

Using financial incentives to improve the performance of hospital clinicians: a rapid review

Anthony Scott
Driss Ait Ouakrim

An **Evidence Check** rapid review brokered by the Sax Institute

saxinstitute

This rapid review was brokered by the Sax Institute.

This report was prepared by Anthony Scott and Driss Ait Ouakrim.

© The Sax Institute, November 2011.

This work is copyright. It may be reproduced in whole or in part for study training purposes subject to the inclusion of an acknowledgement of the source. It may not be reproduced for commercial usage or sale. Reproduction for purposes other than those indicated above requires written permission from the copyright owners.

Enquiries regarding this report may be directed to:

Knowledge Transfer Program
The Sax Institute
www.saxinstitute.org.au
Phone (02) 9514 5950
Fax (02) 9514 5951

Suggested Citation:

Scott A, Ouakrim D. Using financial incentives to improve the performance of hospital clinicians: an Evidence Check rapid review brokered by the Sax Institute (<http://www.saxinstitute.org.au>); 2011.

Disclaimer:

This **Evidence Check** Review was produced using the Evidence Check methodology in response to specific questions from the commissioning agency. It is not necessarily a comprehensive review of all literature relating to the topic area. It was current at the time of production (but not necessarily at the time of publication). It is reproduced for general information and third parties rely upon it at their own risk.

Table of Contents

1 EXECUTIVE SUMMARY	5
<i>Implications for Australia</i>	6
2 Introduction	7
3 Methods	9
4 Identify the range of incentive-based funding models	10
<i>Definition of funding and incentives</i>	10
5 Describe, review and summarise the evidence for each model	13
5.1 <i>The Centers for Medicare and Medicaid Services (CMS) Premier Hospital Quality Incentive Demonstration (PHQID) (US)</i>	13
5.2 <i>Participating Hospitals Agreement (PHA) Incentive Program (US)</i>	0
5.3 <i>Pay-for-Performance in a tuberculosis program (Taiwan)</i>	1
5.4 <i>Technology adoption and financial incentives to improve report signature times in radiology (US)</i>	1
5.5 <i>Commissioning for Quality and Innovation Payment Framework (COQIN) (UK)</i>	2
5.6 <i>Prospective payments to increase quality (Victoria, Australia)</i>	3
5.7 <i>Best practice tariffs for casemix payment (UK)</i>	3
5.8 <i>Financial penalties for "Never Events" (UK)</i>	4
5.9 <i>Financial penalties for "Never Events" (US)</i>	5
5.10 <i>Advancing Quality Initiative (UK)</i>	6
5.11 <i>Clinical Practice Improvement Payment System (CPIP) (Australia)</i>	7
6 Review the models as a group and summarise their strength of evidence	8
<i>Summary of conclusions of other systematic reviews</i>	8
<i>Responses to specific questions on funding models</i>	8
<i>Some general principles and considerations for the design of incentive schemes</i>	10
7 Summary and discussion	14
<i>Implications for Australia</i>	15
8 Appendices	
Appendix 1. Search Strategy	17
Appendix 2. Main characteristics of included evaluations	18
9 References	30

1 EXECUTIVE SUMMARY

The use of financial incentives to improve performance in health care is growing, and this trend is accompanied by an expanding literature. However, evidence for the impact of such schemes remains uncertain due to inadequate research designs and poorly designed incentive schemes. To date, few reviews have focused specifically on the use of such schemes aimed at clinicians working in hospitals.

This review has been commissioned by the Sax Institute as part of their *Evidence Check* series of reports. Its aim is to describe and examine the evidence for the effects of incentive-based funding models for hospital clinicians on performance and quality of care.

The review will:

- Identify the range of incentive-based funding models for the provision of clinical services;
- Review each individual model identified; and
- Review the models as a group and examine the strength of evidence

There is a large literature in this area which is reviewed.

The review is based on:

- An existing overview of systematic reviews;
- More recent systematic reviews;
- A primary literature search; and
- Grey literature, including websites.

A short conceptual framework is presented to help explain the mechanisms by which changes in funding models create incentives that in turn influence behaviour and performance.

Numerous models exist in the United Kingdom, United States and Australia, but there is little rigorous empirical evidence of the effectiveness of changes in funding models for clinicians in hospitals. Those studies that do exist show mixed results, with the more robust studies showing no effect of pay-for-performance targeted at hospitals. One key reason suggested in the literature was that the additional funding was paid to hospitals. There were few reports as to whether and how this funding had been distributed to clinicians or clinical units within hospitals. The way the hospital uses the funds is crucial to the success of such schemes in improving performance and quality of care.

What remains to be implemented and evaluated is a scheme that uses pay-for-performance to reward hospitals, combined with a mechanism that distributes the performance funding directly to clinical units and clinicians.

Pay-for-performance schemes for hospitals include the Premier Hospital Quality Incentive Demonstration (US), the Commissioning for Quality and Innovation payment framework (UK), the Advancing Quality initiative (UK), best practice tariffs in Payment by Results (UK), and the Clinical Practice Improvement Payment System (Queensland Health). There are also

schemes in the US and UK that do not pay hospitals for avoidable adverse events (“Never Events”). Only the American schemes have been evaluated. Many are based on schemes that adjust casemix payments as a reward for higher performance and quality of care.

Internal hospital mechanisms for supporting performance management and budgetary management include Service Line Management and a variety of models from managerial accounting.

Implications for Australia

There is insufficient evidence to determine what type of incentive scheme or method of payment should be used if such a scheme were to be introduced in Australia.

A key issue is that such schemes are usually based on what already exists. The national rollout of casemix payment in Australia, together with initial experience from the UK and US on increasing casemix payments to reflect higher quality care and not paying for “Never Events”, suggest that the adjustment of casemix payments should be considered in Australia as a payment mechanism. This is in addition to the method being used in Queensland Health’s Clinical Practice Improvement Payment System.

However, it is important that financial incentives be delivered to the clinical units or clinicians so that they see some effect of their effort – either in terms of increased funding for the unit or performance-based pay for the health professionals (not just clinicians) involved – and that there exists some internal accounting mechanism to support this.

Rigorous evaluation would be essential to examine the impact of these schemes.

2 Introduction

The purpose of this rapid review is to summarise evidence and describe models of the use of financial incentives to improve the performance of clinicians working in hospital settings. Funding models and financial incentives have the potential to alter clinical performance. There are numerous funding models in existence, and there is now a large literature in this area which has been reviewed a number of times.^{1, 2, 3} The majority of the literature looks at using financial incentives to improve processes and quality of care through pay-for-performance.

The objectives of this review are to:

- 1) Identify the range of incentive-based funding models (such as pay-for-performance) for the provision of clinical services (excluding primary care services such as GPs and Community Health) that have been described and/or evaluated in the last 10 years.
- 2) Describe, review and summarise the evidence for each individual model identified, in terms of:
 - the clinical service in which the model was implemented (for example inpatient hospital, accident and emergency, inpatient clinic, outpatient clinic, ambulance);
 - the specific type of incentive payment system on which the model is based, noting whether the incentives applied to individual clinicians or the clinical unit;
 - the evaluation design used (for example randomised controlled trial, case/control study) in order to judge the likely quality of the evaluation evidence; and
 - the specific outcomes or indicators used to measure the effectiveness of the model, including:
 - patient outcomes and/or satisfaction with care;
 - quantity and/or quality of service delivery, including evidence of reductions in overservicing;
 - costs, including the type of cost analyses undertaken (for example cost-effectiveness, cost-utility);
 - the effect size, or other indicator of the extent to which the model was effective on each outcome variable measured; and
 - the period of time for which the model was evaluated (to comment on issues of ongoing cost and sustainability).

- 3) Review the models as a group, and examine the strength of evidence to identify:
- which incentive-based systems are likely to be most effective, in terms of their initial effect and likely sustainability over time;
 - whether the effectiveness and/or sustainability of the models is improved when clinicians have the capacity to contribute to the design of the incentives;
 - whether some incentive-based models are more likely to create perverse incentives (that is, incentives for clinicians to prioritise payments ahead of patient or service delivery outcomes);
 - whether incentive-based models of funding reduce variation in performance or clinical outcomes between clinical units or clinicians, relative to other strategies such as public disclosure of clinician or clinical unit performance;
 - the extent to which incentive-based models reduce or exacerbate inequities in access to health care;
 - an optimal and/or sustainable balance between financial and non-financial incentives to achieve improved and sustained outcomes; and
 - whether incentive-based models increase budget transparency to funders and the public, and the likely effect on efficiency outcomes.

These objectives are addressed in Sections 4, 5 and 6 respectively.

3 Methods

Our search strategy is summarised in Appendix 1. It was based on three main sources of information:

- 1) An existing Cochrane “review of reviews” that has identified 25 systematic reviews of financial incentives for clinicians.² Each review’s references list was searched for relevant papers.
- 2) A new rapid literature search using key electronic databases, modified from a Cochrane review used by Scott and colleagues (2010).⁴ This aimed to identify more recent papers.
- 3) Grey literature, including websites of research organisations and governments.

Since this is a rapid review, the studies presented and discussed may not include all existing studies or models. Furthermore, the search focused largely on empirical studies. Descriptions of more recent initiatives that have not yet been evaluated are included.

4 Identify the range of incentive-based funding models

The aim of this section is to identify the range of incentive-based funding models that can potentially be used, or have been used, in various contexts.

Each funding model has a range of characteristics that assist with defining each model, and with making comparisons across models. They also assist with designing models that perhaps do not yet exist. The first section below presents a brief conceptual framework highlighting these characteristics, which are then used to help classify and examine the actual models identified in the second section.

Definition of funding and incentives

Outlining a brief conceptual framework is necessary to help understand the mechanism by which funding changes and financial incentives can influence behaviour. It also helps to clarify the terminology used to describe actual incentive schemes. This then helps clarify the key issues to consider when designing effective incentive schemes that require an understanding of such mechanisms.

The existence and nature of financial incentives depends on the different methods of funding and reimbursement. Funding is a transfer of money from a buyer (in our case a third-party funder or insurer) to a supplier of services (in our case a clinician or group of clinicians or hospital). In exchange for the funding the former may require the latter to engage in certain behaviours.

Making funding conditional on expected behaviours or performance creates the financial incentive, and makes clear how the funding change is intended to change behaviour. An incentive or disincentive is a motivation to act in a certain way. There are a number of sources of motivation that create incentives or disincentives to take action or make decisions. Some are defined as intrinsic motivation and are concerned with the value of the task or behaviour itself, altruism and caring for patients' health. Others are extrinsic motivators, such as changes in regulation and funding.

A key issue in examining the effect of funding changes is the relative importance of funding to other sources of motivation for behaviour change. This will vary across clinicians, and so lead to variations in their reactions to changes in funding. Changing behaviour and improving performance also incurs costs and uses scarce resources, and these costs will also vary across clinicians, again leading to variation in their reactions to changes in funding. Though the funder may intend that the funding change will influence behaviour in a certain way, other unintended behavioural consequences may also occur. This may occur because of variation in motivation and the costs of behaviour change, and variation (unknown to the funder) in the context and information available to each clinician.

The extent to which the funding change is sufficient to cover the costs of changing behaviour is the strength of the incentive.¹ If the funding more than covers costs, then changing behaviour will lead to a surplus or profit (revenue minus cost) that may be used as personal income or for other purposes. The size of this surplus reflects the strength of the

financial incentive to change behaviour, conditional on funding being important compared with other sources of motivation.

There are a number of characteristics of different funding schemes that can influence behaviour:⁴

a) Method of payment

- (i) Payments can be made in exchange for the following clinician behaviours:
 - working for a specified time period (for example, salary, sessional payment or block contract);
 - providing specific services/treatments/episodes (fee-for-service);
 - providing care for a patient or specific population (capitation); and
 - providing a pre-specified level, or increase in the level, of quality of care (for example, performance-based pay).

- (ii) Payments can be linear or non-linear:
 - linear, so the same payment is made for each additional unit of payment (for example, service provided); or
 - non-linear, such that payment is conditional on reaching a threshold or target, or a series of thresholds, or that the amount of the payment changes with each additional service provided. This includes payments being based on the marginal cost of each additional unit of activity or outcome.

- (iii) The timing of payments can be either:
 - in advance (prospective payment that provides a fixed overall budget); or
 - after the behaviour has taken place – including:
 - retrospective payment where there is no overall limit on payments; and
 - retrospective payment where there is a cap on the total payments that can be made. Once the cap is reached, either no further payments are made or the amount of the unit payment is reduced.

b) Amount of payment

- (i) The level of payment may be fixed in advance and subject to negotiation (for example in a fee schedule or trade union bargaining of salary increases).
- (ii) The providers' discretion as to the amount of money they can charge for their services may be complete (in a market setting) or partial (for example fee/price controls).
- (iii) The amount of the payment may be reduced or withheld if behaviour does not comply with what is required (that is, a financial penalty).

- (iv) The amount may also vary depending on administrative rules based on the characteristics of the provider (for example whether accredited or qualified to a certain standard or seniority, or practising in a rural area or where the unit costs of providing the service are “high”) or the patients seen (for example through “risk adjustment”, or more complex/costly services, or patients receiving higher payments).

c) Payment target

Payment may be made to:

- (i) individuals;
- (ii) teams or units within organisations; or
- (iii) organisations.

d) Risk sharing

The strength of the incentive also depends on the extent of “risk sharing”. This refers to funding that typically is fixed (for example a fixed department budget) and the extent to which any surplus can be used by the seller or retained by the funder, and the extent to which deficits or overspends are borne by the clinician or shared with funder.

- (i) A “hard” budget is where the clinician can use surpluses as they want (subject to (3) below), but must also bear any deficit. Here incentives are at their strongest but the financial risk of a deficit is borne entirely by the clinician.
- (ii) A “soft” budget is where a proportion of the surplus can be retained by the clinician, and where a proportion of the deficit is funded by the funder. Here surpluses and financial risks are shared, but incentives for the clinician are weaker. This depends on the extent of risk sharing.
- (iii) There may also be administrative rules as to how the surplus can be used:
 - For personal income of the individual;
 - Re-invested in the provision of services; or
 - Re-invested in capital developments (having regard to who owns the capital).

Changes in any of the above will influence the behavioural response to the payment – that is, will either encourage the desired behaviour (have an incentive effect) or discourage it (have a disincentive effect). Economic theory can be used to predict the effects of different types of funding and remuneration on behaviour, with the incentive effects dependent on the changes between the current system of funding and the new system of funding.⁵

In turn, changes in behaviour influence the cost and quality of care provided. Note that changes in payment may also influence quality and cost by influencing recruitment and retention – that is, influencing the mix of providers and their self-selection into the payment scheme, specialty or health plan. Changes in the mix of providers may then impact on quality and costs, rather than the incentive having a direct effect on quality.

5 Describe, review and summarise the evidence for each model

The aim of this section is to describe, review and summarise the evidence for each model identified in the literature. Eleven different models of incentive-based funding are summarised and discussed. From the literature review, there were 12 separate empirical evaluations of six different incentive models. We also summarise five other funding models recently introduced in the UK, US, and Australia that have not yet been evaluated. The summary of each model will cover:

- *Clinical service.* The clinical service in which the model was implemented (for example inpatient hospital, accident and emergency, inpatient or outpatient clinics, ambulance).
- *Type of incentive.* The specific type of incentive payment system on which the model is based, noting whether the incentives applied to individual clinicians or the clinical unit.
- *Details of evaluation.* This includes:
 - the evaluation design used (for example random controlled trial, case/control study) in order to judge the likely quality of the evaluation evidence;
 - the specific outcomes or indicators used to measure the effectiveness of the model;
 - the effect size, or other indicator of the extent to which the model was effective, on each outcome variable measured; and
 - the period of time for which the model was evaluated (to comment on issues of ongoing cost and sustainability).

5.1 The Centers for Medicare and Medicaid Services (CMS) Premier Hospital Quality Incentive Demonstration (PHQID) (US)

This is a major scheme in the American Medicare system that has been implemented and evaluated.

Clinical service. Performance was defined in terms of quality of care for five disease areas – heart failure, acute myocardial infarction (AMI), community-acquired pneumonia, coronary artery bypass grafts (CABG), and hip and knee replacements – using 33 quality measures.

Type of incentive. Hospitals were paid a bonus according to their ranking of performance compared with other hospitals (tournament-based pay):

- Hospitals scoring in the top 10% for meeting pay-for-performance clinical measures would receive a bonus payment consisting of 2% of their annual diagnosis-related group payment.
- Hospitals in the top 20% but not in the top 10% would receive a 1% bonus.
- Hospitals in the top 50% would be acknowledged through public reporting.

Bonuses averaged US\$71,960 per year and ranged from US\$914 to US\$847,227. These additional payments are anticipated to be partially offset by financial penalties ranging from 1 to 2% of Medicare payments for hospitals that by the end of the third year of the program had failed to exceed the performance of hospitals in the lowest two deciles, as established during the program's first year. This was implemented alongside public reporting of performance data.

Details of evaluation. There are six evaluations of the impact of the American PHQID scheme.^{6, 7, 8, 9, 10, 11} Participation by hospitals was voluntary, leading to difficulty in attributing the effects to the incentive scheme. The key characteristics of these studies are summarised in Appendix 2. Glickman⁶ and Grossbart⁷ found little effect of the scheme after following patients for three and two years respectively. However, they used relatively weak case-control study designs that did not account for selection bias or control for differences in hospital characteristics. Using similar outcome measures for AMI, Lindenauer and colleagues (2007)⁸ also used a case-control design but used regression models to control for differences in baseline and other observable characteristics that differed between case and control hospitals. They found modest improvements in quality of care, of up to 4.1% over a two-year period.⁸

The latest published evaluation of the American PHQID scheme is by Ryan.¹⁰ This used a strong difference-in-difference (controlled before and after) study design using panel data econometric techniques that controlled for the voluntary nature of the scheme that can cause selection bias. Six years of data (on over 11 million admissions) were used to compare hospitals in and out of the scheme. The results showed that the scheme did not influence mortality or costs for AMI, pneumonia and CABG, though other process measures used in previous studies were not examined.

Two further studies focused on the effects of the scheme for minority and poor patients respectively. Ryan and colleagues (2010)¹¹ use pooled cross-sectional data to examine whether the likelihood of receiving care at a PHQID hospital for AMI, pneumonia, and heart failure was different between white and minority patients. For AMI patients they also examined the likelihood of patients receiving a CABG. The authors found little evidence that PHQID hospitals avoided minority patients. Jha and colleagues (2010)⁹ examined the impact of the scheme on access by poor patients using 251 PHQID hospitals and a national sample of control hospitals. They examined the association between performance in AMI, pneumonia, and heart failure with the share of poor patients, and found that this was stronger in the PHQID hospitals compared with the control hospitals for AMI and pneumonia. This relationship had weakened substantially after three years. Hospitals with poorer patients were more likely to respond to incentives, though the study did not examine changes over time and so was a weaker design compared with that of Ryan and colleagues (2010).¹¹

Ryan (2009a)¹² also reviewed the American literature and found that although the PHQID scheme improved processes, there was little evidence it improved outcomes. Some reasons for this lack of effect were discussed. This included evidence suggesting that hospitals that received bonuses did not distribute them within the hospital or to physicians, suggesting that incentives to change clinical behaviour were non-existent. Recent changes to the scheme

have also introduced payments for lower performing hospitals that improve quality, rather than just attain a single high threshold.

5.2 Participating Hospitals Agreement (PHA) Incentive Program (US)

This is a small scheme run by Blue Cross Blue Shield of Michigan for 85 hospitals in Michigan. It was introduced in a phased way in 2000.

Clinical service. Heart disease with indicators of:

- 1) the provision of aspirin orders at discharge for patients diagnosed with acute myocardial infarction (AMI);
- 2) the prescription of beta-adrenergic blockers at discharge for patients diagnosed with AMI; and
- 3) the prescription of angiotensin converting enzyme (ACE) inhibitors at discharge for patients diagnosed with congestive heart failure (CHF).

There are also payments for measured performance in the treatment of pneumonia, the use of medication safety practices, the appropriate utilisation of high-cost/high-variation surgical procedures, and the implementation of community benefit programs. In 2005, a surgical infection protection component was added.

Type of incentive. Payments are calculated as a percentage add-on to hospitals' inpatient diagnosis-related group (DRG) reimbursements from the insurer. The maximum possible add-on for heart related care increased from 1.2% of a hospital's inpatient DRG reimbursements in 2000-2002 to 2% of a hospital's inpatient DRG reimbursements in 2003. The actual add-on is determined by multiplying the maximum possible add-on by a hospital's individual performance score. In 2002 and 2003, hospitals only received heart-care-related incentive payments if they achieved an established minimum performance score, or threshold, equal to the median performance level of all participating hospitals. Continuous improvement in hospital quality is encouraged through increases in established thresholds. Prior to 2002 hospital incentives were determined by ranking hospitals in order of overall performance scores, dividing the hospitals into quartiles, and setting fixed incentive rates for each of the resulting groups (tournament-based pay).

Details of evaluation. Nahra and colleagues (2006)¹³ examined the cost per Quality-Adjusted Life Year (QALY) of the program and found a relatively low cost per QALY (US\$30,081), suggesting that incentives were cost-effective. However, they only examined the costs of incentive schemes and not the costs of changes in utilisation, prescribing, or hospital visits. A more fundamental problem was that there was no before period of data collection and no control group, resulting in a very weak prospective observational design. This meant that the measure of effect (changes over time from baseline) is likely to be biased, particularly if the trends in performance would have occurred in the absence of the scheme.

5.3 Pay-for-Performance in a tuberculosis program (Taiwan)

Clinical service. Patients in hospital with tuberculosis (TB).

Type of incentive. The Bureau of National Insurance in Taiwan introduced the pay-for-performance (P4P) demonstration project in 2001. It was scaled up to a national program in 2004. The incentive scheme was designed around four chronological stages of tuberculosis (TB) treatment that lasted 12 months altogether. Points were earned by hospitals, physicians and case managers for the number of cases that were identified (stage 1) and cured in the last three stages of treatment. The value of each point is not stated in the paper.¹⁴

Details of evaluation. Li and colleagues (2010)¹⁴ compared participating hospitals with non-participating hospitals before and after the introduction of the scheme, from 2002 to 2005. Given that the demonstration project ran from 2002 to 2003, the study does not address a period where there was no scheme: hospitals not in the demonstration project may have changed their behaviour in anticipation of the scheme's introduction. Identification of TB cases in Taiwan increased by 30% after the scheme was introduced. They report evidence using data from all hospitals that the cure rate increased from 46.9% to 63% and that the length of treatment fell from 58.3 days to 55.4 days after the scheme was introduced. However, they did not present this before and after data by whether a hospital was in the program or not. They also report that P4P hospitals had a higher cure rate (68.1%) than non-P4P hospitals (48.4%), but did not present this data before and after. The authors did not conduct a difference-in-difference analysis of the data and so it was not possible to say whether the change in cure rates for P4P hospitals were higher than the change in cure rates for non-P4P hospitals before and after the policy was introduced.

5.4 Technology adoption and financial incentives to improve report signature times in radiology (US)

This was a multifaceted intervention combining the implementation of two electronic tools (paging plus speech recognition) with a financial incentive scheme to improve the signature time (ST) on radiologists' reports.

Clinical service. Radiology department in a 751-bed hospital in Boston.

Type of incentive. A US\$4000 bonus was paid twice a year to radiologists if 80% of reports were signed within 16 hours, or if the departmental signing goal of <8 hours was achieved. The intervention was conducted over four years in a 751-bed, urban, tertiary care hospital with a radiology department performing approximately 750,000 examinations per year. The financial incentives, together with speech recognition and paging, lasted only one year and were then removed.¹⁵

Details of evaluation. The design was a prospective observational study and observed the effect on signature times of the introduction of the pager, then speech recognition, and then the financial incentive.¹⁵ There was a short (three-month) "before" period before the pager was introduced, but no concurrent control group – leading to a poor study design that was

unable to attribute the trend changes to the interventions. After the financial incentive was removed, hospitals were followed up for 19 months. The two technology interventions significantly reduced the median radiology report ST from >5 hours to <1 hour, and the 80th percentile ST from >24 hours to between 15 and 18 hours. The introduction of the financial incentive mechanism further improved the signature time, with the 80th percentile ST decreasing from >15 hours to between four and eight hours; but had little additional effect on the median. The largest differences were at the 80th percentile, rather than the median. This trend was sustained up to 19 months following the discontinuation of the financial incentive programme. The multifaceted design of this intervention does not allow us to tell to what extent the improvement in ST is due to the financial incentive, or if the same results could have been reached without the financial incentive.

5.5 Commissioning for Quality and Innovation Payment Framework (CQUIN) (UK)

In 2008 the UK's National Health Service (NHS) set out a plan for quality improvement in the UK in the paper High Quality Care for All. Among its initiatives were linking payments to quality improvements.¹⁶ CQUIN is a national framework that allows local commissioners to withhold a proportion of payments conditional on the provider achieving locally agreed quality and innovation improvements.^{17,18}

Clinical service. All acute, ambulance, community, mental health and learning disability services in England. For acute hospitals, there are also two national goals in 2010/11: reducing the avoidable effects of venous-thromboembolism (VTE) and improving responsiveness to patients' personal needs. These are to account for one-fifth of the value of the local schemes, or 0.3% of provider income.

Type of incentive. Introduced in April 2009, for 2009/10 the size of the incentive paid to each organisation was 0.5% of total contract revenue. In 2010/11 it was 1.5% and continuing increases are likely. Fifty per cent of payments are to be made in advance, to be reconciled with actual performance later in each year. This was to avoid cash flow problems for providers without making the payments automatic. The scheme does not include payments to individual health professionals as personal income.

The focus is on continuous quality improvement beyond the minimum. The aim is to "support a cultural shift by embedding quality improvement and innovation as part of the commissioner-provider discussion everywhere". Since the scheme is mandated for all providers, it will include those with low baseline levels of quality (see section 6.1.9). This includes independent sector providers who contract with the NHS. The framework provides substantial local flexibility in the areas of focus, which can also vary from year to year. It can therefore be integrated with other quality improvement initiatives. It also provides scope for substantial local involvement in the development of the scheme, rather than providers having to accept nationally determined indicators that may not be relevant to the local context. Where providers have more than one commissioner, to avoid duplication the commissioners must jointly set quality improvements so that each provider has only one CQUIN scheme.

Details of evaluation. An independent evaluation has been commissioned that will end in 2012. CQUIN schemes established in 2009/10 are published at http://www.institute.nhs.uk/world_class_commissioning/pct_portal/cquin_schemes.html

5.6 Prospective payments to increase quality (Victoria, Australia)

Clinical service. All hospital inpatient services in Victoria covered by diagnosis-related groups (DRGs).

Type of incentive. Re-distribution of DRG payments between discharges with hospital-acquired complications (diagnoses not present on admission), and those without. This change would be budget neutral and involve lower payments for discharges where such a diagnosis was present, and higher payments for those where it was not.

McNair and colleagues (2009)¹⁹ used data from Victoria to simulate the extent to which DRG payments to hospitals in Victoria could be re-adjusted if they excluded diagnoses that were not present on admission, and showed that this had the potential to provide higher rewards for hospitals that had fewer hospital-acquired diagnoses. The simulation showed that almost 15% of discharges had at least one diagnosis not present on admission removed from the payment rate for that discharge. The most common discharges where this occurred were for coronary bypass, respiratory infections, hip replacement and major bowel procedures. Around 1-2% of discharges were re-classified to a different DRG, and of these, there was a fall in the case weight (the relative cost-weight) of 27.6%, leading to a re-distribution of existing DRG payments.

Details of evaluation. The study was a simulation of the effects on hospital payments and has not been implemented.¹⁹

5.7 Best practice tariffs for casemix payment (UK)

This is an example from the UK of altering casemix payments – known in the UK as “tariffs” – to hospitals to reflect better quality of care. This is in its early stages of implementation.

Clinical service. Best practice tariffs for stroke, cholecystectomy, cataracts and fractured neck of femur were introduced in April 2010.²⁰ There are three different types of payment models to incentivise:

- Pathways for cataract surgery;
- Day case surgery for cholecystectomy; and
- Best practice, including health outcomes, for stroke and hip fracture.

Best practice tariffs for renal dialysis, interventional radiology, primary diabetic medicine, primary total hip and knee replacements, and transient ischaemic attacks are being introduced in 2011. They are also extending the “incentivising day case model” introduced for cholecystectomy to a range of procedures suggested by the British Association of Day Surgery (BADs): breast surgery, hernia repair, orthopaedic surgery and urology. These

changes in 2011 represent a rapid expansion of the scheme.²¹ The areas were chosen on the basis of significant variation in clinical practice, alongside the existence of “clear consensus” about “best practice” based on clinical evidence-based guidelines.²⁰ For example, for hip fracture the best practice tariff will be paid if the following conditions are met:²²

- (a) Time to surgery: within 36 hours from arrival in an emergency department, or time of diagnosis if an inpatient, to the start of anaesthesia; and
- (b) Involvement of an (ortho) geriatrician:
 - 1) admitted under the joint care of a consultant geriatrician and a consultant orthopaedic surgeon;
 - 2) admitted using an assessment protocol agreed by geriatric medicine, orthopaedic surgery and anaesthesia;
 - 3) assessed by a geriatrician (as defined by a consultant, non-consultant career grade (NCCG), or specialist trainee ST3+) in the perioperative period (defined as within 72 hours of admission); or
 - 4) postoperative geriatrician-directed:
 - a. multi-professional rehabilitation team; or
 - b. fracture prevention assessments (falls and bone health).

Type of incentive. Incentives are to be applied to the existing national fixed tariffs for the above HRGs and are mandatory for all hospitals in the NHS. As with the Continuous Quality Improvement Program (COQUIP), they are a result of the broader emphasis on supporting cultural change focused on quality improvement.¹⁶ The prices are set to reflect the average costs of providing quality of care. The previous national tariff has been replaced with a higher best practice tariff and a lower non-best practice tariff. For example, for day case surgery for cholecystectomy the new best practice tariff is higher than for inpatient elective surgery. For hip fracture the best practice tariff is £455 higher than the current base tariff. An average unit treating 350 hip fractures per year with 90% meeting the best practice tariff requirements will result in just over £140,000 of additional income ($0.9 \times 350 \times 455$).²³ In the first two to three years this will represent “real” extra money for those undertaking best practice. After that time, the base tariff will be reduced, reflecting the efficiency gain achieved. In 2011 the additional payment of £455 is being doubled and the base tariff being reduced by the same amount, representing a strengthening of the incentives to adopt best practice.²¹ This will be monitored as behaviour changes.

Details of evaluation. An evaluation has been commissioned using the 2010 clinical areas as case studies and initial, but as yet unpublished, findings have been used to inform the implementation of the 2011 clinical areas.²¹

5.8 Financial penalties for “Never Events” (UK)

Clinical service. In the UK casemix system of “Payment by Results”, HRG payments will not be made for these seven “Never Events” defined by the UK National Patient Safety Agency:²⁰

- 1) Wrong site surgery;
- 2) Retained instrument post-operation;
- 3) Wrong route administration of chemotherapy;
- 4) Misplaced naso or orogastric tube not detected prior to use;
- 5) Inpatient suicide using non-collapsible rails;
- 6) In-hospital maternal death from post-partum haemorrhage after elective caesarean section; and
- 7) Intravenous administration of mis-selected concentrated potassium chloride.

These patient safety events are largely preventable. It is unclear whether they are currently reported publicly.

Type of incentive. The withholding of casemix payment if a “Never Event” occurs.

Details of evaluation. No evaluation has been conducted.

5.9 Financial penalties for “Never Events” (US)

Clinical service. The “Never Events” are deemed to be within the control of hospitals, and include:

- 1) Object left in patient during surgery;
- 2) Air embolism;
- 3) Blood incompatibility;
- 4) Catheter-associated urinary tract infection;
- 5) Pressure ulcer;
- 6) Vascular-catheter-associated infection;
- 7) Mediastinitis after coronary artery bypass grafting; and
- 8) Fall from bed.

In 2009 other “Never Events” were also being considered: ventilator-associated pneumonia, Staphylococcus aureus septicaemia, and deep venous thrombosis or pulmonary embolism.

Type of incentive. The Centres for Medicaid and Medicare Services in 2008 did not pay for eight preventable conditions that were not present at the time of admission.^{24, 25, 26} There are some issues in terms of diagnosing and coding what the complications are and the process in place to support this.²⁴

Details of evaluation. No evaluation has been reported.

5.10 Advancing Quality Initiative (UK)

This program was introduced in 2008 in 24 hospitals in the North West of England and involves Premier Inc, which was also involved in establishing the American PHQID scheme discussed under 5.1 above. Besides reducing hospital costs, the stated objectives of Advancing Quality are to “save lives, reduce re-admission rates, reduce complications in procedures, and significantly reduce the time patients have to spend in hospital”. The scheme is explained at <http://www.advancingqualitynw.nhs.uk/about.php>

Clinical service. The scheme measures performance in the same five clinical areas as the American PHQID program – acute myocardial infarction (AMI), pneumonia, heart failure, hip and knee replacement, and coronary artery bypass grafts (CABG) – and uses 28 quality indicators to produce a composite quality score (CQS).²⁷ There are plans to also include patient reported outcome measures (that is, health status and quality of life) and patient experience measures, and to extend the measures to include stroke, dementia and psychosis.

Type of incentive. The scheme is a tournament-based scheme where rewards are based on hospitals’ rankings of performance – again very similar to PHQID. The top 25% of hospitals are in the top quartile of performance, and the next 25% are the second top quartile of performance. Those in the top quartile receive a 4% additional payment of the NHS fixed price tariff for the relevant HRG, and those in the second quartile receive a 2% payment. Note that these are double the percentage payments paid by PHQID (2% and 1% for top two deciles) and that there are no financial penalties. Rewards in the first year are up to the value of £3.2 million, and between £10,000 and £300,000 per hospital. Twenty-five per cent of the rewards for AMI are passed to the Ambulance Service. All hospitals in the North West are mandated to participate in the scheme, so there is less change of selection effects that were present in the similar American PHQID scheme.

The additional funding is to be used for quality improvements rather than as financial incentives to individual health professionals. At the outset of the scheme hospital managers agreed that rewards should pass to the clinical teams that earned the bonus. For example, case studies on the website document that hospitals have: provided heart failure booklets to patients; provided nurse and patient education on smoking cessation; added prophylactic antibiotics to formularies; revised clinical guidelines; altered referral pathways within hospitals for the relevant clinical conditions; made better use of heart failure nurses; and improved documentation. The website claims strong support from key stakeholders, including clinicians. The scheme was introduced before the CQUIN and Best Practice Tariffs.

Details of evaluation. A five-year evaluation of this scheme is underway, with first results being reported in early 2011.²⁸ They examine data from the 18 months before and 18 months after the scheme was introduced (2007 to 2010) for 24 hospitals, and compare this with all other hospitals in England. This is to examine the effect of the scheme on in-hospital mortality for pneumonia, AMI and heart failure. A difference-in-difference analysis is used that compares the changes in mortality for the 24 hospitals with the change in mortality for other hospitals in England, and also for hospitals in the same area of the scheme (North West of England). For pneumonia and heart failure, results show that the scheme was associated with reductions in

mortality (1.8% and 1.2% respectively) compared with hospitals in England, but not with other performance indicators for hospitals in the North West. There were no effects on AMI mortality.

5.11 Clinical Practice Improvement Payment System (CPIP) (Australia)

This was introduced in 2008 in Queensland and was initially based on seven performance indicators reflecting process and outcomes of care.^{29, 30} This was introduced as part of a range of clinical governance measures after a patient safety scandal in Bundaberg, which included clinical networks and the introduction of casemix payment in 2007. Development of the indicators included extensive consultation over a year with the clinical networks. The aims are to reduce variation in clinical practice, promote evidence-based care, and enable measurement of achievement of improvements in clinical care.³⁰

Clinical service. When CPIP was introduced there were seven indicators: two process indicators for mental health, two for acute stroke, one for emergency department care, one for chronic obstructive pulmonary disease (COPD), and one for discharge medication. The majority of these indicators were for processes of care rather than outcomes.

Type of incentive. There was a fixed payment per patient, ranging from AUD\$30 for the recording of an anti-psychotic injection for schizophrenia to AUD\$20,000 per COPD patient. Total payments represent an average of between 1% and 3% of revenue for these DRGs. The total funding allocated to the program was around 0.1% of Queensland Health's hospital expenditure (around \$8 million). In the 2010 manual the indicators were revised and now relate to outcomes rather than processes. There are now 10 indicators, including new indicators for diabetes (1), intensive care (1), maternity (3) and renal (2); one of the stroke indicators was removed. Payments are made "per indicator" with an annual cap for each. These indicators are quite different from those used in the US and UK. There was an explicit intention of hospital management passing 80% of the payments to the clinical units involved, though ultimately the hospital districts are responsible for allocation of the payments to the clinical unit cost centres. Splitting the funds across multiple clinical units jointly responsible for performance was at local discretion, and the extent to which this has happened is as yet unclear.

The payments cannot be used as personal income by health professionals but should be invested in quality improvement activities such as "clinical improvement activities, projects or professional development" that include IT, HR, equipment and educational material.³⁰

Details of evaluation. A PhD has been conducted on the set-up and operation of the scheme, but there has been no other formal funded evaluation of the scheme. Results have not yet been published.

6 Review the models as a group and summarise their strength of evidence

The aims of this section are to review the models and evidence presented as a whole, and address the specific questions in the tender. Given the scarcity of evidence, some general lessons learned when designing incentive schemes for hospital clinicians and clinical units are also outlined.

Summary of conclusions of other systematic reviews

Existing systematic reviews of the literature conclude that the quality of the evidence surrounding the effects of financial incentives on clinician behaviour is weak, depends heavily on context, and that the reported effects are sometimes positive and moderate in size. Most reviews find no evidence on unintended consequences.

A recent overview of existing systematic reviews on the impact of financial incentives for clinicians found 24 existing systematic reviews of the literature,² of which four were of sufficient quality to include in the draft of the overview.^{31, 32, 33, 34} Other reviews of the literature and issues – some of which were published after the overview of reviews^{35, 1, 36, 37, 3, 38, 39} – are also relevant. The four reviews included were rated “moderate” in quality and included 32 different evaluations of financial incentives. The majority of these studies were based in primary care clinics or other community settings rather than hospitals. Overall, this suggests that the quality of the evidence for hospitals was not sufficient for them to be included in these systematic reviews. A number of more recent reviews have focused on hospitals.^{1, 40, 12}

Christianson and colleagues (2008)¹ summarised the literature separately for financial incentives aimed at hospitals and found five studies, three of which focused on the Centers for Medicare and Medicaid Services (CMS) Premier Hospital Quality Incentive Demonstration (PHQID).¹ Overall, Christianson concluded that it was difficult to draw definitive conclusions about the use of pay-for-performance in hospital settings. The same American studies in hospital settings are discussed in the reviews by Conrad and Perry (2009),³⁶ Mehrotra and colleagues (2009)⁴⁰ and Ryan (2009).¹² The systematic review by Mehrotra and colleagues (2009) identified 40 pay-for-performance programs in the US that focus on hospital inpatient care. Of eight published papers identified, five did not include a control group and so had a weak study design. The three remaining papers were those discussed above,^{6, 7, 8} of which only one showed a modest effect of pay-for-performance.

Responses to specific questions on funding models

6.1.1 Which incentive-based systems are likely to be most effective, in terms of their initial effect and likely sustainability over time?

There is insufficient evidence to recommend which type of incentive scheme is likely to be most effective. Any decisions on the likely effect of the schemes should therefore be driven by careful design based on the context in which the incentive is being introduced, other components of the intervention (for example IT systems to improve reporting of data), theory and careful consideration of anticipated effects. Given the lack of evidence, piloting and

evaluation – including phased rollout – are essential. There are likely to be important lessons to be learned from the experience with the CPIP in Queensland.

6.1.2 Is the effectiveness and/or sustainability of the models improved when clinicians have the capacity to contribute to the design of the incentives?

Although there are no studies that compare involvement of clinicians with no involvement, it is widely accepted in the literature that clinicians need to be involved in the specification of the performance measures and in examining the likely impact of the incentives. Clinicians involved in the design of the incentives can help identify the potential for unintended consequences. However, since clinicians may have interests to secure additional funding whilst not changing their behaviour, the nature of their involvement must be managed carefully, whilst an appropriate degree of risk sharing between clinicians and the funder is maintained to ensure the strength of incentives for behaviour change.

6.1.3 Are some incentive-based models more likely to create perverse incentives (that is, incentives for clinicians to prioritise payments ahead of patient or service delivery outcomes)?

There is little strong empirical evidence of perverse incentives or unanticipated effects reported in the empirical literature. However, issues such as patient selection and distortion of clinical activity are common anecdotes in the literature. Careful design of each scheme should help minimise these outcomes but it may be difficult to eliminate them entirely.

6.1.4 Do incentive-based models of funding reduce variation in performance or clinical outcomes between clinical units or clinicians, relative to other strategies, such as public disclosure of clinician or clinical unit performance?

There have been no “head-to-head” comparisons of funding models versus other methods for improving clinical performance in hospitals. What is emphasised in the literature is the value of multifaceted interventions that include funding and incentives but are focused on other quality improvement strategies such as education, use of IT, feedback or reporting of performance data. It is widely acknowledged that there is not usually “one” intervention that is the panacea. Many studies of financial incentives are actually a combination of different behaviour change strategies.³⁶ For example, some of the models for hospitals are combined with public reporting of performance data.^{10, 28}

6.1.5 Do incentive-based models reduce or exacerbate inequities in access to health care?

This is partly related to the potential unintended consequences of patient selection and hospitals selecting the more healthy patients so that targets can be met. Evidence from two studies discussed earlier^{9, 11} found no evidence that patients from minorities or who were poor had less access to hospital services, and some evidence that there was more access. The specific context through which hospitals (and physicians) might be able to select patients is important to examine.

6.1.6 Is there an optimal and/or sustainable balance between financial and non-financial incentives to achieve improved and sustained outcomes?

This relates to 6.1.4 above. Although there is no evidence on the role of financial and non-financial interventions, a number of papers and reviews suggest that multi-faceted interventions should be used, with suggestions that financial incentives should be used to support and be based around other or pre-existing quality improvement initiatives. The optimal mix is heavily dependent on context, and the assumptions made about the extent to which clinicians and units are motivated by financial incentives relative to other sources of motivation.

6.1.7 Do incentive-based models increase budget transparency to funders and the public - and increase the likely effect on efficiency outcomes?

Some commentators suggest that one benefit of such schemes, even if they do not change behaviour or improve quality, is to enhance the accountability of clinicians for the care they provide through the publication of data on performance, irrespective of whether a financial incentive is linked to improved performance.⁴¹

Some general principles and considerations for the design of incentive schemes

A number of reviews, opinion pieces, and qualitative research propose some general principles about how financial incentive schemes should be designed to enhance their effectiveness, and also provide guidance on how to design and implement such schemes.^{42, 4, 1, 36} Though the context of out-of-hospital settings are quite different from hospital settings, the lessons learned and general principles are similar and should not be ignored.

6.1.8 Build on what exists already

It is crucial to account for contextual factors and the wider framework within which changes in funding are introduced. Funding schemes evolve over time and can involve a number of other interventions that accompany the funding change – for example, education, quality improvement programs (such as guidelines and feedback), investment in IT and data collection systems, professional support, and administrative support structures.¹ These may all be necessary conditions for ensuring that behavioural change occurs. Professional support is identified as a particular issue for the success of schemes.^{1, 3} The evidence suggests that financial incentives are most likely to be effective in influencing professional behaviour when performance measures and rewards are aligned to the values of the staff being rewarded.^{43, 44, 42}

6.1.9 Define clearly the objectives of the funding change

It is important to define clearly the objectives of the scheme in terms of the intended changes in behaviour.^{42, 4} This depends partly on the measures of quality and the quality of the data. For example, if the objective is to improve quality of care, then rewards should be conditional on observed improvements in quality rather than on the attainment of a minimum level or threshold. Single threshold payments are most likely to benefit clinicians who are already providing high quality of care, as the reward represents a windfall financial gain

with no incentive to change behaviour. Rewarding for the achievement of a threshold of quality does not encourage further improvements in quality beyond the threshold. For clinicians with very low baseline levels of quality, the costs of achieving a relatively high threshold may seem very high and thus they may not respond or participate at all. Yet these are the very groups where substantial gains in quality could potentially be made at lower additional cost relative to those with high levels of quality, where the additional costs of further improvement may be relatively high.^{41, 45, 46, 8} These arguments suggest that:

- 1) payments should be made for successive incremental improvements in quality;
- 2) payments should increase for successive improvements depending on the extent to which the marginal costs of improving quality increase; and
- 3) payments should be based on disease areas or be targeted to clinicians with low baseline levels of quality.³³ One way to do this is to make participation in the scheme mandatory.

6.1.10 Consider potential unintended consequences

Possible unintended consequences include a focus on the remunerated areas at the cost of unremunerated areas (multi-tasking); gaming, such as avoiding very sick and costly patients; or manipulating data to receive payment. These need to be thought through and managed carefully. Although there are anecdotal reports of unintended consequences of financial incentives, few empirical studies examine this directly.^{47, 48, 49, 50, 51} All of these studies are in primary care settings. In a study of hospitals, Glickman (2006)⁶ found that non-measured aspects of AMI care did not change as performance increased. Some suggest that payments should not be based on self-reported data which is subject to manipulation.¹² The use of routine administrative data on outcomes such as mortality and patient safety and data on patient experiences is preferred. There is also a risk (though little evidence) that external incentives may crowd out intrinsic motivation (the desire to do a task well for its own sake) – especially if not closely aligned to doctors' views of their role, or if they imply a lack of trust in the clinician and suggest micro-management.³⁶

6.1.11 The size of the payment

Theory suggests that size will matter, though there is little empirical evidence in this context. In pay-for-performance schemes in the US, rewards are usually less than 10% of revenue, which may not have been large enough to cover the costs of changing behaviour.^{52, 45, 53} Incentives were funded from within existing cost or efficiency savings and so for some schemes no new resources were invested, but were funded from physician revenues withheld by the health maintenance organisation (HMO). This may be viewed as an attempt to reduce costs rather than improve quality, which may not align with the goals of clinicians.¹ Larger incentives may also encourage unintended consequences, though again there is little empirical evidence of this.³⁶

6.1.12 Who receives the payment?

The "target" of financial incentives – that is, the person/s or entity to whom the funding is paid – can include the individual clinician or groups of individuals (such as the team, unit or

department), the organisation (such as the hospital), or a mix of these. Most schemes in the US are aimed at large medical groups or hospitals. There is little empirical evidence to suggest that one or the other should be targeted.^{54, 1} However, payments made to groups raise issues of how the group is defined, and how the financial rewards are used by the group. These issues may mitigate the intended effect of the financial incentives. As the group grows in size the amount rewarded to each individual in the group becomes smaller, weakening the strength of incentive. There is also very little information in the literature on how the group uses the reward, either as personal income or investing in further quality improvements.^{37, 46} A focus on rewarding clinicians at the expense of other health care professionals involved in the care of patients may reduce the effectiveness of teamwork.^{55, 5} Any payments to team members should be equitable in terms of reflecting their respective skills and efforts; if they are not, health professionals who feel unfairly treated may lose motivation to change behaviour. This raises issues about the method of remuneration of nurses and others in the group.

For hospitals a particular issue identified in the American PHQID scheme was that the additional funding was paid to the hospitals but was not distributed to the clinical units or departments or individuals.¹² This raises a number of issues about how hospital managers can “pass on” the additional funding or use it to improve quality in the disease areas for which it was intended. This would be easier if hospital budgets are devolved to a clinical or department level so that extra funding can be passed on and used to invest in services at the discretion of clinicians; or the method of remuneration for clinicians and other hospital staff is flexible so that bonuses can be used as personal income or other benefits (for example, attending conferences). Alternatives include the hospital using the funds to invest in education programs or IT systems, or the purchase of specific equipment aimed to change clinical behaviour.

Such systems can be used to underpin pay-for-performance efforts to ensure that additional funding paid to hospitals reaches those whose behaviour is expected to change. No empirical evaluations of pay-for-performance for hospitals describe the internal mechanisms that can be used to support such incentive schemes. This again raises the issue that performance pay needs to be implemented alongside other interventions within hospitals that support the scheme.

6.1.13 Financial incentives are more likely to have the intended effect where there is one single funder

Where hospitals receive their funding from a number of sources, the effects of incentives introduced by any one funder will be less effective.^{1, 4} Experience from the US suggests that where a clinician contracts with multiple insurers or HMOs, any single HMO has less influence on changing their behaviour. Similarly, where an American hospital treats Medicare patients but also patients from a range of other payers including HMOs, then the strength of incentives introduced in the Medicare PHQID program would be weaker where the percentage of Medicare patients is relatively low. For example, large Australian public hospitals are likely to receive funding from private insurers for private patients, from other insurers for certain types of patients (such as the Transport Accident Commission in Victoria for road traffic accidents, or WorkCover). Furthermore, the continuing split between funding

from State governments and from the Commonwealth (40% and 60% for Local Hospital Networks) will be an issue until the Commonwealth takes over 100% of funding. Having multiple funders also means that funding arrangements are more complex, so clinicians may not be aware of the details of particular schemes. This emphasises the role of communication about the funding scheme to those at whom it is targeted.

7 Summary and discussion

This review has found little rigorous empirical evidence of the effectiveness of changes in funding models for clinicians in hospitals. Those studies that do exist show mixed results, with the more robust studies showing no effect of pay-for-performance targeted at hospitals. One key reason identified for this was that the additional funding was paid to hospitals, and there were few reports as to whether and how this funding had been distributed to clinicians or clinical units within hospitals. The way the hospital uses the funds seems crucial to the success of such schemes.

What is missing from the literature is the combination of a well designed pay-for-performance incentive scheme aimed at hospitals, and an internal mechanism to deliver the funding to those whose behaviour needs to be changed so that quality can improve. Do incentives have a greater impact when they are introduced alongside changes in managerial information systems and accounting systems?

There is a large literature in managerial accounting describing how budgetary processes can be used to improve performance with respect to cost control, and can be linked to clinical information and performance systems.^{56, 57, 58, 59, 60, 61, 62, 63} However, controlled evaluations of their effects could not be identified in this rapid review.

Service Line Management (SLM) is one such model that devolves budgetary responsibility, accountability and decision-making down to clinical units, and attempts to make a clear link between resources and performance.^{64,65} This is currently being promoted by the English NHS and was mentioned in the Darzi Review, *High Quality Care for All: NHS Next Stage Review Final Report*¹⁶ as a way to engage clinicians in decision-making. SLM involves significant cultural change and “bottom up” planning. It also involves a clear system of performance management, including rewards such as autonomy and access to a proportion of budget underspends, and could also be linked to individual performance payments as part of the income of health professionals. The performance pay element could be based on a variety of models which have been used in the US, UK and Germany.⁶⁵ Piloting of SLM began in the NHS in 2008 for 35 NHS Foundation Trusts, with a recent survey of 55 NHS organisations (the total number to whom the survey was sent was not reported, so a response rate cannot be calculated) suggesting that 68% of these organisations have adopted it.⁶⁶ The details of this scheme are sketchy and it is not yet clear whether the NHS models of SLM have been linked with pay-for-performance for individual health professionals. There are documents describing what SLM is and how it can be used, but very little formal evaluation of the scheme.^{64,65,67,68} Some documents suggest that the initial focus was on cost and efficiency rather than on outcomes, with a lack of capacity to report such information at a unit or department level.⁶⁹

Evaluations from the US suggest that the effects of SLM are only short-term,⁷⁰ and others suggest that they disrupt services and do not improve outcomes in the longer term.⁷¹ However, the study designs in these papers are very weak and have no control group. There are very few evaluations, and those that exist have mixed results.⁷¹

The role and existence of clinical budgets is potentially important in being able to pass on financial rewards internally, so rewards can be used for improvements in infrastructure, new

staff etc; and in ensuring that the health professionals are involved in how the funding is used, or they are remunerated directly with the additional funding.

One theme emerging from this review is that hospitals and clinical units and departments in the UK and US have more autonomy and discretion to use funds and be accountable for outcomes, and that this is a potentially effective way to engage clinicians in decisions about quality improvements and funding allocation.

A number of key lessons from the literature should be accounted for when designing new funding models:

- 1) Build on what exists already;
- 2) Define clearly the objectives of the funding change;
- 3) Consider potential unintended consequences;
- 4) The size of the payment;
- 5) Who receives the payment; and
- 6) Financial incentives are more likely to have the intended effect where there is one single funder.

Implications for Australia

There is insufficient evidence to determine what type of incentive scheme or method of payment should be used if such a scheme were to be introduced in Australia. Given the national rollout of casemix payments, initial experience from the UK and US on increasing casemix payments to reflect higher quality care and not paying for “Never Events”, and the importance of building on what already exists, then it would appear that the adjustment of casemix payments should be considered in Australia. This is in addition to the method being used in the Queensland Clinical Practice Improvement Payment System. However, it is important that the financial incentives are delivered to the clinical units or clinicians so that they see some effect of their efforts – either in terms of increased funding for the unit or performance-based pay for the health professionals (not just clinicians) involved – and that there exists some internal accounting mechanism to support this.

Both Commonwealth and State governments have the opportunity to introduce such schemes, aimed at Local Hospital Networks. They also have a unique opportunity to rigorously evaluate the impact of such schemes, and to learn from experience in other countries. That experience highlights that hospitals need to be supported through the introduction of devolved budgets and mechanisms whereby additional funding is distributed to those health professionals and clinical units/departments whose behaviour needs to change. This includes the perception that hospitals in the US and UK have much better routinely collected data on costs and outcomes, including patient experience; however, the Queensland experience suggests that a similar system is feasible in Australia. Availability of good information and IT systems is therefore necessary before such schemes are introduced

Most schemes reviewed have not used the additional funding to supplement the remuneration of clinicians or other health professionals in hospitals; rather, they have used it to invest in infrastructure, equipment, educational programs or new staff. The current way that clinicians and other health professionals are paid within hospitals would need to be examined before bonuses were paid directly to hospital staff. Within a salaried structure, progression up the salary scale is probably already linked to a subjective assessment of performance, but it is unclear whether bonuses can be paid and how much flexibility hospitals have to pay their staff differently from other hospitals. This is complicated by the existence of salaried staff working alongside Visiting Medical Officers, who may be paid by fee-for-service or fixed sessional payments. Furthermore, some salaried staff have rights to private practice, with some income from private patients being returned to clinicians as bonuses or for conferences etc. Direct payments to professionals would therefore require further thought.

Though the evidence is mixed, the existence of such schemes for hospitals is growing. The low quality of most evidence needs to be balanced against the fact that existing schemes have been poorly designed. Improved design might make them more effective. This includes the additional payments going directly to the clinical units or professionals involved, and the existence of an internal accounting mechanism to achieve and support this. The further introduction of any such scheme in Australia needs to be thoroughly evaluated.

Appendix 1. Search Strategy

Search strategy term sets

For this literature review two term sets were used in searching for the topic.

- 1) Hospital setting terms
- 2) Payment methods and income terms
- 3) No evidence filter used and no outcome specified at this stage

MEDLINE, CINAHL, ECONLIT, PsycINFO

Database name *	Dates covered	Date searched	References
CINAHL - ECONLIT	2000 – current	07102010	91
Medline	2000 – current	07102010	2313
PsycINFO®	2000 – current	07102010	30
Cochrane Library	2000 – current	07102010	74
TOTAL Unduplicated Endnote refs			2434

Database name	
Medline	<ol style="list-style-type: none"> 1.exp=hospitals medical staff, hospital 2.clinical service* or clinical unit* or clinician*.ti, ab. 3. #1 or #2 4. exp=salaries and fringe benefits 5. exp=reimbursement mechanisms 6. exp=physician incentive plans 7. exp=fees and charges 8. "pay-for-performance" or "p4p" or "pay for quality" or "payment*" or incentive*" or "incentive payment*" or "financial incentive*" or "quality incentive payment system" or "quality incentive scheme" .ti, ab. 9. #4 or #5 or #6 or #7#8 10. #3 and #9 11. limit 10 to yr="2000 -Current"

Appendix 2. Main characteristics of included evaluations

KP Andriole, LM Prevedello, A Dufault, *et al.* Augmenting the impact of technology adoption with financial incentive to improve radiology report signature times. *J Am Coll Radiol* 2010;7(3):198-204.

Country	US
Study main objective	<p>“combination of technology adoption with financial incentive to determine if ST improvement can be augmented and sustained.</p> <p>Two interventions were technological in nature: (1) a paging portal notification application and (2) a real time PACS-integrated SR report generation system. The third intervention involved behavioural modification via a financial incentive (FI) programme.”</p>
Clinical setting	“751-bed, urban, tertiary care teaching hospital with a radiology department consisting of 124 attending radiologists and 100 radiologist trainees providing radiology coverage for several institutions.”
Incentive payment model	“A \$4000 bonus added to the regular salary paycheck was awarded semi-annually to those attending radiologists meeting the departmental signing goal of median ST <8 hours (which could be met by consistent overall performance) or 80% of reports signed within 16 hours during the 6 months period preceding the award date. The bonus stimulus was instituted 3 months after completion of the rollout of SR throughout the department and lasted for 1 year “
Intervention / Design	Interrupted time series analysis
Outcomes	Radiology reports signature time
Results	“Technology adoption (paging plus speech recognition) reduced median ST from >5 to <1 hour (P< .001) and 80th percentile ST from >24 hour to 15 to 18 hours (P< .001). Subsequent addition of financial incentive further improved 80th percentile ST to 4 to 8 hours (P< .001). The gains in median and 80th percentile were sustained over the final 31 months of the study period.”
Authors' conclusion	“Technology intervention coupled with financial incentive can result in synergistic and sustainable improvement in radiologist report-signing behaviour. The addition of financial incentive leads to better performance than that achievable though technology alone.”

1

SW Glickman , F-S Ou, ER DeLong, *et al.* Pay-for-performance, quality of care, and outcomes in acute myocardial infarction. *JAMA* 2007; 297(21):2373-80.

Country	US
Study objective	main "To determine if pay-for-performance was associated with either improved processes of care and outcomes or unintended consequences for acute myocardial infarction at hospitals participating in the CMS pilot project."
Clinical setting	"An observational, patient-level analysis of 105,383 patients with acute non-ST-segment elevation myocardial infarction enrolled in a national quality-improvement initiative."
Incentive payment model	CMS-Premier Hospital Quality Incentive Demonstration (PHOID)*
Intervention Design	/ case/control study "Patients were treated between July 1, 2003, and June 30, 2006, at 54 hospitals in the CMS program and 446 control hospitals."
Outcomes	"6 process measures used by CMS in the care of patients with non-ST-segment elevation acute myocardial infarction including: aspirin at arrival and discharge, beta blocker at arrival and discharge, angiotensin-converting enzyme inhibitor or angiotensin receptor blocker for left ventricular systolic dysfunction, and smoking cessation counselling. 8 additional process measures (ie performing an electrocardiogram within 10 minutes of emergency department presentation, use of unfractionated or low-molecular-weight heparin, use of glycoprotein IIb/ IIIa inhibitors, cardiac catheterization within 48 hours, clopidogrel at discharge, lipid-lowering medication at discharge, dietary modification counselling, and referral for cardiac rehabilitation)"
Results	"Among treatments subject to financial incentives, there was a slightly higher rate of improvement for 2 of 6 targeted therapies at pay-for-performance vs control hospitals (odds ratio [OR] comparing adherence scores from 2003 through 2006 at half-year intervals for aspirin at discharge, 1.31; 95% confidence interval [CI], 1.18-1.46 vs OR, 1.17; 95% CI, 1.12-1.21; P=.04) and for smoking cessation counselling (OR, 1.50; 95% CI, 1.29-1.73 vs OR, 1.28; 95% CI, 1.22-1.35; P=.05). There was no significant difference in a composite measure of the 6 CMS rewarded therapies between the 2 hospital groups (change in odds per half-year period of receiving CMS therapies: OR, 1.23; 95% CI, 1.15-1.30 vs OR, 1.17; 95% CI, 1.14-1.20; P =.16). For composite measures of acute myocardial infarction treatments not subject to incentives, rates of improvement were not significantly different (OR, 1.09; 95% CI, 1.05-1.14 vs OR, 1.08; 95% CI, 1.06-1.09; P=.49). Overall, there was no evidence that improvements in in-hospital mortality were incrementally greater at pay-for- performance sites (change in odds of in-hospital death per half-year period, 0.91; 95% CI, 0.84-0.99 vs 0.97; 95% CI, 0.94-0.99; P=.21)"

Authors' conclusion "Among hospitals participating in a voluntary quality-improvement initiative, the P4P program was not associated with a significant incremental improvement in quality of care or outcomes for acute myocardial infarction. Conversely, we did not find evidence that pay-for-performance had an adverse association with improvement in processes of care that were not subject to financial incentives. Additional studies of P4P are needed to determine its optimal role in quality-improvement initiatives."

SR Grossbart. What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Med Care Res Rev* 2006;63(1 Suppl): 29S-48S.

Country	US
Study main objective	"This study tests the question: Did hospitals participating in the CMS/Premier HQID Project have significantly different rates of improvement in quality performance compared to similar hospitals within the same system that did not participate in the 1st year of the voluntary demonstration project?"
Clinical setting	"The study is based on care provided to 28,925 hospitalized patients with one of three clinical conditions (AMI, heart failure, or pneumonia) who received care during the 24-month period from October 2002 through September 2004 from 10 hospitals within Catholic Healthcare Partners."
Incentive payment model	CMS-Premier Hospital Quality Incentive Demonstration (PHQID)*
Intervention / Design	"The study uses a non random sample of hospitals that were eligible to participate in the CMS/Premier HQID Project to assess the impact of the demonstration. A test group of four acute care hospitals within Catholic Healthcare Partners that are participating in this demonstration project was compared with a control group of six hospitals in the same health care system that chose not to participate in the project."
Outcomes	"The study limited analysis to three of the five clinical areas that are included in the Premier demonstration: AMI, heart failure, and pneumonia."
Results	"The overall composite quality score for the test group of participating hospitals was 89.7 percent compared with 85.6 percent ($p < .001$). For AMI, the composite score for the cohort of participating hospitals was 94.2 percent versus 90.6 percent for nonparticipants ($p < .001$). For heart failure, the composite score for the cohort of participating hospitals was 87.0 percent versus 84.3 percent for nonparticipants ($p = .006$). For pneumonia, the composite score for the cohort of participating hospitals was 87.3 percent versus 80.5 percent for nonparticipants ($p < .001$)."
Authors' conclusion	"Although the incentives are extremely small, the findings show that participation in the pay-for-performance initiative had a significant impact on the rate and magnitude of performance improvement. The project led to marked improvement in the quality of clinical process delivery and accelerated the adoption of evidence-based practices."

AK Jha, EJ Orav, AM Epstein. The effect of financial incentives on hospitals that serve poor patients. *Ann Intern Med* 2010;153:299-306.

Country	US
Study main objective	"To determine how financial incentives for quality performance affect hospitals with more poor patients compared with those with fewer poor patients."
Clinical setting	"251 hospitals that participated in the Premier Hospital Quality Incentive Demonstration program and a national sample of 3017 hospitals."
Incentive payment model	CMS-Premier Hospital Quality Incentive Demonstration (PHQID)*
Intervention / Design	"In our primary analyses, we examined whether disproportionate-share index was associated with baseline performance on the quality indicators, change in performance, and terminal performance among both the pay-for-performance hospitals and the national sample. Both bivariate and multivariate models were used in each analysis to adjust for the important hospital characteristics that we thought might be potential confounders. Next, we directly tested whether the relationship between disproportionate-share index and each of the outcomes varied between the pay-for-performance hospitals and the national sample of hospitals."
Outcomes	Acute myocardial infarction Congestive heart failure Pneumonia
Results	"Among both pay-for-performance hospitals and those in the national sample, hospitals with more poor patients (high disproportionate-share index values) had lower baseline performance than did those with fewer poor patients. A high disproportionate-share index was associated with greater improvements in performance for acute myocardial infarction and pneumonia but not for congestive heart failure, and the gains were greater among hospitals that received financial incentives than among the national sample. After 3 years, hospitals that had more poor patients and received financial incentives caught up for all 3 conditions, whereas those with more poor patients among the national sample continued to lag."
Authors' conclusion	"No evidence indicated that financial incentives widened the gap in performance between hospitals that serve poor patients and other hospitals. Pay-for-performance programs may be a promising quality improvement strategy for hospitals that serve poor patients."

Y-H Li, W-C Tsai W-C, M Khan, *et al.* The effects of pay-for-performance on tuberculosis treatment in Taiwan. *Health Policy Plan* 2010;25:334-41.

Country	Taiwan
Study main objective	"To investigate the effectiveness of pay-for-performance programme in terms of TB cure rate and length of treatment."
Clinical setting	"Hospitals part of TB P4P programme"
Incentive payment model	"Taiwan's pay-for-performance system is based on monetary incentives for physicians and hospitals based on the outcomes of TB treatment."The hospitals in Taiwan could choose to participate in the P4P on TB programme if they satisfied four eligibility criteria: (1) the BNHI has a contract with the hospital for the provision of services; (2) it is approved by the Centers for Disease Control of Taiwan (TCDC); (3) the participating physicians have specialist licensing in infectious disease, tuberculosis, or have related training/certification; and (4) the hospital has more than 100 new cases under treatment (at any point in time), it has a full-time TB case manager in the hospital". The incentive scheme was designed around four chronological stages of TB treatment that last 12 months. Points were earned by hospitals, physicians and case managers for the identification of cases (stage 1) to the number of cases that were cured in the last three stages of treatment. The value of each point is not stated in the paper.
Intervention / Design	"This retrospective study obtained information on all TB cases in the national data sets of Taiwan for the years 2002 to 2005. The number of cases in pre-P4P years (2002 and 2003) was 25,754, compared with 33,536 in the post-P4P implementation years (2004 and 2005). The effectiveness of the programme was evaluated by comparing the TB cure rate and length of treatment before and after the implementation of the P4P programme, and between participating and non-participating hospitals. Logistic regression analysis was conducted to explore the factors affecting TB patients' cure rate within a 12-month treatment period."
Outcomes	TB cure rate Treatment length (months)
Results	"The cure rate and the average length of treatment before the implementation of P4P were 46.9% and 256.24 days, respectively, compared with 63.0% and 249.74 days after implementation of P4P. The cure rate and length of treatment in P4P hospitals were 68.1% and 249.13 days, respectively, compared with 42.4% and 53.71 days in non-P4P hospitals."
Authors' conclusion	"Both the cure rate and average length of treatment for cured cases improved significantly after the implementation of the P4P on TB programme in Taiwan. Compared with non-P4P hospitals, P4P hospitals had significantly better treatment outcomes. These results confirm that P4P on TB is an effective incentive structure for the treatment of TB in Taiwan."

P Lindenauer, D Remus, S Roman, *et al.* Public reporting and pay-for-performance in hospital quality improvement. *New Engl J Med* 2007;356(5):486-96.

Country	US
Study main objective	"To determine the incremental effect of pay-for-performance, we measured improvements in hospital quality that occurred when financial incentives were combined with public reporting and compared these improvements with gains associated with public reporting alone."
Clinical setting	"We measured changes in adherence to 10 individual and 4 composite measures of quality over a period of 2 years at 613 hospitals that voluntarily reported information about the quality of care through a national public-reporting initiative, including 207 facilities that simultaneously participated in a P4P demonstration project"
Incentive payment model	CMS-Premier Hospital Quality Incentive Demonstration (PHQID)*
Intervention/Design	"We compared the P4P hospitals with the 406 hospitals with public reporting only (control hospitals). We used multivariable modelling to estimate the improvement attributable to financial incentives after adjusting for baseline performance and other hospital characteristics."
Outcomes	<p>Acute myocardial infarction</p> <p>Percentage of patients who were given aspirin on arrival</p> <p>Percentage of patients who were given an ACE inhibitor or ARB for left ventricular systolic dysfunction</p> <p>Percentage of patients for whom aspirin was prescribed at discharge</p> <p>Percentage of patients who were given a beta-blocker on arrival</p> <p>Percentage of patients for whom a beta-blocker was prescribed at discharge</p> <p>Heart failure</p> <p>Percentage of patients who were assessed for left ventricular function</p> <p>Percentage of patients who were given an angiotensin-converting enzyme inhibitor or angiotensin-receptor blocker for left ventricular systolic dysfunction</p> <p>Pneumonia</p> <p>Percentage of patients who were assessed for oxygenation</p> <p>Percentage of patients who were given initial antibiotics within 4 hours after arrival.</p> <p>Percentage of patients who were assessed and given pneumococcal vaccination"</p>
Results	"As compared with the control group, pay-for-performance hospitals showed greater improvement in all composite measures of quality, including measures of care for heart failure, acute myocardial infarction, and pneumonia and a composite of 10 measures. Baseline performance was inversely associated with improvement; in pay-for-performance hospitals, the improvement in the composite of all 10 measures was 16.1% for

hospitals in the lowest quintile of baseline performance and 1.9% for those in the highest quintile ($P < 0.001$). After adjustments were made for differences in baseline performance and other hospital characteristics, pay-for-performance was associated with improvements ranging from 2.6 to 4.1% over the 2-year period."

Authors' conclusion

"Hospitals engaged in both public reporting and pay-for-performance achieved modestly greater improvements in quality than did hospitals engaged only in public reporting. Additional research is required to determine whether different incentives would stimulate more improvement and whether the benefits of these programs outweigh their costs."

T Nahra, K Reiter, R Hirth, J Shermer, J Wheeler. Cost-effectiveness of hospital pay-for-performance incentives. *Med Care Res Rev* 2006;63(1 suppl):49S-72S.

Country	US
Study main objective	To examine the cost-effectiveness of a financial incentive to improve quality in hospitals
Clinical setting	"Heart care" discharges in 85 Hospitals
Incentive payment model	Participating Hospital Agreement (PHA) Incentive Program. "Incentive payments are calculated as a percentage add-on to hospitals' inpatient diagnosis-related group (DRG) reimbursements the insurer. The maximum possible add-on for heart related care has increased from 1.2 percent of a hospital's inpatient DRG reimbursements in 2000-2002 to 2 percent of a hospital's inpatient DRG reimbursements in 2003. The actual add-on is determined by multiplying the maximum possible add-on by a hospital's individual performance score. In 2002 and 2003, hospitals only received heart-care-related incentive payments if they achieved an established minimum performance score, or threshold, reflecting at least the median performance level of all participating hospitals. Continuous improvement in hospital quality is encouraged through increases in established thresholds. Although actual performance data may reflect declines in quality, established thresholds are not allowed to decrease." "Prior to 2002, hospital incentives were determined by ranking hospitals in order of overall performance scores, dividing the hospitals into quartiles, and setting fixed incentive rates for each of the resulting groups."
Intervention/Design	Longitudinal follow up of hospitals after the program was introduced with economic modelling of quality-adjusted life years (QALYs).
Outcomes	(1) the percentage of eligible acute myocardial infarction (AMI) patients receiving aspirin orders at discharge; (2) the percentage of eligible AMI patients receiving beta blocker prescriptions at discharge; and (3) the percentage of eligible congestive heart failure (CHF) patients receiving ACE inhibitor prescriptions at discharge. For each indicator, the percentage change was multiplied by the annual number of patients, linked to evidence to estimate additional life years and health related quality of life to estimate QALYs.
Results	Between \$12,987 and \$30,081 per QALY
Authors' conclusion	None

Ryan AM. Effects of the Premier Hospital Quality Incentive Demonstration on Medicare patient mortality and cost. *Health Serv Res* 2009;44(3):821-42.

Country	US
Study main objective	"To evaluate the effects of the Premier Inc. and Centers for Medicare and Medicaid Services Hospital Quality Incentive Demonstration (PHQID), a public quality reporting and pay-for-performance (P4P) program, on Medicare patient mortality, cost, and outlier classification."
Clinical setting	"Data from 3,570 acute care hospitals between 2000 and 2006."
Incentive payment model	CMS-Premier Hospital Quality Incentive Demonstration (PHQID)*
Intervention/Design	"Panel data econometric methods are applied to a retrospective cohort of 11,232,452 admissions from 6,713,928 patients with principal diagnoses of acute myocardial infarction (AMI), heart failure, pneumonia, or a coronary-artery bypass grafting (CABG) procedure."
Outcomes	Three estimators are used to evaluate the effects of the PHQID on risk-adjusted (RA) mortality, cost, and outlier classification in the presence of unobserved selection, resulting from the PHQID being voluntary: fixed effects (FE), FE estimated in the subset of hospitals eligible for the PHQID, and difference-in-difference-in-differences.
Results	"This analysis found no evidence that the PHQID had a significant effect on RA 30-day mortality or RA 60-day cost for AMI, heart failure, pneumonia, or CABG and weak evidence that the PHQID increased RA outlier classification for heart failure and pneumonia."
Authors' conclusion	"By not reducing mortality or cost growth, this study suggests that the PHQID has made little impact on the value of inpatient care purchased by Medicare."

Ryan AM. Has pay-for-performance decreased access for minority patients? *Health Serv Res* 2010;45(1):6-23.

Country	US
Study main objective	"To examine whether the CMS and Premier Inc. Hospital Quality Incentive Demonstration (PHQID), a hospital-based pay-for-performance (P4P) and public quality reporting program, caused participating hospitals (1) to avoid treating minority patients diagnosed with acute myocardial infarction (AMI), heart failure, and pneumonia and (2) to avoid providing coronary artery bypass graft (CABG) to minority patients diagnosed with AMI."
Clinical setting	"Data from 3,570 acute care hospitals between 2000 and 2006."
Incentive payment model	CMS-Premier Hospital Quality Incentive Demonstration (PHQID)*
Intervention/Design	Pooled cross-sectional study
Outcomes	"We test for differences in the conditional probability of receiving care at PHQID hospitals for AMI, heart failure, and pneumonia before and after implementation of the PHQID between white and minority patients. We also test for differences in the conditional probability that white and minority patients diagnosed with AMI receive CABG in hospitals participating, and not participating, in the PHQID before and after the implementation of the PHQID."
Results	"We find little evidence that the PHQID reduced access for minority patients: only "Other Race" beneficiaries had a significant reduction in adjusted admissions to PHQID hospitals in the postperiod, and only for AMI. Only marginally significant (p<.10) evidence of a reduction in CABG was found, also occurring for Other Race beneficiaries."
Authors' conclusion	"Despite minimal evidence of minority patient avoidance in the PHQID, monitoring of avoidance should continue for P4P programs."

9 References

1. Christianson JB, Leatherman S, Sutherland K. Lessons From evaluations of purchaser pay-for-performance programs: A review of the evidence. *Med Care Res Rev* 2008;65:5S-35S.
2. Eccles M, Shepperd S, Scott A, et al. An overview of reviews evaluating the effects of financial incentives in changing healthcare professional behaviours and patient outcomes (Protocol). *Cochrane Database of Systematic Reviews*, Issue 7, 2010.
3. Van Herck P, De Smedt D, Annemans L, et al. Systematic review: Effects, design choices, and context of pay-for-performance in health care. *BMC Health Services Research* 2010;10:247.
4. Scott A, Naccarella L, Furler J, et al. *Using financial incentives to improve the quality of primary care in Australia. Final Report*. Canberra: Australian Primary Health Care Research Institute, Australian National University, 2010.
5. Prendergast C. The provision of incentives in firms. *J Econ Lit* 1999;37:7-63.
6. Glickman SW, Ou F-S, DeLong ER, et al. Pay-for-performance, quality of care, and outcomes in acute myocardial infarction. *JAMA* 2007;297(21):2373-80.
7. Grossbart S. What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Med Care Res Rev* 2006;63:29S-48S.
8. Lindenauer P, Remus D, Roman S, et al. Public reporting and pay-for-performance in hospital quality improvement. *New Engl J Med* 2007;356:486-96.
9. Jha AK, Orav EJ, Epstein AM. The effect of financial incentives on hospitals that serve poor patients. *Ann Intern Med* 2010;153:299-306.
10. Ryan AM. Effects of the Premier Hospital Quality Incentive Demonstration on Medicare patient mortality and cost. *Health Serv Res* 2009;44:821-42.
11. Ryan AM. Has pay-for-performance decreased access for minority patients? *Health Serv Res* 2010;45:6-23.
12. Ryan AM. Hospital-based pay-for-performance in the United States. *Health Econ* 2009;18:1109-13.
13. Nahra T, Reiter K, Hirth R, Shermer J, Wheeler J. Cost-effectiveness of hospital pay-for-performance incentives. *Med Care Res Rev* 2006;63:49S-72S.
14. Li Y-H, Tsai W-C, Khan M, et al. The effects of pay-for-performance on tuberculosis treatment in Taiwan. *Health Policy Plan* 2010;25:334-41.
15. Andriole KP, Prevedello LM, Dufault A, et al. Augmenting the impact of technology adoption with financial incentive to improve radiology report signature times. *J Am Coll Radiol* 2010;7(3):198-204.
16. Department of Health. High Quality Care For All: NHS Next Stage Review final report. London: Department of Health, 2008.
17. Department of Health. Using the Commissioning for Quality and Innovation (CQUIN) payment framework for the NHS in England 2009/10. London: Department of Health, 2008.
18. Department of Health. Using the Commissioning for Quality and Innovation (CQUIN) payment framework – an addendum to the 2008 policy guidance for 2010/11. London: Department of Health, 2010.

-
19. McNair P, Borovnicar D, Jackson T, Gillett S. Prospective payment to encourage system wide quality improvement. *Med Care* 2009;47(3):272-8.
 20. Department of Health. *Payment by Results Guidance for 2010-11*. London: Department of Health, 2010.
 21. Department of Health. *Payment by results: Draft guidance for 2011-12*. Leeds: Department of Health, 2010.
 22. Department of Health. Best practice tariff (BPT) for fragility hip fracture care user guide. London: Department of Health, 2010.
 23. Wilson H, Harding K, Sahota O. Best practice tariff for hip fracture – making ends meet. *British Geriatrics Society Newsletter*, June 2010.
 24. Pronovost PJ, Goeschel CA, Wachter RM. The wisdom and justice of not paying for “preventable complications”. *JAMA* 2008;299:2197-9.
 25. Rosenthal MB. Nonpayment for performance? Medicare's new reimbursement rule. *New Engl J Med* 2007;357:1573-5.
 26. Zhan C, Elixhauser A, Friedman B, Houchens R, Chiang Y. Modifying DRG-PPS to include only diagnoses present on admission: Financial implications and challenges. *Med Care* 2007;45:288-291.
 27. NHS North West. Composite quality score and outcome methodologies: Year One. NHS Northwest and Premier Advancing Hospital Quality Programme, 2010.
 28. Sutton M, Nikolova S, McDonald R. Impact of a pay-for-quality programme in the North West of England on patient mortality risk. Health Economists' Study Group, York, 2011.
 29. Duckett S, Daniels S, Kamp M, et al. Pay-for-performance in Australia: Queensland's new Clinical Practice Improvement Payment. *J Health Serv Res Policy* 2008;13:174-7.
 30. Queensland Health. Clinical practice improvement payment user guide version 2. Brisbane: Queensland Health, 2010.
 31. Akbari A, Mayhew A, Al-Alawi M, et al. Interventions to improve outpatient referrals from primary care to secondary care. *Cochrane Database of Systematic Reviews*, Issue 1, 2008.
 32. Gosden T, Forland F, Kristiansen I, et al. Impact of payment method on behaviour of primary care physicians: A systematic review. *Journal of Health Services Research and Policy* 2001;6:44-55.
 33. Petersen L, Woodard L, Urech T, Daw C, Sookanan S. Does pay-for-performance improve the quality of health care? *Ann Intern Med* 2006;145:265-72.
 34. Sturm H, Austvoll-Dahlgren A, Aaserud M, et al. Pharmaceutical policies: Effects of financial incentives for prescribers. *Cochrane Database of Systematic Reviews*, Issue 3: CD006731, 2007.
 35. Christianson JB, Leatherman S, Sutherland K. *Financial incentives, healthcare providers and quality improvements: A review of the evidence*. QQUIP (Quest for Quality and Improved Performance), The Health Foundation, 2007.
 36. Conrad DA, Perry L. Quality-based financial incentives in health care: Can we improve quality by paying for it? *Annu Rev Publ Health* 2009;30:357-71.
 37. Scott A, Schurer S, Jensen PH, Sivey P. The effects of an incentive program on quality of care in diabetes management. *Health Econ* 2009;18:1091-1108.

-
38. Scott I. Pay-for-performance in health care: strategic issues for Australian experiments. *Med J Australia* 2007;187:31-5.
 39. Frolich A, Talavera JA, Broadhead P, Dudley RA. A behavioral model of clinician responses to incentives to improve quality. *Health Policy* 2007;80:179-93.
 40. Mehrotra A, Damberg CL, Sorbero MES, Teleki SS. Pay-for-performance in the hospital setting: What is the state of the evidence? *Am J Med Qual* 2009;24(1):19-28.
 41. Bokhour BG, Burgess JF Jr, Hook JM, et al. Incentive implementation in physician practices: A qualitative study of practice executive perspectives on pay-for-performance. *Med Care Res Rev* 2006;63:73S-95S.
 42. Custers T, Hurley J, Klazinga N, Brown A. Selecting effective incentive structures in health care: A decision framework to support health care purchasers in finding the right incentives to drive performance. *BMC Health Serv Res* 2008;8:66.
 43. McDonald R, Campbell S, Lester H. Practice nurses and the effects of the new general practitioner contract in the English National Health Service: The extension of a professional project? *Soc Sci Med* 2009;68:1206-12.
 44. Marshall M, Smith P. Rewarding results: using financial incentives to improve quality. *Quality and Safety in Health Care* 2003;12:397-8.
 45. Rosenthal MB, Frank RG, Li Z, Epstein AM. Early experience with pay-for-performance: from concept to practice. *JAMA* 2005;294(14):1788-93.
 46. Rosenthal MB, Fernandopulle R, Song HR, Landon B. Paying for quality: providers' incentives for quality improvement. *Health Affair* 2004;23:127-41.
 47. McDonald R, Roland M. Pay-for-performance in primary care in England and California: Comparison of unintended consequences. *Ann Fam Med* 2009;7:121-7.
 48. Campbell SM, Kontopantelis E, Reeves D, et al. Changes in patient experiences of primary care during health service reforms in England between 2003 and 2007. *Ann Fam Med* 2010;8:499-506.
 49. Steel N, Maisey S, Clark A, Fleetcroft R, Howe A. Quality of clinical primary care and targeted incentive payments: An observational study. *Br J Gen Pract* 2007;57:449-54.
 50. Sutton M, Elder R, Guthrie B, Watt G. Record rewards: the effects of targeted quality incentives on the recording of risk factors by primary care providers. *Health Econ* 2009;19(1):1-13.
 51. Gravelle H, Sutton M, Ma A. Doctor behaviour under a pay-for-performance contract: Further evidence from the quality and outcomes framework. CHE Research Paper 34. Centre for Health Economics, University of York, 2008.
 52. Mullen K, Frank RG, Rosenthal MB. *Can you get what you pay for? Pay-for-performance and the quality of healthcare providers*. National Bureau of Economic Research, Working Paper No.14886, 2009.
 53. Rosenthal MB, Landon BE, Normand S-LT, Frank RG & Epstein AM. Pay-for-performance in commercial HMOs. *N Engl J Med* 2006;355:1895-1902.
 54. Naccarella L, Scott A, Furler J, et al. Narrative literature review on incentives for primary health care team service provision: learning and working together in

-
- primary health care. Canberra: Australian Primary Health Care Research Institute, Australian National University, 2010.
55. McDonald R, Harrison S, Checkland K, Campbell SM, Roland M. Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study. *BMJ* 2007;334:1357-9.
 56. Scarparo S. The integration of clinical and costing information: a comparative study between Scotland and Sweden. *Financial Accountability & Management* 2006;22:133-55.
 57. Lamb SM, David TJ. Playing with fire: An experiment in clinical budgeting. *BMJ (Clin Res Ed)* 1985;290:650-651.
 58. Macinati MS. NPM reforms and the perception of budget by hospital clinicians: Lessons from two case-studies. *Financial Accountability & Management* 2010;26(4):422-42.
 59. Steele R. Clinical budgeting and costing: friend or foe? *Br Med J (Clin Res Ed)* 1984;288(6429):1549-51.
 60. Wickings I, Coles JM, Flux R, Howard, L. Review of clinical budgeting and costing experiments. *BMJ (Clin Res Ed)* 1983;286:575-8.
 61. Abernethy MA, Chua WF, Grafton J, Mahama H. 2006. Accounting and control in health care: Behavioural, organisational, sociological and critical perspectives. In: Chapman CS, Hopwood AG, Shields MD (eds). *Handbooks of Management Accounting Research*. Elsevier 2008.
 62. Abernethy MA. 1996. Physicians and resource management: The role of accounting and non-accounting controls. *Financial Accountability & Management* 12:141-56.
 63. Lapsley I. 2001. The accounting-clinical interface: Implementing budgets for hospital doctors. *Abacus* 37:79-109.
 64. Bury E, Carter KS, Feigelman S, Grant JM. 2007. How service-line management can improve hospital performance. McKinsey. *Health International* 7:54-65.
 65. Monitor 2009. Toolkit 2: Working towards service-line management: organisational change and performance management. London: Monitor.
 66. CIMA 2009. A picture of health: Service cost data in performance management in the NHS. London: Chartered Institute of Management Accountants.
 67. Iannuccillo KH, Butler GA, McGuirl JE. 1998. Developing a service line approach to quality improvement. *J Perinat Neonat Nur* 12(1):31-41.
 68. Jain AK, Thompson JM, Kelley SM, Schwartz RW. 2006. Fundamentals of service lines and the necessity of physician leaders. *Surg Innov* 13(2):136-44.
 69. The Nuffield Trust 2009. Improving productivity and performance in the NHS: Service-line management: Improving efficiency and quality. London: The Nuffield Trust.
 70. Greenberg GA, Rosenheck RA, Charns MP. 2003. From profession-based leadership to service line management in the Veterans Health Administration: Impact on mental health care. *Med Care* 41:1013-23.
 71. Byrne MM, Charns MP, Parker VA, Meterko MM, Wray NP. 2004. The effects of organization on medical utilization: An analysis of service line organization. *Med Care* 42:28-37.