



Review of Incentivised Models of Care

Full Review and Technical Report

Professor Fang Lee Cooke
Dr Alex Veen

July 2016

Report # 096-0716-R02

This report is prepared by:

Project team:

Professor Fang Lee Cooke

fang.cooke@monash.edu

Dr Alex Veen

alex.veen@monash.edu

Department of Management

Monash Business School

Monash University

Melbourne, Australia

Acknowledgements

This project is funded by WorkSafe Victoria and the Transport Accident Commission, through the Institute for Safety, Compensation and Recovery Research.

ISCRR is a joint initiative of WorkSafe Victoria, the Transport Accident Commission and Monash University.

Table of Contents

Table of Contents.....	2
List of Tables	5
List of Figures	5
List of Abbreviations	6
Executive Summary.....	7
Key findings and conclusions	7
1 Introduction	11
2 Objectives of the review study	12
2.1 Research questions.....	13
2.2 Incentivised care: What is it? Why use it?.....	14
2.2.1 Pay-for-Performance (P4P)	15
3 Research methodology	16
3.1 Sampling strategy	16
3.2 Challenges of reviewing incentivised care models.....	18
4 Findings	19
4.1 P4P in five US states	28
4.1.1 Iowa.....	30
4.1.2 Minnesota	33
4.1.3 Oklahoma	37
4.1.4 Utah.....	40
4.1.5 Vermont	43
4.2 Comparative studies of P4P models.....	46
4.3 A competitive tendering P4P initiative in Minnesota.....	49
4.4 An alternative quality-based payment model for Minnesotan nursing home care.....	53
4.5 Successful use of P4P in San Diego in the 1980s – the importance of control groups	56
4.6 Reflections on the nursing home sector case studies	59
5 Summary of key findings and discussion	63
5.1 Existing models and their effectiveness	63
5.2 The purpose of incentivised care models.....	64
5.3 Foundations of incentivised care models.....	65
5.4 What areas of care should be incentivised?.....	66
5.5 Designing incentivised care models	67
5.6 Evaluation	72

6	Recommendations	72
7	Conclusions	73
7.1	Summary.....	73
7.2	Limitations of the current study	75
	List of references.....	77
	Appendices.....	81
	Appendix 1 – Summary of states implementing Medicaid-sponsored P4P in nursing homes 2001-2009	81
	Appendix 2 – Considerations for designing and implementing an incentivised model.....	82
A1	Incentivising care: when is it appropriate?	82
A2	Measuring the quality of care	86
A3	Objectives of incentivised care measures	90
A3.1	Target Incentives	90
A3.2	What areas of care need to be improved?	91
A3.3	What behavioural reactions do incentives target?	91
A3.4	Contrasting characteristics of care incentives	92
A3.4.1	Quality vs. cost	93
A3.4.2	Intrinsic vs. extrinsic motivation.....	94
A3.4.3	Rewards vs. penalties	94
A3.4.4	Financial vs. non-financial incentives	95
A3.4.5	Direct vs. indirect incentives	95
A3.4.6	Long- vs. short-term oriented incentives	96
A3.4.7	Targeted vs. broad improvements in care	97
A3.4.8	Lifting performance across providers.....	99
A4	Types of incentives	101
A5	Effectiveness of incentivised care arrangements.....	107
A6	Risks and unintended consequences of incentivised care	113
A7	Implementing incentivised care models.....	116
A7.1	Pre-conditions for successful implementation	117
A7.1.1	Designing the incentive model	117
A7.1.2	Stakeholder management	122
A7.1.3	Change management processes.....	123
A7.1.4	Investments in the quality of care.....	124
A7.1.5	Legal implications	126

A7.1.6	Monitoring, feedback and evaluation	126
A7.1.7	Bureaucracy.....	127
A7.2	Development of quality management structures.....	127
Appendix 3 – Recommendations for designing, implementing and evaluating an incentivised model		130
A1	Care problem identification.....	130
A2	Co-design	131
A2.1	Collaborative stakeholder co-development.....	131
A2.2	Quality and incentive measures.....	132
A2.3	Compensation for investment and financial sustainability.....	133
A3	Implementation	134
A4	Measurement	135
A5	Evaluation	135
Appendix 4 – Prospects for further research contribution.....		137

List of Tables

Table 1. Description of P4P programs in nursing homes in the United States	22
Table 2. Characteristics of Nursing Home P4P Programs in five US States	28
Table 3. P4P incentive structure under the redesigned Iowa model	31
Table 4. Performance measures and approaches adopted by the States in their P4P models.....	47
Table 5. Quality points system.....	53
Table 6. Proposed payment rate as percentage of facilities per diem cost, by cost and quality tier... 55	
Table 7. What to measure in order to ascertain quality of care.....	87
Table 8. A classification of behaviours activated by different incentive measures.....	91
Table 9. Contrasting characteristics of existing incentives model.....	93
Table 10. Three methods for structuring incentives – how to target high/low quality providers	100
Table 11. Incentive models used in healthcare to change or enable an actor’s behaviour	101
Table 12. Characteristics of financial interventions.....	103
Table 13. How effective are most commonly used incentives across various healthcare systems? . 105	
Table 14. Overview of conclusions of systemic reviews and longitudinal studies on incentivised care arrangements.....	110
Table 15. Typology of behaviours of care providers in relation to incentivised care arrangements . 113	
Table 16. Characteristics of models adopting penalties and/or rewards	119
Table 17. Recommendations for the design of a P4P model.....	122

List of Figures

Figure 1. A step-by-step approach to designing and implementing a P4P model.....	45
Figure 2. Outlines the four phases of the tender-based P4P model in Minnesota	52
Figure 3. Key processes for the introduction of incentivised models of care.....	72
Figure 4. Decision tree for introduction of incentivised care measures.....	83
Figure 5. Steps for the development and implementation of an incentivised model of care.....	130

List of Abbreviations

ADL	Activities of daily living
CPOs	Care Provider Organisations
EBP	Evidence-based practices
FFS	Fee-for-service
HDSG	Health and Disability Strategy Group
ISCRR	Institute for Safety, Compensation and Recovery Research
MDS	Minimum Data Set
NDIA	National Disability Insurance Agency
NDIS	National Disability Insurance Scheme
NH	Nursing home
P4P	Pay-for-Performance
QI	Quality indicator
QOL	Quality of life
SCI	Spinal cord injury
TAC	Transport Accident Commission
TBI	Traumatic brain injury

Executive Summary

The compensable disability care service sector has undergone considerable change in recent years and is likely to encounter radical change resulting from the implementation of the National Disability Insurance Scheme (NDIS). These changes require the realignment of business strategy and care ethos, service delivery and funding models. The Transport Accident Commission (TAC) and Health and Disability Strategy Group (HDSG) are exploring ways in which they can facilitate the best outcomes for their client group, especially maximising independence. This includes identifying or developing innovative models for care services that are cost effective, whilst improving the quality of client experience and outcomes. To facilitate this, a review was commissioned from the research team by the TAC, to identify and assess the transferability of incentivised models of care to the (TAC) Victorian compensable injury context, and how this might be achieved.

The aim of the review is, therefore, to identify and evaluate incentivised models of care both within and outside the disability sector in and outside Australia. More specifically, this aim includes the following:

Aim 1: to evaluate existing models of incentivised care that exist within and outside the disability care sector and how successful they have been in improving care delivery.

Aim 2: to assess what impact these incentivised models may have on the behaviours and outcomes of the various stakeholders.

Aim 3: to uncover the necessary pre-conditions for the successful implementation of incentivised care methods.

Our review reflects the TAC's request that the information to be presented in a specific format. For this reason, it includes comparative tables to highlight models and their key characteristics. Mini case studies are also used to highlight key findings that we wish to draw attention to. In addition, we include extended appendices (Appendix 2 and 3) in which we provide detailed points for consideration in designing and implementing an incentivised model, for reference.

Key findings and conclusions

Our review of the existing literature on the use of incentivised care models reveals that thus far, their use in the disability care sector has been limited. For instance, we found no systematic evaluations of the implementation of such measures within the disability sector, whether in Australia or internationally. Valuable material was, however, uncovered in relation to the healthcare, primary care, and nursing home (NH) care sectors, which informed our review. These findings can

assist the TAC in developing an incentivised care model that will enhance the experience of clients with traumatic brain injury (TBI) and spinal cord injury (SCI), support clients in achieving their goals in relation to more independent living, and improve the cost-effectiveness of service delivery.

The five key learning points of the study are:

1. **Effectiveness:** There is no guarantee that incentivised models of care will be an effective mechanism to improve the quality and/or the efficiency of care.
2. **Tailor-made endeavour:** There is no off-the-shelf (template) model for incentivised care that can be readily implement in the disability care sector. Existing models were all tailored-made to suit the specific contextual circumstances.
3. **Importance of scheme design:** Without careful design, incentive models are destined to fail. Co-design and multi-stakeholder involvement are essential elements for the successful design and implementation of models. Poorly designed models can actually lead to a deterioration, rather than an improvement of care services.
4. **Careful implementation:** Implementation requires a systematic change management process and program; effective, systematic change management involves realising buy-in from all relevant stakeholders.
5. **Need to evaluate:** Both qualitative and quantitative evaluations are necessary to ascertain the impact of incentivised care measures, including short as well as long term assessments of performance. Inadequate evaluations have diminished the credibility of vast numbers of these initiatives.

We draw a number of conclusions from the findings of our review study:

- I. We should not have unrealistic expectations about incentivised care measures and the quality improvements that can be realised with these measures – prior systemic reviews across the other care areas and non-care sectors have all revealed that the success of incentivised arrangements tends to vary. There is only limited evidence to suggest that the quality of care or the cost-effectiveness of care delivery will be improved when care payers decide to implement incentivised care models.
- II. The purpose of an incentivised model needs to be clear to all stakeholders involved. Care payers need to identify clear goals, strategies, and objectives that they want to realise. Care provider Organisations (CPOs) and care professionals further, need to be convinced that these are the genuine reasons for the scheme.

- III. Poor design and implementation, including a lack of (genuine) consultation with relevant stakeholders, and the use of poor quality measures has undermined the viability of numerous prior initiatives.
- IV. Unintended consequences, e.g., undesirable behaviour by CPOs, such as ‘gaming’ the system, threaten the viability of schemes. These unintended consequences include the risk of possible detrimental care outcomes.
- V. When quality measures are not properly designed, this severely diminishes the credibility and validity of findings from evaluation. The inclusion of measures should be based on empirical evidence, i.e., there needs to be proof that evaluation measures relate to the outcomes targeted.
- VI. Narrow quality measures make it difficult to account for all quality improvements. There is a need for broad multi-dimensional measures. Narrow measures may also fail to identify unintended consequences in relation to non-incentivised areas of care.
- VII. Financial and non-financial incentives should be targeted at identified, desired behaviours. While evidence demonstrates that both financial and non-financial incentives can result in improvements in the quality and/or effectiveness of care, neither will guarantee success.
- VIII. The size of financial incentives can either be too small, or too large. Small incentives do not result in the desired behaviours, whereas incentives that are too large become unsustainable for care payers. Careful modelling is therefore required, to identify the appropriate size for these measures. This modelling should be informed by contextual factors, including the required investment by CPOs.
- IX. Non-financial incentive measures that seek to stimulate consumerism amongst care users are largely unproven. Not all care users are able to ‘shop around’ for the best care;
- X. The influence of external factors is often not properly accounted for in the design of incentive programs and quality measures. It needs to be ascertained that incentive measures, and not any other factors, are the cause of changes in the care delivery.
- XI. Incentive models may be based on **the notion that facilities compete with themselves, rather than each other** (Kane et al. 2007). However, by competing with themselves CPOs become more efficient and effective in their services, CPOs will be stronger and more competitive in the market. This has implications for CPOs under the NDIS, which has been set up to encourage competition *between* the providers.

Despite the long-cherished hope amongst politicians and policy makers that, by introducing incentive measures into different areas of care, some of the quality and efficiency challenges that

obtain under existing funding models can be resolved (Werner et al. 2011). Our review of the existing literature on incentivised models of care indicates that these models thus far have not systematically been used in in the disability care sector. Moreover, where these models have been used there is, in practice, limited existing evidence to support this assumption. Incentive measures, such as P4P models, are by no means a magic bullet for fixing quality or efficiency issues.

1 Introduction

The delivery of a total package of disability care services involves the co-ordination of numerous care professionals, support workers and many others who participate throughout the independence plan cycle, i.e., from the acute/rehabilitation phase, through the transition to community, to maximising the independence of service users. Shortcomings in current care models tend to have an adverse impact on client experience, client outcomes and overall scheme viability. These shortcomings are particularly apparent in the implementation of independence plans through the delivery of attendant care by care provider organisations (CPOs).

Fundamental changes to both the delivery of services and their contracting arrangements will be required in the Victorian disability care industry, in order to move towards more client-centred models of care (McLoughlin et al. 2014). Under the current funding model, Victorian CPOs receive a guaranteed income, that is, the Fee-For-Service (FFS) model which enables them to contract staff and provide services to care clients. McLoughlin et al. (2014) indicate that there are few stimulants in place for providers to provide person-centred care under the current FFS model. Current models of care across the sector can be characterised as provider- rather than person-centred, and care is often fragmented. As the case studies will demonstrate, person-centred care be incorporated into P4P models. Moreover, it is noteworthy that moving towards more person-centred care is also one of the key objectives of the NDIS, which will be rolled out from 1 July 2016, across Australia.

Incentivised care arrangements, such as Pay-for-Performance (P4P), are increasingly being introduced across different areas of care, to supplement or replace the traditional FFS model. This is undertaken because there is an increasing conviction amongst policy makers that the traditional model holds few imperatives for improving the quality of care (Trisolini 2011). However, as will be outlined throughout this report, **the introduction of incentivised care arrangements is not a magical bullet (Roland 2012) that will automatically resolve these issues.**

How to incentivise CPOs to provide the most effective care services has become a key issue of strategic and operational concern to the Transport Accident Commission (TAC). Concern is heightened in light of the imminent roll out of the National Disability Insurance Scheme (NDIS) to improve quality of care services and client outcomes (e.g., satisfaction and independence). Specifically, the TAC and the Health and Disability Strategy Group (HDSG) have identified the need for:

- (i) A fuller understanding of how innovative care models used in other contexts have achieved cost-effectiveness (e.g., in terms of the cost of attendant care);
- (ii) How innovative care models have increased client independence, whilst improving the quality of client experience and outcomes; and
- (iii) An assessment of what innovations might be transferable to the (TAC) compensable injury context, and how this might be achieved.

To meet these needs, the TAC has commissioned this research team to carry out a literature review on incentivised models of care service delivery. The research contributes to a number of the TAC's key performance indicators, namely:

- Helping clients achieve their goals;
- Improved efficiency and effectiveness of service delivery (viability); and
- Enhanced client satisfaction by meeting their needs and expectations.

The report contains five main sections, in addition to an extensive appendix section. Section 2 outlines the main objectives of the study. Section 3 outlines our search strategy to locate the relevant literature. Section 4 provides several case study vignettes of incentivised models in the care sector, to illustrate what models have been trialled and their likelihood of success in the context of the State of Victoria, where applicable. In Section 5, we summarise and discuss the literature on incentivised models of care. This is followed by a list of recommendations in Section 6. We conclude the report in Section 7. We also include extended appendices (Appendix 2 and 3), which provides a detailed summary of issues to be wary of when designing and implementing an incentivised model in disability care services.

2 Objectives of the review study

The aim of this review is to identify and evaluate incentivised models of care within and outside the disability sector, in order to identify practices suitable for transfer to the TAC compensable sector in Victoria. This review examines methodologies for the design of incentivised care models, and assesses how models of incentivised care are implemented within and outside the disability sector, both nationally and internationally. It identifies the required pre-conditions for the successful implementation of these models and highlights 'best-practices' and innovations in service delivery across the analysed models.

The objectives of this review study are therefore three-fold:

1. To evaluate which existing models of incentivised care that exist within and outside the disability care sector and how successful they have been in improving care delivery.
2. To assess what impact these incentivised models may have on the behaviours and outcomes of the various stakeholders.
3. To uncover the necessary pre-conditions for the successful implementation of incentivised care methods.

The first two of these objective were specified by the TAC. The third emerged in the process of our review.

2.1 Research questions

This evaluation of incentivised care models is guided by a number of research questions, which reflect the key areas of concerns to the TAC and HDSG. The questions were scoped in close collaboration between the research team and the industry partners, facilitated by the Institute for Safety, Compensation and Recovery Research (ISCRR). They are:

1. What incentivised models of rehabilitation care exist in the published literature?
 - a. What incentives are being paid for?
 - b. What outcomes are measured, if any, and how?
 - c. What are the common components of effective incentivised models of care?
 - d. What is the timeframe of efficacy? (I.e., how long after the incentivised activity are the benefits observed?)
2. What are the purposes of the incentivised care models examined?
3. What behavioural changes are being targeted?
 - a. Are the incentivised models of care effective in promoting intended behaviours /outcomes?
 - b. How to deal with differing client goals?
4. How are incentivised care models designed?
 - a. What steps are followed in the design of incentivised care models?
 - b. What considerations have to be taken?
 - c. How to evaluate the impact of incentivise measures on the quality and effectiveness of care delivery?

2.2 Incentivised care: What is it? Why use it?

In recent times, across a range of care sectors including disability care, there has been a push to introduce new arrangements, with the goal to improve the quality of care, its efficiency, and effectiveness (Bell & Levinson 2007; Nahra et al. 2006; Perry, Engbers & Jun 2009; Rosenthal & Dudley 2007). One possible avenue through which care delivery can be innovated is through the introduction of incentivized measures, which seek to change behaviours and improve outcomes for care user (Mehrotra, Sorbero & Damberg 2010). While the introduction of incentives in care systems differs depending on circumstances, two main rationales for their use can be identified (Custers et al. 2008) as follows:

- 1) Better alignment of service delivery of care providers with the objectives of the wider system; and
- 2) Redesigning the governance system itself.

Traditionally, care 'buyers' (such as the TAC) are providing care 'sellers', or CPOs, with fixed fees for the services that are provided, irrespective of the delivered quality or the experience encountered by the care user (Scott et al. 2011). By introducing incentive measures, care payers can attempt to change the quality and/or effectiveness of the care delivery by rewarding desired or penalising undesired behaviour, practices, or outcomes. Scott et al. (2011) explain that, in its most elementary form, a financial incentive, for instance, is making an additional payment conditional upon the delivery of care at a particular pre-defined level by the care payer. However, not all incentives are in the form of financial rewards and, as further outlined in Section 4.2, incentives can be structured in various ways in order to achieve the desired outcomes.

The behavioural response to an incentive can be positive, resulting in the desired outcomes, or negative, i.e., resulting in no improvements to the delivery of care, or an actual deterioration of existing standards (Scott et al. 2011). Also, the context in which an incentive scheme is introduced may impact its success. There are three key considerations when designing an incentivised care model (Scott et al. 2011):

- The nature of the incentive (e.g. financial and/or non-financial), including how it is structured, targeted and timed;
- Factors that motivate CPOs and care professionals, including how important the (non-) financial factors weigh up against other sources of intrinsic and extrinsic motivation; and
- What the financial and opportunity costs for participation are, as well as how CPOs and professionals rate the cost-benefit ratio of an incentive model.

When incentive models are poorly designed, aside from the chance of not achieving any outcomes or creating adverse ones, incentives might also lead to unintended consequences, by triggering behaviours other than those originally intended (Karve et al. 2008; McDonald & Roland 2009; Weissert & Frederick 2013).

Incentivised care arrangements can take various financial and non-financial forms. Depending on the circumstances and type of care that is being delivered, different incentives can be designed to alter the behaviour of those contracted (indirectly) by care payers. Furthermore, differences in relation to the circumstances under which care are delivered, where the care is provided, and by whom, will affect which incentives might be most appropriate in a given situation. For instance, some incentive models, e.g. in the healthcare system, are designed to affect the behaviour of highly educated professionals who will directly share in any resulting financial proceeds. In other areas of care, however, employees have no direct financial interest in changing their work practices and it will instead, be the CPOs that will benefit from participation in an incentivised care model. Hence different circumstances will require different, tailor-made approaches to the design of an incentivised care model.

Moreover, aligning the interests of different stakeholders in care delivery is not a straightforward exercise and incentivised measures are not a panacea. First, there is not always a clear link between incentives and desired behaviours (Custers et al. 2008). Second, incentives will not resolve the problem of underlying and divergent interests (e.g., level of care quality vs. profitability) between care payers, care users, CPOs and care professionals. In fact, incentives even have the capacity to exacerbate divergence of stakeholder interests (Hackett et al. 2014).

Incentive models are thus a mechanism through which care payers can try to alter the existing behaviours of those providing care to care users. Incentives can be targeted at the provider organisations or at the professionals delivering the (daily) care that care users receive. There are various rationales for using incentive models, including quality improvement, and improving the cost-effectiveness of the care system. The existing evidence presented throughout this report suggests that careful design of a model is a pre-requisite to success, and that there is no certainty that the introduction of a model will result in the desired improvements.

2.2.1 Pay-for-Performance (P4P)

One particular form of incentivised care that has received considerable attention across different areas of care is pay-for-performance models, identified earlier. These seek to improve the performance of care providers and the care experience of care users through the implementation of financial incentives. They have been trialled with varying levels of success across healthcare (de

Bruin, Baan & Struijs 2011; Casalino et al. 2007; Rosenthal & Frank 2006; Werner et al. 2011), primary care (Glidewell et al. 2015; Hackett et al. 2014; Kirschner et al. 2012; Kirschner et al. 2013; McDonald & Roland 2009; Scott et al. 2011), and nursing home care (Arling, Job & Cooke 2009; Bodrock & Mion 2008; Briesacher et al. 2009; Cooke et al. 2010; Dudley 2005; Kane et al. 2007; Miller, Doherty & Nadash 2013; Weissert & Frederick 2013; Werner, Konetzka & Polsky 2013; Werner, Skira & Konetzka 2016). However, no systematic use of these models was uncovered in the literature on the disability care sector.

These models seek to improve the quality of care by providing financial rewards for the implementation of evidence-based performance (EBP) measures and/or reaching particular quality benchmarks (Greene & Nash 2009). Pay-for-performance (P4P) is thus a means to promote quality improvement by altering existing compensation methods and moving away from the FFS model.

Theoretically, as Weissert and Frederick (2013) explain, these models are based on Goal Theory (Perry, Mesch, and Paarlberg 2006), which suggests frequent feedback on performance as well as targets and goals to guide behaviour. Moreover, “P4P is thought to be especially suitable for setting many routine tasks, as typifies some of hospital care and much more of nursing home care” (Weissert & Frederick 2013, p. s141).

3 Research methodology

The researchers completed an evidence review conducted systematically. For this report, we conducted a semi-structured search of the academic and grey literature to inform the TAC about the challenges, alternative pathways, and design decisions that need to be considered when introducing incentivised care arrangements. We conducted an extensive review of the incentivised care literature that went beyond the disability care sector, and incorporated empirical findings about incentivised care measures used across multiple areas of care.

3.1 Sampling strategy

The initial search focus was on literature in the English language, published since 2005, in relevant areas of care. The following databases were systemically searched to identify academic articles from the period of 2005-2015:

- Business Source Complete (EBSCOHost)
- CINAHL
- EMBASE

- PsychInfo

The academic library search engine, Search, and Google Scholar were used to capture highly cited articles related to incentivised care and pay-for-performance. A substantial proportion of the articles identified were found via the reference lists of the high quality articles revealed in the initial Search and Google Scholar results.

In addition to the academic literature, the relevant grey literature was explored. While this report is founded upon academic journal articles, through Google Scholar and Search enquiries, a number of important grey literature publications were identified.

Keywords used for all the search queries included: “incentivised care”; “pay-for-performance”; “disability care”; “aged care”; “home care”; “primary care”; “P4P”; “pay for performance”; “healthcare quality”; “nursing care”; “rehabilitation care”; “provider quality”; “physician”; “vocational rehabilitation care”; “incentivised indicators”; “care providers”; “attendant care”; and “financial incentives”.

Our review showed that the literature on incentivised care and pay-for-performance is most developed in relation to the healthcare and primary care areas. Consequently, the applicability of the findings from these sectors may be less transferable to the circumstances of the disability care sector than they would within the same sector and country context. For instance, a large number of studies focused on the impact of incentive measures on physicians and other highly skilled medical practitioners involved in primary care (e.g. Christianson et al. 2006; Li et al. 2014; Pham et al. 2007; Scott et al. 2011) or hospital-based care (Helm et al. 2007; Werner et al. 2011; Karve et al. 2008; Rosenthal et al. 2006). A notable exception is the study by Bodrock and Mion (2008), which focused more on the implications for the nursing aspect of hospital care.

Therefore, rather than conducting a systemic review of the structure and effectiveness of various incentivised care interventions, it was determined that our review would more usefully highlight some case studies to illustrate salient aspects of selected incentivised models¹ and draw from them implications/recommendations that are relevant to the TAC. We also highlight – in Appendix 2 – those aspects of the literature, e.g., tools and frameworks, that are of most relevance to the circumstances of the TAC and the Victorian compensable sector.

¹ The term ‘scheme’ is normally used in the literature that examines incentivised care models. We use the term ‘model’ in this report when we refer to incentivised schemes as specific schemes to avoid confusion because, as we were informed, the term ‘scheme’ means the overall program for TAC.

3.2 Challenges of reviewing incentivised care models

The purpose of this review is to inform the TAC about the merits of the most relevant incentivised care models and outline the necessary pre-conditions for their utilisation. It must be stressed, however, that the ability to establish causality between incentivised care measures across various areas of care and to link these initiatives directly to improvements in the care quality and/or efficiency is seriously limited.

In particular, these limitations relate to:

- a. The clear formulation of the purpose and goals of the incentivised care model;
- b. The design of the incentive measures; and
- c. Measurement of the care intervention.

As Petersen et al. (2006, p. 269) explain, the link between incentivised measures and quality improvements is difficult to establish because of the “lack of specificity about incentives”. Moreover, Custers et al. (2008) argue that, while incentivised care models have increased in popularity in recent years, the objectively evaluating their effectiveness has proven problematic. This is, in part because, as they explain, most evaluations suffer from weak designs that limit the ability of researchers to establish the causality of effects, in other words, it is not possible to rule out other factors that may be affecting the outcomes that are attributed to the incentive measures. A large proportion of evaluations further suffer from the problem that no proper baseline evidence is collected, which highlights the need to make an assessment of desired care outcomes before an intervention to incentivise care is implemented. In addition, few studies utilise a control group in their research design. The use of a control group strengthens the opportunity of the researcher to validate their findings. In addition to these limitations and the high levels of context specificity that characterise each of the care sectors investigated in the literature, it is difficult to generalise the effectiveness of incentive measures beyond the particular setting and environmental constraints in which they were implemented (Custers et al. 2008).

With respect to the evaluation of P4P interventions, Cromwell and Smith (2011) identified similar challenges. They explain that difficulties arise because there are considerable differences in the ways that models are set up. Most evaluations fail to conduct a proper baseline quality measurement. What is more, the ability to accurately measure the impact of P4P is compromised by the lack of “rigorously constructed comparison groups” (p. 267).

The validity of findings is furthermore, affected by the population of CPOs and care users that participates in any specific incentivised care intervention. In contrast to randomized trials, trials

that are set up in conjunction with selected CPOs inherently risk being invalidated by a biased sample of participants. The impact of selection bias is twofold. First, it may expose the care user to incentive measures, which affects the ability to generalise findings to broader populations. Second, experiences of P4P in one provider's context may not be applicable to the context of other, similar providers due to differences in care delivery (Cromwell & Smith 2011).

Given the above drawbacks of the published literature evaluating incentivised care models, evidence presented in this review should be considered with caution, and in the context of the constraints and limitations outlined in this section.

4 Findings

Our review of the academic and grey literature did not yield any useful cases of the use of incentivised care initiatives, including P4P, in the disability care sector. Therefore, we present instead a number of case studies from the nursing home (NH) sector, which we suggest contain valuable insights and lessons for the design and implementation of incentivised models of care by the TAC and HDSG.

We suggest that several of the characteristics of the sector, as outlined by Arling et al. (2009), are shared with the compensable disability care sector, in particular, the characteristics of the care delivery, the workforce, and the organisational challenges faced by CPOs. In the light of these similarities, we would argue that more valuable lessons may be learned from the implementation of incentivised arrangements in the NH sector than, for instance, from equivalent implementations in the primary care sector or hospitals.

In the NH, sector care payers have considerable capacity to shape initiatives across the industry. For instance, they can lead strategic initiatives that will support quality enhancement. In the case of the US Medicaid system, they further possess considerable market power, as well as some regulatory responsibilities.

Arling et al. (2009) explain that the care delivery itself is one of the reasons that the NH sector lends itself well to the implementation of incentivised arrangements. Care is delivered under controlled conditions. The facilities are already used to collecting data on the application of various (clinical) measures – which can be utilised quality purposes. Moreover, the organizational structures of CPOs tend to be straightforward and not overly complex.

The care user population, however, is more complex. Arling et al. (2009, p. 588) explain that the care user populations in nursing homes tend to be “suffering from chronic health conditions, cognitive impairment, and functional loss.” Moreover, the care users of nursing home facilities can be further divided into two populations (Werner, Konetzka & Polsky 2013):

- Long-stay residents – these are typically chronically ill individuals that receive non-skilled care, such as assistance with ADL: their average stay in the facilities is two years
- Short-term (*post-acute*) residents – this group receives rehabilitative care focused on achieving a healthy discharge into the community after an acute-care hospitalization – their average stay is 25 days.

Another challenge of the NH sector, which makes the findings relevant to the TAC and to HDSG, is that nursing home organisations tend to experience substantial challenges in relation to staff retention and turn-over rates. Many CPOs struggle to retain staff across all levels of the organisation, from management, to support functions, registered nurses, and other care deliverers. There are significant issues related to those staff who deliver the actual care, who tend to undergo limited training, and who turn over at a rate that poses a serious challenge for the industry.

Lastly, policy makers historically placed greatest emphasis on the cost side of care delivery, under which the quality focus was more geared towards preventing poor outcomes. Given that there is a shift in attitudes, towards quality becoming more important, Arling et al. (2009) suggest that both care payers and CPOs should be able to find scope to raise quality across the sector.

We inferred from our discussions with the TAC, HDSG and CPOs that similar circumstances, issues, and challenges are also applicable to the compensable disability sector. Hence, we would suggest that the TAC and HDSG could learn valuable lessons about the design and implementation of incentivised models of care, and particularly P4P models, from the nursing home care sector.

In particular, we present in this section:

- P4P models adopted in five states in the United States (Miller, Doherty & Nadash, 2013);
- Comparative studies of P4P models (Arling, Job & Cooke, 2009);
- A competitive tendering P4P initiative in Minnesota (Cooke et al. 2010);
- An alternative quality-based payment model for Minnesotan nursing home care; and
- The successful use of P4P in San Diego in the 1980s (Weissert & Frederick, 2013).

It has to be emphasized that, while the findings from the nursing home sector are relevant to the TAC and HDSG, it must be understood that the academic research, thus far, has focused on the impact of P4P models with respect to the quality of care received by long-term residents [only] (Werner et al. 2013).

Moreover, Table 1, presented below, offers useful description of various features of P4P programs in US nursing homes across a number of states. However, several caveats must be noted in relation to Table 1. Table 1 was compiled before some of the P4P programs were discontinued, e.g., the Minnesota study. Hence, the material presented in Table 1 is for informational purposes only and does not reflect the most recent state of affairs in relation to these discontinued programs. Additionally, the information presented in this table is not exhaustive, because more recent programs such as the Vermont model, are not reflected in it. The most recent information that we could find is presented in Sections 4.1-4.5. Finally, we do not have information on the outcomes of the programs, or whether their effectiveness has been evaluated at various stages through their implementation.

Table 1. Description of P4P programs in nursing homes in the United States

Program Name	Stated Goal	Participants	Measures	Incentive	Performance Status	Related references
Georgia Nursing Home Quality Initiative, P4P component implemented in 2007	To promote successful measures to monitor quality; raise quality of care	All NHs required to participate	Measures: <ul style="list-style-type: none"> • Family satisfaction • Employee satisfaction • Nursing retention • Nursing assistant retention • Pressure sores • Physical restraints • Pain in long-stayers • Pain in short-stayers 	In 2007, 1% increase for participation (defined as meeting standards for nurse to patient ration and data collection). Up to 2% increase in rate following state review of program participation and completion of goals	In 2007, 78% of NHs received incentive payments	Petersen et al., 2006; Georgia Department of Community Health, 2007a; Georgia Department of Community Health, 2007b; Georgia Department of Community Health, 2007c
Iowa Accountability Measures Incentive Program, implemented in 2002	Achievement suggests that quality is an essential element in the facility's delivery of resident care	All NH required to participate except for measures related to resident satisfaction	10 measures: <ul style="list-style-type: none"> • Deficiency free survey • Regulatory compliance • Nursing hours • Resident satisfaction • Resident Advocate Committee • Employee retention • Occupancy rates • Administrative costs • Special licensure • Medicaid utilisation 	In 2002, US\$2.86 per day Increase of 1%, 2%, or 3% of daily per diem of direct and non-direct reimbursement based on performance.	In 2005, 87% NHs received incentive payments	(Kane, Arling and Mueller, 2007; Kuhmerker and Hartman, 2007)

Kansas Nursing Facility Quality and Efficiency Outcome Incentive Factor, implemented in 2005	To provide a monetary incentive for favourable outcomes	Open to all NHs in the state	6 quality measures: <ul style="list-style-type: none"> • Direct care staffing • Direct care turnover • Staff retention • Operating costs • Total and Medicaid occupancy • Deficiency free survey 	Incentive payments of US\$1.00 to \$3.00 added to daily per diem.	In 2006, 38% of NHS received incentive payments	(Kansas Department on Aging, 2007)
Minnesota Quality Add-on, implemented in 2003	Quality improvement, increase efficiency, and rebalance long-term care	Open to all NHs in State	7 Measures: <ul style="list-style-type: none"> • Staff retention • Staff turnover • Temporary staff • Quality indicators from MDS • Deficiency free survey • Resident quality of life • Consumer satisfaction 	In 2007, bonus payment of up to 2.4% of daily per diem rate. No bonus for quality scores less than 40%	Implementation process encountered opposition from nursing home industry. Design of final measures under negotiation. Preliminary measures in use	(Nursing Facility Rates and Policy Division, 2004; Nursing Facility Rates and Policy Division 2005; Held, 2007; Minnesota Department of Human Services, 2006)
Ohio Quality Add-on, implemented in 2007	N/A	Open to all regular NHs in State	6 measures: <ul style="list-style-type: none"> • Deficiency free survey • Resident/family satisfaction • Nurse staffing • Employee retention • Occupancy rates • Medicaid utilization 	In 2007, US\$3 a day increase to per diem.	N/A	(Ohio Department of Job and Family Services, 2006; Intrator, Grabowski, and Zinn, 2007; Kuhmerker and Hartman, 2007; Rosenthal and Camillus, 2007)
Oklahoma Focus on Excellence, implemented in 2007	N/A	Open to all regular NHs in State	10 measures: <ul style="list-style-type: none"> • Quality of life • Resident/family satisfaction 	1% participation bonus for the first year beginning July 1, 2007, and	N/A	(Oklahoma Health Care Authority, 2007)

			<ul style="list-style-type: none"> • Employee satisfaction • Technician/assistant turnover and retention • State survey compliance • Person-centred care • Clinical outcomes • Direct care staffing hours • Medicaid occupancy and Medicare utilization ratio 	Provider bonuses of up to \$% per diem rate beginning on October 1, 2007.		
Utah Nursing Home Quality Improvement Initiative, implemented in 2004	N/A	Open to all NHs in State	3 measures: <ul style="list-style-type: none"> • Deficiency free survey • Substandard quality of care citations • State-developed Consumer Assessment of Healthcare Providers and Systems-measure 	Payment of between US\$0.50 and \$0.60 per patient per day. State allocated US\$500,000 for bonus payments	Program performance not publicly available	(Kuhmerker and Hartman, 2007)
Colorado, Quality of Care Incentive Payment Program, implemented in 1996 and ended in 2002	Establishes a resident-centred quality improvement program to improve the quality of life in nursing facilities through resident participation	Open to all approved NHs In 1996 164 plan approved, 5 denied	2 measures: <ul style="list-style-type: none"> • Deficiency free survey • Quality improvement plan 	Per diem increase from US\$1 to \$4 depending on # points	Program costs were US\$3 million in FY 96-97, but the program was repealed when State allowable costs increased. In 1998, 30% of incentive payment was based on facility inspection (0-2 deficiencies)	(Office of Policy and Research, 1999)

					and 70% on approved plan	
Florida, Long-term Care Reimbursement Plan, implemented in 1983 and ended in 1996	To encourage high-quality care while containing costs	Open to all NHs in State	2 measures: <ul style="list-style-type: none"> • No class I or class II deficiencies, or uncorrected class III deficiencies – note: criteria used for the classification are unclear • Substantial compliance with state and federal standards 	Incentive factor provided to NHs with per diems that were below class ceiling and had licensure rating of standard or superior quality of care	Program performance not publicly available	(Agency for Health Care Administration, 2006)
Illinois Quality Incentive Program, implanted in 1985 and ended in 1992	Rewarding Facilities for improving resident care beyond minimum State and Federal standards	Open to all NHs in State 91% of NHs participate	6 measures <ul style="list-style-type: none"> • Structure and environment • Resident participation and choice • Community participation and choice • Resident satisfaction • Care plans • Specialized services 	\$0.25 to \$2 added to daily per diem for meeting standards	Over 50% met all or most of standards. State spent US\$20 million on incentives in 1989. Analysis failed to find evidence of influence on overall costs, Medicaid access, occupancy rates or improvement in resident care. Program concludes some	(Geron, 1991)

					measures failed to demonstrate criterion validity.	
Massachusetts, implemented in 1979 and ended around 1983	N/A	N/A	4 measures: <ul style="list-style-type: none"> • Medicaid utilisation • Efficiency • Occupancy rate • Survey compliance score of quality 	Incentives paid to homes scoring 80% or higher on survey compliance costs – actual incentive unknown	Most homes received incentive. Analysis found the incentive did not promote quality. Incentive not well understood and goal of improving quality not clear. Assessment of performance considered subjective	(Willemain, 1983)
Nursing Home Incentive Payment Experiment, conducted in 1980	To test whether monetary incentive can improve the access and health of Medicaid residents in NH while also saving Medicaid money	36 NH in San Diego	2 measures: <ul style="list-style-type: none"> • Improvement in resident health status based on ADLs • Timely discharge and no readmission for 90 days 	Outcome bonuses equal to estimated costs and wages needed to pay for nursing help (\$US126 to \$370 per case). Discharge bonus equal to offset	Average NH costs rose by 5% due to incentive payments. Incentive payments associated with higher probability of discharge to	(Weissert et al., 1983; Jones and Meiners, 1986; Norton, 1992)

				costs of administering discharge and maintaining vacant bed (\$230 to \$60 per case).	home or lower-lever NH, and lower probability of hospitalization or death. Cost savings possible from reduced hospital costs. Experimental design in assessment	
Texas Performance-based Add-on (PBAO) Payment Program, implemented in 2001 and ended in 2003	Facilities should be recognized and rewarded for their performance	Open to all regular NHs in state	3 measures: <ul style="list-style-type: none"> • Regulatory compliance • Resident outcomes • Medicaid utilization 	Incentive payments of US\$1.13 in 2001 added to daily per diem	In FYI 2001, 57.8% received incentive payments. The mean payment was US\$6,780 (range: \$11 to \$38,859)	(Carter, 2002)

Source: adapted from Briesacher et al. (2009, pp. 4-7)

* Briesacher et al.'s (2009) review identified 13 pay-for-performance programs in the nursing home setting between 1980 and 2007: 7 programs were currently active as of 2007 when the review took place, while 6 had been terminated. All of the programs that were current at the time of the review were relatively new, with the oldest program dating back to 2002 (Iowa's Accountability Measures Incentive Program). Nearly all of the terminated programs were short-lived; some lasting no more than 2 years.

4.1 P4P in five US states

We suggest that one of the most valuable sources of inspiration for the TAC and HDSG for designing and implementing an incentivised model of care can be derived from the comparative work by Miller et al. (2013). Their research compared the P4P programs introduced by five US states – Iowa, Minnesota, Oklahoma, Utah, and Vermont (introduced in alphabetical order) – all of which were implemented with the objective to improve the quality of care.

The authors explain that historically the tools available to policy makers, which were mainly based on penalties for grossly failing to provide adequate standards, had yielded disappointing results when it came to quality improvement. This failure of penalty-based systems to improve quality in turn led to the use of more market-based instruments in the sector, including P4P. However, as the authors found, “there is no single design that, as yet, has been demonstrated to achieve best outcomes (p. s162).”

The authors assessed five P4P programs introduced in US Medicaid nursing homes across different states, to better understand the implications for care payers and other stakeholders, of various design and implementation decisions. Of the 14 programs identified by our literature review, only five are compared below, as these cover the widest range of characteristics. The separate case studies below will further illustrate the working of these programs, each of which has a different level of complexity.

Table 2 provides an overview of the characteristics of the P4P programs assessed. It highlights the extent of CPO participation, how the models were financed, what types of measurements were used by the care payer to assess quality, and to award the incentives, what administrative structures were utilised, and how the programs were developed.

Table 2. Characteristics of Nursing Home P4P Programs in five US States

	Iowa	Minnesota	Oklahoma	Utah	Vermont
Implementation Year	2002	2006	2007	2004	2004
Provider Participation					
Program design	Yes	Yes	Yes	Yes	Yes
Program implementation	Yes	Yes	Yes	No	No
Program modification	Yes	Yes	Yes	No	No
Percent participation	0	0	98%	80%	25%
Financing					
Funded with new/existing money	Existing	Existing	New	New	Existing
Program active/suspended	Suspended	Suspended	Active	Active	Active

Measurement						
	Staffing	Yes	Yes	Yes	No	Yes
	Consumer satisfaction	Yes	Yes	Yes	Yes	
	Inspection performance	Yes	Yes	Yes	Yes	Yes
	Clinical Quality Indicators	Yes	Yes	Yes	No	No
	Person-centred/quality of life	Yes	Yes	Yes	Yes	No
	Efficiency	Yes	No	No	No	Yes
	Access	Yes	No	Yes	No	No
	Employee satisfaction	No	No	Yes	No	No
	Quality improvement plan	No	No	No	Yes	No
Administration						
	System complexity	Complex	Complex	Moderate	Simple	Simple
	Relies primarily on existing data	Yes	Yes	Yes	No	Yes
	Industry-wide benchmarks used	Yes	Yes	Yes	No	No
	Composite quality index used	Yes	Yes	Yes	No	No
	Reward structure	% Per diem rate (a)	% Per diem rate	% Per diem rate	Lump sum	Lump sum
	Rewards improvement	No	Yes	No	No	No
	Funds allocated competitively	None	Some	None	Some	All
Development						
	Program phased in slowly	No	Yes	Yes	Yes	No
	Program modified over time	Yes	Yes	Yes	Yes	No
	Program linked to public reports	No	Yes	Yes	No	Yes

(a) Distributed on a lump-sum basis at the end of the rate year.

Source: Adapted from Miller et al. (2013, p. s155)

In terms of the types of incentive payments used, two main reward structures can be identified: (1) lump-sum payments distributed on the basis of the number Medicaid patients served, and, (2) percentages of the per diem rate, whereby the size of the incentive payments received depends, in part, on facility costs. Miller et al. (2013, p. s161) argue that the “fixed per diem add-on approach is preferred because it is dependent exclusively on the basis of facility performance rather than on how much money facilities happen to be paid”. Moreover, the authors found that biases built into the reward calculations affect the perceived fairness of these initiatives by the CPOs.

Below we summarise the five cases in the Miller et al. study, these cases studies illustrate the strengths and limitations of various P4P models within the context of nursing homes in the US Medicaid system. The cases highlight the design of the incentive payments, the quality measures on which they are based, and the utilised models. Key learnings for the TAC and HDSG are indicated.

4.1.1 Iowa

Context:

Initially, the Iowa P4P program was based on rewarding CPOs in anticipation of performance improvements – based on performance in previous years and contingent upon continued similar or improved quality. As it evolved, however, the care payer decided, in response to unintended behaviours, to alter the system and utilise lump-sum retrospective payments instead. The quality dimensions and QIs of the model also evolved over time, to reflect state-of-the-art measures and to better target the desired behavioural responses of CPOs.

Despite the significant progress that the care payers and a broad group of stakeholders made throughout the development and implementation of the model, continued uncertainty about financing threatened to jeopardise the long term viability of the quality initiatives.

Quality measurements:

The model was initially based on performance measures derived from:

- Cost reports (including staffing, retention);
- Inspection data (e.g., deficiency-free surveys, regulatory compliance);
- The Care Ombudsman (e.g., complaints); and
- Resident satisfaction surveys - which were administered by an external party.

Scores up to 10 could be achieved for the first three quality dimensions by achieving QIs above relative performance benchmarks. CPOs could earn an additional point when they scored higher than the fiftieth percentile on the survey. However, this additional point was only awarded if the survey reached a response rate >35 per cent. This threshold was put in place to stimulate CPOs to “seek information about resident experiences whether or not they anticipated good results” (p. s156).

In order to encourage CPOs to participate in the model, the design allowed CPOs to rely on existing administrative systems and structures for reporting on the QIs, wherever possible. Moreover, the model used “industry-wide benchmarks when awarding points, so that any facility could qualify for extra payment. This would not have been possible if, say, points were awarded only to those scoring above a certain percentile on a particular quality measure” (p. s156).

In 2010, after complications (see below), the model was amended (this amendment was not fully implemented or costed at the time of the authors' publication). Under the amendments, CPOs could score up to 100 points based on four quality dimensions. The emphasis shifted towards more 'process' measures, including the inclusion of care user satisfaction data. QIs are related to:

- Quality of life (25 points) – including person-directed care (e.g., dining activities, resident choice) and resident satisfaction (e.g., resident/family survey, complaint resolution);
- Quality of care (59 points) – including state inspection (e.g., deficiency-free surveys, regulatory compliance), staffing (e.g., nursing hours, turnover), and nationally reported quality measures (e.g., high-risk pressure ulcers, physical restraints);
- Access (8 points); and
- Care efficiency (8 points).

Under the model, CPOs with serious inspection deficiencies are excluded from attracting additional payments. Moreover, quality outcomes will be reported publicly.

Incentives used:

Under the original model, overall rewards “were based on the prior year’s performance and calculated as a percentage of the per diem rate, so that those scoring 3-4 would earn an additional 1 percent; 5-6 points, 2 percent; and 7 or more points, 3 percent. Those scoring fewer than 3 points received no reward” (p. s156).

Incentives are awarded under the redesigned model as outlined in Table 3. The incentive payments are estimated to cost the care payer in the range of US\$1.25 to \$6.25 per care user day. One of the conditions is that extra payments be used to support direct care staff, through wages, benefits, and training. Although the model was altered to use retrospective payments, the basic set-up remained the same. Moreover, under the redesigned model, “[e]nd-of –year payments were reduced by 25 percent for each subsequent serious deficiency and withheld altogether for unaddressed deficiencies” (p. s156).

Table 3. P4P incentive structure under the redesigned Iowa model

Points scored by CPOs	Incentive
0-50	No additional payment
51-60	1 per cent
61-70	2 per cent

71-80	3 per cent
81-90	4 per cent
91-100	5 per cent

Complications:

In 2008, several deficiencies in the P4P model were identified by news media, in particular, CPOs that had delivered poor quality received incentives, despite the poor quality of care. This anomaly was, in part, attributed to the reliance on historical data to pay for future care. The authors explain that the subsequent, strong response by the legislator highlights the political sensitivity surrounding P4P initiatives. The model was amended to ensure that incentive payments were only made retrospectively. The new payment is structured on the basis of “a lump-sum payment, using the per diem add-on methodology at the end of the fiscal year, contingent on “good behavior”” (p. s156).

The future of the model:

The Iowa P4P program has been one of the oldest and longest running P4P programs. It dates back to 2002, and was suspended as a result of the global financial crisis. According to Werner, Konetzka & Polsky (2013) it was active in 2013, however on the basis of the information we found on the Department of Health Services from the state (<http://dhs.iowa.gov>) it could not be ascertained what the current status of the program is.

Key learnings:

- The Iowa model highlights the importance of broad stakeholder involvement in the design and implementation of a P4P model. To design the model, the State of Iowa, which was the care payer, set up a workgroup comprised of industry representatives, advocacy groups, state agency personnel, legislative staff, and other interested parties. These stakeholders met annually to review and update the P4P model.
- The Iowa experience highlights especially the problems associated with relying on historical data for awarding incentive payments. Past-performance is no guarantee of future quality of care.
- Iowa’s P4P experience provides useful insights for the TAC and HDSG, as it reveals how P4P models can evolve over time. While the model is a helpful example of the development of

measures and their implementation, another important lesson relates to the sustainability of a model.

4.1.2 Minnesota

Context:

Rather than having one P4P model in place, Minnesota's quality initiatives consist of two complementary models, both of which can attract financial incentives. These are:

- The Minnesota Quality Add-on (P4P model); and
- The Tender-based performance improvement program.

The Minnesota model is rather complex (Miller, Doherty & Nadash 2013). It demonstrates some of the different ways in which 'unintended' consequences can be mitigated. The model seeks to strike a balance between motivating 'high' to 'low' quality CPOs into adopting more quality-focused behaviours. One of the particular successes of the Minnesota experience was that it received broad support from a range of stakeholders, all of which were involved across the various stages. The broad stakeholder group included unions, CPO associations, consumer advocacy groups, and the state's ombudsman, in addition to the CPOs. The Minnesota model is also exemplary in that experts contributed to the development of its quality measures.

While the broad coalition of stakeholders assisted the care payer to reach consensus on the model itself, as Kane et al. (2007, p. 115) explain, they nonetheless "had difficulty agreeing on many of the system's technical details, such as operational definitions, weights, and scoring thresholds for different measures".

The models:

Minnesota Quality Add-on (P4P model)

The first initiative that was deployed in Minnesota was a quality measurement-focused P4P model under which each CPO's facility received a quality score that ranged from 0-100. This score is based on a number of quality dimensions, which can be evaluated and adjusted over time. When the model was first introduced, the following quality dimensions were used:

- 24 clinical quality indicators (40 points);
- Direct staff retention (25 points) – Calculated as the number of nursing staff who left over a year divided by the total number of staff;

- Staff turnover (15 points) – Calculated as the number of nursing staff retained divided by total number of staff;
- Use of pool staff (10 points) – Calculated as the as a percentage of the total nursing hours; and
- Survey deficiencies (10 points).

After the first year, the staffing turnover dimension was replaced by a measure of direct care staffing. In addition, a measure of care user satisfaction and QOL was developed and introduced. The weighting of these quality dimensions was adjusted to assign 35 points for the former, and 20 points to each of the latter measures. Section 4.4 discusses an alternative model for the Minnesota system, which places greater emphasis on the combination of quality and operational efficiency.

The authors of this study explain that the QOL measure was the result of careful consultation and design work, as the policy maker insisted on using quality measures that had a high level of reliability. An external third party was, therefore, engaged to develop this particular instrument. The Minnesota QOL measure has been criticised for being “too detailed, lengthy, and costly” (p. s159), yet one of its strengths is its validity.

Miller, Doherty and Nadash (2013) explain that, to keep the administrative work associated with the P4P to a minimum, the scores were based on measures that could already be derived from the care payers’ existing administrative data systems.

The policy makers preferred quality measures that measured absolute, rather than relative performance, yet they opted to also capture improvements over time. Thus, points were only awarded if certain quality standards were met. This had the result that some QIs were staggered and could attract multiple points for higher levels of quality. At the same time, points could also be accumulated if sustained improvement over time could be demonstrated.

One of the concerns that the Minnesotan policy makers had, when designing the P4P program, was that they wanted to ensure that the program would not result in a deterioration of care for ‘high risk’ care users. Moreover, they wanted to ensure that the CPOs that served high risk users were not disadvantaged by the quality measures. Consequently, the program incorporated a considerable degree of ‘risk adjustment’. Both the resident satisfaction/quality of life and the

clinical quality indicators were, for instance, risk adjusted to account for acuity levels within the populations served.

The quality instruments used by the program are notable for the strong emphasis placed by the policy makers upon outcome measures. However, owing to the resistance of other stakeholders, the program initially placed greater weight on structural and process measures (e.g., the large number of points assigned to staffing issues). This phenomenon highlights the tensions that can exist between care payers and the various other stakeholder groups – ***compromises will need to be made to resolve the tensions.***

Incentives:

Incentive payments were based on the number of points that CPOs scored across the quality dimensions. Payments were calculated as a percentage of a facility's base rate. In the first instance those who scored

- 0–40 did not receive an add-on;
- 41–99 received an add-on proportionate to the summary quality score; and
- 100 received a 2.4 percent add-on.

In the first year, the average add-on was around 1.0 per cent. However, due to budgetary constraints, this was reduced to 0.13 per cent in the second year, and subsequently scrapped altogether.

Tender-based performance improvement program:

Alongside the implementation of the add-on P4P, Minnesota also introduced its Nursing Facility Performance-Based Payment Program. This initiative provides CPOs with the opportunity to receive incentive payments of up to 5 per cent of their base payment rate. This model is set up on a voluntary basis and invites CPOs to put forward quality improvement initiatives themselves, which then go through a tender evaluation. The incentives available to CPOs are for the length of the projects. While projects are tailored to the circumstances of a CPO, they can also involve multiple CPOs who put forward joint proposals. The mechanisms of this incentivised model of care are further outlined in section 4.3.

Financial sustainability:

Despite the sophisticated processes that the care payer set up, one of the biggest shortcomings of this model has been the care payer's ability to secure long term financing for these initiatives. In the context of ever-increasing budgetary constraints, the State of Minnesota – the care payer – had to reduce, and ultimately, halt the funding for these quality endeavours. Minnesota's example thus illustrates the need for a sustainable funding strategy.

Miller, Doherty and Nadash (2013, p. s157) explain that, "[t]he size of the incentive payment declined because, once awarded, it became permanently incorporated into the base rate and unavailable for future add-ons". Although the payment of incentives ceased, it was suggested by the authors that the model continues to have an impact on quality across the industry, since the quality measurement scores continue to be collected and publicly reported.

The future of the model:

While the P4P aspect of the quality program has been suspended, there is a chance that, in the future, the model might be re-instated. The authors point out that there are a number of different scenarios under which it could be reactivated.

One option would be through the "redistribution of existing resources through a rate cut" (p. s159), i.e., relying on penalty measures to punish poorly performing CPOs and utilising these funds to reward those that are improving their quality. An alternative model suggested is to make (a proportion of) future cost-of-living adjustments of the nursing home care budgets dependent on demonstrated quality improvements based on the P4P programs, before these are passed on to CPOs.

Key learnings:

- **The Minnesota experience of incentivised models of care offers a number of meaningful learnings for the TAC and HDSG. These include ways to develop a model, different ways to set up financial incentive mechanisms, and considerations for the sustainability of a model.**

4.1.3 Oklahoma

Context:

The Oklahoma model most closely resembles a P4P model that relies on “best practices criteria, although its strategies for incentivizing lower-performing providers could be strengthened” (Miller et al. 2013, p. s160).

The strengths of the system are that it is only ‘moderately’ complex, while it also has a high level of CPO participation. Active stakeholder engagement further characterises the model. Moreover, the authors hail its phased-in introduction of financial incentives as exemplary.

Measurements – using an external party:

In the case of Oklahoma, the care payer decided to contract out part of the model to a private party, MyInnerView. This organisation was engaged to develop and administer the quality measurement system.

“MyInnerView was chosen largely because the state perceived it to be a well-equipped management company with the expertise necessary to choose the specific quality measures that fell within the guidelines established by the state” (p. s158). MyInnerView was responsible for the development and administration of a care user/family satisfaction survey, as well as an employee survey. The state further delegated the surveys to this party to ensure uniformity in quality measurement across both instruments.

The purpose of the two instruments was to assess employee satisfaction and system-wide culture change (i.e. person-centred care). Survey response rate thresholds were put in place to ensure the validity of findings. For the employee survey, this threshold was set at > 30 percent, while care user/family surveys achieved > 65 per cent return rate were omitted from the results.

While the remaining indicators are not actively measured by MyInnerView, this data is harvested through its online platform. CPOs use the online tool to fill in performance data in relation to staffing, state inspection compliance, and clinical measures (e.g., falls, use of catheters, use of physical restraints, and occurrence of pressure ulcers). The program was designed to minimise the level of additional data collection that CPOs had to undertake, ensuring that the model was not an administrative burden. With the exception of the two surveys, CPOs could simply enter the data from their existing systems into the online platform. No new measures had to be developed. The

authors explain that “[t]he relative simplicity of Oklahoma’s reimbursement system has helped with administration as well” (p. s158).

Incentives used – a phased in approach:

An important thing that can be learned from the Oklahoma experience is the benefit of gradually phasing in a P4P model. The staggered approach to implementation gave CPOs the opportunity to ease into the program and gradually develop the capacity and routines necessary to deal with the additional administrative requirements. The introduction of financial incentives was spread across 3 phases.

In the first phase, CPOs were given additional reimbursement for filling out the surveys. This was done to encourage CPO participation. This step also allowed CPOs to familiarise themselves with the MyInnerView system, and develop the knowledge and capability to deal with the additional administrative requirements. CPOs received a 1.0 per cent incentive payment for participating in the program and completing the surveys. They received the payment quarterly, based on a percentage of the Medicaid per care user day payment rate.

The second phase was that CPOs could attract additional payments based on their quality scores. Performance was measured on 10 QIs. For each of the QIs, CPOs could score a point. Thus up to 10 points could be earned in this way. Points were awarded for each QI score that exceeded the median. In turn, for every two points, CPOs received an additional 1.0 percent of the Medicaid per care user day rate.

In the third phase, the model moved away from a relative performance (median score) measure to an absolute performance measure, i.e., thresholds were set at particular levels of care. If CPOs exceeded these thresholds, they earned points.

The add-on from the model ranged from 1 to 5 per cent of the Medicaid per care user day payment rate, which resulted in an extra US\$1.09 to \$5.45 per day. The authors reveal that the model’s entire annual budget was an estimated US\$560,000.

Strengths:

One of the strengths of the Oklahoma incentive model is that it enjoyed very strong buy-in from the CPOs. 98 per cent of the nursing homes signed up for the model. The authors attributed this,

in part, to the gradual rollout, including the opportunity for the providers to familiarize themselves with the necessary protocols and procedure while being compensated for their efforts.

It further highlights the importance for care payers to build and maintain strong relationships with CPOs. The relationship between the stakeholder groups – including the care payer, CPOs, and the CPO associations – was characterised by such extensive co-operation that even parts of the policy and regulation that set out the workings of the model were co-designed.

Miller, Doherty and Nadash (2013) report that public administrators attributed the success of the model to the minimal additional data collection requirements imposed upon the CPOs. Since all reporting was conducted online and information more easily accessible, the relative administrative burden is limited.

The future of the model:

The authors explained that the introduction of the Oklahoma model was assisted by the fact that the care payer could rely on new funds to set up the model. This meant that CPOs' Medicaid per care user day rates were not placed at risk – i.e., there were no penalties – which further contributed to CPO buy-in. Although the care payer could initially draw upon additional funds, thereafter the level of funding remained static, because the State of Minnesota experienced budget difficulties, which highlights the importance of financial sustainability of a model. The model is currently still active and 298 facilities are participating in the scheme (<http://www.okhca.org/individuals.aspx?id=8135>).

Key learnings:

- **The successful participation of stakeholders helped strengthen the design and reduced resistance to model implementation;**
- **The ability to utilise external parties to develop QIs and monitor performance;**
- **Benefits of a phased-in approach; and**
- **Financial sustainability issues encountered by the care payer.**

4.1.4 Utah

Contexts:

The Utah experience provides insights into securing CPO participation and also into the degree of flexibility that can be incorporated into models. Utah has 2 quality programs in place; QI1 (a P4P model focused on care delivery), and QI2 (which focuses on capital improvements). The high participation rates in the models can be explained by their relative simplicity. They are, in addition, flexible, and allow CPOs to select areas for quality improvement. The Utah model also illustrates the benefits of the gradual phasing in of financial incentives.

The availability of additional funding has further enabled the care payer – the State of Utah – to ensure the sustainability of the quality initiatives.

The Utah experience highlights the advantages and disadvantages of having relatively flexible participation. It is, for instance, an administratively cumbersome form of P4P for the care payer. This feature also meant that a fair degree of in-house expertise had to be developed.

The model's financial sustainability:

The State of Utah was able to draw upon new funds to set up the quality improvement model by introducing a provider tax. This model was adopted because it enabled the care payer – the state – to access additional federal funding.

QI1 (P4P)

QI1 is the State of Utah's P4P program. It has a budget of US\$1m. CPO payment is based on performance and "distributed on the basis of Medicaid patients days", under which payments "range from \$3,000 to \$30,000 per year" (p. s159).

The program was gradually phased in, in to order to secure provider buy-in. Initially, the program was basic. CPOs only had to demonstrate a quality improvement strategy and an implementation plan in order to attract the incentive payments.

From the second year onwards, there were four requirements placed on CPOs. CPOs were obliged to:

- Conduct a customer satisfaction survey;

- Develop a quality improvement action plan detailing how any below-average quality deficiencies were addressed;
- Have in place a cultural change program that was focused on the implementation of person-centred care; and
- Have an employee satisfaction program.

CPOs with serious deficiencies are disqualified from participating in the P4P model, while those delivering a substandard quality of care were only eligible only for half the financial incentives. Instead of using standardized surveys, the CPOs can decide which third parties (e.g., consultants) they engage to develop and conduct these surveys in their organisations, which gives them a degree of flexibility.

Thus, in the first year, CPOs were required to demonstrate that a quality program was in place. From the second year onwards they had to show improvements in relation to their specific quality issues. The authors explain that “[t]his phase-in proved particularly beneficial for less sophisticated facilities by encouraging them to create something more meaningful than what may have existed previously. The slow phase-in was also beneficial where measurement of customer satisfaction was concerned.” (Miller, Doherty & Nadash 2013, p. s159)

QI2

QI2, by contrast, “targets capital improvements that improve quality” (p. s159). The care payer has a separate budget of approximately US\$4.3m to assist CPOs in improving their facilities.

QI2 requires CPOs to prove that they have made capital improvements. The care payer in turn provides them with “a fixed amount per Medicaid certified bed”. Eligible capital improvement initiatives are broadly defined and the list of possible improvement areas continues to be constantly updated and extended. The authors explain that the variety of options includes: nurse call systems, vans, heavy-duty lifts, electronic health records, bathing systems, quality training, and dining systems.

Strengths:

The strength of the Utah model is that QI1 and QI2 are relatively flexible and allow CPOs to focus on those quality initiatives that are most applicable to their particular circumstances. Moreover, the focus is strongly on rewarding quality. The gradual phase-in has further allowed lower quality

providers a period of time to put in place quality improvement measures and structures, which enabled a greater participation rate in this incentivised model.

Weaknesses:

The flexibility of the system is, at the same time, a potential weakness, “because each facility establishes its own areas for improvement, [so] state administrators must assess different material for each facility, which can be time consuming” (p. s159). A downside of QI2 is that is administratively cumbersome for the care payer. The authors explain that CPOs sometime delay submitting documentation of improvement initiatives, thus creating an ‘administrative workload’ for the care payer that has to be swiftly dealt with before looming deadlines.

The future of the model:

The discussed QI initiatives are currently still in place, and are supplemented with a third incentive initiative. The details of the most recent set-up can be found on <http://www.health.utah.gov/stplan/longtermcare.htm>.

Key learning:

- **Sufficient financial support is essential for sustainability of incentivised models – of the 5 P4P models evaluated by Miller et al. (2013), Utah’s is the only one that has not been suspended or put on hold due to financial constraints experienced by the care payer.**

4.1.5 Vermont

Context:

Vermont is a relatively small state with a limited number of nursing homes. The existing incentive models were developed in close collaboration between the care payer – the State of Vermont – and the one nursing home association in the state, which has increased the consensus around the structure of the program.

The state has a P4P award model that is exclusive to the top 5 performing nursing homes in the state and has strict participation requirements. There is also a Gold Star Employer Program, which is open to all CPOs. The focus of this model “is primarily on staffing, under the assumption that the quality and motivation of facility staff is the most important determinant of quality” (p. s160).

Models:

P4P is one of the two models adopted in Vermont. The state has set aside a budget of US\$500,000 annually for the P4P awards, which are open exclusively to the top 5 nursing homes that demonstrate high quality care improvement initiatives. The participation requirements are:

- The CPO must be deficiency-free - no deficiencies may be found during the most recent health & fire safety survey (the authors explain that this disqualifies the majority of CPOs); and
- The CPO must participate in the Gold Star Employer Program.

CPOs are ranked on the basis of ‘objective standards of quality’. If there are more than 5 homes that meet the requirements, then cost-efficiency (allowable costs per day) becomes another consideration.

The awards are subsequently awarded on the basis of “the ratio of its Medicaid days to the total Medicaid days for all qualifying facilities” (p. s159). The care payer allows CPOs to utilise the funds for those quality initiatives that the CPO deems most appropriate, as long as these are geared towards the quality of life of Medicaid-eligible residents. “Typically, the reward is rolled back into the facility to add a garden path or change some structure or to add new programming for the residents” (p. s159). Thus, it can be concluded that these awards, while stimulating high performance, do not directly result in investment in higher levels of care delivery, but are likely to improve the quality of life of care users.

Miller et al. (2013) explain that the findings from Vermont suggest that being associated with these awards is held in higher regard than the actual financial remuneration itself; i.e., the awards function as a market differentiation mechanism through which CPOs can demonstrate that they are delivering 'high levels of care quality', rather than the CPOs placing a high value on the monetary rewards themselves.

The Gold Star Employer program, is the other quality initiative in Vermont. This program is a joint initiative between the care payer and the State's care user advocacy group. It focuses on best practices for the recruitment and retention of direct staff and is built on 3 pillars:

1. A self-assessment in 7 areas (e.g., recruitment, orientation/training, professional development);
2. An implementation plan and strategy for adoption of 'best practices'; and
3. Evidence of progress on the implementation of 'best practices'.

No financial awards were attached to the program. However, in order for CPOs to be eligible for the financial awards, they had to demonstrate successful participation in this quality initiative.

Weaknesses of the Vermont model:

The Vermont P4P initiative is only geared towards 'high quality' CPOs, while lower performing organisations are largely neglected. Furthermore, the fact that only deficiency-free CPOs are eligible to participate in the P4P model excludes those with minor deficiencies and those that may potentially have made the most progress in lifting quality. The model does not recognise this, and Miller, Doherty and Nadash (2013) point out that according to experts, recalibration in relation to this issue might be warranted.

The future and current state of the model:

The current status of the Vermont model could not be ascertained beyond the material found, and presented here, in the academic literature.

Key learnings:

- **The Vermont P4P implementation offers insights for the TAC and HDSG in that it has adopted two rather different, yet simple, quality improvement initiatives. The first of these is a P4P model that appears to drive behavioural change in CPOs based on non-financial motivators as much as financial ones. The second initiative further, functions as a pre-requisite for the first.**

- Questions can be raised, however, about how effective the Vermont initiative is at raising the quality of care outside of ‘high quality’ nursing homes, as the models are skewed in the favour of these leading CPOs.

The conclusions that Miller et al. (2013) draw about the P4P programs analysed provide several useful insights that are relevant to the TAC. For instance, they emphasise the importance of stakeholder input and participation in the design and implementation of incentivised models of care. Care payers have the ability to bring together a broad base of stakeholders who can play a crucial role in the facilitation, design, and acceptance of these programs. Another important implication is that P4P programs are most successful when they are gradually phased in. In other words, models need to be designed so that they can be easily implemented in stages/ with a staged approach in mind, as shown in Figure 1.

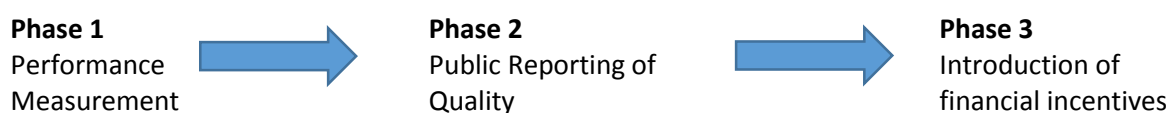


Figure 1. A step-by-step approach to designing and implementing a P4P model

Miller et al.’s (2013) comparison further demonstrates that a range of quality measures and financial incentives can be used, but these features need to be tailored to the contextual circumstances of each model.

The authors highlight that, in all case studies, stakeholder involvement was regarded as essential for the potential success of models. Involving a broad coalition of stakeholders, e.g., in the form of workgroups, helped during the design phase to secure consensus on measures. Broad stakeholder involvement, besides generating greater acceptance, also enabled care payers to tailor the incentive program towards particular industry challenges and needs.

Furthermore, the evaluation of quality measures in the nursing home sector reveals that “a wide range of measures is preferable as it spread[s] the risk” (Miller et al. 2013, p. s160). Another advantage is that it reduces the risk of unjustly penalizing CPOs that perform poorly in one area, but deliver quality in other, while it further limits opportunities for ‘gaming’ and outright fraud.

The findings further suggest that care payers must make sure that the administrative work associated with these models is as minimal as possible for the CPOs. While adequate QIs are essential for the 'fairness' of incentivised care models, care payers should take care to ensure that models do not become an administrative burden. If poorly designed, these models have the potential to be particularly burdensome to care professionals involved in the daily delivery of the care. Hence, it is worthwhile to carefully scrutinize what existing data might be readily available at the care payer and CPO level, and how this can be collected in an efficient way. The authors recommend that it would be beneficial for care payers to consider existing data systems, if possible. This is also preferable from a cost perspective. However, if new measures need to be designed and implemented, care payers will need to consider how they will go about this process. For instance, external consultants can be engaged to set up these measures, and possibly also, to operationalize them.

Lastly, long-term financial sustainability is also worth considering when initially designing a model. The importance of the long-term funding is critical, since it increases the chances that providers will be willing to make necessary investment where required. The authors argue that it is critical that care payers have a long-term funding strategy in place, otherwise the risk is that such incentivised care initiatives will be short-lived only – as is evidenced by the analysed programs, where it was found that only 1/5 (namely, Utah) had not been suspended or (temporarily) placed on hold as a result of funding pressures.

4.2 Comparative studies of P4P models

Similarly to Miller et al. (2013), Arling et al. (2009) compared nursing home P4P models across multiple US states. They focused on models that were operational in 2007, in the States of Georgia, Iowa, Kansas, Minnesota, Ohio, and Oklahoma. The authors explain that a problem with many existing P4P experiments is that a substantial number of initiatives are only short-lived and are not properly evaluated – hence their effectiveness frequently cannot be ascertained. The six models at the time, however, had been operational for a period of time (some as early as 2002). Some of the models investigated in this study were subsequently suspended, due to the financial constraints of the care payer organisations.

Arling et al.'s (2009) study highlights various quality dimensions and QIs used to evaluate the success of the models. It further identifies relevant considerations for the inclusion or exclusion of these measures. In addition, the structure of each of the various financial incentives is briefly discussed.

As depicted in Table 4, the Arling et al. (2009) review reveals that the States under examination utilised Donabedian (structure-process-outcome) measures for the various quality dimensions of P4P initiatives. The authors explain that one notable difference between the States' models was how the QIs were developed and monitored. While some States, e.g. Minnesota, developed a considerable proportion of their measures in-house, others like Georgia and Oklahoma, contracted this function out to the private sector.

Table 4. Performance measures and approaches adopted by the States in their P4P models

Dimension	Measure	Typical data source	Freq. used
Staffing	<ul style="list-style-type: none"> Staffing levels (hours per care user) Staff retention and turnover rates Use of contract and casual staff 	Cost report	6/6
Staff satisfaction	Staff satisfaction with <ul style="list-style-type: none"> Work environment Management Teamwork Training and development opportunities Organisational culture 	Staff satisfaction survey	3/6
Nursing home inspection	Scope and severity of deficiencies in: <ul style="list-style-type: none"> Clinical care Resident QOL Resident rights Dietary services Physical environment Other services 	Nursing home inspection data	4/6
Clinical quality indicators (QIs)	Nursing home QIs and measures <ul style="list-style-type: none"> Pressure sores Physical or chemical restraints Decline or improvement in ADL living 	Minimum Data Set (MDS)	3/6
Resident quality of life	Resident QOL in relation to <ul style="list-style-type: none"> Autonomy Security Environmental adaptations Privacy Dignity Meaningful activity 	Quality of Life Survey	1/6
Consumer Satisfaction	Resident or family satisfaction with nursing home services, environment, staff, and quality of life	Resident Satisfaction survey; family satisfaction survey	4/6
Other	Access to care – high percentage Medicaid, behaviour problem, dementia or high acuity residents Efficiency measures <ul style="list-style-type: none"> low administrative costs, low operating costs, high occupancy rates 	Cost report, MDS, or other administrative system	5/6

Source: Adapted from Arling et al. (2009, p. 589)

Quality dimensions used:

In terms of the 'dimensions' of quality used, it can be observed that some were more frequently utilised across the P4P models than others. For instance, all states tied nursing staff levels to performance. Arling et al. (2009) explain however, that this is in some respects an imperfect

measure, as depending on the level of qualification of the staff and the conditions of nursing home care users, different levels of staffing might be appropriate. Notwithstanding this important subtlety, in practice, facilities were compared on this measure, for which aggregates of staffing levels are used.

Moreover, turnover and retention rates are taken into consideration. Staff satisfaction is another staff-related measure that is used in one half of all models evaluated. This dimension is directly related to the quality of care. The authors, however, stress that staff satisfaction might not be the most appropriate measure, as it can be subjected to ‘gaming’. For example, CPOs may pressure staff to report higher satisfaction than is actually enjoyed, in order to attract incentives.

While inspection data can potentially provide useful information for a care payer, the data itself is often problematic and may be “criticized because of inconsistency in survey practices and rates of citations between states and regions within states” (2009, p. 590). Care deficiencies, for example, are a potentially useful indicator. In practice, however, the risk is that surveys are not carried out consistently across different CPOs. Nonetheless, two thirds of the models studied incorporated various inspection elements. It is further noteworthy that, in line with existing literature, QOL measures were used only sparsely; in only one model was this type of QI incorporated.

The authors further explain that clinical QIs, as set out by regulators in the form of Minimum Data Sets², also have their limitations. In the case of the NH sector in the US context, their usefulness might be limited. These indicators have been criticised from a clinical, as well as a methodological perspective. One of the biggest concerns about MDS measures, according to Arling, Job and Cooke (2009), was that these contained only ‘minimal risk adjustment’ and only utilised data based on 3-month time periods. The authors explain that the Minnesota model, on the other hand, “extensively risk adjusted and derived from data pooled over four calendar quarters to increase their reliability” (p. 590) and was, for these reasons, fairer compared to the MDS clinical indicators.

Despite their subjective nature, surveys of care user and family satisfaction were also commonly used as a quality dimension. This was usually counterbalanced by reliance on more measures of administrative process and outcomes. All states surveyed care users, while family members were surveyed by four states. The experiences of the nursing home sector in the US suggest that response rates to these surveys tend to be low, and that a minimum response might be

² E.g. USA: <https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MDSPubQIandResRep/index.html>
Australia: www.aihw.gov.au/national-minimum-data-sets

warranted before they are used as a performance indicator – for instance, a threshold for use might be placed at 30 per cent. Moreover, the authors flagged that these instruments might be biased towards the healthier, ‘cognitive intact’ section of the care user population, and that this should be kept in mind when weighing the various quality dimensions.

Lastly, although five states also introduced QIs to test the efficiency of care, existing evidence suggests that efficiency is only indirectly related to care quality.

Financial incentives used in the P4P models:

The financial incentives that were most frequently used in the analysed P4P models consisted of “a bonus or an add-on to a facility’s per diem rate, either a relative per cent increase (e.g., 1%-4%) or a fixed dollar amount (e.g., (\$1-\$3)” (p. 590).

In the context of the US nursing home sector, legislative appropriations have the capacity to influence the nature of the incentive payments from year-to-year. With regard to the relationship between the size of financial incentives and the degree of behavioural change by CPOs, the authors argued that the “bonus is generally proportional to a quality of performance score – the higher the score, the greater the bonus” (p. 590).

In most models, the performance scores were a simple sum (5/6 models) or a weighted sum (1/6) of the individual quality dimensions. Minnesota, however, also had a more complex, tender-based model in place, which is described in the next section.

Key learning:

- **If the purpose of P4P is to lift quality, then efficiency measures might be better placed outside of the scope of a P4P model.**

4.3 A competitive tendering P4P initiative in Minnesota

In addition to the standard P4P model used by the Minnesota Department of Human Services (DHS), the State of Minnesota, as the care payer, also introduced a competitive tendering form of incentivised care, which seeks active CPO participation in identifying and improving the quality of nursing home care across most appropriate areas of the organisations. CPOs are invited by the department to submit proposals for the improvement of quality and/or efficiency of care. The Minnesota tendering model seeks to stimulate changes in existing work practices, focussing on collaboration and shared learning.

The Minnesota case highlights the opportunity to put multiple incentive measures in place across an industry. These measures are tailored towards the quality/efficiency needs of a range of CPOs. It needs to be noted, however, the tender-based model is an administratively complex and costly operation to run (Cooke et al. 2010).

Before it was suspended in response to the aforementioned financial constraints on the State of Minnesota, there were 45 Performance-based Incentive Programs in place. 27 of these initiatives were for individual facilities, while 18 were collaborative projects representing multiple facilities. Most of these projects focused on clinical quality and the introduction of new technologies. There was, however, a wide variety of initiatives, ranging from projects focussing on issues from psychosocial problems of the care users to the ability to change the organisational cultures of CPOs.

What is the underlying rationale of the tender-based P4P model?

Rather than the care payer dictating the areas of care that should be improved, the tender-based aspect of the Minnesotan model employs a “bottom-up” approach, which encourages CPOs to identify opportunities to improve quality or care efficiency.

Who is targeted to improve the quality and/or efficiency of care?

CPOs are targeted under this model, which is based on the premise that the CPOs are best positioned to identify the quality gaps in care delivery within their organisations.

What behavioural changes is the model geared towards?

Behavioural responses are dependent upon the initiatives proposed by the CPOs. The care payer organises workshops to assist CPOs with drafting their tender proposals. All proposals must be based on 6 underlying principles:

1. Efforts must improve quality, increase efficiency, and/or shift resources from institutional to community care;
2. Proposals must demonstrate how EBP can improve the quality and/or efficiency of care;
3. Providers should seek to innovate and take risks;
4. Collaboration and information sharing within and between CPOs is stimulated;
5. A business case for investment must be made;
6. The key elements of successful initiatives must be identified and must be disseminated across the industry (adapted from Cooke et al. 2010)

What is the content of a typical CPO tender proposal?

Cooke et al. (2010, pp. 557-558) explain that a tender proposal must contain the following:

1. A detailed description of the project, including the goals, basis for assuming the goals can be achieved, characteristics of residents anticipated to participate, the precise nature of the intervention, and the resources and timeline needed to achieve the goals;
2. For collaborative projects, the rationale for the collaboration and a clear plan of leadership and member involvement;
3. How performance outcomes will be measured, reported, and evaluated, including baseline data, performance period, and target levels to be achieved; and
4. A provision that puts the provider at risk of losing a minimum of 20% of the proposed incentive payment, if the stated outcomes are not achieved.

What incentives are awarded?

According to Cooke et al. (2010, p. 557) , “facilities can request an incentive payment up to 5% of their operating rate per diem for 1–3 years. Incentive payments range from \$.32 to \$13.32 with an average incentive payment of \$5.25 per day. The amount of the incentive payment varies based on project scope and complexity.”

What will be the end goal of the model (or pilot)? What happens when goals are achieved?

Although this initiative was suspended owing to financial constraints, the design of the model is based on the notion of seeking to stimulate ‘continuous improvement’. Hence, learning from every round of the model is incorporated into the next. For instance, most projects conducted later on in the model used an industry-wide report card of standardized quality measures. This change has improved data integrity, increased the scope to compare quality across the industry, and enhanced data management. The model also enables care payers to change strategic priorities over time, because after each funding round the industry’s desired outcomes are evaluated, and the list for subsequent rounds is adjusted to reflect current care and quality needs.

How successful was the implementation of this aspect of the incentivised care initiative in Minnesota?

“Success rates varied by type of project and performance measures targeted by the intervention” (Cooke et al. 2010, p. 562).

Stages of the tender-based P4P model:

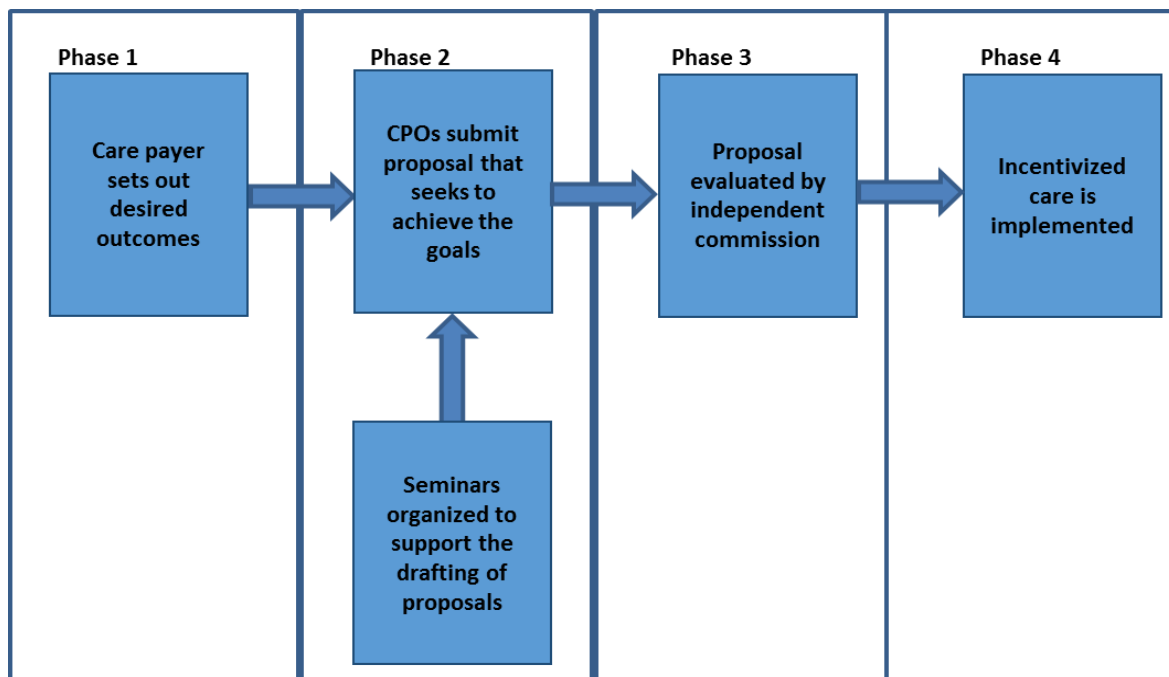


Figure 2. Outlines the four phases of the tender-based P4P model in Minnesota

Phase 1: CPOs are invited annually by the care payer to submit a proposal for the improvement of the quality and/or efficiency of care. This process is called the “Request for Proposal”.

Phase 2: CPOs submit proposals, which may include collaborative projects involving multiple facilities. The CPOs can request the incentive rates, which will be tailored according to the scope and complexity of the proposed quality initiatives.

CPOs are supported by the care payer via seminars to structure their proposals. These seminars “are held to aid the providers [in] understanding the components of a successful proposal, identify valid outcome performance measures, baseline data, and performance targets, and understand the rate impact for both successful projects and those not meeting performance targets” (p. 558).

Phase 3: “An independent selection committee evaluates the proposals and considers how well the project addresses program goals, [the] strength of evidence for the potential success of the project, [the] validity of the outcome measures, and the likelihood of broad-based dissemination and sustainability” (p. 558). The care payer “then specifies selected project details (e.g., intervention, outcomes, performance period, and the willingness to participate in program evaluation) in provider contracts” (p. 558).

Phase 4: “Provider incentive payments are added to their per diem rates. With the contribution of federal participation and private pay residents (whose per diems are specified under Minnesota rate equalization), PIPP [percentage of income payment plan] funding results in approximately US\$18 million available for projects annually.” (Cooke et al. 2010, p. 558)

4.4 An alternative quality-based payment model for Minnesotan nursing home care

The add-on (non-tender-based) P4P model in Minnesota, discussed in section 4.1, is not the model that was originally developed and proposed, as Kane et al. (2007) explain. The incentive model that was initially designed was only partly introduced due to the political lobbying of nursing home trade associations, which, although they supported the model in principal, objected to the proposed pace of implementation.

The key difference between the model initially proposed and the model that as implemented is that the original one took “into account a facility’s quality and its operating efficiency” (Kane et al. 2007, p. 112). It was based on a ‘Donabedian (1988) framework’, in that it included structure, process, and outcomes measures.

The model was set up in such a way that different weightings were attributed to the various quality components, all of which have their own QIs. Moreover, adaptability was built into the model. The authors explain that weightings and the inclusion of various measures could easily be changed when more information on their relevance and/or importance becomes available. Table 5 identifies the various components of the model. It also illustrates the difference between the weightings of the quality indicators used, versus those that were initially proposed.

Table 5. Quality points system

Quality measure	Brief definition	Point assignment	Quality points	Proposed quality points
Direct care staffing level	Case-mix adjusted/wage-adjusted direct care staff hours per day (e.g., registered nurses, licensed practical nurses, nursing assistants, activity directors, and other direct care staff).	Each of three peer groups (standard, hospital attached, and boarding care homes) will be assigned thresholds for achieving maximum and minimum points. Points will be distributed on a straight line between these two points.	0	10
Staff turnover	Number of nursing staff who left between October 1 of one year and September 30	0 points if turnover rate > .70; 15 points if turnover rate < .20; otherwise, points are distributed proportionately	15	0

	of the following year, divided by number of staff	according to rates between .70 and .20.		
Staff retention	Number of nursing staff employed on October 1 who were still employed on September 30 of the following year, divided by number of staff.	0 points if retention rate < .50; 25 points if retention rate > .80; otherwise, points are distributed proportionately according to rates between .50 and .80.	25	20
Use of pool staff	Pool staff hours as a percentage of total nursing hours.	0 points if > .10 pool staff hours; 10 points if no pool staff hours; otherwise, points are distributed proportionately according to rates between .10 and 0.	10	5
QIs from the Minimum Data Set	Summary QI score (range = 0–100) based on facility rates on 24 QIs in care domains such as behaviour or depression symptoms, incontinence, skin care, pain, psychotropic drugs, and nutrition.	0 points if QI score = 0; 40 points if QI score > 40; otherwise, points are distributed proportionately according to rates between 0 and 80.	40	35
Survey deficiencies	Survey deficiencies at Level F or higher for patient care-related F-tags.	0 points if facility had deficiency of H or higher; 10 points if all deficiencies were below F; and 5 points if highest deficiencies were F or G.	10	10
Resident quality of life and consumer satisfaction	Average facility score on a standardized resident interview covering quality of life and consumer satisfaction. Interviews are to be carried out by an independent contractor with approximately 10,000 residents per year.	To be determined.	0	20

Source: Adapted from Kane et al. (2007, p. 111)

The original model took “into account a facility’s quality and its operating efficiency. Facilities that deliver the highest quality at the lowest cost would receive payment in excess of their costs, whereas facilities with low quality and high cost would receive payment below cost. The model assumes that providers of all ownership types will be motivated to avoid operating losses and maximize revenue in relation to cost” (Kane et al. 2007, p. 112).

Therefore, the operating costs of facilities were to be divided into 2 categories:

- *Direct care costs* (e.g., nursing and other care-related services), and
- *Support services costs* (e.g., dietary, laundry, administrative, and other).

The proposed model by Kane et al. (2007) takes into consideration variations between care facilities, as well as regional variations. A facility's case mix (direct care costs) and regional wages would be used to "compare per diem figures across facilities with different resident acuity and regional labor markets" (p. 112). The authors indicated that this information was to be used to categorize facilities on the basis of this composite measure. CPOs were ranked into different tiers from low to high cost, depending on their composite scores. This is subsequently cross-tabulated with similarly constructed quality tiers.

The authors explain that applying this matrix system "is somewhat complicated, although the concept is straightforward. Nursing facilities are paid a higher proportion of their costs (even exceeding 100%), if their quality is better. The proportion depends on their quality score and how their costs compare to those of other nursing homes. No facility with poor quality will receive its full costs" (p. 114). Table 6 represents the matrix as it was proposed – but subsequently not implemented.

Table 6. Proposed payment rate as percentage of facilities per diem cost, by cost and quality tier

Cost Tier	Quality Tier				
	5 (Highest)	4	3	2	1 (Lowest)
1 (Lowest)	115%	110%	105%	95%	95%
2	110%	105%	103%	95%	95%
3	105%	103%	100%	95%	95%
4	103%	100%	97%	90%	90%
5 (Highest)	100%	97%	95%	90%	85%

Source: adapted from Kane et al. (2007, p. 114) – *Note this is not what was implemented!*

The system is based on the notion that facilities compete with themselves, rather than each other.

Thus, measurement is based on an initial 'base score' in the first year, against which progress is measured. Whilst it is not discussed by Kane et al. (2007), this approach could conceivably subject the system to 'gaming'. For example, CPOs might try to ensure that their baseline scores were lower than their actual performance, in order to more easily obtain future incentive payments.

Furthermore, the authors themselves recognise the risk that, if all providers raise their absolute score, there is an associated cost risk for care payers. Given the ultimate unsustainability of the funding of the State of Minnesota's incentivised care arrangements, these kinds of risks should be carefully evaluated by care payers when designing a model.

4.5 Successful use of P4P in San Diego in the 1980s – the importance of control groups

The case study below, on the implementation of P4P in the San Diego nursing home sector was reported by Weissert et al (1983) and summarised by Weissert and Frederick (2013, p. s148). It has been edited to make it relevant to the TAC and HDSG. Weissert and Frederick (2013) explain that the study is valuable because it found successes associated with P4P initiatives, and was, according to them, the only large randomized controlled trial for P4P models that has been undertaken in the context of nursing homes. As mentioned, this lack of properly constructed research and evaluation is a serious deficiency that impeded the informed implementation of most models. The robustness and value of this particular case study is also lauded by other authors (Briesacher et al. 2009; Miller, Doherty & Nadash 2013; Werner, Konetzka & Polsky 2013).

Although the study is dated, it demonstrates the importance of an adequate design process and of strict protocols and procedures to measure the effectiveness of incentivised models of care. Moreover, in terms of the lessons that it offers, it emphasizes the need for proper monitoring and feedback, as well as the use of an experimental control to verify the impact of incentivised care measures.

A Nursing Home Success: San Diego Nursing Home P4P Demonstration

In addition to being the only large, randomized controlled trial of P4P in nursing homes reported in the literature, the authors explain that the San Diego case is unique because it is one of the few studies that also looked closely into the cost-effectiveness of P4P initiatives.

Model characteristics:

The model was based on 36 for-profit Medicaid certified nursing facilities in California. Each nursing home contained at least 30 beds. The study lasted two-and-half-years, and assessed 2,400 care users quarterly.

Quality measurement:

Assessment periods included a six-month baseline and one year of admissions (requiring data collection for an additional 24 months after baseline to capture 12 months of data on each admission). Facilities were first paired for size and other factors, and then pairs were randomly split between a control group ($N = 18$) and a treatment group ($N = 18$).

Incentives:

CPOs in the treatment group were offered three bonuses:

1. A bonus payment for admitting severely dependent, heavy-care care users requiring expensive, skilled nursing care;
2. A bonus for discharging care users who were medically ready, provided that they remained in the community for at least 90 days; and
3. A bonus for achieving specified outcome goals in selected care users who require expensive, special care to either maintain or improve their health status (e.g., skin condition, or function after a stroke).

Control group:

The control group was ineligible for the bonus payments and was only compensated for the administrative costs associated with participation.

Care user group:

Care users were categorized by nurses, by their health status. The classification range was from type A, “the healthiest,” to type E, “the sickest”. The classification was based on six ADL (activities of daily living) measures—bathing, dressing, eating, using the toilet, transferring into and out of a bed, and continence—as well as the need for a range of skilled nursing services, such as wound care. Specific outcome targets were selected from a list initially identified by nursing home inspection staff as particularly difficult problems.

Theoretical justification:

The authors explain the justification of the study as follows: the way in which nursing homes are paid affects their effort. The effort made by nursing homes in turn, affects the health of residents. Since all nursing homes included in the study were for-profit, the study therefore assumed that these participating nursing homes would seek to maximize their profit.

Bonus:

Bonuses were structured to maximize revenues for appropriate performance:

- The admission bonus covered extra costs of expensive-to-care-for care users.
- The outcome payment paid facilities double their costs for each successful outcome, but paid nothing when success was not achieved. The goal of this arrangement was to encourage the nursing staff to try their skills on any acceptable candidate, but to give up when their efforts

were not working. Long experience in nursing home payment shows that simply paying for effort results in waste.

- The discharge payment was awarded only if a care user stayed alive and remained out of any nursing home or hospital for 90 days. Facilities were paid double their costs of effecting the discharge, encouraging home visits to check for safety concerns, assessing adequacy of caregiver support, and so on.

Findings:

The incentives offered were found to indeed trigger the type of behavioural responses from the CPOs that were expected (Norton 1992). It is noteworthy that each one of the performance goals was reached.

Clinical outcomes

- Admission of a more complex case mix, improved care users' health status and longevity, and increased discharges (resulting in shorter lengths of stay in both the nursing homes and the hospitals serving them).
- In all types of residents except the sickest, most costly, group (type E), the mean and median lengths of stay were much shorter in the experimental group. The probability that residents would either go home or go to an intermediate care facility was higher in the experimental group, and the experimental group was significantly more likely to admit more heavy-care type D and E care users, than the control group.

Cost-effectiveness

- Due to the significantly shorter lengths of stay, the costs per day in the treatment group were considerably less for all types, except for type E (the sickest care users). The authors explain that the study's goal was to induce better care for these sickest care users.
- In 1982 dollars, the average cost savings per stay was US\$3,000. The average daily cost to Medicaid rose by about 5 percent owing to administrative costs and incentive payments. Yet, these costs were small, compared with gains in improved health and lower hospital expenditures.
- The earlier discharge of healthier care users resulted in the freeing up of room in the participating nursing homes for the use of sicker care users. Therefore, those in hospital waiting for a nursing home bed could be discharged sooner, which had the knock-on effect of saving hospital costs.

Norton (1992) attributed the success of the model to two key factors: (1) external monitoring and feedback, and (2) continuous assessment of the untreated control group. The control group was used as a benchmark for the maintenance of effort on care users for whom no bonuses were paid. The authors explain that the control group design of the implementation avoid the undesired behaviour in which these care users might have had care drawn away from them by the nursing home, in favour of investing effort in care users who were participating in the incentive model.

While Weissert and Frederick (2013) suggest that the model was cost-effective, it also has to be noted that other authors, such as Briesacher et al. (2009), argue that the 5 percent increase in administration and incentive payments was quite a substantial cost. It was not possible for the researchers to further scrutinise whether or not this particular case was cost-effective. It is, however, strongly advocated throughout the literature as the only properly designed and evaluated example of P4P in the NH sector.

4.6 Reflections on the nursing home sector case studies

The above care studies in the NH sector demonstrate that incentivised models of care can be implemented in more than one way. Care payers made diverse decisions about the design and implementation of the models, despite the fact that the rationale for most of these models was to improve the quality of care. As highlighted throughout Appendix 2, design and implementation decisions, such as the extent of stakeholder involvement, had a direct impact on the mechanism of the cases discussed.

While the above cases illustrate some of the implications of design and implementation decisions for the operation of the models, it is not possible to ascertain that one of the models studied is necessarily better for improving the quality of care than another. The literature indicates that there is limited empirical evidence to support the argument that P4P actually improves the quality of care in the NH sector (Briesacher et al. 2009). Moreover, while there is some evidence that particular measures may be effective, this evidence is limited and, at times, contradictory (Weissert & Frederick 2013).

“... P4P in nursing homes did not result in consistent improvements in nursing home quality. Expectations for improvement in nursing home care under P4P should be tempered”
(Werner, Konetzka & Polsky 2013, p. 1393).

The inability to establish empirically the effectiveness of P4P in the NH care sector can in part be explained by the poor design of existing endeavours, which limited the opportunity to properly evaluate them. This limitation also applies in other care sectors that have used incentivised care measures. The reviews by Briesacher et al. (2009) and Arling et al. (2009) of the empirical literature on NH P4P models highlight that few of these studies could be properly evaluated; see Table 1 for a summary overview of the Briesacher studies.

Of the 13 programs that Briesacher et al. (2009) reviewed from 1980 to 2007, only four studies could be evaluated. The San Diego case study – as outlined in section 4.5 – was the “only one [that] provided sufficiently robust evidence for drawing any clear conclusions” (p.3). Hence, the Briesacher et al. (2009) suggest that new P4P initiatives should try to learn from the strengths of this study and even attempt to replicate it. Most P4P models reported in the literature thus far are only short-lived quality improvement exercises, with little demonstrable evidence about their effectiveness. It is currently unclear whether the short-lived improvements in outcomes are due to the lack of efficacy of the incentive, or the funding implications which have limited the duration of implementation.

An observation about the studies reviewed by Briesacher et al. (2009) about the studies that is equally applicable to the case studies in this report, is that the NH sector tends to rely on eight quality dimensions in the P4P models that are highly similar to the quality dimensions discussed by Arling et al. (2009). However, none of the programs relied on all of these dimensions. The Minnesota and Oklahoma models were the most refined – these were also evaluated by the authors – and relied on six out of the eight quality dimensions. The eight quality dimensions were:

1. Staffing (e.g., low staff turnover);
2. Performance on certification survey (e.g., deficiency-free inspection);
3. Quality indicators from the Minimum Data Set (e.g., no physical restraints or no new pressure ulcers);
4. Facility efficiency (e.g., occupancy and operating costs);
5. Service to Medicaid enrollees;
6. Resident and family satisfaction;
7. Quality of life (e.g., privacy and comfort); and
8. Other outcomes (e.g., timely discharge and improved functioning).

Of the incentive payments used, it has already been mentioned that these tend to be lump-sum payments distributed on the basis of Medicaid patients served, and percentages of the per diem rate. What is often missing, however, is how the care payers arrive at the rates that are being paid.

Frequently, “the rationale for the level of the incentive payment (whether there was any relationship to the costs of improving care) was poorly documented, and the proportion of providers receiving the incentive payment varied considerably” (Briesacher et al. 2009, p. 9).

Moreover, Miller et al. (2013) highlighted the importance of the financial sustainability of the model. Werner et al. (2013) explain, however, that in general, the sources for the incentive payment are poorly described across most studies. Hence, the studies presented here were useful, because they illustrated the importance of ensuring that incentivised models of care are designed to be financially sustainability.

The work of Werner et al. (2013) is the only large-scale evaluation of quality in nursing homes that compared facilities across states, both with and without P4P models. This study compared data from 17,579 nursing homes, of which 3,513 (20 percent) were located in states where P4P models were in place. The authors accounted for differences across facilities and care populations, and concluded that, on the basis of the available quality measures (see Table 13 in Appendix 2³), P4P had a mixed impact on care outcomes. One of the interesting aspects of this study was that it compared the impact of P4P over time. That is, they tested for differences at 1 year, as well as 2 years, after the implementation of incentive measures, and compared the quality results of the P4P group with the non-P4P group.

What Werner et al. (2013) found was that those NHs in the P4P group scored higher on 3 measures (decline in proportion of residents who were physically restrained; in the proportion of residents in moderate to severe pain; and decline in the development of pressure sores), after the first year. At the same time, they also found that the quality of care worsened in the P4P group on 2 measures (the percentage of patients with a bladder catheter, and the incidence of unexplained weight loss). After the first year, however, “the changes in resident outcomes were in the same direction and magnitude” (p. 1403) for both the P4P and non-P4P nursing homes. This would therefore suggest that the impact of the evaluated P4P models was relatively short-term.

Moreover, P4P did not impact on the staffing levels of NHs. From this finding it can be inferred that P4P was not used as a measure to increase or reduce staffing levels.

³ NB: The information in the Table 1 may appear to contradict what was outlined in Section 4.1, i.e., that the Oklahoma and Utah schemes were placed on hold. It was not possible for us to ascertain whether the schemes had been re-activated or not.

A concerning finding was that the “total number of nursing home deficiencies increased under P4P” in the first year after the implementation, when compared to non-P4P nursing homes (Werner et al. 2013, p. 1403).

One possible explanation put forward for the lack of success of P4P initiatives in the NH context is that care payers, when designing these programs, did not take into account the ability of CPOs to make investments. Medicaid facilities - those facilities at which the bulk of P4P measures in US States were targeted - are often in a poor financial position to make investments. The authors further suggest that financial incentives in this case might not have been sizeable enough to motivate the CPOs to make the behavioural changes necessary to improve quality (Werner et al. 2013). However, they observe that one of the models that appeared quite effective, in Georgia, had relatively small financial incentives in place. For this reason, they suggest that more careful scrutiny of the non-financial side of incentivised models of care would also be advisable.

Another challenge is to create a clear link between financial incentives, feedback, and performance. It is common for models to award incentive payments at the end of the financial year, thereby watering down the direct link between pay and performance. As Werner et al. (2013) explain, in theory it might be preferable to ensure that performance feedback and payment are structured on a monthly basis, yet they point out that, in practice, this might not be feasible – particularly from an administrative point of view. For instance, small facilities might not be able to generate a sufficient amount of data over such a short time frame to allow meaningful analysis and benchmarking.

A further explanation offered for the limited effectiveness of P4P in nursing homes was that financial incentives are geared at the CPOs, rather than the care professionals. Werner et al. (2013, p. 1409) explain that:

Providing payments to organizations, rather than individuals, has conceptual appeal as quality deficits are thought to be system based. Payments to organizations can be used to help improve system failures by investing in large-scale approaches to quality improvement that would be expensive or infeasible for individuals to implement. Targeting payments to organizations may also be helpful if individuals are risk averse. Organizations, however, are difficult to motivate and to hold accountable for direct effects on patient care. Thus, providing payments to individuals may have a larger effect (Rosenthal and Dudley 2007). Bonus payments based on facility-level performance could be at least partially redistributed

to managers or front-line providers (directly or by facility managers themselves) to increase personal motivation to improve performance.

5 Summary of key findings and discussion

The aim of this review was to identify and evaluate incentivised models of care both within and outside the disability sector, in order to identify practices suitable for transfer to the TAC compensable sector in Victoria. We therefore examined methodologies for the design of incentivised care models, and assessed how models of incentivised care were implemented. We have identified the underlying fundamentals of these models, assessed the effectiveness of existing models, and outlined the pre-conditions for their implementation. More detailed suggestions on each stage of introducing an incentivised model may be found in Appendix 2.

5.1 Existing models and their effectiveness

The review also assessed the effectiveness of incentivised models of care across various areas of healthcare, both domestically and internationally. We relied on systematic reviews that had previously been conducted, to assess whether introducing incentive measures is an effective mechanism to improve the quality and/or effectiveness of care. Our conclusion, on the basis of the existing literature, has to be that **there is no convincing evidence that would suggest that incentivised models of care, such as P4P, are an effective way to raise the quality of care.** We recognise, however, that – in theory – incentivised models may result in positive outcomes, if:

- The model has a clear goal;
- The measures are carefully designed;
- All stakeholders are brought along in the process;
- A proper baseline analysis is conducted; and
- Thorough evaluation takes places – preferably including the use of a control group.

That being said, Greene and Nash (2009) explain that P4P plans constantly fail to deliver on their potential. They identified that poor design, implementation, and evaluation were the biggest errors committed by care payers across different care areas. Often, the utilisation of incentivised models appears to be a case of coercive isomorphism driven by policy makers, rather than being driven by the purposeful desire of care payers to tackle specific problems in the quality or effectiveness of care delivery. In such instances, the instruments that are used are not necessarily the mechanisms that would be most effective to resolve particular issues but, merely those that are most popular.

Our analysis of the literature further highlighted a number of recurring issues. Firstly, poor model design and implementation, including the lack of consultation and the use of poor quality measures undermine the viability of models that were assessed across the various systematic reviews. Secondly, incentives were not always properly designed. Consequently, causality between outcomes and incentives could not be validated empirically, and this therefore diminished the credibility of these models. Poor design was apparent in cases when, for example, the size of incentives was not commensurate to the desired behavioural outcomes, or where, the influence of external factors was not properly accounted for. It needs to be demonstrated that incentive measures, and not any other factors, are the cause of changes in the care delivery. Thirdly, studies lacked a proper baseline analysis, which diminished the ability of the researchers to assess the effectiveness of the measures utilised. For instance, in the nursing home care sector, very few studies utilise control groups as a validity measure. Fourthly, cost-effectiveness is difficult to account for without information on all associated costs, and often, the necessary data to evaluate all aspects of a model were unavailable.

In closing, there is unfortunately no ‘best practice’ approach for incentivised models of care. The literature has not provided any conclusive evidence as to what arrangements are most successful. The outlined recommendations, however, provide a roadmap for the most likely route to success. It has to be stressed, however, that incentive measures are not always the most appropriate remedy to resolve concerns about care quality. Whether or not incentives will be appropriate an appropriate means to improve care quality will depend on the type of quality improvements that are sought, contextual factors, and the ability of the care payer and other stakeholders to properly design and implement a model.

5.2 The purpose of incentivised care models

We identified that the purpose of the majority of incentivised models of care examined was to improve the quality of care across different areas, either broadly or focussing upon specific issues. In some instances, the models were also tailored towards raising the effectiveness of CPOs. However, research thus far has failed to conclusively demonstrate that incentive programs do in fact have the capacity to deliver on either of these objectives.

The rationale behind incentivised models is that existing FFS funding models provide limited motivation for CPOs to improve the quality of care beyond the level what is regarded as essential. By introducing incentives, both monetary and non-monetary, the goal of a care payer is to stimulate improved service delivery in a care sector, and possibly to move towards more person-centred forms

of care. Person-centred care is what, for example, the NDIA seeks to achieve through the NDIS scheme. The literature indicates that incentive measures can be used to target three areas of care:

1. Clinical care
2. Organisational systems
3. Care user experience

5.3 Foundations of incentivised care models

In order for incentivised models of care to have any reasonable chance of success it is essential that such initiatives are built on solid foundations. The models are a mechanism to drive behavioural change amongst CPOs, yet without clear direction, they are bound to fail. Problems, for instance, arise when a care payer fails to formulate clear *Goals*, *Strategies*, and *Objectives*. Since challenges differ from sector to sector we were not able to identify a template model that could be easily imported into the Victorian disability sector. As Petersen et al. (2006, p. 269) explain, the challenge is to design incentives with an intended goal in mind.

An incentivised model of requires that a care payer formulates clear Goals, Strategies, and Objectives. A model cannot motivate providers ‘to do the right thing’ when it is not clear what the right thing is.

Thus, before introducing incentive measures, the underlying problems that a care payer seeks to address should be clearly identified. Even when the healthcare quality and/or efficiency problem to be addressed has been identified, it still needs to be established that introducing incentive measures is actually the most effective remedy to resolve the issues. EBPs are warranted to justify the take up of incentive measures. The literature highlighted that any organisation wanting to design a P4P scheme should consider:

- Is it possible to properly measure the identified issue?
- Can the quality measure be used with a reasonable level of validity and robustness?
- Are incentives likely to trigger the desired CPOs behaviours?
- Can broad stakeholder buy-in be realised?
- Can the model be designed in such a way that unintended consequences and undesirable behaviour by CPOs and care professionals are mitigated?
- Can a model be set-up on a financially sustainable basis for both care payers and CPOs?

The literature points out, for example, that incentivised care models that seek to improve the quality of care require very careful design, as they will otherwise be susceptible to unintended

consequences and gaming of the system by care providers. Further, it is preferable that models should be gradually introduced, so that CPOs and care professionals have time to familiarise themselves with the model and its expectations. A contemporary example is the staged approach currently being taken to implement the NDIS across the Australian disability care sector.

Although incentive programs have recently increased in popularity, both in Australian healthcare (Boxall 2009; Partel 2014) and overseas, it has to be noted that the use of these programs is often based on a resurging interest in new public management (Perry, Engbers & Jun 2009). It is further, the result of political decision-making, rather than convincing empirical evidence about the effectiveness of incentive models (Partel 2014). Incentivised care models, and particularly P4P models, were already used throughout the 1980s and 1990s with inconclusive results, after which their implementation dropped off. The recent, increase in use of these quality initiatives is notable in multiple areas of care, including nursing homes (Briesacher et al. 2009) and healthcare (Rosenthal & Frank 2006). The literature thus sounds significant warnings about where and why the implementation of a P4P scheme may be appropriate, and more importantly, when it is not.

“Don’t adopt conventional pay-for-performance systems simply because everyone else is doing it. Consider the contextual contingencies and adapt accordingly.”

(Perry, Engbers & Jun 2009, p. 46)

5.4 What areas of care should be incentivised?

Incentivised models of care can be employed by care payers to address specific – targeted – quality issues in an industry, while they can also be utilised to improve care more broadly. The literature highlights that, while from a cost perspective it might be favourable to use a relatively narrow incentive model, the risk is that narrowly targeted incentives result in distorted quality delivery. Narrow models additionally tend to be more susceptible to undesired behaviour by CPOs and other actors. Moreover, previous research has demonstrated that by themselves, incentivised models of care such as P4P, are likely to have a more limited impact than when they form a part of broader quality improvement initiatives throughout a sector (Arling et al. 2009; Weissert & Frederick 2013).

Incentivised models of care are most successful when they are part of broader quality improvement initiatives

We identified from the literature that incentive measures have been used to improve different dimension of care. Appendix 1 highlights the various quality measures and incentives used in P4P model in the US nursing home context.

As previously mentioned, there was very limited empirical material on incentivised care in the disability sector. In the case study section of this report we argued, however, that valuable lessons could be learned from the nursing home care sector. There are a number of similarities between the disability and nursing home care in relation to care delivery, the organisational structure of providers, and the HR challenges of the sector. Hence, it is worth noting that those quality dimensions that were most frequently utilised in P4P models in the NH sector were:

- Staffing;
- Performance on certification survey;
- Clinical quality indicators;
- Provider efficiency;
- Services to care users;
- Resident and family satisfaction;
- Quality of life; and
- Other outcomes.

While QOL is one of the measures that has been included in a number of models, the operationalisation of this measure has been problematic, contentious, and has therefore been limited.

5.5 Designing incentivised care models

“Creating effective incentives is difficult. The design of the incentive must reflect the values and goals of the healthcare purchaser, be well matched to the performance objectives and reflect a range of contextual factors that can influence the effectiveness of even well-designed incentives. Incentive design can be likened to building a house: form follows function and the design of an incentive model depends on what you want from it.” (Custers et al. 2008, p. 12)

In order for the TAC to introduce an effective incentivised care model into the compensable sector a second critical element, after first identifying for what purpose it will be used, is to determine which quality measures and incentives can be used. Together with the implementation process, this is one of the most critical steps towards implementing effect incentivised care. As the case studies highlight, from an administrative perspective, it may be advantageous to rely as much as possible on

existing measures to minimise burdensome work, however the links between quality improvements and incentives need to be evidence-based.

Properly designed incentive measures have the ability to, on the one hand, achieve desired outcomes, but similarly, are able to mitigate undesired behaviours by CPOs and have no unintended consequences in other areas of care. Measures need to be put in place that minimise the opportunity for undesired CPO behaviours such as gaming, adverse selection, multi-tasking problems, and demoralisation.

The appropriateness of the quality measures utilised depends on their being an empirical basis to support their relevance to the desired outcomes. Where this empirical evidence exists incentives may then be linked to these quality measures. EBPs are important to convince stakeholders of the validity of measures. However, it is not always possible to directly link quality measures to measured outcomes. Hence other types of quality measure with an established relationship to improved quality can also be employed. Most incentivised models of care use a combination of different Donabedian structure-process-outcome measures, while composite measures based on a combination of these can also be included in the model's design. Outcome measures, for instance, are frequently affected by other external factors, so CPOs can be reluctant to be performance-managed on these terms. Consensus about which measures to include and exclude should therefore be reached before proceeding with implementation.

Appropriate quality measures can further be assessed in a variety of ways. Performance can be measured on an absolute or a relative scale, although the literature suggests that absolute measures are preferable. By utilising absolute measures that are 'tiered', a care payer can ensure that those CPOs that start from a lower base will be less likely to become demotivated, whereas an absolute performance threshold may have this effect because it can be perceived as unattainable.

Another consideration concerns how a model incorporates 'risk adjustment'. Risk adjustment needs to be undertaken to ensure that neither vulnerable care users, nor care users with disadvantages, nor the CPOs that serve them will be adversely impacted by the implementation of incentive measures. This model again is dependent on contextual circumstances. There is no clear template as to how care payers and other industry stakeholder should go about resolving this problem, however, through careful stakeholder consultation, it is possible to address such issues before implementation.

The variety of financial and non-financial incentives was found to be numerous and diverse. The most frequently used incentive model was, however, P4P. The characteristics of the various

incentives will be dependent upon the overarching purpose of a model and the contextual circumstances of the care sector for which it is designed. Incentives can have a variety of characteristics depending on the type of outcomes that are sought, the type of CPO behaviours that a care payer wants to trigger, the types of motivators (penalties and/or reward based) that can be employed, whether monetary and/or non-monetary rewards are used, whether changes are directly or indirectly pursued, whether the focus of the scheme is long- or short-term focused, and to what extent the desired quality improvements are either targeted, or broad.

Several different models of incentives were uncovered in the literature. These included schemes based on bonuses, performance based withhold, performance-based fee schedules, pay for activities, share saving contracts, linking regular payment to increases in performance, quality grants and performance funds, and financial awards. The non-financial measures included public reporting, earned autonomy, managerial replacement, cost differential for beneficiaries, and public reporting/recognition.

Financial incentives thus differed substantially, and were constructed in varying ways. It was outlined in this report that multiple choices can be made in relation to (1) the type of CPO behaviour that is rewarded; (2) the nature of the payments; (3) the structure of payments; (4) payment timing; and (5) payment quantity. The composition of the incentives found across the literature was often described as being the outcome of negotiations between stakeholders. With the P4P models in the NH sector, however, it was found that the nature of the payments themselves was most commonly either:

- Lump-sum payments based on the number of patients serviced; or
- Percentages of the per diem rate.

Incentive payments that CPOs received were further awarded on either their relative performance against other CPOs, or on their absolute performance. This arrangement in turn meant that the amount of the incentive payment is affected in the first case by the performance of industry peers, whereas in the latter case, it is fully at the discretion of the CPO. The literature suggests that from the perspective of CPO buy-in perspective, it is preferable to design incentives such that it is fully at the discretion of the CPO whether they achieve the incentive payment.

The importance of effective stakeholder management by the care payers was often cited as one of the key pre-requisites for the successful introduction of incentive measures. Ensuring a broad buy-in from all relevant stakeholders was shown to be essential, across the analysed areas of care, and was repeatedly illustrated throughout the case studies. The cases from the nursing home care

sector demonstrated that, aside from the CPOs that will participate, a range of other stakeholders can also be involved. The involvement of these other stakeholders should be seriously considered in the process of designing the model, to ensure that resistance to incentive measures is minimised. In the case of the TAC and HDSG for the Victorian compensable sector, this list of stakeholders may include:

- *Care users*: This group can include the care users and/or their family, both directly or indirectly through advocacy groups;
- *Care professionals*: Some form of worker representation, for instance involving trade unions;
- *Policy makers*: In the Victorian context, the possible inclusion of the Department of Health and Human Services (DHHS) might be appropriate.
- *Relevant Ombudsmen*: similarly, the inclusion of the Victorian ombudsman could be considered since this public official has investigated issues across the sector.⁴

The introduction of incentivised models of care is, in effect, the introduction of a program of change, often subconsciously. Poorly designed and implemented, or hastily implemented change management initiatives run the risk of resulting in negative outcomes, rather than the intended positive ones (Reichers, Wanous & Austin 1997). Moreover, change programs often fail to realise significant and lasting change (Beer 2003). In order to accept a model, CPOs need to understand them at three different levels, as Kane et al. (2007, p. 114) explains:

1. They must accept the measures and the calculations of quality as valid;
2. They must understand the way they are used to create a payment rate; and
3. They must be able to develop and implement strategies to improve quality, as measured through investments that will result in rate increases sufficient to justify those expenses.

Co-design of the quality measures and of how various measures are weighted is one way through which acceptance can be facilitated. Furthermore, the relevant stakeholders should be given the opportunity to familiarise themselves with the expectations of the change initiatives. CPOs and care professionals should additionally be provided with an opportunity to develop the right capabilities, skills and knowledge to adequately respond to the targeted quality improvements that are stimulated through incentivisation.

⁴ <https://www.ombudsman.vic.gov.au/getattachment/c6499f78-0eec-4e4a-8e94-e4cd716a64f8>

Another means by which care payers may employ to bring along CPOs and other stakeholders in the process is through clear, transparent and open communication, e.g., through workshops and other platforms.

Incentive measures seek to change the existing forms of care delivery. However, frequently, investment by CPOs is necessary in order to achieve quality improvements (e.g., improvement in clinical care, or improved administrative processes and systems). Therefore, as part of a scheme, a care payer will need to decide how it will compensate CPOs for any investments that they have to make in order to lift their performance. The literature highlights that there is a range of pre-conditions that affect the willingness of CPOs to do this, namely:

- The ability to make an ROI on investments;
- Certainty surrounding the sustainability of a model and the ability to recover costs; and
- Perceived fairness of the QIs in relation to measuring the performance.

With respect to the fairness of QIs, it has to be noted that when these measures are perceived to be subjective, or heavily influenced by external factors outside the discretion of CPOs, this will likely result in decreased CPO willingness to undertake the necessary investments.

From a care payer's perspective, on the other hand, it will be essential that compensation for investments is clearly coupled with improvement in quality. Thus, while it might be appropriate to reimburse a proportion of costs upfront, the literature advises that full reimbursement should be made contingent upon improved performance.

The literature further highlights that the legal implications of incentivised care need to be considered. While not providing clear guidance on the matter, the important point raised is that the legal implications of the model for relations between care payers, CPOs, care professionals, and care users must be given serious consideration. In particular, it must be ensured that appropriate remedy measures are in place when any issues are encountered, e.g.:

- Who is responsible if care delivery deteriorates?
- How to deal with a dissatisfied care user?
- How to resolve disputes between the care payer and CPOs arising from the model, such as disagreements about quality results or incentive payments?

Hence, some of these legal issues need to either be resolved through separate agreements, or be regulated on top of existing care contracts.

5.6 Evaluation

The success of models will also be determined by the degree and quality of the feedback that CPOs receive on their performance. This requires proper monitoring systems and feedback mechanisms. The literature, highlights, for example that there needs to be a clear relationship between demonstrated behaviours and incentive payments. As Figure 3 reveals, evaluations are an integral, and continuous aspect, of incentivised care models. Incentive measures need to 1) serve clear objectives, 2) be then properly designed, and, 3) measure performance carefully. It must then 4) be evaluated whether the desired outcomes are realised, and subsequently a model needs to 5) be adjusted where necessary. This adjustment includes modifying any QIs and incentive measures that are driving undesired behaviours or unintended outcomes. These five steps are illustrated in Figure 3, below.

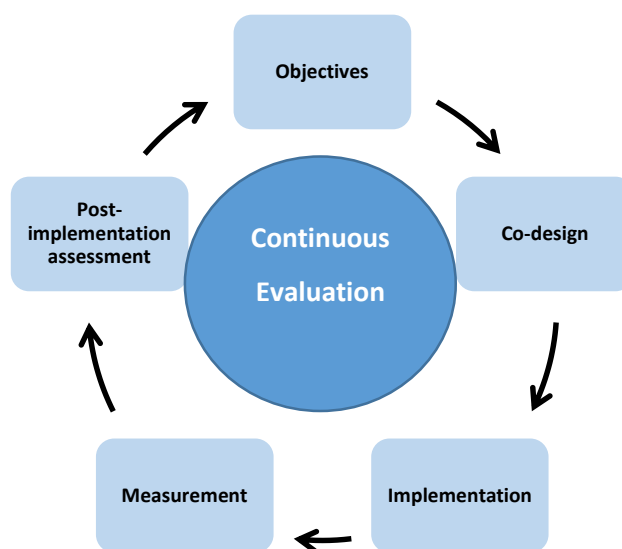


Figure 3. Key processes for the introduction of incentivised models of care

6 Recommendations

On the basis of the review, we make the following recommendations:

1. The TAC, as the dominant care payer in the Victorian compensable disability care industry, can actively shape the care experience of care users and drive the development of targeted quality improvement initiatives throughout the sector. The introduction of incentivised care measures might assist the organisation in achieving this objective. It is not, however, a proven panacea for quality issues.

2. It may be worthwhile for the TAC to consider the potential of incentivised measures to address issues across the broader industry, before experimenting with it in the context of a pilot study with one provider. In particular, we would suggest that it needs to be identified which goals, strategies and objectives the introduction of an incentive model will fulfil.
3. If the TAC decides to proceed with the introduction of incentivised care measures, it should be willing to take the leadership in driving this initiative. While buy-in from providers is essential, it will be important that this is a care payer driven process that focuses on outcomes. For instance, the financial interests of CPOs should not be privileged over the quality of the care delivered to care users.
4. The TAC should identify what aspects of an incentivised model of care can be developed in-house, and what external expertise needs to be brought in.
5. The TAC will need to be aware that a change management processes associated with the introduction of incentivised care measures needs to be carefully managed. If this is not properly undertaken, the change initiative can be resisted and frustrated by different stakeholders.
6. In order to successfully implement an incentivised model of care, it will be essential for the TAC to identify specifically what areas of the care that need to be improved and the timeframes in which these outcomes should be achieved for key client groups so that targeted measures can be designed.

We have formulated further recommendations for the TAC and HDSG with respect to the design, implementation and evaluation of an incentivised model. Our recommendations are outlined in Appendix 3. Here we highlight our recommendations for each of the five phases that we identified. These additional recommendations provide direction on broad strategic and operational considerations only, since the literature clearly suggests that a model needs to be tailored to contextual circumstances.

7 Conclusions

7.1 Summary

It has been a long-cherished hope amongst politicians and policy makers that, by introducing incentive measures into different areas of care, some of the quality and efficiency challenges that obtain under existing funding models can be resolved (Werner et al. 2011). Our review of the existing literature on the use of incentivised models of care revealed that these models thus far have

not systematically been used in in the disability care sector. We found no (systematic) evaluation of the design and/or implementation of such measures within the disability sector, whether in Australia or internationally. While incentivised models as solutions may, theoretically, sound hopeful (Mehrotra, Sorbero & Damberg 2010; Rosenthal & Frank 2006; Scott et al. 2011), findings of systematic reviews in the nursing home care sector have revealed that, despite this hope, there is, in practice, limited existing evidence to support this assumption. This finding should function as a warning to care payers such as the TAC. Incentive measures, such as P4P models, are by no means a magic bullet for fixing quality or efficiency issues. We draw a number of conclusions from the findings of our review study, as follows:

- I. We should not have unrealistic expectations about incentivised care measures and the quality improvements that can be realised with these measures – prior systemic reviews across the other care areas, and non-care sectors, all revealed that the success of incentivised arrangements tends to vary. There is only limited evidence to suggest that the quality of care or the cost-effectiveness of care delivery will be improved when care payers decide to implement incentivised care models.
- II. The purpose of an incentivised model needs to be clear to all stakeholders involved. Care payers need to identify clear goals, strategies, and objectives that they want to realise. CPOs and care professionals further need to be convinced that these are the genuine reasons for the scheme.
- III. Poor design and implementation, including the lack of consultation with relevant stakeholders, and the use of poor quality measures undermined the viability of numerous prior initiatives.
- IV. Unintended consequences, e.g. undesirable behaviour by CPOS such as ‘gaming’, threaten the viability of schemes. This includes the risk of possible detrimental care outcomes.
- V. When quality measures are not properly designed this severely diminishes the credibility and validity of findings. The inclusion of measures should be based on empirical evidence, i.e. there needs to be proof that measures relate to targeted outcomes.
- VI. Narrow quality measures have difficulty accounting for all quality improvements, there is a need for broad multi-dimensional measures. Narrow measures may also fail to identify unintended consequences in relation to non-incentivised areas of care.
- VII. Financial and non-financial incentives need to be targeted at identified desired behaviours. While evidence demonstrates that both can result in improvements in the quality and/or effectiveness of care, neither will guarantee success.

- VIII. The size of financial incentives can either be too small or too large. Small incentives do not result in the desired behaviours, whereas incentives that are too large become unsustainable for care payers. Careful modelling is therefore required to identify the appropriate size of measures. This should be informed by contextual factors, including required investments by CPOs.
- IX. Non-financial incentive measures which seek to stimulate consumerism amongst care users are largely unproven. Not all care users are able to 'shop around' for the best care;
- X. The influence of external factors is often not properly accounted for in the design of incentive programs and quality measures. It needs to be ascertained that incentive measures and not any other factors are the cause of changes in the care delivery.
- XI. Incentive models may be based on **the notion that facilities compete with themselves, rather than each other** (Kane et al. 2007). However, by competing with themselves and become more efficient and effective in their services, CPOs will be stronger and more competitive in the market. This has implications for CPOs under the NDIS which has been set up to encourage competition between the providers.

7.2 Limitations of the current study

This study reports on an in-depth assessment of the various models of incentive care that could be found within the academic and grey literature, in Australia and abroad. Nonetheless, it is subject to several limitations which should be acknowledged.

Firstly, it is recognised that there is more literature available on incentivised care, particularly on P4P initiatives in the hospital and primary care sectors, than could be scrutinized within the timeframe of this study. Due to time and resource constraints, we limited our analysis to incentivised models of care within the nursing care setting. However, we relied on highly cited academic literature and thoroughly conducted systemic reviews – all of which had similar outcomes and conclusions. Consequently, we are confident with the presented findings. There might, however, be some relevant material that was not captured by our existing search queries.

Secondly, another limitation of our findings was that only limited material on the disability care sector was uncovered, and no systematic reviews of careful evaluations of incentive measures were found from this sector. While our search efforts were methodical and included the most commonly used databases, it might be that possibly useful material from other sources was overlooked.

Thirdly, only limited material on incentivised models of care in the Australian context was included in the analysis. As described, incentivised models of care are context-sensitive and need to

be tailored to the specific circumstances of a care payer and CPOs, including adjustments to incorporate sectoral and regulatory factors. For instance, the outlined case studies were all conducted in the context of the US Medicaid system, which promoted the use of P4P models. Hence, the material presented should be considered within these limitations. Although this does not diminish the findings, readers need to be wary of this. The findings from the grey literature (Boxall 2009; Partel 2014) would suggest, however, that incentivised models of care, including P4P, are not likely to operate vastly differently within the Australian context.

Fourthly, due to time and resources constraints it was not feasible to explore the non-English language literature.

In conclusion, despite these limitations we are convinced that our findings can make a meaningful contribution to the TAC and HDSG's discussions on the merits and limitations of incentivised care models. We also identified a number of prospects through which further research could possibly assist the TAC and HDSG in the next steps of developing and implementing an incentivised model of care, these are outlined in Appendix 4.

List of references

- Arling, G, Job, C & Cooke, V 2009, 'Medicaid Nursing Home Pay for Performance: Where Do We Stand?', *The Gerontologist*, vol. 49, no. 5, pp. 587-595.
- Beer, M 2003, 'Why Total Quality Management Programs Do Not Persist: The Role of Management Quality and Implications for Leading a TQM Transformation*', *Decision Sciences*, vol. 34, no. 4, pp. 623-642.
- Bell, CM & Levinson, W 2007, 'Pay for performance: learning about quality', *Canadian Medical Association. Journal*, vol. 176, no. 12, pp. 1717-9. Available from: ProQuest Central.
- Bodrock, JAMRNC & Mion, LCPRNF 2008, 'Pay for Performance in Hospitals: Implications for Nurses and Nursing Care', *Quality Management in Health Care April/June*, vol. 17, no. 2, pp. 102-111. Available from: Journals@.
- Boxall, A-m 2009, *Should we expand the use of pay-for-performance in health care?*, Parliamentary Library.
- Boyd, CM, Darer, J, Boulton, C, Fried, LP, Boulton, L & Wu, AW 2005, 'Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: Implications for pay for performance', *JAMA*, vol. 294, no. 6, pp. 716-724.
- Briesacher, BA, Field, TS, Baril, J & Gurwitz, JH 2009, 'Pay-for-Performance in Nursing Homes', *Health Care Financing Review*, vol. 30, no. 3, pp. 1-13. Available from: PMC.
- Casalino, LP, Elster, A, Eisenberg, A, Lewis, E, Montgomery, J & Ramos, D 2007, 'Will Pay-For-Performance And Quality Reporting Affect Health Care Disparities?', *Health Affairs*, vol. 26, no. 3, pp. w405-w414.
- Cashin, C 2014, 'P4P programme design', in *Paying for Performance in Health Care Implications for Health System Performance and Accountability: Implications for Health System Performance and Accountability*, eds C Cashin, YL Chi, P Smith, B M. & S Thomson, OECD Publishing / World Health Organization.
- Christianson, JB, Knutson, DJ & Mazze, RS 2006, 'Physician pay-for-performance', *Journal of General Internal Medicine*, vol. 21, no. 2, pp. S9-S13.
- Cooke, V, Arling, G, Lewis, T, Abrahamson, KA, Mueller, C & Edstrom, L 2010, 'Minnesota's Nursing Facility Performance-Based Incentive Payment Program: An Innovative Model for Promoting Care Quality', *The Gerontologist*, vol. 50, no. 4, pp. 556-563.
- Cromwell, J & Smith, KW 2011, 'Evaluating Pay for Performance Interventions', *Pay for Performance in Health Care: Methods and Approaches*, p. 267.
- Custers, T, Hurley, J, Klazinga, NS & Brown, AD 2008, '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 Services Research*, vol. 8, no. 1, pp. 1-14.
- de Bruin, SR, Baan, CA & Struijs, JN 2011, 'Pay-for-performance in disease management: a systematic review of the literature', *BMC Health Services Research*, vol. 11, pp. 272-272. Available from: PMC.
- Doran, T, Fullwood, C, Gravelle, H, Reeves, D, Kontopantelis, E, Hiroeh, U & Roland, M 2006, 'Pay-for-Performance Programs in Family Practices in the United Kingdom', *New England Journal of Medicine*, vol. 355, no. 4, pp. 375-384.
- Dudley, R 2005, 'Pay-for-performance research: How to learn what clinicians and policy makers need to know', *JAMA*, vol. 294, no. 14, pp. 1821-1823.
- Edwards, ST, Bitton, A, Hong, J & Landon, BE 2014, 'Patient-Centered Medical Home Initiatives Expanded In 2009–13: Providers, Patients, And Payment Incentives Increased', *Health Affairs*, vol. 33, no. 10, pp. 1823-1831.
- Eijkenaar, F 2011, 'Key issues in the design of pay for performance programs', *The European Journal of Health Economics*, vol. 14, no. 1, pp. 117-131.
- Eijkenaar, F 2012, 'Pay for Performance in Health Care: An International Overview of Initiatives', *Medical Care Research and Review*, vol. 69, no. 3, pp. 251-276.

- Emmert, M, Eijkenaar, F, Kemter, H, Esslinger, AS & Schöffski, O 2011, 'Economic evaluation of pay-for-performance in health care: a systematic review', *The European Journal of Health Economics*, vol. 13, no. 6, pp. 755-767.
- Gillam, SJ, Siriwardena, AN & Steel, N 2012, 'Pay-for-Performance in the United Kingdom: Impact of the Quality and Outcomes Framework—A Systematic Review', *The Annals of Family Medicine*, vol. 10, no. 5, pp. 461-468.
- Glidewell, L, West, R, Hackett, JE, Carder, P, Doran, T & Foy, R 2015, 'Does a local financial incentive scheme reduce inequalities in the delivery of clinical care in a socially deprived community? A longitudinal data analysis', *BMC Family Practice*, vol. 16, no. 1, pp. 1-13.
- Greene, SE & Nash, DB 2009, 'Pay for Performance: An Overview of the Literature', *American Journal of Medical Quality*, vol. 24, no. 2, pp. 140-163.
- Hackett, J, Glidewell, L, West, R, Carder, P, Doran, T & Foy, R 2014, 'Just another incentive scheme': a qualitative interview study of a local pay-for-performance scheme for primary care', *BMC Family Practice*, vol. 15, no. 1, pp. 1-11.
- Hamblin, R 2007, 'Publishing 'quality' measures: how it works and when it does not?', *International Journal for Quality in Health Care*, vol. 19, no. 4, pp. 183-186.
- Helm, C, Holladay, CL, Tortorella, FR & Candio, C 2007, 'The Performance Management System: Applying and Evaluating a Pay-for-Performance Initiative', *Journal of Healthcare Management*, vol. 52, no. 1, pp. 49-62; discussion 62-3. Available from: ProQuest Central.
- Kane, RL, Arling, G, Mueller, C, Held, R & Cooke, V 2007, 'A Quality-Based Payment Strategy for Nursing Home Care in Minnesota', *The Gerontologist*, vol. 47, no. 1, pp. 108-115.
- Karve, AM, Ou, F-S, Lytle, BL & Peterson, ED 2008, 'Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients', *American Heart Journal*, vol. 155, no. 3, pp. 571-576.
- Kilpatrick, KE, Lohr, KN, Leatherman, S, Pink, G, Buckel, JM, Legarde, C & Whitener, L 2005, 'The insufficiency of evidence to establish the business case for quality', *International Journal for Quality in Health Care*, vol. 17, no. 4, pp. 347-355.
- Kirschner, K, Braspenning, J, Akkermans, RP, Jacobs, JEA & Grol, R 2013, 'Assessment of a pay-for-performance program in primary care designed by target users', *Family Practice*, vol. 30, no. 2, pp. 161-171.
- Kirschner, K, Braspenning, J, Jacobs, JA & Grol, R 2012, 'Design choices made by target users for a pay-for-performance program in primary care: an action research approach', *BMC Family Practice*, vol. 13, no. 1, pp. 1-10.
- Korda, H & Eldridge, GN 2011, 'Payment Incentives and Integrated Care Delivery: Levers for Health System Reform and Cost Containment', *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, vol. 48, no. 4, pp. 277-287.
- Kristensen, SR, Siciliani, L & Sutton, M 2016, 'Optimal price-setting in pay for performance schemes in health care', *Journal of Economic Behavior & Organization*, vol. 123, pp. 57-77.
- Lester, HE, Hannon, KL & Campbell, SM 2011, 'Identifying unintended consequences of quality indicators: a qualitative study', *BMJ Quality & Safety*, vol. 20, no. 12, pp. 1057-1061.
- Li, J, Hurley, J, DeCicca, P & Buckley, G 2014, 'Physician response to pay-for-performance: evidence from a natural experiment', *Health Econ*, vol. 23, no. 8, pp. 962-78.
- Maynard, A 2012, 'The powers and pitfalls of payment for performance', *Health Economics*, vol. 21, no. 1, pp. 3-12.
- McDonald, R & Roland, M 2009, 'Pay for Performance in Primary Care in England and California: Comparison of Unintended Consequences', *The Annals of Family Medicine*, vol. 7, no. 2, pp. 121-127.
- McLoughlin, I, Bayati-Bojakhi, S, Purushothaman, K & Sohal, A 2014, 'Informational Requirements and Client-centred Disability Care: Issues, Problems and Prospects in Australia', *Social Policy and Society*, vol. 13, no. 04, pp. 609-621. [2014].

- Meacock, R, Kristensen, SR & Sutton, M 2014, 'The cost-effectiveness of using financial incentives to improve provider quality: A framework and application', *Health Economics*, vol. 23, no. 1, pp. 1-13.
- Mehrotra, A, Sorbero, MES & Damberg, CL 2010, 'Using the lessons of behavioral economics to design more effective pay-for-performance programs', *The American journal of managed care*, vol. 16, no. 7, pp. 497-503. Available from: PMC.
- Miller, EA, Doherty, J & Nadash, P 2013, 'Pay for Performance in Five States: Lessons for the Nursing Home Sector', *Public Administration Review*, vol. 73, no. s1, pp. S153-S163.
- Misfeldt, R, Linder, J, Lait, J, Hepp, S, Armitage, G, Jackson, K & Suter, E 2014, 'Incentives for improving human resource outcomes in health care: overview of reviews', *Journal of Health Services Research & Policy*, vol. 19, no. 1, pp. 52-61.
- Mor, V 2005, 'Improving the Quality of Long-Term Care with Better Information', *Milbank Quarterly*, vol. 83, no. 3, pp. 333-364.
- Nahra, TA, Reiter, KL, Hirth, RA, Shermer, JE & Wheeler, JRC 2006, 'Cost-Effectiveness of Hospital Pay-for-Performance Incentives', *Medical Care Research and Review*, vol. 63, no. 1 suppl, pp. 49S-72S.
- Partel, K 2014, *Can we improve the health system with pay-for-performance?*
- Perry, JL, Engbers, TA & Jun, SY 2009, 'Back to the Future? Performance-Related Pay, Empirical Research, and the Perils of Persistence', *Public Administration Review*, vol. 69, no. 1, pp. 39-51.
- Petersen, LA, Woodard, LD, Urech, T, Daw, C & Sookanan, S 2006, 'Does Pay-for-Performance Improve the Quality of Health Care?', *Annals of Internal Medicine*, vol. 145, no. 4, pp. 265-272.
- Pham, HH, Schrag, D, O'Malley, AS, Wu, B & Bach, PB 2007, 'Care Patterns in Medicare and Their Implications for Pay for Performance', *New England Journal of Medicine*, vol. 356, no. 11, pp. 1130-1139.
- Pope, GC 2011, 'Overview of Pay for Performance Models and Issues', in *Pay for Performance in Health Care: Methods and Approaches.*, eds J Cromwell, MG Trisolini, GC Pope, JB Mitchell & LM Greenwald, Research Triangle Institute, Research Triangle Park, NC, pp. 33-76.
- Reichers, AE, Wanous, JP & Austin, JT 1997, 'Understanding and managing cynicism about organizational change', *The Academy of Management Executive*, vol. 11, no. 1, pp. 48-59.
- Roland, M 2012, 'Pay-for-Performance: Not a Magic Bullet', *Annals of Internal Medicine*, vol. 157, no. 12, pp. 912-913.
- Rosenthal, MB & Dudley, R 2007, 'Pay-for-performance: Will the latest payment trend improve care?', *JAMA*, vol. 297, no. 7, pp. 740-744.
- Rosenthal, MB & Frank, RG 2006, 'What Is the Empirical Basis for Paying for Quality in Health Care?', *Medical Care Research and Review*, vol. 63, no. 2, pp. 135-157.
- Rosenthal, MB, Landon, BE, Normand, S-LT, Frank, RG & Epstein, AM 2006, 'Pay for Performance in Commercial HMOs', *New England Journal of Medicine*, vol. 355, no. 18, pp. 1895-1902.
- Scott, A, Sivey, P, Ait Ouakrim, D, Willenberg, L, Naccarella, L, Furler, J & Young, D 2011, 'The effect of financial incentives on the quality of health care provided by primary care physicians', *Cochrane Database Syst Rev*, no. 9, p. Cd008451.
- Trisolini, MG 2011, 'Quality Measures for Pay for Performance', *Pay for Performance in Health Care: Methods and Approaches*, p. 99.
- Van Herck, P, De Smedt, D, Annemans, L, Remmen, R, Rosenthal, MB & Sermeus, W 2010, 'Systematic review: Effects, design choices, and context of pay-for-performance in health care', *BMC Health Services Research*, vol. 10, pp. 247-247. Available from: PMC.
- Weissert, WG & Frederick, LF 2013, 'Pay for Performance: Can It Help Improve Nursing Home Quality?', *Public Administration Review*, vol. 73, no. s1, pp. S140-S151.
- Werner, RM, Kolstad, JT, Stuart, EA & Polsky, D 2011, 'The Effect Of Pay-For-Performance In Hospitals: Lessons For Quality Improvement', *Health Affairs*, vol. 30, no. 4, pp. 690-698.

- Werner, RM, Konetzka, RT & Polsky, D 2013, 'The Effect of Pay-for-Performance in Nursing Homes: Evidence from State Medicaid Programs', *Health Services Research*, vol. 48, no. 4, pp. 1393-1414.
- Werner, RM, Skira, M & Konetzka, RT 2016, 'An Evaluation of Performance Thresholds in Nursing Home Pay-for-Performance', *Health Services Research*, pp. n/a-n/a.

Appendices

Appendix 1 – Summary of states implementing Medicaid-sponsored P4P in nursing homes 2001-2009

States in US	Dates of P4P Program	<u>% of Residents Who</u>						Regulatory Deficiencies	Staffing Ratios
		Had Bladder Catheter Inserted	Were Physically Restrained	Had Moderate to Severe Pain	Had Falls	Developed Pressure Sores	Had Unexplained Weight Loss		
Colorado	(7/2009 to present)		X	X		X		X	
Georgia	(7/2007 to present)		X	X		X		X	X
Iowa	(7/2002 to present)							X	X
Kansas	(7/2005 to present)								X
Minnesota	(10/2006 to 9/2008)	X	X	X	X	X	X	X	X
Ohio	(7/2006 to present)							X	X
Oklahoma	(7/2007 to present)	X	X		X	X	X	X	X
Utah	(7/2003 to present)							X	

Source: adapted from Werner et al. (2013, p. 1396)

Appendix 2 – Considerations for designing and implementing an incentivised model

Our review presents a relatively extensive assessment of existing incentivised care arrangements and P4P models across various areas of care. While the TAC is considering experimenting with some form of incentivised care, it has not yet determined what models to adopt. Therefore, in this extended appendix, we provide a detailed roadmap outlining important considerations and decisions that need to be taken when designing and implementing these initiatives in order to improve the quality of care and/or the effectiveness of care delivery. In particular, we contemplate the following issues:

1. Incentivising care; when is it appropriate?
2. Measuring the quality of care
3. Objectives of incentivised care measures
4. Types of incentives
5. The effectiveness of incentivised care arrangements
6. Risks and unintended consequences of incentivised care
7. Implementing incentivised care models

A1 Incentivising care: when is it appropriate?

Before the TAC proceeds with the design of an incentive pilot, it would be worthwhile to reflect, in the light of the articulated broad, high level, key performance indicators, on whether incentivising care arrangements are the most appropriate remedy to improve the quality of care.

A useful tool to identify whether incentivised measures are a suitable instrument to improve the quality and effectiveness of care delivery was the 7-step decision framework by Custers et al. (2008). The framework contains tailored incentivised care measures to particular contexts. Figure 4 shows how this framework can be adapted for the TAC and can inform the design of any utilised incentive measure. The 7 steps are as follows:

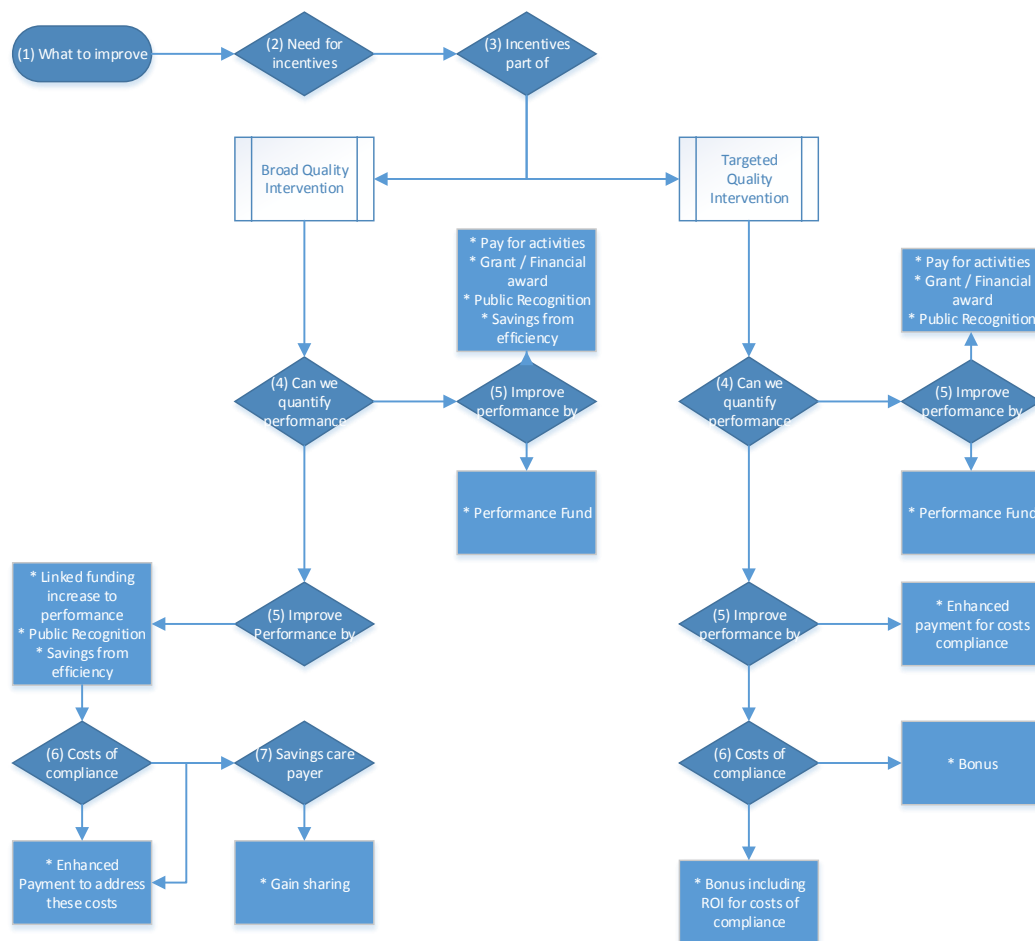


Figure 4. Decision tree for introduction of incentivised care measures.

Source: adapted from Custers et al (2008, p.8)

Step 1 – Firstly, the goals that a care payer, such as the TAC, would like to achieve by implementing incentivised measures need to be clearly defined – i.e., decide **what needs to be improved**.

Particularly at a lower level, there has to be a clear identification of the problem(s) that require(s) improvement. Custers et al. (2008, p. 8) explain that it needs to be clear from the outset which particular objectives a model seeks to achieve, as it cannot motivate CPOs to “do the right thing” when it is unclear what a care payer considers this ‘right thing’ to be.

Step 2 – Secondly, after establishing the specific objectives, **a critical evaluation** should be undertaken to reveal whether incentives are the most suitable instrument to improve *performance for those particular areas of care that require improvement*. There can be several underlying causes for poor performance. First, however, it has to be ascertained whether the current system is actually underperforming, and if so, whether incentives might help to address identified issues. It should also be evaluated whether there are any existing incentives in place. If so, are these properly aligned? Are there any inefficient practices that are currently rewarded? Thus, before opting to implement

incentivised measures, it is important to conduct a root cause analysis at the operational level, to identify whether or not incentives are the best mechanism to improve the quality of care. This includes the following steps:

- a. Define the problem(s)
- b. Assess 'why' the problem(s) occurs, identify all underlying causes of the problem
- c. Identify the most appropriate corrective action, i.e., are incentive measures appropriate?
- d. Ensure corrective actions to prevent recurrence without resulting in new unintended problems.

Step 3 – Thirdly, a care payer has to ***make a decision as to what purpose the incentives will serve***. Is it the objective of the incentivised model to target a single aspect of care, or is it aimed at a broader improvement of care delivery across the board? Thus, a decision needs to be made about whether multiple areas of care should be incentivised, or if the intervention will be targeted at a particular aspect of the care delivery.

There are direct costs and quality trade-offs associated with this decision. While broad schemes have the capacity to improve the quality of care across the board, they have associated costs, as the more dimensions that need to be measured, the more costly the exercise will become (Custers et al. 2008). Models directed towards specific 'targeted measures' are cheaper. The moral hazard of such interventions, however, is that only those behaviours that get measured and incentivised will be undertaken. In the most extreme case, targeting measures narrowly can even lead to the deterioration of care in other areas, as resources are shifted to those areas that get measured and rewarded (Lester, Hannon & Campbell 2011; Mehrotra, Sorbero & Damberg 2010; Rosenthal & Frank 2006).

Another matter that needs to be considered is how the incentivised model is organised with the CPOs. Can incentives be integrated in existing care agreements, or will they require additional agreements? Targeted measures are easier to implement on top of existing arrangements, whereas a broader system of incentivised care might require a more thorough re-negotiation of the existing care contracts (Custers et al. 2008).

Step 4 – The types of incentives that can be utilised to reward or penalise provider behaviour should further be informed by the nature and quality of the performance measures. For instance, ***can the desired performance be quantified?*** Incentives are linked to what is measured, and the objectivity or subjectivity of these measures will have a direct impact on their acceptance by various

stakeholders (Cooke et al. 2010, p. 10 Rosenthal & Frank 2006). Key validity criteria for performance measures are, according to Custers et al. (2008, p. 10):

- a. Validity – does the measure capture what it is supposed to?
- b. Reliability – how replicable are the measurements when taking at different times or in different circumstances?
- c. Responsiveness to changes – will the measure discriminate between good and poor quality and be able to detect small, but worthwhile, improvements?

Step 5 – Next, it must be decided *how incentives will improve performance*. Are incentives aimed at rewarding, penalising, or recognising particular sets of CPO behaviour? Rewarding entails that care providers should be financially better off when demonstrating the required behaviour, whereas recognition focuses on promoting the reputational aspects, and penalties seek to punish undesired behaviour. Custers et al. (2008, p. 10) suggest that recognition tends to be a one-off effect, whereas rewards have the capability to ensure “reinforced change” and lead to more structural improvements in the quality of care. Penalties, on the other hand, can be an effective alternative, yet face the limitation that stakeholder buy-in will be more difficult to achieve (Maynard 2012).

Incentivised care models can be a mechanism to improve the quality of care. Investments might be required by CPOs in order to realise higher levels of quality. These investments will also need to be accounted for in the design of the model. Otherwise, CPOs might be reluctant to invest in the organisational capacities required to achieve ‘high performance’, due to upfront costs and a lack of return of investment (ROI).

These investments can be related to the clinical side of care, e.g., implementation of new practices, or infrastructure and organizational design, such as new information technology systems for record keeping (Rosenthal & Dudley 2007).

If CPO performance can be quantified, then one avenue to deal with investments is to incorporate these into the incentive payments, once a required level of quality has been achieved. If, however, the performance cannot be measured appropriately, an alternative would be that providers are compensated for their investment through other means, such as an industry performance fund (Custers et al. 2008).

One way that these investments can be accommodated, according to Rosenthal and Dudley (2007), is by structuring incentive payments on a performance contingency basis, whereby a substantial proportion of the upfront costs are reimbursed to providers by the care payer. However, to prevent any moral hazards from arising, the remaining compensation for the initial investment is

integrated into the incentive payments. In this case, reimbursement can be reduced (or withheld) in the case of non-performance by the care provider. While thus far it has not been a common practice to integrate investments in this way in incentivised care models, Rosenthal and Dudley (2007) point out that similar arrangements are standard practice in health plan contracts.

Step 6 – A further consideration that needs to be incorporated in the design of the model is: ***are there compliance costs?*** CPO behaviour will also be affected by administrative costs that are incurred as a result of the introduction of incentive measures and associated quality management. “To be effective and to prevent undesired behaviour like gaming or tunnel vision, the reward needs to address these additional costs in its design” (Custers et al. 2008, p. 11).

Step 7 – Besides improving the quality of care, incentive models can also result in cost savings for CPOs and care payers. A final decision relates to ***how any realised savings are distributed amongst the parties***. For instance, if there are anticipated savings, these can be used to inform the size of any bonuses that can be obtained by CPOs (p. 11).

If there is, however, too much of an emphasis on the cost savings side of the model, there is the added risk that, instead of an emphasis on performance improvement, providers will shift their focus to efficiencies and program saving. This unintended behaviour can in turn be detrimental to the quality of care. Therefore, Custers et al. (2008) pointed out that an emphasis on the savings side of an incentive model might be less appropriate if the goal is to improve person centeredness or the quality of care.

A2 Measuring the quality of care

Three of the stages in the decision-making process of the incentive model (namely, Steps 3, 4 & 5) require further elaboration, as they will determine the kind of incentives that to be put in place. Besides considerations such as costs, duration of the model, and target population, another key factor to consider will be what particular quality measures can and will be used. The ability to measure changes in quality as a result of incentivised interventions is key to determining the success of a model. Hence, the decision as to how quality can be measured should inform what incentivised are implemented. Depending on the type of care system that is evaluated, different quality indicators should be utilised (Scott et al. 2011).

Therefore, an important consideration is what particular aspect of the delivered quality of care is focused on when measuring any changes as the result of the introduction of incentives. There are a number different existing care quality frameworks, such Donabedian (1966) and IOM (2001), that can be used to inform the adopted quality measures (Trisolini 2011). IOM, for example, outlines

six areas that can improve the quality of care: (1) safety; (2) effectiveness; (3) person centeredness; (4) timeliness; (5) efficiency; and (6) equity (Scott et al. 2011; Trisolini 2011). Across the different frameworks, however, three common types of measures can be identified (Trisolini 2011, p. 100):

1. **Structure measures:** Inputs in the care process;
2. **Process measures:** Clinical and organisational procedures used to diagnose and deliver care
3. **Outcomes measures:** Outcomes achieved for care users, which can be compared with similar care users.

These three types of measures focus on different aspects of care in order to evaluate the quality (see Table 7 for an overview). Each group has its strengths and limitations. Trisolini (2011) explains that, due to the limitations of the various measures, a combination is often used to assess the quality of care, since they can be complementary to each other.

Table 7. What to measure in order to ascertain quality of care

Element	Emphasis	Advantage	Disadvantage
Structure	Focus on inputs to care delivery	Easiest to measure	When inputs are unavailable, or when providers consider inputs to be non-optimal, difficult to assess quality
Process	Focus on delivered treatment	Less likely to be considered unfair by care providers, as they have greater discretion over the process than the outcomes	Treatments must be statistically associated with outcomes in order to valid. Detailed data collection needed, which is costly. Emphasis on process measures can result in underuse, misuse, or overuse of test & treatments
Outcome	Focus on realised outcomes	These are the ultimate goal of care, such as improving Quality of Life (QOL)	Due to time lags, difficult to measure Other external factors might also affect outcomes, hence can be perceived by providers to be unfair

Source: adapted from Trisolini (2011, pp. 100-102)

Structure measures are often considered by care professional and policy makers to be inferior quality indicators, compared to process and outcome measures, as structure measures are furthest removed from the end goal: i.e., care outcomes. Structure measures only indicate *potential* for quality. Nonetheless, some structure measures can be effective indicators. For example, investments in electronic records can be a useful indicator, as it signals attempts to reduce manual recording

errors. Investments in organizational development can be another insightful structural measure, since certain organisational designs can ensure that a higher quality of care is delivered.

Process measures of care quality, on the other hand, focus on the actual procedures and treatments that CPOs and care professionals provide (Trisolini 2011). Process measures tend to be dichotomous variables that indicate whether a certain process was performed for a population of care users. Trisolini (2011) explains that, for process measures to be used, their effectiveness in lifting the quality of care needs to be demonstrated, as the use of unproven measures to assess provider performance has become increasingly unacceptable. Therefore, one of the benefits of process measures is that, before their implementation, the underlying root causes of quality issues will need to be identified, and this approach subsequently leads to targeted and informed interventions. Process measures should be implemented on the basis of EBP on care user outcomes, e.g., improvements in length or QOL, however, as Maynard (2012) explains, this is often not the case.

Trisolini (2011) observed four other characteristics of process measures that make them more useful than outcome measures. Firstly, processes occur more frequently than outcomes. Hence, from an evaluation perspective, the availability of more data to analyse is regarded favourably. Secondly, processes are directly observable, while outcomes tend to require time. Therefore, by using process measures it can be directly recorded whether or not the desired quality was achieved. Thirdly, from the perspective of CPOs and care professionals, another benefit of process measures is that they focus on the actions that these stakeholders undertake, i.e., over which they have discretion. Outcomes, on the other hand, can be influenced by other external factors beyond the stakeholders' control. Process measures do not necessarily contain this risk. Finally, "Significant improvements in processes are generally larger in relative terms than improvements in outcomes, which makes it easier to measure the former and easier to identify significant changes. This aspect enables P4P programs to base incentive payments on more statistically reliable data" (Trisolini 2011, p. 108).

At the end of the day, however, outcomes are what matter most to care payers and care users, and any quality improvement efforts are aimed at achieving outcomes. While most existing incentive and P4P models have focused on process measures rather than outcomes, outcomes are preferable as they are the objective (Trisolini 2011). Furthermore, while the desired outcomes themselves do not necessarily change over time, innovations in care delivery can affect the processes through which they are realised. Thus, process measures will require constant re-evaluation and updating, which, for long term models, can be an additional cost.

In terms of outcomes measures, a differentiation can be made between:



CPO reported outcomes can be further broken down, into intermediate and final outcomes. Intermediate outcome measures sit halfway between process and final outcome measures (Trisolini 2011). These outcomes are measured throughout an intervention and, hence, have a number of advantages. Firstly, in contrast to final outcome measures, intermediate outcome measures are not time lagged. Yet, they are closer to the final outcome than process measures. Secondly, they provide more data to robustly analyse the effectiveness of the incentivised care model than final outcomes measures do. Thirdly, they provide the opportunity to structure intermediate payments to care providers in case of satisfactory progress. These interim payments can assist in stimulating CPO engagement throughout the duration of an incentive model.

Final outcomes, on the other hand, are the objective for which incentive models were introduced. These outcomes can include a number of objectively verifiable results, such as a reduction in hospital related complications. One form of final outcome that can be useful for measuring the success of incentivised measures in the disability sector is functional outcomes. Functional outcomes include measures of activities of daily living (ADL) for care users, including independence measures (Trisolini 2011). The limitation of these measures in relation to an incentive model, however, is that they “require risk adjustment”, as they can be affected by factors unrelated to the incentives (Trisolini 2011, p. 113). Moreover, as care givers have to evaluate these outcomes and conduct their administration, there might be vulnerabilities in the system in relation to the potential for gaming.

Care user reported outcomes, on the other hand, are obtained directly from care users. These measures of care user well-being can include a number of different patient reported outcomes, including health, quality-of-life (QOL), and care experience (Scott et al. 2011). The benefit of these measures is that they ascribe an important role to care user experience and minimise the ability of CPOs or care professionals to game the system. The risk is, however, that care providers might object to the utilisation of some of these instruments, due to their subjective nature (Trisolini 2011). A possible solution, therefore, is to integrate these measures with other more objectively verifiable measures.

What is noteworthy for the TAC is that, in the category of care user reported outcomes, quality of life measures is rarely used in P4P models. Trisolini (2011) explains that the reason for this

is that there are many external variables aside from the incentive measures that can affect QOL, so relying on these outcomes is considered undesirable from the CPO perspective.

Based on the IOM performance measurements, two further measures can be useful building blocks for an incentive model for the Victorian disability care sector. Firstly, composite measures are a combination of process and intermediate outcomes. The advantage of composite measures is that they “may show greater room for improvement than individual measures and may be more closely related to final outcomes than single measures are” (Trisolini 2011, p. 121). By attributing different weights to various measures, trade-offs can be made by a care payer as to whether the emphasis of composite measures is on the quality or the cost side of a model. Secondly, population-based measures can be used to make incentives dependent on outcomes by population subgroups, rather than on the basis of individual delivery of care.

A3 Objectives of incentivised care measures

Incentivised care models can be found across different areas of care, and depending upon their purpose, will differ in their form and shape. Before outlining the types of incentives that can be found in the existing literature, it is first necessary to assess how the rationale of a model further informs what the actual incentives will look like. Four questions can guide the TAC to clarify the purpose of introducing incentive care arrangements, as follows:

1. Who is targeted to improve the quality and/or efficiency of care?
2. What area of care needs to be improved?
3. What behavioural responses are desired?
4. What are the characteristics of incentive measures (these can be contrasting depending upon the targeted outcomes and behaviours)?

A3.1 Target Incentives

An overarching commonality of the wide range of different incentivised care models internationally is that they seek to change the behaviours of actors involved in care delivery. Incentivised care models can be targeted at different stakeholders in a care system (Petersen et al. 2006), including:

- Care users
- CPOs
- Care Professionals; or
- A combination thereof

Depending on which stakeholders are targeted, the design of an incentivised model will need to be adapted accordingly. On the basis of the briefs from the TAC and HDSG to the research team, it was

evident that the emphasis for this review should be on the care delivery provided by CPOs and their staff, hence the incentivised frameworks explored in this review are deliberately geared towards altering CPO and care professionals conduct and behaviour.

A3.2 What areas of care need to be improved?

Incentives have the potential to address quality gaps in existing care service delivery. Incentive models can be targeted at different quality gaps in the care service delivery and experience. Depending on which particular objectives a model seeks to address, different incentive measures will be more appropriate. The 'care experience' can be broken down into three elements (Trisolini 2011).

- a. Clinical Care
- b. Organizational Systems
- c. Care user Experience

Each of the three areas of care experience can be included in an incentive model through the utilisation of varying types of incentive measures. However, an initial assessment of where problems associated with the current care service delivery exist can help to inform and prioritise the areas of care that will be part of the model.

A3.3 What behavioural reactions do incentives target?

Care payers have the ability to structure incentives in different ways to stimulate different desired behaviour among care providers and professionals. A useful typology of the different types of behavioural motivators that can be activated by the introduction of different types of incentives is shown in Table 8 (Hamblin 2007).

Table 8. A classification of behaviours activated by different incentive measures

Types of incentives	Subtypes	Descriptions
Intrinsic motivation	Altruism	This is the desire to perform as well as possible in order to increase the common good as a result of one's actions. It may also reflect a desire to conform to one's self image as someone who does good.
	Professionalism	Achievement of belonging to the group by behaving according to a shared set of values
Implicit incentives	Competition	Simply, the desire to be seen as better than one's peers (regardless of any other incentive).
	System approval / kudos	The perceived likelihood that performing well will lead to increased respect and other non-financial rewards from other players in the system. This motivator operates at an

		institutional, as well as an individual level. The prestige of a CPO is an important motivating force in recruiting and retaining staff and ensuring a flow of referrals and income.
	Censure - avoidance	The belief that being seen as performing badly will have negative consequences not directly related to payment; a fear of 'name blackening'
Indirect incentives	Personal promotion	Being seen as a successful performer leads to more responsible position, national recognition, or long-term increases in salary and importance
	Benefits in kind	The expectation, rather than the explicit statement, that successfully delivering a programme will contribute to some kind of desired benefit, e.g., good performers experience less management oversight, which removes some irksome bureaucracy.
	Market advantages	Good performance is reported or rewarded in ways designed to make the provider more attractive to consumers.
Direct incentives	Direct payment	X amount of money associated with Y level of performance or Z level of improvement (or some combination thereof)
	Access to investment funds	(might include non-repayment of investment funds for, e.g., electronic medical records)
	Avoidance of catastrophic consequences	Typically, this motivator refers to avoidance of events such as, dismissal of the management, potential closure of a CPO or catastrophic loss of contracts
	Direct non-monetary rewards	Some form of higher performance rating or accreditation, e.g., introduction of 'CPO of the year' award or quality certificates.

Source: adapted from Hamblin (2007, p. 184) for disability care

The desired behavioural effects that a care payer wants to activate should inform what incentive measures are used and how the model is designed, particularly since different measures are not necessarily complementary. Moreover, if the wrong behaviours are activated, then unintended consequences have the capacity to undermine the effectiveness of the model (Hamblin 2007).

A3.4 Contrasting characteristics of care incentives

The literature review suggests that the design of the incentive measures may be informed by a number of contrasting considerations, which are related to: 1) the outcomes that are desired; 2) the behaviours that are targeted; 3) the quality measures that can be utilised; and 4) the incentives that care payers implement, for example, cost consideration might exclude the use of some financial incentives. These considerations are from opposite ends of the same spectrum, and depending on circumstances, it might be preferable to adopt one or the other. Each of these contrasting aspects, as highlighted in Table 9, will be detailed below.

Table 9. Contrasting characteristics of existing incentives model

Quality- focused	Cost-focused
Intrinsic Motivation	Extrinsic Motivation
Rewarding	Penalising behaviour/practices
Financial	Non-financial
Direct	Indirect
Long-term oriented	Short-term focus
Targeted improvement in care quality	Broad improvement in care quality
Aimed at 'high performance' providers	'Lower-end of the spectrum' providers

Source: Compiled from the literature

A3.4.1 Quality vs. cost

In designing the incentive model, it first should be considered whether the core objective of a model is to lift the quality of care, or to reduce costs in the care service delivery. Incentive measures can be designed to achieve either outcome. While not being exclusive, the different objectives are likely to trigger various CPO behaviours (Scott et al. 2011). The literature suggests that incentive models, such as P4Ps, can be used to achieve both objectives, but there is no certainty that both of these objectives will be achieved (Mehrotra, Sorbero & Damberg 2010). For example, the cost effectiveness of P4P models has not been conclusively established (Nahra et al. 2006; Meacock, Kristensen & Sutton 2014; Kilpatrick et al. 2005; Van Herck et al. 2010).

While in many instances quality might be the objective of the introduction of incentive models, it also cannot be ignored that substantial costs are associated with the development of these endeavours. “Agreeing and benchmarking the process and outcome measures together with the transaction costs for individuals and organisations and the cost of bonuses make these innovations expensive” (Maynard 2012, p. 9). There is, however, only limited evidence in relation to the cost-effectiveness of these models (Meacock, Kristensen & Sutton 2014), and existing evaluations would suggest that evidence is “scarce and inconclusive” (Emmert et al. 2011, p. 755).

Meacock, Kristensen and Sutton (2014), for example, emphasize the need to consider costs beyond the incentive measures themselves. They estimate that to focus solely on the costs of the incentive, would be to ignore as much as 40 per cent of the costs of the model and its development. Kristensen and Sutton (2014) stress that resources spent on the development of an incentive model also have an opportunity cost.

Moreover, in terms of cost analysis, limited research has been conducted so far on ROI considerations (Greene & Nash 2009). However, both CPOs and care payers will be concerned as to

whether or not the investments required in order to raise the quality of care to levels that attract incentives will generate an appropriate return. If the ROI case for incentivised measures – and their development – is not there, then it might be worthwhile to consider alternative remedies to raise the quality of care.

A3.4.2 *Intrinsic vs. extrinsic motivation*

Another important consideration is whether an incentive model is built on the notion of triggering intrinsic or extrinsic motivators of CPOs and care professionals. If the goal is to utilise intrinsic measures to motivate CPOs and care professionals to demonstrate the behaviours associated with the objectives of an incentive model, then incentive measures should be constructed around *identity* and *recognition* (Cashin 2014; Emmert et al. 2011). On the other hand, if the model seeks to stimulate behaviours that the other stakeholders would not necessarily undertake out of their own volition, then incentive measures should focus on extrinsic motivators, such as *reward* or *penalties*.

Table 8 already indicates those incentives measures that best correspond with various intrinsic and extrinsic motivators. If the goal of a care payer organisation, such as the TAC, is to stimulate different practices through extrinsic measures (i.e., financial incentives), then that care payer should also be wary of unintended consequences of such measures. Research demonstrates that, particularly in relation to care professionals, financial incentives can detrimentally impact intrinsic motivation, e.g., for physician (McDonald & Roland 2009). In the most extreme form, the introduction of financial stimulants can even lead to demoralization of staff (Custers et al. 2008).

A3.4.3 *Rewards vs. penalties*

A third decision that needs to be made when designing incentive measures is whether or not the desired behaviour will be achieved by rewarding actors for desired behaviour or by penalising undesirable behaviour and practices, i.e., through positive or negative incentives.

In practice, most incentive models are based on ‘rewards’, such as bonuses, rather than penalties. The literature that was assessed suggests that, across the various care industries, there is the perception that penalties can result in undesired outcomes, as they can alienate both CPOs and care professionals.

‘Penalty’ models are also less attractive for providers to participate in voluntarily. Maynard (2012) explains that, from the point of view of care payer, such as the TAC, it will be easier to get provider participation in bonus rather than penalty models, as there is no direct threat to income flows, but the potential for gains.

Moreover, where penalty models have been utilised, their effects have been limited. Custers et al. (2008) explain that, in the few instances where penalties were used, thresholds were set relatively low in order to ensure that the bulk of providers avoided a penalty. This design decision, in turn, meant that there was limited stimulus for providers to actually improve their care service.

A3.4.4 Financial vs. non-financial incentives

Behavioural changes can be triggered by pecuniary incentives (e.g., bonuses) and non-pecuniary incentives (e.g., reputation and peer pressure). It has to be recognised, however, that these two types of incentives are aimed at different motivators to achieve the desired outcomes.

When designing an incentivised care model, careful consideration should be given to what intrinsic or extrinsic factors are tapped into – as there is the risk of unintended consequences. For instance, intrinsic motivators of care givers/professionals can be undermined by placing an overemphasis on achieving certain targets to unlock financial incentives (Mehrotra, Sorbero & Damberg 2010; McDonald & Roland 2009).

Moreover, there can be a duality between the intrinsic motivations of the care professionals and the introduction of an extrinsic motivator, i.e., a financial incentive, for the CPO. The CPO might be more strongly motivated to achieve particular objectives, yet this can directly result in undermining the intrinsic motivation of the care professionals (Eijkenaar 2011; Hamblin 2007). Therefore, a carefully designed incentive model needs to assess how various financial and non-financial incentive are likely to change the behaviour of all stakeholders involved, not solely the CPOs.

Incentivised care models can, thus, use financial or non-financial incentives, or a combination of both (see Table 11). Where both are used, it can be difficult to ascertain whether the “effects are complements or substitutes and how these effects can be separated and quantified” (Maynard 2012, p. 8). The difficulty is not just in measuring the relative impact of various incentive measures; as Maynard (2012) explains, another challenge is mitigating the risk that financial incentives may erode non-pecuniary incentives, such as duty and trust.

A3.4.5 Direct vs. indirect incentives

A further distinction that can be made is between direct and indirect incentives. Custers et al. (2008, p. 5) explain how different types of incentives target outcomes through different mechanisms:

- ***Direct incentive models*** tend to be used to change the behaviour of care providers directly;

- ***Indirect incentive models***, on the other hand, enable care users to make more informed choices about care providers and thereby mediating the behavioural change of care providers.

One of the underlying rationales of the *indirect incentive models* is that by creating better market transparency and more informed care users, this will help choose those providers that deliver the highest quality of care. The development of care users as informed consumers should consequently raise the standards of providers throughout the industry, including those that are providing lower levels of care services.

While more informed ‘consumers’ are expected to shop around for the best deal for care services (Mor 2005), evidence from the literature (Kane et al. 2007; Custers et al. 2008) suggests that, in most cases, care users are not typical consumers. They often lack the ability or willingness to shop around. Other restraining factors, such as geographical location, form further constraints that limit their choices in the ‘market’ for care. An important consideration thus, for the utilisation of indirect incentives, is to what extent care users have the discretion to choose their care providers.

For the purpose of this review, and of the TAC pilot study, it is assumed that the TAC’s objective is to consider the implementation of direct care incentives. The emphasis of this study is on direct forms of incentivised care mechanism.

A3.4.6 Long- vs. short-term oriented incentives

A further consideration is whether or not incentives are targeted at changing behaviours in the short-term, or geared towards the long term, e.g., towards more structural changes.

The length of a model will, in part, be informed by the ability of the care payer to fund an incentive model. Petersen et al. (2006) observed that there is limited evidence to guide the design of an incentive model in terms of the “optimal duration”. The length of existing models can vary considerably. Emmert et al. (2011) study found, for example, that the length of models varied between four months and seven years – with some long-term models still ongoing at the time of publication in 2011.

It should, however, be noted that some clinical outcomes are subject to considerable time lags (Pope 2011), and therefore, the ability to evaluate the success of incentive measures over short time periods might be limited. Maynard (2012) also highlights that, while incentives might change behaviours in the short term, it is, at times, questionable whether these effects will be sustained in

the long term. Thus, when designing a model, the length of the incentive period should be given careful consideration. As Meacock, Kristensen and Sutton (2014, p. 7) explain:

Although we may expect to observe performance improvements when P4P is first introduced, these may reach a ceiling after which little or no further quality improvements are achieved. It may then be relevant to consider the consequences of removing the financial incentives currently in place if they are failing to induce additional benefits. The effect of this removal will depend upon whether quality improvement is a transitory or investment activity. Quality could fall, perhaps even to levels below those observed before the introduction of P4P (Lester et al., 2010). Alternatively, incentivised behaviours may have become routine and therefore continue even after payments are withdrawn. If some of the benefits are sustained beyond the period of cessation of the incentive payments, the cost-effectiveness of the model will be underestimated by a restricted evaluation period.

Moreover, the timing of the payment over shorter or longer periods will have a direct effect on the behaviour of CPOs and care professionals. Time lags are also an issue when evaluating the performance of CPOs. There tends to be a considerable time period between the actual delivery of care – including improvements – and when the incentives are awarded. Mehrotra, Sorbero and Damberg (2010) argued that these time lags are undesirable, as they decouple the behaviour from the reward/penalty. Hence, during the design of an incentive model, a care payer should consider structuring (financial and non-financial) incentives in such a way that they can be awarded, in full or in part, relatively soon after services have been provided. This will, in turn, ensure that, in the perception of CPOs and care professionals, there is a clear direct link between incentives and behaviour. Various timings of incentives that can be considered include:

- End of year
- Concurrent fee
- Intermittent bonus

A review by Petersen et al. (2006) suggests that “end-of-year” incentives may not always be effective, particularly where care providers do not receive regular feedback on their performance, or in cases where awareness of the incentives may reduce over time.

A3.4.7 Targeted vs. broad improvements in care

Are quality initiatives developed to improve care as a whole, or targeted at particular measures? The work by Boyd et al. (2005) demonstrates that when implementing incentivised care arrangements,

careful consideration should be given as to whether “single” issues are incentivised, or whether the emphasis is on a “holistic” improvement of the quality of care.

Hamblin (2007, p. 185) explains that, when incentives are targeted at improving particular aspects of care, the effect might be that other areas get neglected. “[M]easurements that concentrate on only one part of a system of care, may force rigid ‘management to measure’ responses which have negative effects elsewhere in the system, creating situations in which overall quality of care or outcome actually declines despite an apparently useful measure being achieved”.

According to Hamblin (2007), targeted behaviours can be measured in two different ways:

1. **Single measures** – have greater risk of failure or being gamed
2. **Multiple measures** – allow for verification of actual achievements. There are however likely to be greater costs.

“Clinical Practice Guidelines” can, for instance, be utilised to develop single care measures when structuring incentive payments. While it might be appealing to develop performance measures on the basis of these practice guidelines, there are associated risks. Issues of ‘over-care’ or ‘misaligned care’ may arise in relation to care users with multiple care needs. Where care users are receiving treatment for multiple conditions that are financially incentivised, the chance of over- or mis-aligned care occurring may increase. Another issue can be that, if incentives are paid on the basis of single measures, then quality improvements in areas that have a relatively small impact on a care provider’s bottom line may receive a lower priority (Trisolini 2011).

Hence, when designing the incentivised care model careful consideration should be given to whether financial incentives are paid out on the basis of achieving a “single” measure, or whether a “composite” indicator, or “overall” indicators should instead be used.

Cost considerations can further inform the number of measures that are introduced. Introducing a greater number of measures, while more comprehensive, will require additional data verification and audit costs (Trisolini 2011). A possible remedy to overcome this cost burden would be to selectively audit providers and/or measures, however, the overall cost base of the incentive program will increase when a larger number of care areas are incentivised.

Trisolini (2011) explains that, for incentive models with a limited number of measures, the advantage is that costs can be better controlled, the administrative burden will be smaller, and the link between measures and outcomes will be easier to establish. The downside, however, is the

aforementioned chance of cherry picking in care delivery by providers, which a proper design should seek to prevent.

A3.4.8 Lifting performance across providers

A further consideration for a care payer regarding the design of an incentivised model is whether it is targeted at a limited number of CPOs or that is introduced throughout an entire sector. If a model is to be rolled-out across an entire sector, a care payer should be aware about that some segments of the industry will be more, or less receptive to particular incentive measures used within the model.

A question that needs to be addressed is: are incentives geared towards lifting quality across the industry, or aimed at getting high quality providers to innovate, or intended to lift the quality of care services of the lower-end providers? The answer to this question will inform the different behavioural responses that eventuate among high quality providers and among those at the lower end of the spectrum.

There might be room for improvement across various areas of care services in an industry, however, depending on how they are structured, incentives are more effective in lifting the quality of care in some areas than others. Hence, it needs to be identified **where** in the industry there is an opportunity to improve the quality of care (Trisolini 2011).

Moreover, from an effectiveness perspective, care payers should select those incentives which have the greatest likelihood of translating into the greatest improvements in quality of care (Trisolini 2011). In other words, incentive measures should be aligned with the level of care delivery by provider across the sector.

From a financial incentive perspective, the impact of 3 types of incentive measure structures should be considered: (1) threshold targets, (2) improvement-over-time targets, and (3) comparison with other providers. Table 10 outlines the merits of the three incentive types.

It should be considered whether incentives are paid out on **relative** or **actual** performance. There is a trade-off between these approaches. Payment on relative performance is more attractive to low-performing organisations, as there is greater room for improvement, whereas threshold- or ranking -based performance will be more appealing to those at the higher end of the band.

Yet, the issue with actual thresholds for low-performance providers is that, when targets are perceived to be unachievable, the motivation and efforts of these organisations to participate in an incentive model will deteriorate. Similarly, if thresholds are perceived as low by high-performance providers, they may become relaxed, and care may even deteriorate. The work by Werner, Skira and Konetzka (2016) in the context of the US Nursing Home (NH) sector would at least suggest that

decline in care quality may occur in the case of with respect to high-performance providers. In contrast to the theory, however, Werner et al. found that those low-performance providers who were furthest removed from quality thresholds managed to achieve the greatest performance improvements.

Table 10. Three methods for structuring incentives – how to target high/low quality providers

Structuring bonus payments	How it works	Possible structure	Advantages	Disadvantages
Threshold targets	Most commonly used method. Threshold target must be met in order to receive payment, can be set in multiple ways.	(1) By agreement between providers and care payer (2) Set by payer as reasonable level of care (3) By benchmarking providers (4) Derived from other care quality initiatives	Simple to set-up and easy to understand for providers	(1) Difficult to achieve for low-base providers (2) Once threshold is met, little incentive to continue improving care*
Improvement-over-time targets	Comparison of improved performance against baseline performance	(1) Based on percentage or percentage point improvements (2) Reductions in performance gaps	Ensures that low performance providers can achieve improvement goals and are motivated to participate in incentivised care	(1) High performing providers might struggle to improve (2) Care payers may be unwilling to reward low-performance providers at a higher level for comparatively lower quality of care
Comparison with other providers	Quality of care by providers is compared. Incentive payments are only provided to high performers. **	Predetermined group of high quality providers will be eligible for bonus payment (e.g., the upper quartile)	Costs for care payer are certain and minimised, as only a proportion of providers will be rewarded.	(1) Ignores absolute levels of performance. Low levels of care may still yield incentives.

				(2) Providers do not have clear goals to strive for.
--	--	--	--	--

Source: adapted from Trisolini (2011)

* Thresholds can be set-up on the basis of a “stair steps model”, which entails that incentive payments are staggered and paid out as incremental improvements are achieved. This should ensure that low-base providers stay motivated to achieve higher levels of care, even when the end goal threshold might be out of reach.

** An alternative to this approach would be to impose penalties on low-performance providers. This will likely result in better levels of care. However, the challenge of this option would be to get buy-in from providers, since negative incentives tend to demotivate, rather than motivate.

Comparing provider performance will lead to increased competition. While the advantage of this approach is that it will stimulate a number of providers to innovate and improve their care service delivery, the downside of the resulting competition will be a reluctance to collaborate and share best practices and information across the sector (Rosenthal & Dudley 2007, p. 742).

A4 Types of incentives

As discussed earlier, incentives measures can be made contingent upon meeting quality criteria related to different aspects of the care service delivery (structure, process, and outcome measures). Additionally, the types of performance incentive that are at the disposal of care payers to design incentive models can be structured in more than one way.

Incentives can be financial or non-financial. Further, they can be directly awarded to CPOs, or structured in such a way that CPOs benefit indirectly from meeting the stipulated/agreed upon quality criteria. Table 11 provides an overview of the relevant types of incentive measures that were identified by Custers et al. (2008) in the healthcare sector. Contingent upon the objectives of a model, on contextual factors, and on the behaviours that a model seeks to generate, some incentives will be more appropriate than others.

Table 11. Incentive models used in healthcare to change or enable an actor's behaviour

	Financial	Non-Financial
Direct	<ul style="list-style-type: none"> • Bonus • Performance-based withhold • Performance-based fee schedule • Pay for activities • Shared saving contracts • Link regular payment (rate) increase to performance • Quality grants/ performance fund 	<ul style="list-style-type: none"> • Public reporting/ recognition (<i>appeals to intrinsic motivation</i>) • Earned autonomy • Managerial replacement

	<ul style="list-style-type: none"> • Financial award • Auto assignment 	
Indirect	<ul style="list-style-type: none"> • Cost differential for beneficiaries 	<ul style="list-style-type: none"> • Public reporting/ recognition (<i>appeals to care users who base their choice of provider on quality</i>)

Source: adapted from Custers et al. (2008, p. 3)

Financial incentives are thus, just one sub-category of a larger pool of incentive measures that are at the disposal of care payer organisations in the design process. Existing evidence suggests that there is a substantial level of uncertainty in relation to the effectiveness of the various types of incentive measures. Table 13 for instance, outlines the findings of the systemic review by Custers et al. (2008) on how incentive measures can drive performance.

Bonus payments, for instance, are often found to be ineffective. Causes of ineffectiveness include the size of the bonuses, inability of CPOs to meet performance thresholds, high complexity of the behaviour required to obtain them, and other contextual factors. Moreover, bonus payments were frequently found *not* to be “cost-effective” measures for improving the quality of care delivery (Custers et al. 2008).

When utilising an incentive model based on bonuses, one of the critical elements will be the setting of the size/ proportional payment of the bonus. If an incentive is not substantial enough for it to be considered worthwhile by CPOs and/or care professionals, then it is likely to have a limited impact on the quality of care. By contrast, if the bonus is overly generous for the achieved quality of care improvement, it will be a substantial expense for the care purchaser, and may invite debate about whether the additional investment is warranted or worthwhile (Kristensen, Siciliani & Sutton 2016).

Werner et al. (2011) assessed the effect of the size of bonuses upon hospital behaviours and found that larger bonuses tend to generate larger improvements. The study therefore advocated an approach that tailors the size of incentives to the level of quality increase desired. As Rosenthal and Dudley (2007, p. 741) explain, “economic theory suggests that the reward should be commensurate with the incremental cost of the quality improvement required, including the lost revenue that the provider could generate in other activities, such as seeing more” care users.

Financial incentives can further be structured in a number of ways. Table 12 reveals four considerations that help to determine how to shape an incentive measures; depending on circumstances, certain combinations will be more appropriate than others. Firstly, a care payer should determine what type of behaviour it wants to reward. Secondly, the nature of the payment

has to be considered. For example, will incentives payments be a fixed sum per care user? Will providers have a level of discretion in setting the payment level? Will the incentive be in the form of a penalty model? Thirdly, the structure of payments needs to be considered. Payments can be linear or non-linear. The timing of the payments – as already indicated – needs to be considered, as different behaviour can be anticipated, depending on when payments are made available. Lastly, it needs to be decided in the design phase whether or not a financial incentive model will be capped. Budget constraints, for instance, might affect this decision.

Table 12. Characteristics of financial interventions

Type of incentive payment based on provider behaviour	Nature of incentive payment	Structure of payment	Timing of payment	Quantity of payment
<ul style="list-style-type: none"> Specified time 	<ul style="list-style-type: none"> Fixed 	<ul style="list-style-type: none"> Linear: same payment for each additional unit (service per unit) 	<ul style="list-style-type: none"> In advance: prospective payment for anticipated results 	<ul style="list-style-type: none"> Uncapped
<ul style="list-style-type: none"> Providing specified services (fee-for-service) 	<ul style="list-style-type: none"> As charged by provider (complete or partial discretion for provider) 	<ul style="list-style-type: none"> Non-linear, payment is conditional upon particular targets or thresholds 	<ul style="list-style-type: none"> After delivery of the service, based on actual performance 	<ul style="list-style-type: none"> Capped at a particular total. Where the cap is reached, no further incentives are available
<ul style="list-style-type: none"> Care for specific population 	<ul style="list-style-type: none"> As penalty (i.e., deductions by funding body for non-compliance) 			
<ul style="list-style-type: none"> Pre-specified level, or increase, in the quality of care 	<ul style="list-style-type: none"> Flexible arrangements (based on provider characteristics, e.g., training & development status) 			

Source: adapted from Scott et al. (2011, pp. 3-4)

Increasing costs for care payers can threaten the long term viability of incentive models. The most financially sustainable financial incentive that was identified by Custers et al. (2008) was shared

saving contracts, i.e., dividing up anticipated saving to be realised by the introduction of innovative care delivery methods between a payer and the CPOs, as well as linking regular payment increases to performance. In terms of linking regular payment increases to performance it has to be stressed that this was a relatively new method at the time of Custer et al.'s review and further investigation into its use would be warranted.

In terms of non-financial incentives, public reporting was identified as a sustainable quality improvement incentive. Hamblin (2007) outlines that there is an increasing trend in developed countries to measure the quality of care and publish the results. The underlying assumptions of increased transparency about the quality of care are that, firstly, publication of results will increase awareness amongst care consumers, resulting in more sophisticated consumer behaviour, whereby clients may actively engage the best performing care provider; secondly, providers, by contrast, might also take note of quality initiatives and thereby try to market themselves differently, or seek to improve their service delivery.

Whether the publication of quality performance results will prompt CPO action is contingent upon the explicit and implicit incentives that are in place. Hamblin (2007) also stresses that CPOs' responses to incentivised care quality models have thus far, been mixed. For instance, unintended consequences have arisen due to unforeseen design implications of the model.

Table 13 highlights that some incentive measures are more likely to face resistance from CPOs and care professionals than others. One way to overcome such issues is to involve all intended users, including care professionals, in the design of a program from the outset. Studies by Kirschner et al. (2012) and Kirschner et al. (2013) reveal that involving them (all stakeholders) from the outset can assist in reaching consensus about the design of a model.

Relying on one particular incentive measure when designing an incentive model, introduces the risk of gaming and other unintended consequences. Therefore, it is common to utilise a combination of different incentive measures to mitigate some of the undesired behaviours that single measures are more likely to attract (Trisolini 2011, p. 129).

Petersen et al. (2006), for instance, suggest that the utilisation of incentives that are based on a combination of process-of-care measures and outcome measures might be one way to develop a more robust incentive model. Such a model would be able to minimize some of the gaming that can occur under models based on process-of-care measures. It would also mitigate some of the imbalances that el that would affect an outcomes-based mod, such as the influence of external factors beyond the direct control of providers.

Table 13. How effective are most commonly used incentives across various healthcare systems?

Incentive Type	Model	How they work	Effectiveness as found by Custers et al. (2008)
Direct			
1.1 Financial incentive models	Bonus	<p>Bonuses are add-on money; provider can continue underperforming without negative consequences to the 'bottom-line'.</p> <p>Bonuses are not suitable for the long term: - funding with new money in times of rising healthcare costs is not realistic</p>	Some evidence of effectiveness, in majority of instances where this type of incentive is used.
	Performance-based withhold	<p>Non-performance leads to loss of income; as such, it could be a stronger incentive than bonuses because, on average, people place more importance on losses than equivalent gains.</p> <p>Risk of creating resistance from providers, which might lead to dysfunctional outcomes</p> <p>Advantage is that the model is budget-neutral, as it places the provider's funding at risk.</p>	Limited evidence
	Performance-based fee schedule	<p>The primary difference between performance-based fee schedules and bonuses is that the payment is ongoing, rather than one-time or periodic. (p.5)</p> <p>Similar issues to the use of bonuses, described above.</p>	Very limited evidence
	Pay for activities	There is a risk that this approach might freeze existing practices. There must be evidence that the activities that are subject to payment would lead to improved outcomes.	Limited evidence
	Shared savings contract	Funding by anticipated savings as a result of the improvement is difficult because the saving might be too small to generate a meaningful incentive; differences between anticipated and actual spending could vary significantly from year to year; and some improvements may not result in savings.	Limited evidence
	Link regular payment (rate) increase to performance	<p>In principle, this approach does not need additional money or reallocation of existing funds across providers, although providers may demand a higher reimbursement rate for putting their income at risk. The amount that is 'at risk' is relatively small, minimizing the risk for care of unintended consequences. Due to the relatively small amount of funding in question, resistance from providers may be minimal.</p> <p>This model has benefits in comparison to the bonus and withhold models (p. 6) It does, however, require a long-term commitment in order to make funding more performance-oriented.</p>	Limited evidence (was relatively new method at the time)

	Quality grants/ Performance fund	This approach is often used as a tool to encourage innovation, or to promote infrastructure investment and capacity building. It is not simple to administer. Resource intensive for healthcare purchaser (e.g., eligibility assessment and evaluation)	No evidence
	Financial award	Used as a tool to stimulate innovation and superior performance. Because it is a one-time recognition, it doesn't overcome the misalignment caused by institutional arrangements.	No evidence
	Auto assignment	--	Not explored
1.2 non-financial incentive models	Public reporting	CPO and care professional concerns for their public image appear to be a key motivator for improvement. Although professional pride is a motivator, more concrete financial incentives could also be affected by changes in hospital reputation, such as, a hospital's ability to raise funds or to recruit and retain qualified physicians and nurses. Despite some doubts at the hospital management level in the NHS about the validity of the performance rating system, they found it useful as a lever to influence staff behaviour. In general, hospitals don't take many actions after reporting performance, particularly if the performance is adequate in the eyes of a hospital. Poor performing hospitals will more frequently take action.	Evidence of improved performance, however only in relation to reported findings
	Earned autonomy	Stakeholders appreciate a greater degree of freedom, but their effectiveness is highly dependent on: - the nature of that freedom, which might lead to a conflict with a healthcare purchaser's effort for alignment - the actual degree of freedom or reduced oversight, which is dependent on legislative requirements or other stakeholders, such as unions or professional bodies. - the commitment of the care payer to award only those organizations that perform well	No evidence
	Managerial replacement	--	No evidence
Indirect			
2.1 Financial incentive models	Cost differential for beneficiaries	--	Not explored
2.2 Non-financial incentives	Public reporting/ recognition	---	No evidence; care users do not find it particularly useful in the context of hospital

Source: adapted from Custers et al. (2008, pp. 4-5)

The number of quality measures used across incentivised care models varies widely (Trisolini 2011). While employing a greater number of measures might improve the robustness of a model, at the same time, other associated factors need to be considered. A larger number of measures will likely increase the administrative burden on CPOs and care professionals, which might distract from the actual care service delivery (e.g., Cooke et al. 2010). Moreover, the more complex incentives measures become, the more difficult they are to interpret for those delivering the care, and this opacity can result in care professional alienation. Also, where the causality of the effectiveness measures is not evident, the inclusion of a larger number of measures is not advisable.

A5 Effectiveness of incentivised care arrangements

Our review, and the existing systemic reviews, show that the evidence to support the introduction of incentivised care arrangements as a quality enhancing or cost-effective solution to care delivery is rather weak – see Table 13. This finding holds true not only in the care sector, but additionally in other areas where organisations have experimented with incentivised measures, such as P4P models, where the results have also been mixed (Rosenthal & Frank 2006).

Summary:

Our analysis of the literature highlights a number of recurring issues:

- Poor model design and implementation, including the lack of consultation and the use of poor quality measures undermine the viability of models;
- Incentives are not properly designed, and the causality between outcomes and incentives cannot empirically be validated which diminishes the credibility of findings;
- Narrow quality measures have difficulty accounting for all quality improvements, therefore there is a need for broad multidimensional measures;
- Viability is jeopardised by the size of incentives, e.g. too small or too large;
- Studies lack a proper baseline analysis, also few studies utilise control groups as a validity measure;
- The influence of external factors is often not properly accounted for, it needs to be ascertained that incentive measures and not any other factors are the cause of changes in the care delivery;
- There is a lack of data available to evaluate all aspects of a model, e.g. cost-effectiveness is difficult to account for without information on all associated costs.

It should, however, be noted that our ability to assess actual interventions in more detail was constrained by the nature of the literature. Petersen et al. (2006) observed that a problem with the existing literature is that a large number of prior studies are descriptions of the introduction of P4P

initiatives, rather than evaluations of these programs. In their systemic review, for instance, a large number of studies were excluded because a “concurrent comparison group was not analysed or groups were not compared at a baseline on the quality indicator” (Petersen et al., 2006, p. 266). They also suggest that their review may have been affected by negative publication bias. Unsuccessful attempts at incentivised care are less likely to have been published; *i.e., the literature would appear to provide a rosier outlook for the successful implementations of incentivised care model to improve the quality of care than may be warranted.*

A serious problem identified in the literature is that the design choices in incentive models are often made arbitrarily by care payers and that no external experts with the necessary experience in designing these instruments to alter the quality of the care delivery are involved in their development (Eijkenaar 2012).

Another matter to consider when designing an incentive model is the current and possible future influence of the institutional and regulatory regimes on care payers and CPOs’ expectations. Custers et al. (2008) pointed out that, although incentives can have a positive impact on care service delivery, they often do not yield the desired outcomes because of a poor understanding by the designers of the impact of the institutional environment on their effectiveness. Hence, there is a need to clearly identify and understand the institutional setting in which any incentives developed will be required to operate.

In the US Medicaid (hospitals) system, for instance, P4P initiatives are mandated by law and the program is set up under the guise of “value-based” purchasing. The expectation is that these initiatives are cost neutral (Weissert & Frederick 2013). For instance, initially 1% of current payments are withheld, so that these may be distributed in the form of bonuses amongst the best performing organisations.

The US context is considerably different from the Australian context – where, although experiments with incentivised measures have been made, these were not mandated by legislation. Boxall (2009) explains, that currently there are no signs that Australian Federal or State governments have the intention to mandate P4P in the hospital sector. Yet, at the same time “the decision to expand pay-for-performance and other types of financial incentive programs in Australia remains a political call because there is not enough high-quality evidence to guide policymaking” (Partel 2014, p. 10). Thus, awareness of policy developments, and of the emerging outcomes from the introduction of the NDIS across the disability care sector, should also inform the pathway that is taken by the TAC.

Finally, a proper business case needs to be made for the introduction of incentive measures, as Kilpatrick et al. (2005, p. 348) argued:

A business case for a healthcare improvement intervention exists if the entity that invests in the intervention realizes a financial return on its investment in a reasonable time frame, using a reasonable rate of discounting. This may be realized in 'bankable dollars' (profit), a reduction in losses for a given program or population, or avoided costs. In addition, a business case may exist if the investing entity believes that a positive indirect effect on organizational function and sustainability will accrue within a reasonable time frame.

Table 14. Overview of conclusions of systemic reviews and longitudinal studies on incentivised care arrangements

Authors	Area of care	Findings about the ability to raise quality and effectiveness of models	Focus of review	Period of literature reviewed
Rosenthal & Frank (2006)	Healthcare	"...the empirical foundations of pay for performance in healthcare are rather weak." (p. 135)	Healthcare quality	All until late 2003
Weissert & Frederick (2013)	Nursing homes	"Overall, findings from the full literature reviewed and the 22 studies synthesized show that pay for performance can work for some performance targets when it is properly designed and implemented. But the findings also show that beneficial effects are likely to be small, and in some cases – more than 11 percent of target outcomes – results achieved were negative even in studies with some positive outcomes." (p. s145)	Nursing home quality	2002 – 2012
Emmert et al. (2011)	Healthcare	"...evidence on the efficiency of P4P is scarce and inconclusive. P4P efficiency could not be demonstrated." (p. 755)	Economic evaluation	Jan 2000 – April 2010
Custers et al. (2008)	Healthcare	"Overall, the literature suggests that there is no single best approach to create incentives yet and the ability of financial and non-financial incentives to achieve results depends on a number of contextual elements." (p. 1)	Effectiveness of improving CPOs performance	Jan 1996 – May 2006
Scott et al. (2011)	Primary care (physicians)	"The use of financial incentives to reward [Primary Care Physicians] for improving the quality of primary healthcare services is growing. However, there is insufficient evidence to support or not support the use of financial incentives to improve the quality of primary healthcare. Implementation should proceed with caution and incentive models should be more carefully designed before implementation. (emphasis added) " (p. 2)	Quality of healthcare	Jan 2000 – Aug 2009
Petersen et al. (2006)	Healthcare (including physicians)	Link between quality improvements difficult to establish, due to "lack of specificity incentives as well as [due to] the observational design" (p. 269)	Quality of healthcare	Jan 1980 – Nov 2005
Meacock, Kristensen & Sutton (2014)	Healthcare	"Despite much research by economists on this topic, there remains remarkably little evidence on the cost-effectiveness of such schemes." (p. 1)	Cost-effectiveness	Jan 2000 – Sep 2012

Rosenthal & Dudley (2007)	Healthcare	“Despite purchasers’ enthusiasm for pay-for-performance, it has become clear that it should not be a foregone conclusion that these programs will benefit patients or even significantly assist providers who want to improve care.” (p. 740)	Quality of healthcare	
Van Herck et al. (2010)	Healthcare	“The effectiveness of P4P programs implemented to date is highly variable, from negative (rarely) or absent to positive or very positive. ...To maximize the probability of success programs should take into account recommendations such as attention for baseline room for improvement and targeting at least the individual provider level.” (p. 9)	P4P effects, including clinical effectiveness, access and equity, coordination and continuity, person-centeredness, and cost-effectiveness.	Jan 1990 – July 2009
de Bruin, Baan & Struijs (2011)	Healthcare (disease management)	“Most studies showed positive effects of P4P on quality of care delivered. It should, however, be noted that the observed differences between schemes hinders comparability of their effects on healthcare quality and as a consequence drawing conclusions on the effectiveness of P4P to stimulate delivery of chronic care through disease management in general. It is therefore not possible to determine the characteristics of P4P schemes that may contribute to improved healthcare quality.” (p. 12)	Healthcare quality and costs	Jan 2000 – Jan 2010
Kilpatrick et al. (2005)	Healthcare	“Although payers, purchasers, and policy-makers increasingly accept that quality of care is now a necessary rather than a discretionary attribute of healthcare delivery, the lack of data to illuminate the cost and revenue impacts of investments in various [Quality Indicators] inhibits the speeds and predictability of their adoption and diffusion.” (p. 354)	Business case for quality	Jan 1980 – Aug 2004
Misfeldt et al. (2014)	Healthcare	“Overall, evidence of effective strategies for improving outcomes is mixed. While financial incentives play a key role in enhancing outcomes, they need to be considered as only one strategy within an incentives package. There is stronger evidence that improving the work place environment and instituting mechanisms for work-life balance need to be part of an overall strategy to improve outcomes for healthcare practitioners.”	Improving HR performance of care organisations (recruitment, retention, job satisfaction, absenteeism, turnover, intent to leave)	2000 – 2012
Eijkenaar (2012)	Healthcare	“P4P is now widely applied in many healthcare systems and there are no indications that this will change in the near future. However, the current evidence base suggests that designing an effective P4P program is a highly	International overview of characteristics P4P Initiatives	2000 – 2011

		complex undertaking. Given our limited knowledge about “what works” in P4P, it may then not be very surprising that the design of current programs is lacking in several respects (at least in theory) and that purchasers struggle with developing effective programs.” (pp. 18-19)		
Edwards et al. (2014)	Patient-Centred Medical Home Initiatives	“... patient-centered medical home initiatives are highly heterogeneous, and it is likely that some approaches will be more successful than others. Most initiatives now include evaluation plans. Therefore, current initiatives likely will highlight what elements of the patient-centered medical home model are most important to achieving measurable success.	Patient-centred medical home initiatives that include payment reform incentives	Jan 2000 – Dec 2012
Briesacher et al. (2009)	Nursing homes	“We found little empirical evidence that pay-for-performance programs increase the quality of care of residents or the efficiency of that care in nursing homes.” (p. 10)	Quality of care and efficiency care delivery	Jan 1980 – Aug 2007
Werner, Konetzka and Polsky (2013)	Nursing homes	“... P4P in nursing homes did not result in consistent improvements in nursing home quality. Expectations for improvement in nursing home care under P4P should be tempered.” (p. 1393)	Quality of care	2001 – 2009
Gillam, Siriwardena and Steel (2012)	Primary care	“Observed improvements in quality of care for chronic diseases in the framework were modest, and the impact on costs, professional behavior, and patient experience remains uncertain. Further research is needed into how to improve quality across different domains, while minimizing costs and any unintended adverse effects of payment for performance schemes. Healthcare organizations should remain cautious about the benefits of similar schemes.” (p. 461)	Quality of care	Jan 2004 – July 2011

A6 Risks and unintended consequences of incentivised care

While the objective of introducing incentivised measures is to realise behaviours that improve the quality of care delivery, it has to be noted that poorly designed models can lead to a number of undesirable behavioural outcomes.

Undesirable outcomes may have different causes, and can be explained in part by the type of CPOs and care professionals that are involved in a model. The typology of behaviour summarised by Bevan and Hood (cited in Hamblin 2007) – Table 15 – is also of relevance to CPOs and individual care professionals in the compensable disability care sector.

Incentive models should be designed to be sensitive to the type of CPOs or care professionals working in a particular care context. Hence the need for carefully designed quality measures that can objectively verify the outcomes of changes as a result of the incentive measures. Hamblin (2007) explains that issues that are frequently encountered include gaming of the system and dishonesty in data recording – often caused by weaknesses in the measures themselves.

Table 15. Typology of behaviours of care providers in relation to incentivised care arrangements

Typology	Description
Saints	“May not share mainstream goals, public service ethos so high that they voluntarily disclose shortcoming to central authorities, their agenda will be motivated by their intrinsic motivations, and if a common goal conflicts with these then external incentives will not change their behaviour”
Honest triers	“Broadly share mainstream goals, do not voluntarily draw attention to their failures, but do not attempt to spin or fiddle data in their favour, this group will report honestly what they have done, and will endeavour to improve in response to reported poor performers (as opposed to gaming or manipulating figures). However, broadly sharing the mainstream goals implies that improvement depends upon something being incentivized. They are less likely than ‘saints’ to have an internal drive for improvement (or, perhaps, more fairly put, they do not have the space to do other than respond to that which is incentivized)”
Reactive gamers	“Broadly share mainstream goal, but aim to spin or fiddle data if they have a motive or opportunity to do so. Like honest triers, gamers will concentrate their efforts on services which are being measured and incentivized. Unlike them[,] their responses will not always be positive. If it is easier for them to give the appearance of doing well than actually doing well, and if there are no apparent sanctions for giving false impressions (or the chance of such sanctions are perceived as minimal) then they are likely to game”
Rational maniacs	“Do not share mainstream goals, aim to manipulate data to conceal their operations, rational maniacs act entirely in self-interest and respond to incentives in unpredictable ways. In some instances, this group will be unethical, even criminal, in their behaviours”

Source: Adapted from Hamblin (2007, p. 184)

In the design of the incentive model, opportunities for this kind of behaviour should be eliminated as much as possible, as, “if such behaviour is not identified and discouraged, it will be seen to be rewarded, and a race to the bottom ensues where a majority of organizations see it as being in their interests to act as reactive gamers. As a consequence the measurements themselves lose credibility” (Hamblin 2007, p. 183).

Overall four risks of unintended behaviour could be identified across the literature (Custers et al. 2008; Doran et al. 2006; McDonald & Roland 2009; Petersen et al. 2006; Rosenthal & Frank 2006):

1. **Gaming:** some providers will attempt to gain the system in order to comply with the performance measures, therefore, while compliance with the performance measures is achieved, the quality of care is not improved. For example;
 - Misclassification of care user care needs, e.g., a provider may attempt to over-emphasize care needs of care users to achieve better results.
 - Issues with accurate documentation of delivered care, e.g., working on basis of scheduled rather than delivered care
2. **Adverse selection (or cream skimming):** CPOs may try to manipulate their performance by excluding those care users that may have had a negative impact on their rating. Selective sampling is a serious problem since, while it is likely to contribute to a positive outcome in relation to the incentive for CPOs, those care users who could potentially benefit the most would be unfairly excluded.
3. **Multi-tasking problem:** the goals of care payers can be multi-dimensional, yet performance measures may not capture these dimensions in full. As a consequence, only those outcomes that are measured will be realised, since these will receive the primary attention of CPOs.
4. **Demoralisation:** Incentives have the capacity to undermine morale of the care professionals, in particular. Thus, by replacing intrinsic motivation with extrinsic goals, the effectiveness of care delivery might be detrimentally affected.

Rosenthal and Frank (2006) observed that most studies only measure intended outcomes and ignore unintended consequences. Yet issues such as gaming can pose a serious risk to the viability of a model as “incentive produces improvements in documentation rather than a change in the quality of care delivered to care users” (Petersen et al. 2006, p. 268). There are, however, a number of ways to mitigate these risks, as identified by Custers et al. (2008, p. 9):

1. Involving care providers in the development of incentive models
2. Utilisation of independent benchmarks where possible
3. Use local and independent experts to verify the quality of utilised performance measures
4. Evaluations during and post implementation, as well as auditing of data

Undesirable outcomes are not only caused by CPO and care professionals behaviour, but can also be related to the behaviour of care payers and the design of the model (Weissert & Frederick 2013).

Further issues that can distort the intentions of care payers include:

- **‘The slippery slope’:** This refers to care payer leniency in relation to awarding incentive payments.
 - If care payers start to compensate CPOs before the CPOs achieve outcomes, this renders incentives in part ineffective and meaningless;
 - Measures need to represent a worthwhile outcome. There is no purpose to having easy to measure incentives which do not contribute to the desired ‘final outcomes’;
 - Outcomes must be demonstrably different from performance at the baseline level;
 - Better administration should not, in itself, attract incentive payments.
- **‘The distortion effect’:** Incentive models should be designed in such a way that they do not result in a decrease of care in non-incentivised areas.
- **‘Design failures’** include:
 - Too many conditions that need to be met in order to receive incentives – leading to, staff confusion;
 - Incentive targets that can be influenced by external factors, without risk adjustment;
 - Puny incentives that do not trigger behavioural change, or incentives that are too great and become a cost burden for care payers;
 - Incentives that are out of reach for the majority of CPOs, i.e., only accessible for ‘high performance’ providers;
 - Incentive arrangements that only benefit management, and do not flow through to care givers;
 - Non-transparency of care payers to CPOs in relation to baseline measurements and/or performance;
 - ‘Free rider problems’ for care payers, e.g., where multiple care payers are involved with shared care users, the burden of the incentive measures should be shared.
- **‘Implementation failures’**
 - Lack of training;
 - Inadequate monitoring.

Extensive monitoring and evaluation are identified as the most effective remedy against undesirable behaviours, while a careful design process should be able to mitigate and prevent a considerable number of these concerns from the outset (Weissert & Frederick 2013).

Finally, it has to be stressed that **the introduction of incentive measures can have a number of detrimental implications for some of the most vulnerable care users, as well as for the CPOs that deliver their care.** A study by Karve et al. (2008) examines whether P4P increases socio-economic disparities. Where an incentive model is set up such that an FFS payment is reduced to create a bonus pool to reward high performers across the industry these findings (based on the US Medicare system) would suggest that there might be an increased risk that care users with poorer socio-economic backgrounds receive poorer care services as an unintended consequence. As Karve et al. (2008, p. 575) explain:

In [P4P], hospital profitability could suffer double blows: first, from financial penalties due to poor performance; and second, due to potential reductions in patient volume after public reporting of poor quality. Patients may avoid lower ranked hospitals, opting for those that rank higher, potentially creating a drain on already foundering hospitals. To salvage their reputations, poorly performing hospitals may cherry-pick more healthy patients as well as those with higher socioeconomic status.

Furthermore, CPOs with socio-economically disadvantaged care users may also be disadvantaged by incentive models, because incentive measure do not necessarily accurately reflect and account for the additional work needed when dealing with users from disadvantaged backgrounds (Glidewell et al. 2015).

Thus, when the TAC is designing its incentivised model(s), consideration should be given beyond the pilot project, to the implications for care users of varying socio-economic backgrounds, as well as for care users in geographically remote areas.

A7 Implementing incentivised care models

Incentivised care models can be used as a mechanism to improve the quality and/or effectiveness of care delivery. These initiatives are not usually all-in-one solutions. Instead, they need to be combined with other quality improvement initiatives, e.g., consumer reports, and regulatory improvements (Arling et al. 2009). Incentivised measures must be part of broader quality improvement initiatives. The literature reviewed indicates, for instance, that P4P models are most

successful when they form but one element of more extensive initiatives to improve quality (Weissert & Frederick 2013).

The literature also indicates that incentivised models require careful design and implementation, while to some degree, incentivised models are a ‘learning-by-doing-process’ – i.e., they require continuous measurement and evaluation. Care payers, such as the TAC or the NDIA, will need to identify beforehand how they can identify whether a program is working and whether it is yielding any unintended consequences.

The design of an incentive must reflect the values and goals of the healthcare system, be well matched to the performance objectives, and reflect a range of contextual factors that can influence the effectiveness of even well-designed incentives. The remainder of this section will firstly, outline several pre-conditions for the successful implementation of incentivised care measures and secondly, address considerations for the development of quality management measures.

A7.1 Pre-conditions for successful implementation

In order to introduce both financial and non-financial incentives into a healthcare system these measures must, according to Custers et al. (2008):

- 1) Reflect the values and goals that are set out to be achieved;
- 2) Be matched against quality objectives; and
- 3) Take into consideration contextual factors that will influence their effectiveness; i.e., risk adjustment.

This means that models must be tailor-made and take into consideration important contextual factors that will impact on their effectiveness. Arling et al. (2009) explain that financial incentives measures, such as P4P, may not always be the most appropriate remedy to resolve particular care quality concerns. For instance, the needs of a specific population with particular care needs might be better addressed by “raising payment rates for these populations rather than bundling these conditions into the p4p system” (p. 593). Moreover, efficiencies are more easily realised by rate setting measures, e.g., by fixed price or limiting levels of reimbursement, rather than a P4P model.

A7.1.1 *Designing the incentive model*

Systematic steps must be followed to properly design an incentivised model of care. For example, following the steps of the decision-making tree (Figure 4) can guide a care payer in the design process. Eijkenaar (2011) formulated a number of useful guidelines for the adequate design of P4P models, which will also be useful for the design of the TAC’s incentivised care model. The guidelines deal with three key questions:

- *What to incentivise?*
- *Whom to incentivise?*
- *How to incentivise?*

The first question, *what to incentivise* deals with the issue of ‘how performance is defined.’ Eijkenaar (2011) explains that performance can be related to different outcomes, e.g., the quality of care or its cost-effectiveness. The author suggests that it is preferable that a model should seek to identify the most “high-impact” measures. Moreover, a model can seek to target a “broad and deep” number of performance targets or, alternatively, to focus on “narrow and shallow” ones – e.g., on one type of clinical treatment only. The risk with narrowly defined performance objectives is that these can result in the selective distribution of resources. At the same time, it needs to be recognised that, due to resource constraints, a large number of prior initiatives have targeted a relatively small number of performance outcomes. A narrow model, in particular, will need to incorporate additional quality measures to mitigate unintended behavioural responses.

With respect to what is incentivised, another consideration will be how differences in the needs of care user populations are accounted for. Where a model will seek to compare provider performance across a sector, then it will be essential that there is ‘risk adjustment’ for “characteristics that directly or indirectly affect providers’ performance, but cannot be influenced by providers, including demographic characteristics, socioeconomic status, and severity of disease.” (Eijkenaar 2011, p. 119). As the author explain, the need for ‘risk adjustment’ will, in part, be determined by the quality measures that are used. The point is also made that process measures tend to require less adjustment than outcome measures, owing to the level of discretion that CPOs and care professionals have over them.

Cooke et al. (2010) stress the need to build a solid foundation for the introduction of incentivised care measures. This is realised by building the necessary structures to collect data on targeted quality issues. Additionally, it will be crucial to bring along all the stakeholders including staff, care users (and their relatives) in the quality initiatives. This can also assist in the ‘education process’, which should commence well before the implementation of the measures.

In terms of the second question, *whom to incentivise*, Eijkenaar’s (2011) study discusses whether incentives should be geared towards individuals or groups. This question is more applicable to the situation of primary physicians than it is to the circumstances of the disability care sector. Nonetheless, the existing evidence indicates that group incentive measures are preferable. In the case of the disability care industry, the question is whether incentives are geared towards the CPOs

(e.g., towards the bottom-line) or to care professionals (e.g., towards training and development opportunities).

The third question of Eijkenaar (2011) is *‘how the scheme is structured’*. Firstly, there is the issue of whether models should be based on rewards, on penalties, or on a combination thereof. As mentioned, penalties are, generally, perceived negatively by CPOs. What is more, the ability to introduce penalties depends on the bargaining power of care payers; if CPOs are dependent on care payers for their work then the capacity to impose such a measure is greater – although such measures are still likely to be viewed unfavourably.

Penalties have a higher probability of inspiring opportunistic or undesirable behaviour from CPOs (e.g., ‘gaming’). On the other hand, a model based purely on rewards can be financially unsustainable for care payers. An alternative model is to base a model on a combination of penalties and rewards – as outlined in Table 16. One of the combinations, depicted as model design number 4 in the table, can provide a care payer with useful information about providers’ own assessment of their capability to lift performance. By allowing CPOs to select which model they regard as preferable, a care payer can infer a CPO’s assessment of its own capability to raise performance. Weissert and Frederick (2013) reveal that, despite the fact that penalties are controversial, they can, in some instances, be an effective part of a ‘carrot and stick’ approach.

At the same time, it is important to realise for care payers that frequently “.... performance improvement will, at least in the short term, often be accompanied by cost increases because a substantial share of quality problems is related to under treatment” (Eijkenaar 2011, p. 123).

Table 16. Characteristics of models adopting penalties and/or rewards

Model	Income increase or decrease possible?	Incentive strength	Likelihood of negative reaction
1. Penalties for poor performance only	Decrease only	High	High
2. Rewards for good performance only	Increase only	Moderate	Low
3. Penalties for poor performance, (larger) rewards for good performance	Both	High	Moderately high
4. Choice between 2 and 3, provided that the potential increase in income is larger in 3 than in 2	Depends on choice	Moderately High	Moderately low

Source: adapted from Eijkenaar (2011, p. 123)

The second aspect of the P4P models relates to the size of incentives. As discussed, the size of incentives can affect the extent of behavioural change by CPOs, as well as influencing the sustainability of a model. The size of incentives must be substantial enough to motivate CPOs and care professionals, but not so large that they become an undue burden for care payers (Weissert & Frederick 2013). Eijkenaar (2011, p. 124) explains that:

All else equal, the higher the revenue potential for providers, the larger their response and the impact on performance, up to a certain point.... In general, the relationship between incentive size and performance will be positive with diminishing marginal increases in performance above a certain payment level. This is because the marginal utility of income generally diminishes and because every unit of performance improvement will be harder to attain than the previous unit.

Another aspect related to the size of the incentives that must be considered is whether these are going to result in undesired behaviours by CPOs and/or care professionals. The greater the incentives, the more rewarding these undesired behaviours can be. The greater the incentives, the greater the need to have effective monitoring in place, although monitoring can be costly. Thus, a trade-off must be made between the size of the incentive, the likely behavioural impact, and the monitoring and compliance costs.

The third consideration about the structure of a model is related to the way in which performance is measured. Performance can be measured on the basis of 'absolute' or 'relative' performance. Absolute performance targets are preferable, as they are more transparent. These targets are, however, problematic, in that they tend to favour the 'high performers' in an industry, while the goals should be to promote improvement across the industry. Relative performance, on the other hand, can be utilised to create an environment of continuous competition and improvement initiatives. However, it carries the risk that willingness to share knowledge across the industry may disappear in response to an increasingly competitive mindset amongst CPOs.

An alternative is the use of multiple tiered – stepwise – absolute targets, “possibly combined with additional fees for each appropriately managed” care user. This approach is “preferable over a uniform, single threshold system and models using relative targets” (Eijkenaar 2011, p. 126). Mehrotra, Sorbero and Damberg (2010) also argue that, from a behavioural economics perspective, tiered incentives are preferable because when absolute targets are too difficult for CPOs to accomplish they will put minimal effort into achieving them.

In terms of the frequency of payments, it was mentioned that there should be a clear relationship between the behaviours and payments, to reinforce the link. Aside from the behavioural link, another reason that staggered payment are preferable to lump sum payments is that “people tend to discount future gains by a certain rate, which increases with the length of delay” (Eijkenaar 2011, p. 126). The challenge, however, is that time lags and administration processes will affect the ability to directly link the payments to behaviours. Nonetheless, the objective of a model should be to ensure that the timing of the payments is structured in such a way as to ensure that CPOs and care professionals are appropriately motivated (Weissert & Frederick 2013).

Furthermore, financial incentives need be predictable and achievable for CPOs. In order to persuade CPOs to invest in the quality of care, the incentive must be obtainable and predictable; CPOs will otherwise be reluctant to make the necessary investment. However, as Arling, Job and Cooke (2009) point out, despite being obtainable, the achievement of performance targets need to remain ‘at risk’, otherwise the desired behavioural changes are less like to be demonstrated.

Lastly, certainty in relation to the performance targets should remain stable over time (3-5 years), so that outcomes for providers are predictable and allow for investment in quality (Arling et al. 2009). A degree of financial certainty for the duration of incentivised care programs will also have a number of implications for the realisation of desired behavioural changes. Firstly, CPO expectations of the continuation of a model will directly influence the behaviours that are demonstrated. Where expectations are that models will not be sustained, their willingness to invest and reallocate the necessary resources will decrease.

Aside from duration, the stability of the performance targets is also crucial. Although performance measures should be constantly evaluated, and should be adjusted where appropriate, if targets are changed too frequently, this can also be demoralizing for CPOs and care professionals. Moreover, due to time-lags, this may also result in CPOs failing to realise improvements that would be achievable over time. Hence, the targets should in the first place be informed by existing evidence, and expectations about the period in which they can be realised should be realistic and informed by evidence.

CPO expectations about the sustainability of an incentivised model of care will affect their willingness to allocate the necessary resources, to invest in quality improvement, and will impact the degree of behavioural change.

The entire list of Eijkenaar’s (2011) recommendations for the design of P4P models can be found in Table 17.

Table 17. Recommendations for the design of a P4P model

What to incentivize

- Performance is ideally defined broadly, provided that the set of measures remains comprehensible;
- Concerns that P4P encourages “risk selection” and “teaching to the test” should not be dismissed;
- Outcome and resource use measures should be included provided that risk adjustment is sophisticated and sample size is sufficient;
- Other strategies to minimise incentives for risk selection may still be necessary;
- Measure sets should at least incorporate “high-impact” measures; the more indeterminate aspects of care such as care user satisfaction and continuity of care are ideally also included or monitored;
- P4P incentives should be aligned with professional norms and values; it is vital that providers are actively involved in program design and in the selection of performance measures;
- Monitoring, structured feedback, and sophisticated information technology will remain important in preventing undesired provider behaviour.

Whom to incentivise

- On balance, group incentives are preferred over individual incentives, mainly because performance profiles are then more likely to be reliable;
- Individual or small-group incentives as well as using measures with small sample size will become increasingly feasible as methods for constructing composite scores evolve;
- Caution should be upheld in applying hybrid models;
- Participation is ideally voluntary provided that broad participation among eligible providers can be realized.

How to incentivise

- Whether rewards or penalties should be used is context-dependent;
- Offering providers a choice among models also including penalties may be considered;
- Increasing the size of the incentive increases their strength up to a certain point. Yet, relatively low-powered payments are preferred, provided that providers’ costs of improving performance are covered;
- Differentiated absolute targets across groups and/or a tiered series of absolute targets, possibly combined with additional “piece rates” for each appropriately managed care user, are preferred over single targets and models using relative targets;
- The time lag between care delivery and payment should be minimised;
- P4P should be a permanent component of compensation and is ideally decoupled from base payments. Measures should be re-evaluated periodically and be replaced or updated as necessary;

Source: adapted from Eijkenaar (2011, p. 127)

A7.1.2 Stakeholder management

Stakeholder management by care payers can make a substantial impact on the successful design and implementation of an incentivised model of care. It is important to clearly communicate, consult and

involve the various stakeholders affected by such initiatives, including CPOs, care professionals, care users and their families, and relevant advocacy groups. Arling et al. (2009), for example advise to involve all key stakeholders throughout the various stages in the design and implementation of these undertakings, in the context of nursing home care in the US.

Proper stakeholder management can result in two important outcomes. On the one hand, it can help to improve the design of the model. It can assist in identifying the potential for unintended consequences at the outset. Further, the financial concerns of different stakeholders can be addressed (e.g., a model may be developed for the distribution of any arising cost-savings or investment needs). Moreover, it is helpful if the incentive measures can be de-politicized, as a level of consensus is required for the success of the model.

Moreover, stakeholder management can smooth the implementation process by building consensus around measures that are implemented, and reducing possible resistance. As Greene and Nash (2009) explain, when some stakeholders are not fully informed about the anticipated implications of an incentive model, they are particularly likely to approach it with scepticism, and may even have the capacity to derail the measures. Therefore, careful consultation and education of all parties involved will be required. Stakeholder management hence forms an integral part of the change management process.

A7.1.3 Change management processes

In addition to the design of the incentives and quality measures, another critical aspect of the successful implementation of incentivised models of care relates to the change management process that care payers and CPOs adapt (Cooke et al. 2010; Weissert & Frederick 2013). The following issues, outlined below, must be properly accounted for.

Change management plan:

- A plan requires a realistic assessment of the required resources and realistic deadlines;
- Care payers and CPOs need to be flexible and adjust where necessary;
- A plan must be adapted to the local context, needs, and circumstances;
- Be aware of the effect of uncertainties associated with change for staff & care users;
- Aim to achieve small success early on, to realise momentum.

Communication is essential. It was found to be one of the key aspects of successful P4P programs.

Motivation:

- Managers and leaders stressed the importance of the incentivised program for implementing innovative care practices. The additional funding associated with models enables CPOs to realise

the quality outcomes, moreover the introduction of a model is also an important motivator and stimulant to take risks;

- Open communication is important;
- There is a need to facilitate communication between various carers (and also the cleaning staff) who may not necessarily be targeted by incentive measures;
- A culture of joint problem-solving, patience and flexibility amongst care professionals will need to be fostered.

Staff buy-in:

- Critical for success of projects;
- Targets must be regarded as realistic and feasible by staff, in order to generate the required behavioural change;
- Evidence of likely success, and underlying motivation, is essential for buy-in;
- “Staff became more engaged in the project when they could see improvements in resident quality of life and when they observed that better clinical outcomes such as fewer falls, less pain, and improved continence resulted in less paperwork and a reduced workload” (Cooke et al. 2010, p. 561).

Supportive Management

- Necessary for leaders from care payers and CPOs to build trust amongst different stakeholder groups;
- At CPOs, encourage all levels of the organisation to take ownership of the quality improvement efforts;
- Involve and empower care professionals in the quality improvement process.

Sustainability

- What happens when care goals are achieved? Should the incentive payment be phased out?
- Need to have clear post-implementation (business) plans in place;
- Is the incentive a one-off initiative, or a stepping stone to the rolling out of further quality improvement initiatives?

A7.1.4 Investments in the quality of care

Investments may be necessary to raise the quality and/or effectiveness of care. These investments must be incorporated into the design of the model. Both care payers and CPOs will likely want to realise a ROI, with the result that the latter will be particularly reluctant to invest in the absence of proper compensation for any associated costs.

Investments may be required for the clinical – care delivery – side of care (e.g., training care professionals) or be necessary to support the broader organisational infrastructure (e.g., administration, IT systems, etc.).

P4P models, for instance, should encourage investment from CPOs, but constraints need to be understood by care payer (Arling, Job & Cooke 2009). As the authors explain, care payers will have to assist CPOs with the development of methods and tools to lift their performance, since “simply offering a financial incentive for better performance will probably not be sufficient for many providers; they may have the will to perform better but not know the way.” (p. 591-2). Therefore, care payers should play an active role in the development of innovative practices of care.

They further explain that an associated risk of P4P models is that they can result in profit-maximizing behaviour by CPOs, rather than investment in the quality of care. Incentive measures must be constructed in such a way as to counter this behaviour.

A limitation of much of the existing literature is that it is impossible to verify what the total costs of the incentives, their design, and implementation have been. Meacock, Kristensen and Sutton (2014) explain that frequently, a large proportion of the costs is unaccounted for. They suggest that, in order to ascertain the costs of an incentivised care model, it is necessary to include more inputs than just the incentive payments. They estimate that these only roughly cover 40% of total costs.

A number of cost categories need to be considered (Meacock, Kristensen & Sutton 2014) – as indicated below (although this not an exhaustive list):

- Set-up and development costs – e.g., staff time, infrastructure investment, etc.;
- Running costs— e.g., administration;
- Incentive payments themselves;
- Costs to CPOs participating in the various model—e.g., staff & management time;
- Cost savings—e.g., reduced complications, or hospital readmissions. The cost savings may be to the benefit of CPOs or the care payer, so it is important to consider how these are distributed.

Korda and Eldridge (2011) explain, for example, that if care payers want to introduce new payment arrangements, then adequate IT systems are essential for facilitating the introduction of incentive measures. Similarly, (Greene & Nash 2009) argue that, within P4P models, IT systems are an

essential tool for properly reporting on the delivery of care and measuring progress in relation to meeting quality thresholds. They take the example of electronic records and explain that, at the very least, these will result in better report keeping. Yet, the introduction of electronic records can be costly and it needs to be decided how such costs will be accounted for in the design of an incentivised care model.

Compensation for investment can be dealt with in a number of ways. The literature reveals, for instance, that it can be incorporated in the incentive payments by upfront reimbursement or paid from a dedicated industry performance fund. It is, however, advisable that performance improvements are demonstrated before CPOs are fully compensated for investments.

A7.1.5 Legal implications

Legal implications for the introduction of an incentivised care programs should also be considered (Greene & Nash 2009). How does it affect the position of the care payer and the CPOs vis-à-vis the care user? Who will be responsible if the care user is aggrieved by any changes to the care delivery as a result of the incentivised programs/pilot?

A7.1.6 Monitoring, feedback and evaluation

To generate changes in the behaviour of CPOs and care professionals the design and implementation of a model must be complemented by an education process (Greene & Nash 2009). Care payers need to pay careful attention to how they will be able to provide the tools and information required to guide the other parties. It is also important that incentivised models of care, including P4P models, are designed and implemented in a transparent way and are subject to continuous scrutiny and evaluation (Arling et al. 2009). Extensive monitoring and evaluation are identified as the most effective remedy against undesired behaviours (Weissert & Frederick 2013).

Support and feedback, including the promotion of best practices and technical assistance, will be instrumental to align the interest of the parties. These elements will assist care deliverers to acquire the necessary skills and practices. Care payers, therefore, should set up structures that enable them monitor performance and engage in “possible corrective action through dialogue between the parties” (Custers et al. 2008, p. 11).

It has already been flagged, for instance, that “end-of-year” incentives may not always be effective. This is particularly true when care providers do not receive regular feedback on their performance, or when awareness of the incentives has fallen over time (Petersen et al. 2006). Arling et al. (2009) suggest that proper monitoring, evaluation and feedback mechanism are characterised by the following:

- A model should be subject to scrutiny and be transparent:
 - Information about the mechanism of a model should be available and understandable to all relevant stakeholders;
 - De-identified information should be available for public scrutiny, where appropriate.
- Models must be continuously evaluated
 - Models will evolve over time, and it is more than likely that correction will need to be made. By conducting ongoing evaluations, care payers and CPOs should be able to assess whether measures are resulting in the envisioned outcomes.
- CPOs must be provided with an understanding of the performance measures, and supported to understand outcomes and opportunities for improvement.
- CPO performance information should not remain privileged. Information about CPO performance should be available to care users; and should also be effectively communicated to care users so that they can make better-informed decisions about their care needs.

A7.1.7 Bureaucracy

A downside of increased monitoring, feedback, and evaluation of care can be that it may result in increased bureaucracy. One possible pitfall of an incentive model is that it becomes a bureaucratic hurdle, rather than an enabler of care delivery. Any required (additional) documentation that is necessary to monitor the implementation of P4P should not detract from the actual care delivery itself (Cromwell & Smith 2011; Cashin 2014).

A7.2 Development of quality management structures

“Ongoing measuring of incentive programs is critical to determine the effectiveness of financial incentives and their possible unintended effects on quality of care” (Petersen et al. 2006, p. 265).

A considerable threat to the validity of incentivised care models is that the quality instruments developed and implemented fail to capture either the realised results or the unintended outcomes (Cromwell & Smith 2011). The challenges of designing instruments relate to:

- Data collection;
- Imperfect measures; and
- Inconsistent collection methods.

To prevent these issues arising, it will be essential to follow a proper design process to develop the quality measures. Where there are existing quality measures, it must be evaluated whether these

are accurately capturing the information required to assess the implications of the introduction of incentives. If not, it will be necessary to develop quality measures that do capture the required information. Quality measures must be designed to account for external factors that are beyond the control of CPOs. The proper design of quality measures is essential to ensure the successful implementation of incentivised care, and to secure the ability to evaluate the model effectively (Custers et al. 2008).

A further challenge facing care payers is whether to use standardized measures, or whether tailor-made measures will be developed. Cooke et al. (2010) explain that the former has advantages in relation to benchmarking, while the latter can enable management to capture improvements in areas of the client experience such as health, functioning, or quality of life, which otherwise might be missed.

Cromwell and Smith (2011) indicate that measuring changes in care quality that result from pay-for-performance measures is more difficult than assessing the implications for costs. They explain that “quality is a latent multidimensional construct that must be quantified to be evaluated” (Cromwell and Smith 2011, pp. 274-5). However, because of the multi-dimensional character of quality, multiple measures need to be constructed in order to adequately evaluate it. When unrefined quality measures are utilised, this is likely to result in misleading conclusions.

One of the challenges of assessing the impact of P4P models is that simple pre- and post-implementation measurements of performance can be distorted by other temporal changes and exogenous factors (Cromwell & Smith 2011). Hence, it is preferable to use control groups, where possible (Custers et al. 2008). Alternatively, a trend on the basis of a proper baseline study can be utilised.

In terms of the measurements themselves, incentivised care models regularly seek to capture improvements in the quality of care. However, since quality is a multi-dimensional concept, it may be that the performance measures adopted do not accurately capture all improvements in quality (Meacock et al. 2014). Consequently, it is preferable to have a broad range of quality measures in place.

“Performance measures should be comprehensive, valid and reliable, risk-adjusted where appropriate, and communicated clearly to providers and consumers.”

(Arling et al. 2009, p. 587)

Arling et al. (2009) formulated a number of useful guidelines for the development of quality /performance measures that can guide the development of a P4P model. These include:

Performance measures need to be credible and have a solid research basis

- Quality measures should be based on fact and evidence. They should not be contentious.
- Measures should be constructed on an EBP basis, and preferably, be statistically validated.
- Acceptance of measures can be increased through guidance from experts.
- Risk adjustment may be needed in order to compare outcomes between providers.
- Data collection procedures should be developed to prevent sampling- or measurement biases.

Broad and multiple performance measures are preferable to targeted and selective performance measures

- Despite the associated costs, broad measures that are included in primary data collection (e.g., surveys or QOL measures) are more likely to result in the desired quality of care improvements.

P4P should seek to motivate both high- and low-quality providers

- Incentive payments need to be structured in such a way that they motivate both relative performance improvement (e.g., from a baseline) as well as rewarding continuous improvements over time.

The risks of unintended consequences should be mitigated as much as possible

- It is advisable that areas of care not directly related to incentivised measures are monitored as well. This can aid a care payer in monitoring whether there is a deterioration of care in non-incentivised areas.
- Measures to minimize the ability of CPOs to 'game' the system also need to be put in place. Procedures must be implemented to audit procedures to monitor CPO behaviour and administration.

P4P should be part of broader quality measures

P4P is not the only, nor necessarily the best, way to improve quality and change behaviours.

Therefore, it is advisable that P4P models should be part of a broader quality improvement drive.

Appendix 3 – Recommendations for designing, implementing and evaluating an incentivised model

We have broken down our recommendations for the design, implementation and evaluation of an incentivised model into the five key phases, as illustrated in Figure 5, below. We will outline our recommendations for each of the phases on the basis of the literature reviewed. The literature has, however, clearly indicated that each model needs to be tailored to contextual circumstances. Therefore, while our recommendations provide direction on broad strategic and operational considerations, they should not be regarded as a template that can be readily implemented without adaptation.

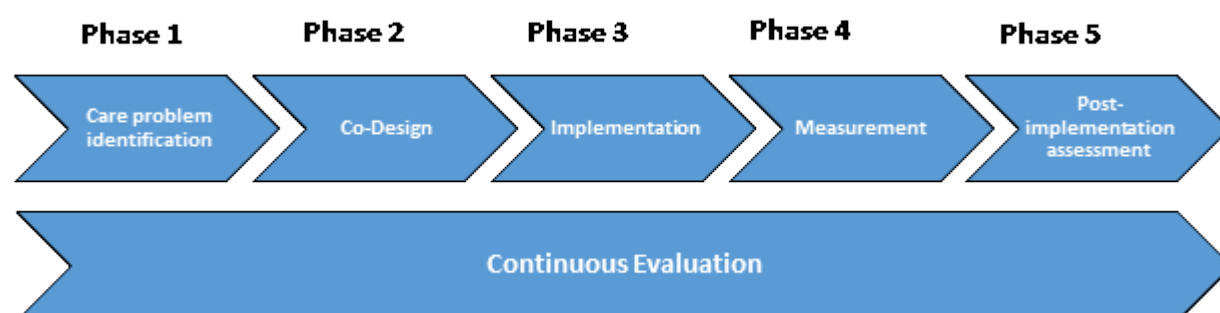


Figure 5. Steps for the development and implementation of an incentivised model of care

A1 Care problem identification

We suggest that a next critical step in the process of developing an incentivised care model must be a more in-depth consideration of a number of fundamental questions. The answers to these questions will, in turn, inform the strategies that will be used by the TAC and HDSG, and will also directly help to identify the goals and objectives that the incentive model seeks to achieve. These questions are:

- What area of care needs to be improved?
- Who is targeted to improve the quality and/or efficiency of care?
- What behavioural responses are desired?
- What to incentivise?
- How to incentivise?
- What will be the end goal of the model (or pilot)?
- What happens when goals are achieved?

When it is clearer what the specific goals and objectives of the model are, then appropriate quality measures can be identified to enable the independent evaluation of the effectiveness of any incentive measures.

A2 Co-design

On the basis of the existing evidence, we would strongly advocate a broad stakeholder engagement and collaborative approach for the design of an incentivised model of care. Bringing along a broad coalition of stakeholders in the design process should reduce resistance and improve the quality and acceptance of the quality- and incentive measures used.

Quality measures must be evidence -based, and some may need to be tailored-made for the model. Yet, there is a fine balance to be found between measuring too little, and measuring too much. Moreover, the (perceived) financial sustainability of a model will also affect CPO buy-in and willingness to invest in quality.

A2.1 Collaborative stakeholder co-development

The design of quality and incentive measures forms a second critical step before an incentivised care pilot can be launched. The literature strongly suggests broadening the pool of stakeholders involved in the design of an incentive model and a pilot project. If the purpose of the pilot is to function as a first steps towards incentivising care across the industry at a later stage than it would be warranted that the TAC and HDSG should include other CPOs, as well as a broader range of stakeholders.

Furthermore, failing to consult care users and/or their advocacy groups might jeopardize their buy-in. Communication will be key throughout the change management process. For instance, it is a risk that care users may perceive an incentive model as a cost-reduction initiative rather than a quality improvement initiative.

Some form of worker representation should also be considered. Introducing measures unilaterally can result in employee resistance. Thus, management will need to prioritise communication with, and addressing any concerns from, care professionals will as a part of the change management strategy.

Bringing along all stakeholders in the development of quality measures and incentives (including care professionals, care users, their families, and/or advocacy groups) can have very positive effects. It will, for example, further help to strengthen the quality of the measures that are used, aid in the acceptance of the model, and can start the 'education process', well before implementation. We, therefore, recommend the following to the TAC and HDSG:

- Consider broadening the number of stakeholders involved in the development of the model;
- Adopt a co-design approach for the quality measures that is founded upon empirically; reliable data, to ensure that the effectiveness of the pilot can be objectively verified;
- Design all the quality and incentive measures before a pilot project is initiated;
- Identify which aspects of the incentivised care model can be developed in-house and where expertise is not currently available within the organisation; and
- Determine whether external experts will be required to facilitate this process.

A2.2 Quality and incentive measures

Appropriate quality measures can be based upon different types of measures, all of which have strengths and limitations.

The literature suggests the use of a broad range of quality measures – this can include measures of structure, process and outcome, as well as composite measures that are related empirically to the desired outcomes and behaviours. Moreover, it is strongly advisable that quality in areas of care that are not incentivised also be monitored, to minimize adverse selection behaviour by CPOs and care professionals. For example, this monitoring will prevent the redistribution of resources to those areas of care that are incentivised at the expense of other areas. Since measures can be objective (e.g., the administration of particular treatments) or subjective (e.g., care user experiences), it is necessary to negotiate with all relevant stakeholders beforehand how the results from different measurements will affect the ways in which incentives are awarded.

Quality measures will need to be carefully constructed in order to create a solid foundation for a model. This careful construction should include an assessment of it is possible to collect data on the quality issues to be targeted. Picking the right quality standards will be critical, and care payers may rely on standardized quality measures from existing quality frameworks in making these decisions, or seek to develop tailored measures. The former approach has advantages in relation to benchmarking, whereas the latter enables management to capture improvements that are adjusted to a particular situation.

In addition, care payers should work to prevent the creation of a model that becomes an administrative burden for CPOs and professionals. Therefore, it is advisable that the TAC and HDSG evaluate which existing quality measures are readily available at prospective CPOs, for a pilot project, as well as throughout the industry. This information about the availability of data might, in turn, inform what types of incentive measures can be deployed. The literature highlights that it may be advantageous to rely as much as possible on existing measures, however there needs to be clear evidence that the measures that are available do, in fact, capture quality improvements in the areas

of care that are incentivised. Otherwise, it is strongly advisable that new measures should be designed.

It is further advisable that measures are only included if appropriate risk adjustment can be undertaken. The number of available measures and the sample size should be other determinants for the in- or exclusion of specific measures. It is preferable that quality measures are able to motivate CPOs from the entire spectrum to improve their care service delivery, i.e., that measures are applicable to participating CPOs, from 'high' to 'low' performers.

Regarding incentive measures themselves, we suggest that it must first be established what types of behavioural changes the TAC and HDSG wants to pursue among the CPOs and care professionals. This decision will, in turn, inform what types of incentives are most appropriate.

A2.3 Compensation for investment and financial sustainability

In the case of both financial and non-financial incentives, any investment by CPOs has to be accounted for. Appropriate compensation measures need to be put in place. Investments in relation to care delivery (e.g., staff training) or infrastructure (e.g., administration and IT systems) are likely to be required in order to achieve higher levels of quality. CPOs will be focused on an ROI, and hence a care payer will need to decide how some form of compensation will be structured. This compensation can for instance, be incorporated into the incentive payments by upfront reimbursement, or paid from a dedicated industry performance fund. It is strongly advisable that performance improvements are required before CPOs are fully compensated for any investments.

CPOs will also need certainty surrounding the sustainability of a model, and the ability to recover costs will also be essential in order to stimulate providers to make the necessary investments in the first place.

There are a couple of considerations which will inform the incentives that are attached to reaching particular quality thresholds. For financial incentives (e.g., bonuses), a range of matters should be clarified, in addition to the size of the payment, as follows:

- **How will incentives be funded?** Bonus incentives, for example, can be based on making a proportion of a current fee-for service arrangement dependent upon performance, or can be a complementary payment from an additional source of funding;
- **How will funds be allocated?** Will financial incentives in the long term be based on performance across the industry, or will each provider be measured on its own relative performance? Although it is preferable that providers should have as much discretion as

possible over reaching their performance thresholds, measurement of CPOs against their own individual performance contains an associated financial risk for care payers;

- **Will multiple or single quality performance targets be used, and how will payments be structured?** While multiple measures are superior to single targets, they do have a greater associated cost in development and maintenance. Moreover, when utilising multiple performance targets, it must be decided how incentive payments are structured if some targets have not been achieved;
- **When will incentives be paid?** The timing of payments is crucial to ensuring that there is a clear relationship between behaviour and outcomes. Therefore, it is preferable that payments are incentives are paid as soon as possible. It is, however, possible that time lags with measurement and analysis might prohibit this: – a balance must be found;
- **Will the incentives match the expectation of care providers?** The predictability of incentives is important for CPOs' and care professionals' motivation and financial planning. It is further advisable that the ability to meet incentivise thresholds is at the discretion of the care deliverers, and that this is not affected by external factors.

A3 Implementation

The literature, and in particular the case studies, highlighted that it can be beneficial for a care payer to opt for the gradual implementation of incentivised care arrangements. Phasing in quality and incentive measures gives CPOs and care professionals the opportunity to familiarise themselves with the expectations of the model.

For the implementation of incentive measures, the TAC and HDSG should identify whether they currently possess the capabilities to guide CPOs on the required change management process associated with the introduction of incentives, and if not, how they will develop or acquire this expertise. One possible avenue will be to bring in external experts to guide the organisations through this process. Matters that need to be part of the change management strategy include:

- The education process for CPOs, care professionals, and care users on incentivised care measures;
- The development of a communication strategy;
- The creation of support structures for adopting new and innovative care practices; and
- The implementation of appropriate evaluation and feedback structures on performance, and of dispute resolution mechanisms to resolve any bumps along the way.

As a first step, we recommend the development of a detailed change management plan for the next phases of an incentivised care.

- This plan will need to include realistic deadlines for the identification and development of quality measures, as well as the design of incentives;
- The length of the pilot project should be linked to the types of incentive measures that can be put in place, as well as the length of time before outcomes can be measured – this requires closer scrutiny.

A4 Measurement

Proper measurement of quality will be crucial to ascertain the effectiveness of incentive measures. Besides accurately capturing quality, there needs also to be a demonstrable, empirical relationship between the incentives and the behaviour that they purport to measure.

In addition to identifying which data is currently already available across the sector, some measures will need to be developed from scratch.

The TAC and HDSG can consider outsourcing the design and measurement of some of these instruments to external parties, or develop the expertise in house. As highlighted in the NH case studies, for instance, some care payers opted to measure care user satisfaction in-house, whilst others engaged external consultants to conduct this. A choice will thus need to be made between developing aspects of the models, such as quality measures and tools for the model (e.g., online administrative systems) in-house or outsourcing to the private sector or other independent experts.

A5 Evaluation

One of the major problems that was flagged throughout the literature was the inability of researchers to evaluate whether or not incentive models had been effective as a result of poor model design and implementation.

In terms of the evaluation, a proper baseline analysis needs to be undertaken at any participating CPO. It will not be possible to empirically demonstrate whether incentivised measures resulted in quality improvements without this baseline assessment. Thus, an AS IS analysis is required before moving forward to the TO BE.

We suggest that the TAC and HDSG adopt an evidence-based approach towards the implementation of incentives in the disability care sector. This approach should include an undistorted baseline analysis, which requires that no incentive measures should be put in place until a thorough baseline analysis with empirically validated quality measures is conducted.

We suggest that TAC and HDSG proceed with great care at this stage with the implementation of any incentive measures as part of a pilot study. We make this recommendation

because moving too fast might compromise the ability of the care payer, or an independent third party, to robustly assess whether the introduction of incentive measures has been the driver for any change in the care service delivery. The NDIS scheme, for instance, which is being gradually phased-in, is subject to a clearly formulated quality assurance and safeguards working arrangements, is an example of good practice.⁵

Additionally, it would be strongly preferable if a pilot project were to utilise a group of care users with similar care needs, in which case it is advisable that a control group be utilised as well. These pre-conditions would enable a more robust and valid analysis of the implications of incentivising care in one or more areas. It is also needs to be ascertained that there are no other quality improvement initiatives in place at any of the CPOs that may participate in any Incentivised care pilot project, since this might distort – and even could possibly invalidate – the findings about the impact of incentive measures.

Finally, as we stress, the desired outcomes (e.g., improvements in clinical care, organisational systems, and/or care user experience) need to be formulated in greater detail, so that an evidence-based model can be developed. This activity includes clearly identifying the goals, strategies, and objectives of the prospective incentive model.

⁵ For the safeguards that are in place for Victoria see <http://www.ndis.gov.au/document/quality-assurance-and-safeguards-workin-3>

Appendix 4 – Prospects for further research contribution

Finally, we identified four opportunities through which academic research could further assist the TAC and HDSG in the next steps of developing and implementing an incentivised model of care, should the organisation wish to proceed with this.

Firstly, we recognise the possibility for the care payer to engage further expertise to develop the quality measures that will be required to determine whether incentivising particular areas of care results in quality improvements. Providing assistance in evaluating existing quality measures that are available in the disability care sector is one avenue through which researchers could assist the organisation. Another opportunity arises in relation to the development of new measures. Expertise can be provided in relation to the ‘state-of-the-art’ empirical evidence on particular quality issues, ranging from clinical expertise to organisational systems. This may include academic experts from different disciplinary backgrounds, including the health sciences, health economics, and management.

Secondly, the TAC and HDSG may wish to engage assistance when setting the financial incentives. Modelling to identify the optimal size for incentives may be warranted. In particular, health economists may be able to assist the TAC in identifying the most suitable sizes of incentives within the compensable disability sector. The work by Kristensen, Siciliani and Sutton (2016), for instance, demonstrates that it is possible to identify the optimum price levels in a P4P model. At the same time, however, it needs to be recognized that, due to the high level of contextual factors that apply, the model would need to be tailor-made, should the TAC wish to base the size of incentives on such a model.

Thirdly, academic researchers can assist the TAC with the development of a change management program for the implementation of incentivised care initiatives. This assistance can include the co-development of the change plan, or facilitation of workshops with various stakeholders.

Lastly, researchers can be involved throughout the process to constantly monitor, document and evaluate the performance of CPOs to evaluate the effectiveness of implemented measures. This monitoring and evaluation can include qualitative and quantitative methods, including short- and long term assessments of performance.