Submission to the Independent Hospital Pricing Authority Discussion Paper on the Pricing Framework for Australian Public Hospital Services 2017–18

Professor Anthony Scott
Associate Professor Jongsay Yong

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Level 5, FBE Building, 111 Barry Street
Parkville, Victoria 3010
Correspondence: a.scott@unimelb.edu.au

Introduction

The aim of this submission is to examine some of the issues surrounding pricing and funding for safety and quality as set out in the consultation document Consultation Paper on the Pricing Framework for Australian Public Hospital Services 2017–18. The emphasis of this submission is on how design and implementation of pricing and funding schemes may impact on the economic incentives for behavioural changes that lead to desirable quality and performance outcomes.

We welcome the move towards linking funding to the provision of value-based health care. There is much to be learned from experience overseas. In addition to issues of design and implementation, there should be ongoing evaluation and methodological work on the evolving design of the scheme. The details on the design and implementation of incentive schemes matter, especially the way incentives are targeted. An evaluation framework should form an integral part of the design of any incentive schemes. Transparency and fairness are critical for risk adjustment to be acceptable to all stakeholders. Lack of socio-economic status and health behaviour information in current data infrastructure is a major gap in developing risk adjustment models. Improving data infrastructure should be a priority, not only for purposes of risk adjustment but also for improving understanding on a range of social and economic issues on health care utilisation and equities. Quality of health care is a multifaceted concept involving multiple dimensions of health care provision. It is important to develop a set of comprehensive quality indicators, including patient reported outcomes measures (PROMs). Focusing on a narrow set of quality indicators can lead providers to only perform to incentivised metrics while ignoring areas of care that are not incentivised.

Summary of recommendations.

1. The size of incentives should <u>not</u> always be related to costs if one wants to move towards a true value-based payment system.

2. Payments/changes in funding should be targeted at the lowest level of disaggregation that the data allow – at least to hospital level.

3. We strongly recommend against using a national benchmark for measuring relative performance. Rewards are for an improvement in quality (HACS or re-admissions) between a baseline period and some future period at the level of the hospital.

4. A carrot-and-stick approach on pricing and funding of hospital services be adopted, rather than focusing on non-payment or payment reduction for low-quality episodes.

5. Incentive payments and penalties be provided at both episode and hospital levels. In this regard, on approaches to reduce HACs, Option 3 of blending a quality-adjusted NEP with funding incentives at hospital level (pp. 46–48) is preferred to Options 1 and 2.

6. Hospital volume should be accounted for when comparing hospital performance. Hospital volume can be included as a hospital characteristic in risk adjustment, or it can be used as the main criterion for classifying peer groups for comparison and benchmarking performance.

7. Perverse incentives need to be anticipated when introducing reduced or non-payments for specific episodes such as episodes with an adverse event. Hospitals may have an incentive to reduce or withdraw care for patients in these episodes ex post if the marginal costs of care are far in excess of the anticipated revenue. Hospitals may, for example, transfer patients to other hospitals or discharge patients earlier than planned and readmit them later on separate episodes. Such actions could contribute to higher costs of care and lower health outcomes for the patients involved.

8. The effects of incentives on coding and gaming need to be considered. Hospitals may be less likely to report sentinel events or HACs if funding is at stake, and so observed reductions may reflect changes in coding. A system of audit should be put in place, or independent data collected and used (eg from registries, or for sentinel events data from AHPRA could be used on medical complaints or litigation).

9. Risk adjustment should be expanded to explore the possibility of supplementing the current administrative data with survey data on health behaviour (smoking, exercise, diet, etc.), and socio-economic status (income, education, employment, etc.).

10. The robustness of different risk adjustment methods should be thoroughly investigated with the view to arrive at models that have good statistical properties and acceptable to all stakeholders.

11. The number of measures of hospital quality should be expanded to include more indicators to guard against hospitals performing to incentivised metrics. Methods to integrate a large number of quality indicators should also be investigated.

12. Further research should be undertaken to examine the relationship between hospital quality and cost efficiency or productivity.

13. We recommend IHPA considers establishing an evaluation from the outset, and to set up and evaluation baseline dataset, and consider carefully the best study design to be used (eg difference in difference or randomised stepped wedge design).

Context

Continuing rise in health care expenditure has prompted governments to focus on improving value, cost efficiency and productivity of the health care system. A range of funding and financial initiatives have been trialled in many countries. These include pay for performance schemes as well as schemes that focus on both quality and cost. In the U.S. many incentive payment models have been in force under the banner of Accountable Care Organisations (ACOs). In the UK schemes such as the Advancing Quality Initiative and Best Practice Tariffs provide financial incentives aiming at improving quality of hospitals.
Systematic reviews of these schemes have found mixed effects (e.g., Scott et al. 2016; Eijkenaar et al. 2013). The consensus is that details on how a scheme is designed and implemented matter, especially on whose incentives are targeted; for example, whether physicians or clinical teams were able to use rewards either as direct income supplement or for quality improvement activities. Another important consideration is how rewards or penalties are set—whether there is a fixed target threshold, or whether performance is measured relative to peers according to the distribution of chosen indicators, or whether payments are tied to performance between two time points and relative to own baseline.

Value-based incentives

Several implementation issues surrounding incentive payment schemes have been highlighted in a recent review of 80 evaluation studies in 10 countries (Scott et al. 2016). Contrary to expectations, the size of the incentives as a percentage of revenue was not associated with the likely effectiveness of a scheme, although it should be noted that the number of schemes was small (22 schemes). A particular issue in the consultation paper is that the reductions in funding across the options were in relation to costs, not value (the proportionality criterion). For example, p43: "In contrast, the national funding impact of not paying for episodes with a HAC would be about $3.1 billion. Such an approach would not meet the proportionality assessment criterion as it would result in hospitals being penalised for costs that are greater than the actual costs of HACs. Accordingly, the next option limits the funding adjustment so that it is, more closely, commensurate with the additional costs of HACs."

We recommend that the size of incentives should <u>not</u> always be related to costs if one wants to move towards a true value-based payment system.

For example, though sentinel events are few in number and cost, one could impose large 'fines' if they occur. Similarly for HACS: do some HACS have larger impacts on patients' health outcomes and quality of life than others? At the moment they are given equal weight: one HAC is the same as another. Perhaps reductions in these should attract larger rewards? If one wants to promote value-based health care, then one could impose penalties even if the impacts on costs are quite small.

Salience of incentives.

There is some evidence that the salience of incentives matters, e.g., the extent to which clinicians or clinical teams were aware of the rewards or were able to influence how the rewards were used (Scott et al. 2016). Evidence suggests that schemes allowing incentive funding to be used for specific (but non-physician income) purposes were more likely to be effective compared to physicians being allowed to use incentive funding as income. It should be pointed out that scant data exists on how incentive payments were actually used in practice.

We recommend that payments/changes in funding should be targeted at the lowest level of disaggregation that the data allow – at least to hospital level. Transmitting these funding changes to hospital departments and clinicians requires that the funding changes are regarded as a separate funding stream, or made salient in some other way. This requires the co-operation of LHNs and hospitals through existing national quality improvement initiatives.

Incentive dynamics and benchmarking

A critical design issue is how rewards and penalties are structured—whether there should be fixed target thresholds or providers should be ranked according to the distribution of performance measures. Fixed target thresholds can have weak incentive effects, e.g., providers who are already meeting targets are unlikely to change their behaviour, whereas low performing providers may find the thresholds too high to be worth attempting. The behavioural effects depends on a number of specific factors, including the distance between target thresholds if there are multiple targets, the payment structure (e.g., increasing with distance from thresholds rather than simply meeting targets), and the way thresholds are chosen via-a-vis the distribution of performance measures.

Schemes that reward for improvements in performance over time were found to be less effective than those that do not tie rewards to past performance, although it is worth noting that the sample of schemes were small (Scott et al, 2016) and that the dynamics of incentive schemes are complex and much remains to be understood. For example, comparisons made against the national average assume the average is 'best' and this is not the case.

We strongly recommend against using a national benchmark.

This is less meaningful to hospitals as their relative position depends on the behaviour of other hospitals, not only their own – they will feel they have less control of the impact of changing performance on financial rewards and so have less of an incentive to change behaviour. Hospitals with low HACs will get rewarded even if they do not change their behaviour. If one is interested in quality improvement, then national benchmarking doesn't necessarily support this key objective.

If the data allow, *we recommend rewards are for a reduction in HACS between a baseline period and some future period at the level of the hospital.* Only hospitals which *reduce* HACS (presumably those with high incidence rates) or re-admissions will be rewarded, thus more effectively targeting the rewards. Incentives for quality improvement would exist regardless of where a provider is located in the performance distribution.

The costs of improving quality

A key reason cited for poorly designed incentive scheme is the failure to consider a theoretical or conceptual framework that underpins the production of health care services. In the context of hospital care, one can think of hospitals as production units that use inputs to produce outputs in the form of hospital care services. Hospitals are to a large extent able to choose to produce outputs of varying quality. All else equal, higher quality outputs are more costly to produce. It is reasonable to think that this difference in costs is reflected in both the cost of each additional unit of output (i.e., marginal cost) and the average cost of a unit of output. The latter can differ from the former due to the presence of fixed costs such as setup costs and lump sum costs that have to be incurred regardless of output volume. For example, to reduce health care associated infections, hospitals can implement patient-specific measures that prevent the transmission from carriers to other patients, as well as general measures that promote standardised practices such as enhancing hand hygiene and improving environmental cleaning to reduce the risk of infections (e.g., Septimus et al. 2014; Huang et al. 2013). The former measures raise the cost of care for specific patients (i.e., marginal cost) while the latter type of measures could incur large fixed costs.

Recognising that producing higher quality outputs incur both higher marginal and fixed costs has important implications on the design of incentive schemes. Incentives should be provided so that providers are able to cover the costs of implementing measures for quality improvement. Further, incentive payments should be directed toward marginal as well as fixed costs associated with quality improvements. The presence of fixed costs implies economies of scale, which means small hospitals will be at a disadvantage compared to large hospitals in implementing quality improvement measures. This implies that hospitals should be compared with their peers of comparable size on risk adjusted outcomes, or the risk adjustment process should account for hospital volume (e.g., Gutacker et al. 2013).

We recommend that IHPA:
  1. *adopts a carrot-and-stick approach on pricing and funding of hospital services, rather than focusing on non-payment or payment reduction for low-quality episodes;*
  2. *provides incentive payments and penalties at both episode and hospital levels. In this regard, on approaches to reduce HACs, Option 3 of blending a quality-adjusted NEP with funding incentives at hospital level (pp. 46–48) is preferred to Options 1 and 2;*
  3. *accounts for hospital volume when comparing hospital performance. Hospital volume can be included as a hospital characteristic in risk adjustment, or it can be used as the main criterion for classifying peer groups for comparison and benchmarking performance;*
  4. *anticipates and monitors perverse incentives when introducing reduced or non-payments for specific episodes such as episodes with an adverse event. Hospitals may have an incentive to reduce or withdraw care for patients in these episodes ex post if the marginal costs of care are far in excess of the anticipated revenue. Hospitals may, for example, transfer patients to other hospitals or discharge patients earlier than planned and readmit them later on separate episodes. Such actions could contribute to higher costs of care and lower health outcomes for the patients involved.*
  5. *Consider the effects of incentives on coding and gaming. Hospitals may be less likely to report sentinel events or HACs if funding is at stake, and so observed reductions may reflect changes in coding. A system of audit should be put in place, or independent data collected and used (eg from registries, or for sentinel events data from AHPRA could be used on medical complaints or litigation).*

Risk Adjustment

Risk adjustment is an integral component of any incentive scheme that aims to reward good performance. Hospitals treat different types of patients. Some patients are inherently difficult to treat and prone to adverse outcomes due to the complexity of their conditions or comorbidities. Hospitals with a good reputation and specialised facilities tend to attract more complex and difficult patients. Without risk adjustment these hospitals will show poor performance and hence be unfairly penalised under the incentive scheme.

Ideally a risk adjustment model should remove all patient complexity and characteristics from the outcome measures so that incentive payments or penalties are applied purely based on factors within the control of providers. The process of doing so should be transparent and seen as fair by all stakeholders. Iezzoni (2009) identified five categories of factors that are outside the control of the hospitals.

  1. Clinical factors, e.g. diagnoses, comorbidities, mental health.
  2. Demographics, e.g. gender, age, ethnicity.
  3. Socio-economic status, e.g. employment, occupation, income, neighbourhood characteristics.
  4. Health behaviour and activities, e.g. smoking, drinking, diet and nutrition.
  5. Attitudes and perceptions, e.g. religion, care preferences, motivation and expectations.

Hospital administrative data such as those in IHPA's collection typically only contain information on the first two categories. Information on socio-economic status, health behaviour and attitudes is usually poor or non-existent. Yet these are important factors that are known to affect patient health outcomes and are typically beyond the control of hospitals and clinicians (e.g., Glance et al. 2016). The omission of relevant variables can cause serious

problems in the estimation of risk adjustment models and in econometrics the problem is known as omitted variable bias (e.g., Wooldridge 2016).

There is no consensus on the methodology of risk adjustment. Different methods are used by the US Centers for Medicare and Medicaid Services, the UK NHS and Canadian provincial health systems (COPSS-CMS White Paper Committee 2012; HSCIC Clinical Indicators Team 2016; Spencer et al 2008). There are different views on the appropriate statistical methodology for risk adjustment (Normand and Shahian 2007). Opinions also differ on whether a single risk adjustment model is appropriate for all disease and treatment groups, or whether separate disease- and treatment-specific risk adjustment models should be developed, and for different indicators (Zhou et al. 2016).

We recommend that IHPA
1. *explores the possibility of supplementing the current administrative data with survey data on health behaviour (smoking, exercise, diet, etc.), and socio-economic status (income, education, employment, etc.). The survey can be modelled along the line of the Medical Expenditure Panel Survey (MEPS) in the U.S. The survey data will fill the gap in the current data collection. Further, the survey data can be linked to administrative records to enable research on important social issues such as health care utilisation, health inequities and health behaviour in the Australian society.*
2. *Systematically investigate the robustness of different risk adjustment methods with the view to arrive at models that have good statistical properties and acceptable to all stakeholders.*

Hospital Quality
Hospital quality is a multifaceted concept involving many dimensions. Besides sentinel events, HACs and avoidable readmissions as outlined in the consultation document, many more quality measures can be constructed. The Alternative Quality Contract (AQC) in Massachusetts in the U.S. determined quality payments based on 64 measures on processes, outcomes and patient experiences (Song et al. 2014). Similarly, more than 30 quality measures were used by the Premier Hospital Quality Incentive Demonstration program (Shih et al. 2014) implemented in the U.S. in 2003 and later adopted in one region in England (Kristensen et al. 2014) in 2008; some of these measures are broad outcome measures such as risk adjusted in-patient mortality and readmission rates for selected conditions. In England Patient Reported Outcome Measures are used to measure health status before and after acute care, and are now being used as part of the process of risk adjustment.

It is important to measure quality using many indicators and as comprehensively as possible to avoid hospitals performing to incentivised metrics and neglecting on areas that are not incentivised. With many indicators, a practical issue arises on how to structure payments based on many quality measures. A common approach is to compute average scores and compare specific hospitals vis-a-vis the averages. Alternatively, it is possible to integrate different quality measures using a regression-based approach (Cheng et al. 2016a). This approach has the advantages of being simple to implement and allowing for further risk adjustment at the hospital level, for example, allowing for hospital volume to be included as a risk factor.

The focus on quality and quality improvements may have implications on hospital costs and productivity. Hospital focussing on quality improvements may do so at the expense of efficiency and productivity. However, Street et al. (2014) find that better patient outcomes in

UK NHS hospitals were associated with lower costs for hip and knee replacement, but no significant correlation was found for other treatment groups. A recent study using hospital administrative data from Victoria also could not find any relationship between quality and productivity (Cheng et al. 2016b). The findings are tentative due to the small number of observations. Understanding the relationship is important for the design of an incentive scheme which should in principle provide adequate financial resources for providers to act on and introduce measures to improve outcomes. For example, if better outcomes are associated with lower costs, the design can take this into account by withholding financial resources for poor performing providers.

We recommend that IHPA:
1. *expands the number of measures of hospital quality to include more indicators to guard against hospitals performing to incentivised metrics;*
2. *explores ways to integrate a large number of quality indicators using regression analysis;*
3. *undertakes further research verifying the relationship between hospital quality and cost efficiency or productivity.*

Evaluation and methodological research

Overall few previous studies have employed strong evaluation designs, though there has been a growth in difference-in-difference study designs (Scott et al, 2016). Weak evaluation frameworks have resulted in weak evidence on whether specific incentive schemes work, and how they can be changed over time to maintain behaviour change. This is in part due to a lack of emphasis on evaluation during the design and implementation of most incentive schemes. It is of utmost importance that the evaluation of impacts of changes to pricing and funding on quality improvements be regarded as an essential component of the incentive scheme, in particular on providing essential feedback to the evolving design of the scheme.

We recommend IHPA considers establishing an evaluation from the outset, and to set up and evaluation baseline dataset, and consider carefully the best study design to be used (eg difference in difference or randomised stepped wedge design).

References

Cheng, C, A Scott, V Sundararajan, J Yong (2016a). "On measuring the quality of hospitals: A simple method of constructing a composite quality index," *Unpublished Working Paper*, University of Melbourne.

Cheng, C, A Scott, K Sethuraman, V Sundararajan, J Yong (2016b). "The effect of raising hospital quality on productivity: Evidence from system GMM estimation of hospital production function," *Unpublished Working Paper*, University of Melbourne.

COPSS-CMS White Paper Committee (2012), "Statistical issues on assessing hospital performance," *CMS White Paper*, U.S.: Centers for Medicare and Medicaid Services.

Gutaker, N, Bojke C, Daidone S, Devlin N, Parkin D, Street A (2013). "Truly inefficient or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes," *Health Economics* 22: 931–947.

Eijkenaar, F, M Emmert, M Scheppach, and O Schoffski (2013). "Effects of pay for performance in health care: a systematic review of systematic reviews," *Health Policy,* 110 (2-3):115–130.

Glance, LG, AL Kellermann, TM Osler, Y Li, W Li, AW Dick (2016). "Impact of Risk Adjustment for Socioeconomic Status on Risk-adjusted Surgical Readmission Rates," *Annals of Surgery*, 263(4): 698–704.

Gutacker, N, C Bojke, S Daidone, NJ Devlin, D Parkin, A Street (2013). "Truly inefficient or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes," *Health Economics,* 22(8): 931–947.

HSCIC Clinical Indicators Team (2016*), Indicator Specification: Summary Hospital-level Mortality Indicator*, U.K.: Health and Social Care Information Centre.

Huang, SS, E Septimus , K Kleinman , et al. (2013). "Targeted versus universal decolonization to prevent ICU infection," *New England Journal of Medicine*, 368: 2255–2265.

Iezzoni, LI (2009) "Risk adjustment for performance measurement," in PC Smith, E Mossialos, I Papanicolas, S Leatherman (eds.) *Performance Measurement for Health System Improvement*, Cambridge, U.K.: Cambridge University Press.

Kristensen, SR, R Meacock, AJ Turner, R Boaden, R McDonald, M Roland, M Sutton (2014). "Long-term effect of hospital pay for performance on mortality in England," *New England Journal of Medicine*, 371(6): 540–508.

Normand, SLT and DM Shahian (2007). "Statistical and Clinical Aspects of Hospital Outcomes Profiling" *Statistical Science*, 22(2): 206–226.

Scott, A, M Liu, J Yong (2016). "Financial incentives to encourage value-based health care," *Medical Care Research and Review*, forthcoming.

Septimus, E, RA Weinstein, TM Perl, DA Goldmann, DS Yokoe (2014). "Approaches for Preventing Healthcare-Associated Infections: Go Long or Go Wide?" *Infection Control & Hospital Epidemiology*, 35(S2): S10–S14.

Shih, T, LH Nicholas, JR Thumma, JD Birkmeyer, JB Dimick (2014). "Does pay-for-performance improve surgical outcomes? An evaluation of phase 2 of the Premier Hospital Quality Incentive Demonstration," *Annals of Surgery*, 259(4): 677–681.

Song, Z., S Rose, DG Safran, BE Landon, MP Day, ME Chernew (2014). "Changes in Health Care Spending and Quality 4 Years into Global Payment," *New England Journal of Medicine*, 371:1704–1714.

Spencer, G, J Wang, L Donovan, JV Tu (2008), "Report on Coronary Artery Bypass Surgery in Ontario, Fiscal Years 2005/06 and 2006/07," *Technical Report*, Institute for Clinical Evaluative Sciences.

Street A, N Gutacker, C Bojke, N Devlin, S Daidone (2014). "Variations in outcome and costs among NHS providers for common surgical procedures: econometric analyses of routinely collected data," *Health Services Delivery and Research*, 2(1). DOI: 10.3310/hsdr02010.

Wooldridge, JM (2016). Introductory Econometrics: A Modern Approach, 6th ed., Cengage Learning.

Zhou, H, PR Della, P Roberts, L Goh, SS Dhaliwal (2016). "Utility of models to predict 28-day or 30-day unplanned hospital readmissions: an updated systematic review," *BMJ Open*, doi:10.1136/bmjopen-2016-011060.