

Unpacking quality indicators: how much do they reflect differences in the quality of care?

Jill Tinmouth

Department of Medicine,
Division of Gastroenterology,
Sunnybrook Health Sciences
Centre, 2075 Bayview Avenue,
Room HG40, Toronto, ON M4N
3M5, Canada

Correspondence to

Dr Jill Tinmouth, Sunnybrook
Health Sciences Centre, 2075
Bayview Avenue, Room HG40,
Toronto, ON M4N 3M5, Canada;
Jill.Tinmouth@sunnybrook.ca

Accepted 8 September 2017
Published Online First
26 September 2017

Just over 50 years ago, Avedis Donabedian published his seminal paper, which sought to define and specify the ‘quality of health care’, articulating the now paradigmatic triad of structure, process and outcome for measuring healthcare quality.¹ In recent years, we have seen the rapid expansion of increasingly inexpensive information technology capability and capacity, facilitating the collection and analysis of large healthcare data sets. These technological advances fuel the current proliferation of performance measurement in healthcare.² Increasingly, in an effort to improve care, many cancer health systems, including those in England,³ the USA⁴ and Canada,^{5,6} are publicly reporting performance indicators, generally derived from these large data sets. Not surprisingly, differences in prevention, early detection and/or treatment of cancer are often used to explain the observed differences in performance across jurisdictions.^{6–9}

Given the considerable effort and resource invested in performance measurement as well as potential adverse consequences if done poorly,¹⁰ it is important to get it right. Determining the effectiveness of healthcare performance measurement is challenging,¹¹ particularly at the health system level. Often, performance measurement is implemented uniformly across an entire system, making well-designed controlled analysis less feasible or impossible^{12,13} and leaving evaluations vulnerable to secular trends.¹⁴ At the physician level, audit and feedback studies report variable results: meta-analyses show a modest benefit overall,^{15–17} but an important proportion of interventions were ineffective or minimally effective with a few studies suggesting a negative effect on performance.¹⁶ Likely, this heterogeneity is due to the complexity of

the endeavour and its many moving parts, which include the behaviour targeted, the recipients of the feedback, their environment, the use of cointerventions and the components of the audit and feedback intervention itself.¹⁸ The latter generally comprises performance indicators, often derived from large healthcare data sets; however, who reports these indicators and how they are reported are also critical components of audit and feedback.^{15,16,19}

In this issue of the *BMJ Quality & Safety*, Abel *et al*²⁰ illustrate empirically some of the complexities of measuring performance. They examine the properties of 16 primary care ‘diagnostic activity indicators’ related to cancer (ie, performance indicators for diagnosing cancer) among 7000+ large general practices in England. Performance indicators are used to assess the quality of healthcare delivery, sometimes by examining outcomes such as survival or recovery of function and at other times, measuring processes, that is, the intermediate steps that have been shown or are felt to be important in terms of achieving the desirable outcomes. In this interesting analysis, the authors challenge the assumption that variation in indicators necessarily reflects underlying differences in the quality of cancer care. Using mixed models, they parse the aetiologies of the observed variation in the cancer diagnostic activity indicators across the practices in their cohort. They then assessed the reliability or ‘rankability’ (whether a practice can be meaningfully distinguished from others using a particular indicator) of the indicators.

This study used data from the Cancer Service Public Health Profile,³ which is one of 30+ thematic National Public Health Profiles published in an interactive web format by Public Health England. The Cancer Service Profile reports cancer



► <http://dx.doi.org/10.1136/bmjqs-2017-006607>



To cite: Tinmouth J.
BMJ Qual Saf 2018;**27**:4–6.

services-related indicators at the general practice level for practices of at least 1000 patients as well as at the health region level. The Profile is intended to assist health system planners to make decisions about services and to stimulate 'reflective practice' among providers. While use of these data in this way is supported by links between some of the indicators and outcomes such as cancer survival,^{21 22} there have been concerns regarding small sample sizes and the impact of underlying case mix as causes of the observed variation.²³ This is an increasingly familiar refrain in the 'big data' era—similar concerns have been articulated in other contexts and health systems.^{2 11}

In their study,²⁰ Abel and colleagues found that an important proportion of the observed variation in indicators across practices is related to factors other than quality of healthcare. Depending on the indicator, chance alone accounted for 7%–85% of the observed differences in practice. They then examined the role of case mix, finding that age and sex differences explained an additional 5%–75% of the observed variation across practices beyond the role of chance. Chance played a larger role for indicators that were identified a priori as outcome indicators than for those considered process indicators. Outcome indicators also tended to be less rankable than process indicators. As the authors point out, the findings for the outcome indicators may have resulted from smaller sample sizes. However, these findings might also be expected on the basis of fundamental differences in the nature of these two types of indicators. As has been noted,^{1 24 25} clinicians and the health system influence healthcare processes to a greater extent than they do patient outcomes, as the latter are subject to factors other than the quality of medical care (and this may be reflected as chance). Put simply, the hospital that a patient attends has a greater impact on the tests and treatments they receive than on whether they live or die.

Touted by some as more sensitive and less vulnerable to differences in case mix,²⁴ the paper²⁰ by Abel *et al* demonstrates empirically that process measures can also be affected by chance and case mix, although to a lesser extent than outcome measures. This finding contributes to what is already known about the limitations of process measures,^{26–29} including the need for evidence linking processes to outcomes, challenges introduced when defining eligible patient populations and lack of comprehensiveness. In the case of the Cancer Service Profile used by Abel *et al*, only a few of the process indicators in the Profile have been shown to be linked to outcomes such as survival.²⁰ Variation in process indicators, unrelated to underlying quality of care, can be introduced during measurement if exclusions are applied variably when defining eligible populations.²⁹ Finally, many existing process indicators are limited to measuring a specific step in the care process; as such, they may not comprehensively reflect the full

process that is tied to outcomes of interest.²⁸ This issue is illustrated in the study by Abel *et al*, where process indicators are limited to practice screening coverage, rates of endoscopy, and referral, and as such do not reflect all the care steps leading to diagnosis of or survival from cancer.

Abel and colleagues' work does have some limitations. Although they were able to quantify the proportion of the variation that could be attributed to case mix, they used age and sex only. There are other aspects of case mix, such as comorbidity and disease severity, which were not included and could also be a cause of the observed variation across practices. For instance, clinicians might understandably not pursue a new cancer diagnosis as assiduously in a patient who already has advanced medical conditions than a patient of the same age and gender who has no serious medical problems. In addition, social determinants of health could also contribute to the observed variations in care. Although the analysis did not account for these factors, the underlying message would remain the same if it did: much of the observed variation in Cancer Service Profile performance indicators (both outcomes and processes) is unrelated to the underlying quality of healthcare.

A previous editorial³⁰ on hospital standardised mortality ratios suggested that we should consider appending cautionary notes to performance indicators, just as we publish warnings for patients on medication labels. Inappropriate use or interpretation of performance indicators might be considered less worrisome than some medication side effects; however, as pointed out in the editorial, resources could be directed towards non-existing problems and complacency induced among providers or hospitals that do have problems. The study by Abel *et al*²⁰ adds to the list of caveats needed and helps us understand the properties of different types of indicators in a nuanced way. In general, process indicators for diagnostic performance can be used to reliably discern between primary care practices while outcome indicators cannot. An important reason for this is that chance plays too great a role in the variation we see in outcome indicators. While chance plays this role more commonly with outcome indicators, it is not always the rule—chance variation can be an issue for any type performance indicator in primary care if the number of cases per practice remains relatively small.

As Abel and colleagues²⁰ suggest, further empirical work may address some of the specific shortcomings of these indicators. No doubt some of what we call chance variation today will likely turn out to involve identifiable characteristics of patients, providers or health systems. But, it will probably also remain the case that, just as Donabedian proposed some 50 years ago, no single type of indicator will capture the entirety

of the quality of care. We will always need a balanced set of indicators involving structure, processes and outcomes.

Competing interests None declared.

Provenance and peer review Commissioned; internally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

REFERENCES

- Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966;44:166–206.
- Smith PC, Mossialos E, Papanicolas I. Performance measurement for health system improvement: experiences, challenges and prospects. *Copenhagen* 2008.
- Public Health England. Cancer Service Public Health Profile. 2016 <https://fingertips.phe.org.uk/profile/cancerservices>
- Ryerson AB, Massetti GM. CDC's Public health surveillance of cancer. *Prev Chronic Dis* 2017;14:160480.
- Cancer Quality Council of Ontario. Cancer system quality index (CSQI). 2016 <http://www.csqi.on.ca/>
- Rahal R, Chadder J, DeCaria K, *et al.* How different is cancer control across Canada? Comparing performance indicators for prevention, screening, diagnosis, and treatment. *Curr Oncol* 2017;24:124–8.
- Allemani C, Weir HK, Carreira H, *et al.* Global surveillance of cancer survival 1995-2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). *Lancet* 2015;385:977–1010.
- Coleman MP, Forman D, Bryant H, *et al.* Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the international cancer benchmarking partnership): an analysis of population-based cancer registry data. *Lancet* 2011;377:127–38.
- De Angelis R, Sant M, Coleman MP, *et al.* Cancer survival in Europe 1999-2007 by country and age: results of EURO-CARE-5-a population-based study. *Lancet Oncol* 2014;15:23–34.
- Shahian DM, Normand SL, Friedberg MW, *et al.* Rating the raters: the inconsistent quality of health care performance measurement. *Ann Surg* 2016;264:36–8.
- Shahian DM, Normand SL. What is a performance outlier? *BMJ Qual Saf* 2015;24:95–9.
- Ellegård LM, Dietrichson J, Anell A. Can pay-for-performance to primary care providers stimulate appropriate use of antibiotics? *Health Econ* 2017:1–16.
- Craig P, Dieppe P, Macintyre S, *et al.* Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;337:a1655.
- Chen YF, Hemming K, Stevens AJ, *et al.* Secular trends and evaluation of complex interventions: the rising tide phenomenon. *BMJ Qual Saf* 2016;25:303–10.
- Hysong SJ, Best RG, Pugh JA. Audit and feedback and clinical practice guideline adherence: making feedback actionable. *Implement Sci* 2006;1:5–3.
- Ivers N, Jamtvedt G, Flottorp S, *et al.* Audit and feedback: effects on professional practice and healthcare outcomes *Cochrane Database Syst Rev*. 2012:6.
- Jamtvedt G, Young JM, Kristoffersen DT, *et al.* Audit and feedback: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2006:CD000259.
- Ivers NM, Sales A, Colquhoun H, *et al.* No more 'business as usual' with audit and feedback interventions: towards an agenda for a reinvented intervention. *Implement Sci* 2014;9:14.
- van der Veer SN, de Keizer NF, Ravelli AC, *et al.* Improving quality of care. A systematic review on how medical registries provide information feedback to health care providers. *Int J Med Inform* 2010;79:305–23.
- Abel G, Saunders CL, Mendonca SC, *et al.* Variation and statistical reliability of publicly reported primary care diagnostic activity indicators for cancer: a cross-sectional ecological study of routine data. *BMJ Qual Saf* 2018;27:21–30.
- Møller H, Gildea C, Meechan D, *et al.* Use of the English urgent referral pathway for suspected cancer and mortality in patients with cancer: cohort study. *BMJ* 2015;351:h5102.
- Shawihdi M, Thompson E, Kapoor N, *et al.* Variation in gastroscopy rate in English general practice and outcome for oesophagogastric cancer: retrospective analysis of Hospital Episode Statistics. *Gut* 2014;63:250–61.
- Murchie P, Chowdhury A, Smith S, *et al.* General practice performance in referral for suspected cancer: influence of number of cases and case-mix on publicly reported data. *Br J Cancer* 2015;112:1791–8.
- Mant J, Hicks N. Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *BMJ* 1995;311:793–6.
- Green J, Wintfeld N, Sharkey P, *et al.* The importance of severity of illness in assessing hospital mortality. *JAMA* 1990;263:241–6.
- Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? *J Am Coll Surg* 2004;198:626–32.
- Mant J. Process versus outcome indicators in the assessment of quality of health care. *Int J Qual Health Care* 2001;13:475–80.
- Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. *Int J Qual Health Care* 2001;13:469–74.
- Bruckel J, Liu X, Hohmann SF, *et al.* The denominator problem: national hospital quality measures for acute myocardial infarction. *BMJ Qual Saf* 2017;26:189–99.
- Marang-van de Mheen PJ, Shojania KG. Simpson's paradox: how performance measurement can fail even with perfect risk adjustment. *BMJ Qual Saf* 2014;23:701–5.