# Measure to Improve, Not to Rank: A discussion paper on including quality metrics on the Medical Costs Finder

Trevor M Sutton PSM – Python Hill Advisors

June 2026

Avant
By doctors for doctors

# Contents

# 1. Executive Summary

## Purpose of this report

The *Health Legislation Amendment (Improving Choice and Transparency for Private Health Consumers) Bill* 2026 has signalled that national datasets on the quality of individual practitioners may, in future, complement the cost information already published on the Medical Costs Finder (MCF) website.

Using the MCF website to publish information about 'quality outcomes' of individual Australian doctors is not, on the evidence available today, a practical or appropriate way to guide either referrals from general practitioners or patients' choice of specialist. It is not even a way to improve patient outcomes and it would risk doing more harm than the problem it seeks to solve.

While this report assesses whether that step is even feasible, let alone appropriate, it's also important to highlight that metrics can be beneficial in all manner of health care settings. What matters is the context in which they are used rather than the metric itself. Metrics such as complication, infection and readmission rates are valuable tools for improving care. The problem is publishing them against named individuals rewards favourable-looking numbers, which, in turn, discourages clinicians from taking on complex patients, and would potentially steer GP referrals toward the top of a statistically fragile table. What can fairly be shown publicly is not a performance score, but an indicator that a specialist is engaged in the quality-improvement programs through which good practice is sustained.

## Approach/methodology

The key approaches to support this assessment included:

- Consultation and engagement with the Department of Health, Disability and Ageing (DoHDA) and with non-government bodies including specialist groups.
- Desktop review of publicly available documents and statements from a range of organisations including the Consumers Health Forum of Australia (CHF), the Australian Medical Association (AMA), the Grattan Institute, the Australian Commission on Safety and Quality in Health Care (ACSQHC), some clinical registries and their leaders, and various other relevant sources.
- Review of relevant international experience, including the United States and United Kingdom.

## Key findings

1. **Quality is already assured through many mechanisms, not just a single score.**

   Specialist quality in Australia is already protected by many overlapping assurance mechanisms designed to detect adverse performance over time. A public, individualised, ranking layer would add little value to existing assurance mechanisms and pose a serious risk to patient health and wellbeing.

2. **Measurement that improves outcomes already works.**

   Clinical quality registries, as exemplified by the Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR), improve outcomes because they are confidential, peer reviewed, risk adjusted and protected by qualified privilege to encourage practitioners to participate and share their data.

3. **Measures in contemporary use are not robust enough for individual comparison in a public setting.**

   Many indicators of quality today are not sufficiently valid, risk adjusted or statistically stable to compare individuals fairly in publicly-facing settings.

4. **Crude metrics can cause harm rather than improve quality.**

   Unintended potential consequences include risk avoidance, as clinicians decline complex or high-risk patients to protect their published scores; patient delays as patients avoid affordable care wrongly presented as inferior and postpone treatment as a result; erosion of participation in the voluntary registries that currently improve care; and gaming, as effort shifts from improving outcomes toward managing how they are measured.

## Recommendation

There is a substantive opportunity to invest in the maturity and reach of the clinical quality registries in order to improve outcomes rather than just acting as a reporting mechanism. Such a system would encourage specialist participation by protecting their disclosures with qualified privilege, being led by specialists inputting their data.

It is therefore recommended that State, Territory and Federal health ministers prioritise and provide ongoing funding and support for the development and sustainability of Clinical Quality Registries (CQRs) within the Australian Framework for National Clinical quality registries under the National Strategy for Clinical Quality Registries and Virtual Registries 2020-2030.

Once nationally consistent and supported, the MCF website could then display a simple indicator of whether a specialist participates in a nationally consistent quality improvement program. Such an indicator would provide patients with meaningful contextual information on specialist quality without creating the risks outlined in this report that come from measuring individual clinician performance on a public website. The illustrative indicator below is provided as an example of what it could look like on the MCF website.

**In context, on a listing:**
**Dr David Smith**
Orthopaedic Surgeon – Sydney NSW

✓ **Quality Improvement Program**
Participants in a recognised program

# 2. Specialist fee reform extends quietly from price to quality

The Australian Government's MCF website was launched in December 2019 with the aim of assisting patients compare typical fees and out-of-pocket costs for private medical specialist services. Despite significant consultation with key stakeholders prior to its introduction, uptake by specialists was poor, prompting post-implementations reviews and further stakeholder consultations.

The *Health Legislation Amendment (Improving Choice and Transparency for Private Health Consumers) Bill,* introduced in February 2026, sought to mandate the inclusion of specialist cost data on the website. The proposed mandating of data inclusion marks a major policy shift in two key respects. Firstly, the website will no longer depend on voluntary uploading of cost information by specialists. Secondly – and the focus of this report – its Explanatory Memorandum signals that national datasets relating to the quality of individual practitioners may in future serve as a "valuable complement" to cost information

When considering the legislative amendments some groups – including patient advocates, health economists, and policy academics – support pairing quality metrics with out-of-pocket costs data, arguing that price alone can mislead, regardless of whether those prices are high or low.

Other groups agree that, while patient outcomes matter, repurposing quality measures intended for improving patient care and system safety to publicly identify and judge individual specialist's standard of care is not only potentially misleading but carries risk. Such outcomes data are inherently difficult to adjust for complexity, meaning that raw comparisons may punish clinicians dealing with a more complex case mix.

## The question this report answers

Can patient and referrer-oriented "simple metrics" reliably and accurately capture the quality of care provided by an individual medical specialist, such that an average consumer could make a genuinely informed decision from them if they were published on the MCF?

The approach taken, the bodies consulted and the sources reviewed are set out in Appendix 1.

# 3. Quality is already assured through many overlapping mechanisms, not just a single score

The Australian Commission on Safety and Quality in Health Care (ACSQHC) and the Australian Institute of Health and Welfare (AIHW) adopt a standard definition for quality in healthcare - the degree to which health services increase the likelihood of desired outcomes and are consistent with current professional knowledge. The definition itself has been adapted from the US Institute of Medicine and the World Health Organisation.

Quality is captured through a range of complementary measures. Clinical outcomes such as mortality, surgical complication, readmission and infection rates; compliance with evidence-based clinical guidelines; patient-reported outcome and experience measures (PROMs and PREMs); access and waiting times; procedure volumes; and registry-derived outcomes. Each answers a different question, and none is a complete measure of quality on its own.

The more important structural point is that assurance in Australian specialist care is distributed, not singular. No single organisation assures quality, given it emerges from multiple overlapping mechanisms, summarised in Exhibit 1.

## Exhibit 1: Quality assurance in Australian healthcare system

| Line of Assurance | Principal mechanisms | Why it provides assurance |
|---|---|---|
| **Patient-reported outcomes** | PROMs, PREMs and value-based measurement | Captures the outcomes and experience that matter to patients, from the patient's own perspective |
| **Clinical quality registries** | Confidential benchmarking, outlier detection, variation analysis and peer review across specialists | Hold the richest clinical detail at the specialist level and drive improvement through confidential benchmarking, not public individual league tables |
| **Care co-ordination** | GP referral and care co-ordination as a judgement-based quality filter in private specialist care | GPs act as expert gatekeepers, drawing on years of observed outcomes, communication and patient feedback – an influential quality filter in private specialist care |
| **Hospital governance & quality systems** | Morbidity and mortality meetings, audit, credentialing, peer review, accreditation, theatre access, clinical indicator programs | Embeds continuous, real-time scrutiny of practice at the point of care |
| **Professional standards & training** | Specialist colleges and societies - training, continuing professional development, guidelines; Medical Board registration standards and guidelines including the professional performance framework | Sets and maintains the standard of practice across a specialist's whole career |
| **Regulation, accreditation & national frameworks** | Ahpra, Medical Boards, ACSQHC (NSQHS Standards), accrediting agencies (e.g. ACHS Clinical Indicators Program), Professional Services Review | Statutory enforcement of compliance with standards; complaints management and regulation; accreditation enforces standards system-wide |
| **Policy, funding & system oversight** | DoHDA, States and Territories, AIHW, ABS, private health and indemnity insurers | Detects fraud and non-compliance; indemnity insurers monitor safety and reduce risk; national bodies measure and report system performance |

These assurance mechanisms are largely reinforcing and complementary. Together they create the reasonable expectation that the majority of poor performance will be identified and addressed, and that the quality and safety of care will improve over time. Against this backdrop, a public layer that ranks named individuals would add little assurance the system does not already provide, while introducing significant new risks. It is also important to distinguish system- and hospital-level public reporting, which is well established and broadly defensible, from the public ranking of individual specialists which is what some stakeholders are currently canvassing.

Public reporting is complicated by varying levels of health literacy which will impact patients' ability to adequately assess the quality of care of an individual specialist, based on a single score.

# 4. Measurement that improves outcomes works

Of all the assurance mechanisms, clinical quality registries hold the richest clinical and outcome detail at the specialist level. A registry collects standardised clinical and patient-outcome data for a defined condition, procedure or device, and compares performance across individual specialists, healthcare teams, clinical units and hospitals over time. Models range from hospital-based registries to community and outpatient registries and device or procedural registries, and they are operated by specialist colleges, universities and other government-supported agencies.

The Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR) is regarded internationally as a leading example and is among the most successful quality-improvement initiatives in Australian healthcare. It has identified poorly performing implants, reduced revision rates, benchmarked surgeons and hospitals, and monitored long-term patient outcomes. Its clearest success measure is the fall in the "revision burden". This has resulted in fewer patients needing repeat joint-replacement surgery because of implant failure or complications, and a recent economic evaluation found the registry to be highly cost-effective (de Steiger et al, 2013; Babazadeh et al, 2022).

The model is not confined to orthopaedics. Other Australian registries have driven improvements in trauma care, intensive-care outcomes, stroke management, cardiac intervention and hip-fracture care. Many less-mature registries still have clear room to improve which is precisely why the approach represents such a significant opportunity: with further investment and support, the gains already demonstrated in orthopaedics could be extended across far more of specialist care.

The ACSQHC is the national policy lead, maintaining both the register of recognised registries and the framework for how they should operate (ACSQHC, 2024). The framework includes guidance on how to identify the causes of variation and outlier performance, allowing clinicians to recognise better practice or correct sub-optimal practice. The DoHDA currently funds sixteen registries, and the Australian Medical Association (AMA), in its 2019 submission on a national strategy for clinical quality registries, supported broadening both their range and participation in them.

Importantly, registries publish openly on outcomes such as implant and device performance, surgical techniques, and de-identified aggregate outcomes.  These data have driven change (AOANJRR, 2014). The registry also provides feedback to the individual clinician under legal protections such as qualified privilege so that errors, complications and opportunities for improvement can be discussed frankly without fear the discussion will later be used in litigation. These design features, recognised across stakeholders and reinforced in the ACSQHC's framework, are what make registries work, and they are what public, individual-level reporting on a consumer website would remove, as set out in Exhibit 2.

### Exhibit 2. Impacts of public individual reporting on the effectiveness of registries.

| What makes a registry effective | What public-level individual reporting does to it |
|---|---|
| Specialist leadership of measure design → clinically meaningful indicators | Externally imposed indicators, some seen as clinically inappropriate |
| Voluntary participation built on trust | Participation becomes exposure; engagement and data quality fall |
| Legal protections that encourage candid reporting (qualified privilege) | Legal exposure discourages reporting and open discussion of problems |
| Robust risk adjustment for fair comparison | Raw comparison penalises those treating complex cases |
| Two-stage, time-tested handling of outliers | Snapshot comparison (without robust review) mistakes natural variation for poor care |
| Each clinician's own results fed back against benchmarks | Public ranking shifts focus to reputation, not improvement |

# 5. Crude metrics can create patient harm and cannot support fair public comparison of individual specialists

The great majority of stakeholders contacted during the compilation of this report agreed that quality metrics can improve patient outcomes. The questions they raised were whether the measures available today are sufficiently valid, reliable, risk-adjusted, clinically meaningful and statistically robust to be displayed publicly at the individual specialist level adjacent to fee information. The concerns raised in consultation fall into four groups.

## No major stakeholder has advocated ranking individuals

It is significant that no specialist organisation, Australian patient body, government agency or mainstream health-policy group has formally advocated for a public website that ranks individual specialists from best to worst. Many advocates of greater transparency are careful to avoid that model and for good reason. The Consumer Health Forum argues for better information on quality and outcomes but not for public, specialist-level outcome rankings. The ACSQHC strongly supports the use of registries that measure outcomes, PROMs, PREMs but to drive benchmarking, reduce unwarranted variation and support evidence-based reporting, not to publish league tables. The Grattan Institute and health economists press for better quality as well as price information yet caution against clinical-level rankings because risk adjustment is difficult, case mix between clinicians varies and small samples mislead. Registries already produce clinical-level benchmarking and outlier detection which remains confidential and is used for quality improvement and not consumer shopping. There is a substantial difference between giving patients some quality information and presenting them with a public ranking of individual specialists.

## Methodological validity

Patient outcomes are difficult to risk-adjust fairly for case mix, co-morbidity, geography, socio-economic status and demographics, all of which independently affect results. Several structural problems compound this. Outcomes are often team-based rather than attributable to one specialist. For example, in hip-fracture and stroke care, the result reflects a whole team's effort. They are also frequently shared across multiple clinicians over the course of a patient's care, with patients often seeing several GPs or, in the case of mental health, moving between clinicians over many years. This makes meaningful attribution of outcomes to any single practitioner inherently unreliable. Many specialists perform a range of services that resist a single indicator.  A typical example is that of an obstetrician-gynaecologist who may provide genetic counselling, IVF treatment, antenatal care and delivery. Comparing across different specialties is harder still, since a psychiatrist and a procedural specialist require entirely different measures, and heterogeneity within a specialty, such as the type of psychiatrist, the complexity of the patient, means appropriate treatment varies widely.

Frequency is an issue that compounds the problems. Annual procedure volumes for many specialists are low, making comparisons statistically unstable. Some of the most clinically important events are simply too rare to serve as indicators. In gastroenterology, post-polypectomy haemorrhage, the key adverse outcome of colonoscopy, has an average incidence of about 0.08%. Factors outside the specialist's control intrude as well. For example, where post-operative physiotherapy is needed but unaffordable or unavailable, the outcome suffers for reasons unrelated to the specialist. International work reinforces the caution and shows low procedure numbers can produce false complacency (Walker et al, 2013), and reliably detecting genuine individual outliers is difficult (Harrison et al, 2016).

## Perverse incentives

Without robust risk adjustment, public comparison rewards doctors who avoid complex patients - the cream-skimming effect (Werner and Asch, 2005). The risk is not hypothetical. In both the US and the UK, publishing crude, poorly adjusted indicators can produce a misleading picture in which worse outcomes simply reflected more complex patients and drive patients toward services that may not meet their needs, and discouraged specialists from undertaking more complex surgery (Behrendt and Groene, 2016). Referral patterns compound the distortion. GPs acting as co-ordinators of care steer their most complex cases to the specialists they judge best able to handle them. The data penalises the clinicians taking the hardest work. Some measures depend on voluntary participation and can carry selection bias, meaning those who do not participate in a registry or dataset may be the very performers of most concern, which could make a low participation-based score self-undermining as a quality signal.

## Data quality and participation

Stakeholders raised serious concerns about the access, completeness and timeliness of the underlying data. Many datasets cover only an incomplete population of specialists and are therefore likely to be unrepresentative, and some suffer from poor data quality such as incomplete fields. Access to data itself is a problem. Specialist groups reported that obtaining state or hospital data could take more than 12 months, meaning that the data could already be three to four years old by the time it is published. In some cases, data custodians refused access altogether and historical data may not reflect a specialist's current performance. There also are concerns about the clinical appropriateness of indicators developed without specialist input. Procedure volume, for instance, can signal over-servicing or inappropriate use of a service and not necessarily experience. Finally, there is a direct risk publishing registry-derived data would erode the high voluntary participation on which those registries depend, degrading the very datasets a public website would draw upon.

### Limited influence in GP referral pathways

Research, commissioned by Avant (The Navigators, 2026) into how GPs refer patients to specialists confirms that referral already functions as a quality filter and that the people best placed to judge specialist quality are confident in it. Both GPs and patients report high confidence in the quality of specialist care. Among GPs surveyed, 94 per cent were satisfied with the quality of specialists available in their area and 95 per cent were confident in the quality of those they refer to. Their judgements are informed by patient feedback, clinical outcomes, the quality of communication from specialists, and the professional networks and referral experience built up over time.

The same research found little appetite for a government-sponsored specialist quality rating website. GPs were markedly more negative than positive about the concept, and their intended use of it was low. Their reservations mirror the concerns set out in this report including quality is hard to measure fairly, that risk adjustment is difficult, that public ratings create perverse incentives and are easily misinterpreted.

### Direct patient harm

Published scores could raise patient anxiety and prompt patients, particularly in regional areas, to delay or avoid care. For example, a patient who can only afford a lower-priced specialist presented as lower quality due to misleading indicators may put off a needed procedure. Aggregated patient experience metrics are also particularly misleading, because a good experience is not the same as good clinical care. A patient may feel very satisfied yet receive clinically inappropriate treatment, while a clinician with excellent outcomes may score poorly precisely because they deliver difficult news or decline inappropriate requests including unnecessary medications or investigations. Such data is also likely to be unrepresentative, since the most and least satisfied patients are the most likely to respond while outliers must be assessed over a sufficient period, with root-cause analysis, to distinguish a genuine problem from natural variation in the data.

## 6. Public reporting can be safe but only under certain conditions

It would be wrong to conclude that public reporting of clinical outcomes is categorically harmful. A study of colorectal cancer surgeon outcomes in the United Kingdom (Vallance et al, 2018) found no evidence of widespread risk-averse behaviour because the underlying data were properly risk-adjusted. The lesson is public reporting can be safe and useful but only when specific conditions are met.

Those preconditions for success include:

- Outcomes robustly risk-adjusted for case mix and co-morbidity
- Attribution controlled where team-based care or shared over time
- Datasets representative, complete, timely and accurate
- Indicators reliable, unambiguous and not based on rare events and small samples
- Patient behaviour accounted for where it materially affects outcomes
- Specialists engaged in indicator design to ensure clinical meaning and trust.

# 7. Key recommendation

The most productive investment for government, the profession, insurers and Australian patients would be to improve the maturity and reach of clinical quality registries, in line with what the joint replacement registry has achieved and which demonstrably proves how to improve patient outcomes. Properly extended, this approach addresses most of the concerns raised about quality measurement at the source.

To be effective, that investment must protect the conditions that make registries work such as the legal protections that encourage participation and candid reporting (qualified privilege); specialist leadership of indicator design, sustainable funding and custodianship, and a clear sequence in which clinician-level feedback is well established before any form of public reporting is contemplated. Framed this way, the recommendation is not an argument against transparency but a roadmap to the conditions under which meaningful quality information can eventually be produced to create the dual benefit of outcomes-based measurement and public reporting of quality indicators.

It is therefore recommended that State and Federal health ministers prioritise and provide ongoing funding and support for the development and sustainability of Clinical Quality Registries within the Australian Framework for National Clinical quality registries under the National Strategy for Clinical Quality Registries and Virtual Registries 2020-2030.

Once nationally consistent and supported, the MCF website could then display a simple indicator of whether a specialist participates in a nationally consistent quality improvement program. Such an indicator would provide patients with meaningful contextual information on specialist quality without creating the risks outlined in this report that come from measuring individual clinician performance on a public website. The illustrative indicator below is provided as an example of what it could look like on the MCF website.

**In context, on a listing:**
**Dr David Smith**
Orthopaedic Surgeon – Sydney NSW

✓ **Quality Improvement Program**
Participants in a recognised program

## Keeping measurement in proportion

Used well, metrics and indicators deliver real and targeted benefits to patient care. But measurement is a tool for specific jobs, not a remedy for the larger challenges facing healthcare. No indicator, however well designed, will by itself secure the viability of services, the availability of specialists to provide care, the structural reform the system needs, or an easier path for patients through its complexity. These are questions of access and capacity, not measurement. The risk is that a quality indicator on a price website is asked to deliver progress on problems it was never designed to solve. Measurement should be invested in for what it does well and kept in proportion to what it cannot do. The constructive path is not to publish individual quality scores the data cannot yet support, but to strengthen the measurement that already improves care.

# 8. Appendix 1

## Approach

The assessment drew on consultation and engagement with key stakeholders across the sector with a particular focus on the risks and concerns related to measuring and reporting patient safety and quality. This was supported by a desktop review of publicly available documents and statements, and a review of relevant international experience, principally from the US and UK.

## Bodies consulted and sources reviewed

Specialist groups: the Australasian College of Dermatologists; the Royal Australian and New Zealand College of Psychiatrists; Australian Society of Psychiatrists; the Australian Orthopaedic Association; and the Royal Australian and New Zealand College of Obstetricians and Gynaecologists. Government and statutory body: the Department of Health, Disability and Ageing (DoHDA). Public statements reviewed included those of the Consumers Health Forum of Australia, the Australian Medical Association, the Grattan Institute, and clinical quality registry leaders and academics. Informal discussions were held with a range of specialists, including general practitioners.

## A note on balance

Consultation was weighted toward specialist and government bodies, with consumer and insurer positions drawn substantially from public statements rather than interview. This asymmetry is acknowledged. It does not affect the central findings, which rest on the statistical and design properties of the measures themselves, not on stakeholder preference.

## Sources

1. Australian Commission on Safety and Quality in Health Care (ACSQHC), "The state of patient safety and quality in Australian hospitals", 2019.
2. Australian Medical Association, Submission to Department of Health, Disability and Ageing Draft National Clinical Quality Registry Strategy Consultation, 2019.
3. Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR), "Annual Report, Hip and Knee Arthroplasty", 2014.
4. Babazadeh S et al. "Cemented Polished Tapered Stems Have Lower Revision Rates Than Commonly Used Cementless Implant up to 17 Years of Follow-Up: An Analysis of 201,889 Total Hip Replacements From the Australian Orthopaedic Association National Joint Replacement Registry", J Arthroplasty, 37(1), 2022.
5. Behrendt K and Groene O, "Mechanisms and effects of public reporting of surgeon outcomes: A systematic review of the literature", Health Policy, 120(10), 2016.
6. de Steiger RN et al, "Joint Registry Approach for identification of outliers protheses", Orthop, 84(4), 2013.
7. "GP Referral Drivers Survey", research conducted by The Navigators commissioned by Avant, June 2026.
8. Harrison EM et al, "Individual surgeon mortality rates: can outliers be detected? A national utility analysis", BMJ Open, 2016.
9. Vallance AE et al, "Effect of public reporting of surgeons' outcomes on patient selection, "gaming," and mortality in colorectal cancer surgery in England: population based cohort study", BMJ, 361, 2018.
10. Walker K et al, "Public reporting of surgeon outcomes: low numbers of procedures lead to false complacency", Lancet, 382(9905), 2013.
11. Werner RM, Asch DA, "The unintended consequences of publicly reporting quality information", JAMA, 293(10), 2005.

avant.org.au | 1800 128 268

Avant

By doctors *for* doctors