



## Beyond Targets: Measuring Better and Rebuilding Trust Comment on “Gaming New Zealand’s Emergency Department Target: How and Why Did It Vary Over Time and Between Organisations?”



Richard Hamblin\*, Carl Shuker<sup>1</sup>

### Abstract

Tenbenschel and colleagues identify that a target for emergency department (ED) stays in New Zealand met with gaming in response from local hospitals. The result is in line with studies in other jurisdictions. The enthusiasm for targets and tight performance measurement in some health systems reflects a lack of trust in professionals to do the right thing for altruistic reasons. However such measurement systems have failed to address this loss of trust and may, ironically, have worsened the situation. A more promising approach for both improving performance and restoring trust may depend upon collaboration and partnership between consumers, local providers, and central agencies in agreeing and tracking appropriate local responses to high level national goals rather than imposing tight, and potentially misleading measures from the centre.

**Keywords:** Targets, Gaming, Performance Measurement, Performance Improvement

**Copyright:** © 2020 The Author(s); Published by Kerman University of Medical Sciences. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Citation:** Hamblin R, Shuker C. Beyond targets: measuring better and rebuilding trust: Comment on “Gaming New Zealand’s emergency department target: how and why did it vary over time and between organisations?” *Int J Health Policy Manag.* 2020;x(x):x–x. doi:10.34172/ijhpm.2020.38

### Article History:

Received: 1 February 2020

Accepted: 3 March 2020

ePublished: 11 March 2020

### \*Correspondence to:

Richard Hamblin

Email: [richard.hamblin@hqsc.govt.nz](mailto:richard.hamblin@hqsc.govt.nz)

A new paper by Tenbenschel and colleagues<sup>1</sup> has found for the first time evidence that gaming of New Zealand’s national emergency department (ED) wait time target of six hours was rife in at least four district health boards.

The authors found that, as in the United Kingdom, patients were discharged from ED to short stay units that existed only in name, the clock was stopped on patients who had not been admitted or discharged, and for some the clock was simply stopped to protect the institution’s performance.

New Zealand’s set of ten varied targets for healthcare performance were instituted in 2007. This set has evolved and mutated over time and changes in government since, before retirement of public reporting of performance against these targets in 2018.<sup>2</sup>

To date, study of the issue of gaming in response to targets within the public sector has tended to show gaming occurs where either strong financial incentives were attached to the target<sup>3</sup> or strong, formal threats to organisations and individuals existed – particularly for their continued employment or organisational viability. These were the regimes of so-called “targets and terror.”<sup>4,5</sup> New Zealand did not embrace such a formal pay-for-performance structure, and while targets were incentivised through publication and strong informal pressure on management, New Zealand’s district health boards retained considerably greater autonomy than National Health Service (NHS) trusts under the United Kingdom’s target regime of the early 2000s. In spite of this,

Tenbenschel’s paper suggests that even the weaker incentives of New Zealand’s regime were sufficient to encourage gaming behaviour.

The target for ED wait times has operated consistently throughout and is the most studied of New Zealand’s target regime, and is the target, after that for immunisations, perhaps viewed most positively by staff and the public.<sup>6–8</sup> These new findings that gaming of performance was widespread provide a useful counterpoint and context to how we understand the effects of targets on provider behaviour.

### What’s the Point of Targets?

So, is there any point in using targets in health systems? Their negative and distortionary effects and unintended consequences have been identified and are increasingly studied, if little known to the wider public. Yet understanding how a health system is performing is essential to ensure its quality; and the studies identifying perverse effects of target regimes also recognised the genuine improvements they engendered. Further, making information about quality available to the users and funders of public health systems seems a moral imperative, and one increasingly expected as a matter of course.

How can we resolve these tensions?

### The Letter Versus the Spirit

To answer this question some consideration of both targets

themselves, their strengths and limitations, and the context in which they were introduced, is helpful.

It is vital to distinguish between a measure and a target. Measuring the distribution of time spent inside EDs is essential to running these safely, effectively and efficiently. However, an associated target is not the measure itself but an externally imposed constraint upon the system which uses the measure to promote a particular aim.

Herein lies the first limitation of targets. Targets are only rarely a direct expression of the aim they are promoting – the aim of the ED target is to prevent inappropriately long waits in ED, not give everyone a six-hour wait there. Thus, the target measure is often only a proxy which has, hitherto, been associated with the aim.

The trouble with fixating on proxies is summed up by what is known as “Goodhart’s Law,” named for the economist Charles Goodhart. In plain English this states that “once a measure becomes a target it ceases to be a good measure.”<sup>9</sup> Or, more formally, the relationship between the measure and the aim breaks down once the achievement of the measure, rather than the aim itself, becomes the focus of those charged with achieving it.

Yet, as the paper demonstrates, targets tend to stimulate actions that produce progress towards the underlying aim, at least at first. The phenomenon of gaming associated with targets might best be described as “gilding the lily” of initial, genuine improvement, as seen in both ED waits and ambulance response times in the United Kingdom.<sup>10,11</sup>

Here too, falsification of data or the arrangement of services to meet the letter but not spirit of the target has been shown to produce telltale signs of “management to measure” – the distributional discontinuities and terminal digit preference bias (opportunistic rounding) Tenbesel and colleagues identify, among other evidence.

It is therefore attractive to seek a way of keeping the benefits of targets while limiting the opportunity for gaming. This paper, like others before it<sup>12</sup> promotes the value of independent validation of measures, and this has an intuitive technocratic appeal. However, this approach also has limitations. Data collection has costs; data collection about the collection of data (which is what validation amounts to) still more so.

### The Evolution of Professional Performance in Public Services

An alternative approach might be to think about what targets represent within public systems. To do this, it may help us to revisit two concepts, one half a century old and the other from the 1990s.

Michael Lipsky, in his concept of “street-level bureaucracy,”<sup>13,14</sup> sketches a picture of front-line work that many health professionals would recognise: staff interact with citizens, operate under resource constraint, and have considerable independence in how they undertake their job. They have the potential to affect considerably the lives of those receiving their services, and yet face ambiguous expectations about job performance. Lipsky explicitly recognises that the unavailability of appropriate performance measures limits the ability of managers to control the application of policy at the

ground level.

This need for control, and the use of targets as a mechanism to gain this, chimes with Le Grand’s 1997 insight into the changing perception of public servants.<sup>15</sup> From being seen as “knights” acting altruistically for the public good, public servants (including healthcare professionals in a publicly funded health system) became “knaves,” primarily motivated by self-interest. What was necessary therefore was a mechanism to harness this tendency to act in self-interest.

The so-called New Public Management that emerged in the 1990s reflected this belief in the self-interested public servant. As originally conceived, rigorous monitoring of services would support market and quasi-market mechanisms that would drive improvement of services through the self-interest of the provider (incentives being protection of their service and thus budget, increased income, personal kudos and so forth).

Over time, as evidence emerged that publication of performance data was far more likely to change provider behaviour because of its potential to harm reputations than it was to stimulate a market of informed consumers voting with their feet,<sup>16</sup> measurement and publication in and of itself became a central policy thrust.<sup>17</sup>

### The Loss of Trust

Yet this development has failed to address what is implicit in both Lipsky’s and Le Grand’s insights: a mutual loss of trust between central government, public services and, crucially, the public.

To counteract this loss of trust, quantitative measurement (apparently objective and precise) is given the task of restoring trust (“one version of the truth” or “a shared understanding of reality”) and targets are given the task of providing accountability: from public service agencies to central government, and from central government to the public.

But measures and targets have been unable to bear the weight placed upon them, and how they have been used may even have served to intensify the mistrust. In our view, this has happened for two reasons: the perception of imposition of measure and the response that this engenders; and over-interpretation of a limited range of measures.

If front-line services believe that targets have been imposed with little understanding of the mechanisms of giving care, what is clinically meaningful, or even what is most pressing and important in an area, they will lose trust in central government’s genuine commitment to actual (as opposed to apparent) performance of the service. This will be exacerbated when, for reasons of practicality, a small number of access measures are presented as an overarching judgement on the overall quality of a service. This imposition and misrepresentation reduces the credibility of the measures themselves and providers respond by what Lipsky describes as a “simplification”: doing what is necessary to hit the target as easily as possible in order to devote more resources to actual local priority. In this way, according to Bevan and Hood’s useful classification, ‘honest triers’ who do not attempt to spin or fiddle data in their favour become ‘reactive gamers’ who do.<sup>4</sup> All incentives are to hit the target, and inevitably the

threshold will arise when doing so means missing the point.

As a result, Goodhart's law comes into play and the target no longer works as anticipated. When central government becomes aware of the disconnect between reported and actual performance this further reduces its trust in front-line services, and historically the response to this loss of trust has been increased reporting requirements that are more complex, more directive and more onerous. Thus, a spiral of mistrust ensues. Meanwhile the public tend to lose trust in reporting by both local services, and especially central government.

### A New Way Ahead

Another approach is needed to address this loss of trust.

One that we have advocated elsewhere<sup>18</sup> involves central government, local services and the public working together to agree necessary local measures to deliver high-quality services. There is some precedent and a considerable literature regarding this approach with regard to local public service agreements introduced under New Labour in the United Kingdom in the early 2000s. In this example, a changed relationship between central and local government to a model where dialogue and negotiation set agreed focuses for improvement was associated with improved outcomes.<sup>19-21</sup> While the literature suggests this new co-operative, trust-based approach was not a panacea for all ills,<sup>19</sup> and successful implementation was crucial, there is much to learn from the experience, and from the wider literature of performance management in complex systems. In particular, recent work on performance management in environments that are characterised by change and uncertainty points to a shift from performance management to a learning rather than control mechanism, and the need for performance management to be more flexible and devolved.<sup>22,23</sup> In the United Kingdom, the nature and spirit of the co-operative approach being reflected in actual negotiations appears key, as does a coherent and consistent narrative in central government departments to guide negotiating behaviours with local authorities in order to survive staff churn and a "regression to the mean" of traditional central command-and-control practices. 'Working in partnership with government rather than in tension' was identified as the result – and goal – of a successful implementation.<sup>19</sup>

In a public health system, especially one primarily tax funded, democracy demands that government should have the right to set high-level aims for the system, and accountability demands that progress towards these aims should be reported publicly. However, the necessary actions to best advance these aims will vary between different hospitals, services and locale. In response to this, local services need to work with their local populations to co-produce plans for local improvement aligned to the high-level objectives, including appropriate, focused measures to track progress (including accepted tools such as statistical process control, cumulative sum analysis, etc, to monitor and direct improvement). These plans should be agreed with central government, be flexible to changes in the environment (whether these changes be successful improvement or addressing emergent issues), and again progress against these measures should be publicly reported.

This approach makes central government and local providers partners in delivering high-quality services.

In our view there are several advantages to this approach:

- Aims that are agreed, rather than targets that are imposed, have a greater likelihood of local professional ownership and support, and are more likely to lead to genuine, clinically and locally relevant change.
- Because of this the incentive to game measures is reduced – technical responses to discourage gaming are important and have a role, but certainly when embedded within a culture of local ownership and trust seem likely to be more effective.
- Mutually agreed aims are more likely to generate trust across the system. Without this health services will not be able to address the challenges they face in the 21st century.

In terms of the challenges noted above in the UK local public service agreement experience, the New Zealand health sector's largely positive experience with the co-developed "System Level Measures" programme may have primed the pump for spread and scale of a truly national, co-operative, trust-based approach to setting and agreeing local contributions to national aims.<sup>24</sup>

Some form of monitoring of health services is now inevitable. Equally inevitable is the risk that these regimes create perverse unintended consequences such as the gaming that Tenbensen and colleagues identify. The choice open to us is whether we respond purely technically to this risk or think deeper about how monitoring and targets can strengthen systems. The latter, with particular reference to how to encourage trust between different parts of the system, is likely to be a more successful strategy.

### Ethical issues

Not applicable.

### Competing interests

Authors declare that they have no competing interests.

### Authors' contributions

RH and CS both contributed to the drafting of the manuscript. Both authors read, reviewed and approved the final manuscript.

### References

1. Tenbensen T, Jones P, Chalmers LM, Ameratunga S, Carswell P. Gaming New Zealand's emergency department target: how and why did it vary over time and between organisations? *Int J Health Policy Manag.* 2020;9(4):152-162. doi:10.15171/ijhpm.2019.98
2. Kirk S. How's your DHB doing? Govt does away with National Health Targets. Stuff; 2018. <https://www.stuff.co.nz/national-politics/104976776/how-your-dhb-doing-govt-does-away-with-national-health-targets>. Accessed January 6, 2020.
3. Kontopantelis E, Doran T, Gravelle H, Goudie R, Siciliani L, Sutton M. Family doctor responses to changes in incentives for influenza immunization under the U.K. Quality and Outcomes Framework pay-for-performance scheme. *Health Serv Res.* 2012;47(3 Pt 1):1117-1136. doi:10.1111/j.1475-6773.2011.01362.x
4. Bevan G, Hood C. What's measured is what matters: targets and gaming in the English public health care system. *Public Adm.* 2006;84(3):517-538. doi:10.1111/j.1467-9299.2006.00600.x
5. Levitt SD, Dubner SJ. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. William Morrow; 2005.
6. Ardagh M. The 'six hour target' in New Zealand is associated with

- reduced mortality and greater efficiency. *N Z Med J*. 2017;130(1455):12-14.
7. Ardagh M. How to achieve New Zealand's shorter stays in emergency departments health target. *N Z Med J*. 2010;123(1316):95-103.
  8. Jones P, Wells S, Harper A, et al. Impact of a national time target for ED length of stay on patient outcomes. *N Z Med J*. 2017;130(1455):15-34.
  9. Strathern M. 'Improving ratings': audit in the British University system. *Eur Rev*. 1997;5(3):305-321. doi:10.1002/(SICI)1234-981X(199707)5:3<305::AID-EURO184>3.0.CO;2-4
  10. Locker TE, Mason SM. Analysis of the distribution of time that patients spend in emergency departments. *BMJ*. 2005;330(7501):1188-1189. doi:10.1136/bmj.38440.588449.AE
  11. Bevan G, Hamblin R. Hitting and missing targets by ambulance services for emergency calls: effects of different systems of performance measurement within the UK. *J R Stat Soc Ser A Stat Soc*. 2009;172(1):161-190. doi:10.1111/j.1467-985X.2008.00557.x
  12. Mears A, Webley P. Gaming of performance measurement in health care: parallels with tax compliance. *J Health Serv Res Policy*. 2010;15(4):236-242. doi:10.1258/jhsrp.2010.009074
  13. Lipsky M. Toward a theory of street-level bureaucracy. Institute for Research on Poverty Discussion Papers. University of Wisconsin-Madison; 1969:48-69.
  14. Lipsky M. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. Cambridge: MIT Press; 1980.
  15. Le Grand J. Knights, knaves or pawns? human behaviour and social policy. *J Soc Policy*. 1997;26(2):149-169. doi:10.1017/S0047279497004984
  16. Marshall MN, Shekelle PG, Leatherman S, Brook RH. Public disclosure of performance data: learning from the US experience. *Qual Health Care*. 2000;9(1):53-57. doi:10.1136/qhc.9.1.53
  17. Dunleavy P, Margetts H, Bastow S, Tinkler J. New public management is dead—long live digital-era governance. *J Public Adm Res Theory*. 2005;16(3):467-494. doi:10.1093/jopart/mui057
  18. Health Quality & Safety Commission. A Window on the Quality of New Zealand's Health Care 2018. <https://www.hqsc.govt.nz/our-programmes/health-quality-evaluation/publications-and-resources/publication/3364/>. Published June 14, 2018.
  19. Sullivan H, Gillanders G. Stretched to the limit? the impact of local public service agreements on service improvement and central-local relations. *Local Government Studies*. 2005;31(5):555-574. doi:10.1080/03003930500293450
  20. Young K. Local public service agreements and performance incentives for local government. *Local Government Studies*. 2005;31(1):3-20. doi:10.1080/0300393042000332828
  21. Boyne GA, Law J. Setting public service outcome targets: lessons from local public service agreements. *Public Money Manag*. 2005;25(4):253-260. doi:10.1080/09540962.2005.10600128
  22. Bourne M, Franco-Santos M, Micheli P, Pavlov A. Performance measurement and management: a system of systems perspective. *Int J Prod Res*. 2018;56(8):2788-2799. doi:10.1080/00207543.2017.1404159
  23. Melnyk SA, Bititci U, Platts K, Tobias J, Andersen B. Is performance measurement and management fit for the future? *Management Accounting Research*. 2014;25(2):173-186. doi:10.1016/j.mar.2013.07.007
  24. Ministry of Health. System Level Measures Framework. <https://www.health.govt.nz/new-zealand-health-system/system-level-measures-framework>. Last updated February 8, 2018.