Because of the rigor introduced by the statistical component of the analysis (Dennis, 1998), Q-methodology is also seen as a more robust technique than alternative methods for measuring people’s attitudes and subjective opinion (Akhtar-Danesh et al., 2008).

3. Results – the four factors

Burke and Litwin (1992) present a model of organisational performance and change based on the difference between transformational and transactional change variables. They contend that the external environment has the greatest impact on an organisation (eg, governmental policy changes), followed by internal transformational variables such as mission/strategy, leadership and culture because it is these variables that affect the whole system and require entirely new behaviour sets from organisational members. The transactional or operational variables (structure, management practices and systems) carry less weight than the transformational variables because changing the structure or changing management practices may not necessarily change the system.

With the above distinction between transformational and transactional variables in mind, the following four perspectives emerged from the Q-analysis.

The first perspective was a '**robust systemic understanding**', which was coined as a phrase that captured a disposition towards factors that were focused on organisational culture, leadership and innovation, all of which were deemed to be transformational in nature.

Similarly, a '**sound information infrastructure**' was coined to capture a second disposition of factors pointing towards the more transactional or operational aspects mentioned above. Participants were of the opinion that having a good information system, good quality data and other good information were critical to their ability to deliver on the service improvement goals that are inherent in the benchmarking approach.

The third coinage '**people, values and principles guiding action**' captured the disposition toward interpersonal factors such as trust and how this trust needs to be cultivated and maintained to enable people to work together to create 'impact'. Rousseau (1990) described this part of organisational culture as the 'inner invisible layer'.

While 'trust' is a complex construct, Jones et al. (2010) maintain that there are two core dimensions. Firstly, trust is confidence that another team or organisation will not act opportunistically or exploit vulnerabilities. Secondly, the partners must have the intent and ability to 'perform to promise'. The amount of congruence between organisational purpose, values and day-to-day practice (ie, walking the talk) pervaded the interviews with participants in stage one of the study, so it is not a surprise to see a perspective emerge from the Q-analysis that was focused on the people, the principles and the values (including trust) that underpin the benchmarking process.

The fourth perspective '**workforce preparation**' was added to the results on the basis that one participant, who was new to the organisation, produced a dramatically different Q-sort to everyone else. This result led to the hypothesis that the organisation socialised aspects of the benchmarking process amongst its staff, but that it did not do this deliberately or consistently. If this hypothesis was correct, then any 'new entrant' would learn to adopt a particular perspective on the process only after they had been exposed to it over a period of time. Essentially it combined organisational and interpersonal effects.

The Q-method, combined with stages one and two of the evaluation, allowed me to develop the final criteria into a rubric against which the DHB was evaluated (see appendix one).

4. Implications

The implications of this evaluation are many and varied.

Evaluation field

Firstly, complex subjects of enquiry may require a more adaptive response from evaluators at all stages of the evaluation. The final criteria for the evaluation would not have been

produced without the use of Q-methodology. Secondly, evaluators may need to consider utilising a wider range of methods, such as Q-methodology, to help reveal information previously obscured by the 'noise' that is generated within complex systems.

Health services involved in benchmarking activity

Thirdly, health care organisations will need to consider how to integrate and balance the four factors in order to sustain an optimal benchmarking practice in their local service (Ling, 2012).

Funders

Finally, funders will need to consider a wider range of methods when commissioning the implementation and evaluation of initiatives that operate in a complex and dynamic context. In addition, in those cases where the mix of the four factors is weak, funders will need to decide if the costs associated with supporting health agencies to participate in benchmarking activity outweigh the potential benefits.

5. Conclusion

While the evaluation could not determine the extent to which the benchmarking process directly contributed to service improvements at this particular DHB, the participants were very clear that the benchmarking process had offered them a useful vehicle for holding different sorts of conversations with one another about service improvement.

Benchmarking activity engages people in conversations in some really different and good ways and networking in different ways (participant).

Conversation as a core means for improving organisational performance is not new (Brown and Isaacs, 2005; Tsoukas, 2009), but it is not a valued commodity in the current language of performance measurement in health, probably because it challenges traditional ideas about the production of knowledge and possibly because there are issues to do with professionals sharing their perspective with others without prejudicing their disciplinary integrity (Smith, 2009; de Bruijn, 2011).

The implications of the health sector adopting more of an inter-subjective approach (Plaskoff, 2011) towards benchmarking are huge. If benchmarking was approached as a vehicle for developing a democratic 'community of practice' that engaged people in reflexive conversations, adaptive learning and action (Reed et al., 2006) instead of it being implemented as a series of steps (Anand and Kodali, 2008) then DHBs might be better positioned to create real 'impact' at the local community level and, at the same time, enhance their level of accountability to various stakeholders.

In conclusion, benchmarking is not a recipe; it is a complex process that is embedded within complex systems. Evaluators might like to consider the use of methods such as Q-methodology to explore the intersubjective components of processes like benchmarking so that other voices and influences are brought into the conversation.

Acknowledgements

I would like to express my gratitude to the people that were interviewed as part of this evaluation. In addition I would like to thank my support team - Dr Robin Peace, Dr Iris Hutchinson and Barbara Wallace.

APPENDIX 1 - THE EVALUATIVE CRITERIA FOR THE BENCHMARKING EVALUATION

Rating	Systemic understanding	Information infrastructure	People, values & principles	Workforce preparation	Clear evidence of service improvements
Excellent	<ul style="list-style-type: none"> Many transformational organisational characteristics are in evidence – e.g. leadership, learning from experience, innovation. 	<ul style="list-style-type: none"> Data quality efforts are supported. Processes & information systems exist to support staff to use the data. Service improvement activities are hard-wired into the overall system and form a core part of BAU. 	<ul style="list-style-type: none"> People trust one another and work together as a whole system of care with a common purpose. Reflective dialogue is seen as a core process for change. The benchmarking process fits well with other related QI activities. 	<ul style="list-style-type: none"> The organisation is proactive and orientates all staff to all aspects of the benchmarking process (i.e. from the selection of KPIs to using the reports). Staff use basic statistical rules to identify practice variation. Staff routinely use QI concepts & tools. 	<ul style="list-style-type: none"> Significant improvements in service practice have occurred which are evidenced by quantitative measures and positive service user and family/whānau feedback. The improvements are systemic in nature and are able to be sustained over time.
Average	<ul style="list-style-type: none"> Some transformational organisational characteristics are in evidence. 	<ul style="list-style-type: none"> Ad hoc arrangements exist to improve the quality of the data. Limited processes & information systems exist to support staff to collect, access and use the data. Service improvement activities occur as one-off projects rather than as BAU. 	<ul style="list-style-type: none"> People talk with one another but are hesitant to trust one another. People work together on defined projects, but not as a whole system of care. The benchmarking process is out of step with other related QI activities. 	<ul style="list-style-type: none"> Information on the benchmarking process is available, but it is not routinely provided to all staff. Staff are familiar with the concept of process variation but struggle to apply the basic statistical rules in practice. QI concepts & tools are known but are not routinely applied by all staff. 	<ul style="list-style-type: none"> Some improvements in service practice have occurred which are evidenced by both quantitative and qualitative information. Service improvements may not be systemic in nature and are vulnerable to changes in personnel and/ or other competing service priorities.
Poor	<ul style="list-style-type: none"> There are no (or very few) transformational organisational characteristics in evidence. The service is managed rather than led. 	<ul style="list-style-type: none"> Data quality is not important. Processes & information systems to support staff to collect, access and use the data are virtually non-existent. The organisation does not prioritise service improvement activity. 	<ul style="list-style-type: none"> People do not trust one another. People avoid conflict and do not talk about the things that matter. People work on their own priorities. Benchmarking is perceived to be a compliance activity rather than an opportunity to identify and make service improvements. 	<ul style="list-style-type: none"> Staff do not receive any orientation to the benchmarking process. Staff are not familiar with either basic statistical rules or the concept of process variation. Staff do not use any QI concepts or tools (e.g. PDSA cycle). 	<ul style="list-style-type: none"> No improvements in service practice have occurred or, in those cases where a few improvements have been made, they are very sporadic, specific to certain staff and are not very likely to be sustained over the longer term.

References

- Akhtar-Danesh, N., Baumann, A. & Cordingley, L. (2008) Q-methodology in Nursing Research: A Promising Method for the Study of Subjectivity. *Western Journal of Nursing Research*, 30, 759-773.
- Amaral, P. & Sousa, R. (2009) Barriers to internal benchmarking initiatives: an empirical investigation. *Benchmarking: An international Journal*, 16, 523-542.
- Anand, G. & Kodali, R. (2008) Benchmarking the benchmarking models. *Benchmarking: An International Journal*, 15, 257-291.
- Balestracci, D. (2009) *Data Sanity: A Quantum Leap to Unprecedented Results*, Medical Group Management Association
- Brown, J. & Isaacs, D. (2005) *The World Café : Shaping our futures through conversations that matter*, San Francisco, CA, Berrett-Koehler Publishers.
- Brown, S. R. (1996) Q Methodology and Qualitative Research. *Qualitative Health Research*, 6, 561-567.
- Burke, W. W. & Litwin, G. H. (1992) A Causal Model of Organizational Performance and Change. *Journal of Management*, 18, 523-545.
- Codling, S. (1996) *Best Practice Benchmarking: An International Perspective*, Gower Publishing Ltd.
- Counties Manukau District Health Board (2007) The Key Performance Indicator Framework for New Zealand Mental Health and Addiction Services.
- de Bruijn, H. (2011) *Managing Professionals*, Routledge.
- Deming, W. E. (1994) *The New Economics for Industry, Government, Education*, Cambridge, Massachusetts, MIT.
- Dennis, K. E. (1998) Q-methodology: New perspectives on estimating reliability and validity. IN Waltz, C. F. & Strickland, O. L. (Eds.) *Measurement in nursing outcomes*. New York, Springer-Verlag.
- Donner, J. C. (2001) Using Q-Sorts in Participatory Processes: An Introduction to the Methodology. *Social Analysis: Selected Tools and Techniques (World Bank Social Development Papers 36)*. Washington D.C., The World Bank, Social Development Department.
- Jones, S. L., Fawcett, S. E., Fawcett, A. M. & Wallin, C. (2010) Benchmarking trust signals in supply chain alliances: moving toward a robust measure of trust. *Benchmarking: An international Journal*, 17, 705-727.
- Ling, T. (2012) Evaluating complex and unfolding interventions in real time. *Evaluation*, 18, 79-91.
- Longbottom, D. (2000) Benchmarking in the UK: an empirical study of practitioners and academics. *Benchmarking : An International Journal*, 7, 98-117.
- Mann, R. & Grigg, N. (2004) Helping the Kiwi to Fly: Creating World-Class Organisations in New Zealand Through a Benchmarking Initiative. *Total Quality Management*, 15, 707-718.
- McKeown, B. & Thomas, D. (1988) *Q methodology*, Newbury Park, California, Sage Publications.
- Moriarty, J. P. (2011) A Theory of Benchmarking. *Benchmarking : An International Journal*, 18, 588-612.

- NMHWG Information Strategy Committee Performance Indicator Drafting Group (2005) Key Performance Indicators for Australian Public Mental Health Services. ISC Discussion Paper No. 6. Australian Government Department of Health and Ageing, Canberra.
- Northern DHB Support Agency Ltd (2012) Key Performance Indicator Framework for New Zealand Mental Health and Addiction Services - Phase III: Implementation of the framework in adult mental health services. NDSA.
- Patton, M. Q. (2008) *Utilisation - focused evaluation*, Thousand Oaks: Sage Publications.
- Plaskoff, J. (2011) Intersubjectivity and Community-Building: Learning to Learn Organisationally. IN Easterby-Smith, M. & Lyles, M. A. (Eds.) *Handbook of Organisational Learning and Knowledge Management*. UK, John Wiley & Sons.
- Reed, M. S., Fraser, E. D. G. & Dougill, A. J. (2006) An adaptive learning process for developing and applying sustainability indicators with local communities. *Ecological Economics*, 59, 406-418.
- Rousseau, D. M. (1990) Assessing Organisational Culture: The case for multiple methods. IN Schneider, B. (Ed.) *Organisational Climate and Culture*. San Francisco, Jossey-Bass.
- Smith, G. C. (2009) From consultation-liaison psychiatry to integrated care for multiple and complex needs. *Australian and New Zealand Journal of Psychiatry*, 43, 1-12.
- Tsoukas, H. (2009) A Dialogical Approach to the Creation of New Knowledge in Organisations. *Organization Science*, 20, 941-957.
- Watson, G. H. (1993) *Strategic Benchmarking: How to Rate Your Company's Performance Against the World's Best*, New York, Wiley.
- Webler, T. & Tuler, S. (2001) Public Participation in Watershed Management Planning: Views on Process from People in the Field. *Human Ecology Review*, 8, 29-102.