



The Holy Grail of Quality Measurement

Richard Lilford, ARC WM Director

Writing in JAMA, Austin and Kachalia argue for automation of quality measurements.[1] We ourselves have argued that the proliferation of routine quality measures is getting out of hand.[2]

The authors argue, as we have argued, that using quality measures to incentivise organisations is a blunt tool, subject to gaming. Far better, is to use quality measures in real-time to prompt doctors to provide high quality care.

In fact, this is what computerised decision support offers. There is considerable empirical support for use of this type of decision tool. Working with Prof Aziz Sheikh and colleagues NIHR ARC West Midlands has investigated decision support for prescribing [3] and we are now investigating its use in antibiotic stewardship.[4] We are entirely in support of the use of decision support to improve care in real-time.

However, we question the idea that the majority of healthcare can be guided by online decision support. Working with Prof Timothy Hofer in Michigan, ARC WM co-investigators have shown that the measurement of the quality of hospital care is extremely unreliable.[5] Kappa measures of agreement between reviewers were about 20%. This means that seven reviewers would be needed for each case note, to achieve a reliability of 80%.

That is to say, that for much of medical care, there is no agreed standard. Truly, the majority of medical care is more art than science.

We think that the time has arrived to abandon hubristic notions about standardising and quality assuring the generality of clinical care. Medicine is not like aviation. Commercial aviation is almost entirely computerised. Emergencies aside, the whole process can be guided algorithmically. Our paper in *Milbank Quarterly*, shows quite clearly that this is not the case for medicine.[5]

Working with Prof Julian Bion, the ARC WM Director had an opportunity to audit numerous case notes from patients with sepsis.[6] The idea was to observe quality of care against a package of evidence-based criteria. Many of these criteria was based on actions that should be carried out within a specified time from diagnosis. The exercise proved almost impossible, since the point of diagnosis was ephemeral. In most cases there was no clear point to start the clock and the very diagnosis of sepsis had to be reverse-engineered from the time at which a sepsis-

associated action took place! This exercise provided eloquent testimony to the judgemental, rather than rules-based, nature of much medical practice. We should use algorithmic decision support where clear rules exist, but we must stop pretending that the whole of medicine can be guided in this way. Perhaps we should just stand back a little, and accept some of the imperfections in our systems. Like a harm-free world, perfection will always lie beyond our grasp.[7]

References:

1. Austin JM, Kachalia A. The State of Health Care Quality Measurement in the Era of COVID-19. The Importance of Doing Better. *JAMA.* 2020.
2. Lilford RJ. Measuring Quality of Care. NIHR CLAHRC West Midlands News Blog. 21 April 2017.
3. Yao GL, Novielli N, Manaseki-Holland S, et al. Evaluation of a predevelopment service delivery intervention: an application to improve clinical handovers. *BMJ Qual Saf.* 2012;**21**:i29-38.
4. Usher Institute. ePrescribing-Based Antimicrobial Stewardship. 2020.
5. Manaseki-Holland S, Lilford RJ, Te AP, et al. Ranking Hospitals Based on Preventable Hospital Death Rates: A Systematic Review With Implications for Both Direct Measurement and Indirect Measurement Through Standardized Mortality Rates. *Milbank Q.* 2019;**97**(1):228-84.
6. Lord JM, Midwinter MJ, Chen YF, et al. The systemic immune response to trauma: an overview of pathophysiology and treatment. *Lancet.* 2014;**384**(9952):1455-65.
7. Meddings J, Saint S, Lilford RJ, Hofer TP. Targeting Zero Harm: A Stretch Goal That Risks Breaking the Spring. *NEJM Catal Innov Care Del.* 2020; **1**(4).

ARC WM Quiz

Born on 24 July 1914, Dr Frances Oldham Kelsey famously refused to authorise which drug for market in the USA due to concerns about its safety?

email your answer to:
ARCWM@warwick.ac.uk

Answer to previous quiz: The wet bulb temperature is that measured by a thermometer covered in water-soaked cloth over which air is passed.

Congratulations to those who answered correctly.

