



OPEN ACCESS

Measurement with a wink

Marleen Kunneman,^{1,2} Victor M Montori,^{1,3} Nilay D Shah^{1,4}

¹Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, Minnesota, USA

²Medical Decision Making, Leiden University Medical Center, Leiden, Netherlands

³Division of Endocrinology, Mayo Clinic, Rochester, Minnesota, USA

⁴Health Care Policy and Research, Mayo Clinic, Rochester, Minnesota, USA

Correspondence to

Dr Marleen Kunneman, Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, Minnesota, USA; kunneman.helena@mayo.edu

Received 13 April 2017

Revised 14 June 2017

Accepted 19 June 2017

Published Online First
21 July 2017

The Smartest Person is a popular Dutch television quiz show in which three contestants receive a few seconds to answer trivia questions. In one of the rounds, to answer the question, contestants must hit certain key words, for example, ‘Apartheid’, ‘Prison’, ‘Nobel Peace Prize’ and ‘South Africa’ for the question ‘What do you know about Nelson Mandela?’ Contestants who mention one of these key words, regardless of context, hear a rewarding ‘ting!’ and receive 20 extra seconds. The most efficient strategy to win is to mention the four key words without a linking sentence; answering with a complete and cogent sentence costs precious seconds. In fact, contestants can win even when they get the context wrong, for example: ‘Nelson Mandela is an Italian actor starring in a film about the Apartheid (ting!). He recently spent a night in Prison (ting!) after he egged the house of a Nobel Peace Prize winner (ting!) from Zimbabwe... Namibia... well, somewhere in South Africa (ting!)’.

The show is advertised as a ‘knowledge quiz with a wink’. The show’s ‘smartest person’ is not the most knowledgeable, but the one who knows *how* best to play the game.

Similar challenges affect performance measurement in healthcare. For example, consider the proportion of patients with 10-year cardiovascular risk >7.5% prescribed statins, as a measure of good preventive care. Or the proportion of patients with type 2 diabetes with haemoglobin A1c (HbA1c) levels <7.5%, as a measure of good diabetes care. Or the proportion of patients getting low-dose CT (LDCT) to screen for lung cancer who participated in shared decision making, as a measure of patient-centred care. These measures are set up to improve the quality of care for patients, but if you know how best to play this game, you would be rewarded for prescribing statins (ting!) to patients who are not willing or able to take them, or for getting a HbA1c of 6.9% (ting!) in a patient living in fear of

recurrent hypoglycaemic episodes, or for giving a brochure and ‘engaging patients’ (ting!) but only to determine their eligibility for Medicare reimbursement of LDCT. Clinicians are gearing up to play a high-stakes version of this game, as they try to make the 60% of Medicare’s Merit-based Incentive Payment System or MIPS by acing (ting!) quality measures of their choice.¹

Playing the measurement game to win has important unplanned consequences beyond unintended patient harm. Clinicians spend about twice as much time on documentation and billing tasks interfacing with the medical record than on facing patients.² Health systems in the USA spend over US\$15 billion each year collecting, processing and reporting performance measures.³ It is not self-evident that the upside of this attention, time, work and cost is getting care right, that the care improves because of measurement and reporting, or that these care improvements are sustained.⁴ Nor is it self-evident that this improvement proceeds without being distracted by efforts to improve the scores—to get more ‘tings’—and to secure the resulting incentives, that winning is not at the expense of addressing issues not subject to measurement but of greater importance to a patient.⁵

The impact of measurement on the nature of clinical care and on clinicians has received limited attention.⁶ What effect does securing ‘tings’ and winning prizes have on the soul of the contestants? To have caring turned into the unending pressing of levers for sweet rewards, rewards that barely reflect evident improvements in patient well-being. To check HbA1c levels until, by chance, one falls in the rewards area (ting!), or, if not, to add another antihyperglycaemic agent, even in patients for whom this is likely to be harmful or futile.⁷ To prescribe a statin (ting!) to a patient ‘just in case you change your mind’ after the patient reasons that their risk is too low to justify one. To simply send a decision aid through the



CrossMark

To cite: Kunneman M, Montori VM, Shah ND. *BMJ Qual Saf* 2017;**26**:849–851.

patient portal (ting!), and assume that it will help a patient participate in decisions about their healthcare. We are training health professional and organisations to respond to quality gaps others value and to respond in a manner that ensures the rewards.⁸ This is not the smartest healthcare. This is not care.

Quality must always improve and quality improvement is a professional obligation and a healthcare competency. It was important to the early efforts to improve quality to reliably measure and lower HbA1c levels, to demonstrate that high-risk patients were less likely to receive statins, and to describe patients who receive unnecessary care or care that is inconsistent with their informed preferences. Like those pioneering efforts, improving care must result from our commitment to careful and kind patient care. We should continue to use measures, like an accurate map, to know we are moving north. But we must also improve quality where the terrain is complicated or when only inaccurate maps exist. In these cases, we should be able to imagine improvements in care that reflect and respond not to incentives, but to what is pertinent and important to each patient's situation, particularly when this cannot be quantified.

It is difficult to improve care without the risk of falling into the kinds of decontextualised measures and incentives that erode rather than promote better care. Personal, social and comorbid complexities force increasing levels of sophisticated tailoring of clinical action, a simultaneous attention to biology and to biography. Today, that attention often places clinicians in a soul-killing bind to either be patient-centred or to gain quality points and rewards. This conflict must be avoided; quality programmes must reward deep individualisation,⁹ variability as signal rather than noise. McGlynn and colleagues, for example, proposed a measurement system that embraces such variation, measuring quality as the extent to which patients' needs are met and evidence is used to do so.¹⁰ To date, measurement has been almost exclusively numerical. Numbers have undeniably desirable practical advantages—for example, easy to extract from medical or billing records, enabling statistical adjustments and comparisons with minimal-to-no subjectivity. Yet the evaluation of care in context is fundamentally a subjective judgement best aided by descriptive 'a-metrical' assessments. The need to involve human judgement in making meaning of the descriptions, of the words in these descriptions, as Iona Heath stated, 'is perhaps the one attribute of words that make them so peculiarly appropriate for judging quality within healthcare'.¹¹

Our central point, however, is not about the nature of quality measures, but rather about the perverse effect of being rewarded by obtaining a desirable result on a metric without demonstrating a true improvement in the quality of care. Narratives that make judgements about quality may inhibit the simple paths connecting

acing the measure and receiving the 'ting' without actual improvements. Deploying other measures that triangulate observations about the underlying care patients receive may also serve as balancing procedures (akin to balancing measures) to uncover when 'measurement with a wink' is affecting a quality programme. Examples of potentially helpful balancing procedures include accounting for the number of HbA1c measures in a period of time (as clinicians are trying to get under a quality threshold), or obtaining nurse or patient accounts of how statin prescriptions are prepared or dispensed (as clinicians try to ace statin use metrics by prescribing stains without shared decision making). These balancing procedures, like their linked improvement programmes, in their complexity, should not crowd out care. As such, before widespread adoption, measurement and improvement programmes should be tested for safety and efficacy and for minimally disrupting care.

Measurement misleads when it fails to award the best score to the best care, but instead rewards the best player. High scores on the measures provide the illusion of good, better or improved care, while favouring measurable care that is predominantly standard, technical, mechanical and context-blind.¹² In the TV show, key words can be hit without context, in perfect context or in the wrong context, all leading to the same scores. Playing this game can help healthcare organisations and professionals learn the basics of improvement science, but mastering this level is not the endgame. We can get it completely right, ace it and yet be completely wrong. The best quality programmes—deploying multidimensional assessments that closely track with underlying improvements and are linked to intrinsically rewarding care experiences and better patient outcomes—may be immune to this concern. Focusing on aching the test instead of improving the underlying care that should be careful and kind dehumanises care and de-professionalises. It is time to move beyond measurement with a wink.

Contributors MK came up with the idea for this article and wrote the first draft; discussion with VMM and NDS led to revisions to include specific examples and discussions. All authors approved the final version.

Competing interests None declared.

Provenance and peer review Not commissioned; internally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2017. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- 1 Berenson RA. If you Can't measure Performance, can you improve it? *JAMA* 2016;315:645–6.
- 2 Sinsky C, Colligan L, Li L, *et al.* Allocation of physician time in ambulatory practice: a Time and Motion Study in 4 specialties. *Ann Intern Med* 2016;165:753–60.
- 3 Casalino LP, Gans D, Weber R, *et al.* US Physician Practices spend more than \$15.4 Billion annually to Report Quality measures. *Health Aff* 2016;35:401–6.
- 4 Chauhan BF, Jeyaraman M, Mann AS, *et al.* Behavior change interventions and policies influencing primary healthcare professionals' practice-an overview of reviews. *Implement Sci* 2017;12:3.
- 5 Kronick R, Casalino LP, Bindman AB. Introduction. Apple Pickers or Federal Judges: Strong versus Weak Incentives in Physician Payment. *Health Serv Res* 2015;50:2049–56.
- 6 Flodgren G, Eccles MP, Shepperd S, *et al.* An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Syst Rev* 2011;7:CD009255.
- 7 McCoy RG, Lipska KJ, Yao X, *et al.* Intensive treatment and severe hypoglycemia among adults with type 2 Diabetes. *JAMA Intern Med* 2016;176:969–78.
- 8 Lester H, Schmittiel J, Selby J, *et al.* The impact of removing financial incentives from clinical quality indicators: longitudinal analysis of four Kaiser Permanente indicators. *BMJ* 2010;340:c1898.
- 9 <https://innovation.cms.gov/initiatives/Million-Hearts-CVDRRM/> (accessed June 2016).
- 10 McGlynn EA, Schneider EC, Kerr EA. Reimagining quality measurement. *N Engl J Med* 2014;371:2150–3.
- 11 Heath I. “Arm in arm with righteousness”. *Philos Ethics Humanit Med* 2015;10:7.
- 12 Menninger WW. “Caring” as part of health care quality. *JAMA* 1975;234:836–7.