Reviewer's report

Title: Payment for performance (P4P): any future in Italy?

Version: 1 Date: 7 September 2010

Reviewer: Robert Fleetcroft

Reviewer's report:

1. Is the question posed by the authors well defined?

Major Compulsory Revision

The aims of the study need to be more clearly defined in both the abstract and the paper. At the moment in the abstract in the conclusions a loosely defined aim is expressed (“this paper explored the feasibility of a P4P in the Lombardy region”). In the ‘background’ section, last paragraph it states “The purpose of this paper is to evaluate whether data is available and conditions for the implementation of P4P programs exist in Italy, in particular in the Region of Lombardy. We focused our attention on hospital inpatient care.” In the conclusion the paper states that it “explored the feasibility of a P4P program in the Lombardy region, Italy”.

If the aim is to explore the feasibility of a P4P programme, it needs to consider addressing the following points which have been identified in a recent systematic review as being important. [Van Herck P et al, Systematic review: Effects, design choices, and context of pay-for-performance in health care. BMC Health Serv Res. 2010 Aug 23;10(1):247] These are:

(a) select and define P4P targets on the basis of baseline room for improvement,
(b) make use of process and (intermediary) outcome indicators as target measures,
(c) involve stakeholders and communicate information about the programs thoroughly and directly,
(d) implement a uniform P4P design across payers,
(e) focus on both quality improvement and achievement, and
(f) distribute incentives to the individual and/or team level.

This paper provides answers based on the activity in 3 hospitals for a) and b), but not for the other outcomes. However for the rest of my review I will use the authors second definition of their aim, which was “to evaluate whether data is available and conditions for the implementation of P4P programs exist in Italy, in particular in the Region of Lombardy. We focused our attention on hospital inpatient care.”

2. Are the methods appropriate and well described?

Major Compulsory Revision
The method was to 1] identify quality measures that could be used in the hospital setting for a P4P scheme, and 2] examine the quality of this data recording in the electronic records in 3 hospital settings. The authors state that “Following a review of the international literature thirteen quality measures were identified regarding four clinical conditions”. This is an important exercise, but no further information was given as to how this literature review was conducted. It is important that the correct measures of quality are selected if we accept that P4P will make a difference, so it is therefore important that the reader knows that the selection of these measures of quality are robust and comprehensive, and this isn’t currently clear in this paper.

With respect to the setting of the study, the reader will need to know why 3 hospitals were chosen for the site of this study. For example how many hospitals are there in Lombardy, and are these 3 hospitals representative of others not included? Are they generalisable to the whole of Italy (the abstract states this as an objective, “It demonstrates that a P4P program could be implemented in Italy.”

Minor Essential Revision

With respect to the data quality, I assume patients were selected using the ICD-9-CM diagnostic coding. If this assumption is correct we need to know how accurate this coding is in these hospitals. Do the hospitals fail to record patients with a diagnosis? Are codings sometimes inaccurate? Both of these are needed to calculate the sensitivity and specificity that are mentioned in the results section. We need a clearer explanation as to how the hospital discharge letters were assessed for quality of their data. In pay for performance this data quality is crucial to the success of the scheme.

3. Are the data sound?

Minor Essential revision

There are missing data in table 5 for hospital N.2 in quality measures 3, 9, 11 and 12 (and also missing data on mortality in quality measure 6). Does this mean the discharge letter did not continue this information? If so then quality measure 3 should be 0/345, in 9 should be 0/1277 and in 12 should be 0/420.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?

Minor Essential Revision. I don’t know whether ethics approval is required for this type of study in the Italian setting, this need to be stated.

5. Are the discussion and conclusions well balanced and adequately supported by the data?

Major Compulsory Revisions

The “conclusions” section heading is really “discussion and conclusions”. It could be better structured, starting with a discussion of the results (currently it starts with commentary on aims and method, before discussion).
Currently discussion covers the availability of data and quality of data. This is quite well balanced and is supported by the data that has been presented. When the aims are more clearly defined then the discussion section could be better structured, and refer to what quality measures were selected, the disadvantages of areas of care not included in the proposed pay for performance scheme and possible unintended consequences, and what percentage of overall hospital activity these selected quality measures involved.

The discussion then goes on to discuss in more detail the implications of the quality measures in these 3 hospitals, and why performance was different between hospitals. This is really outside any of the 3 aims that the paper has stated and is probably a step too far for this paper. Using this type of data to accurately assess what performance can be accepted as high/low is important if payment stages for financial incentives are to be set in a proposed pay for performance scheme, but would require much more robust data that that returned from 3 hospitals such as in this study.

6. Are limitations of the work clearly stated?

Major Compulsory Revision

Some limitations of this work are discussed in the conclusions- 3rd paragraph from the end. However these do refer to some limitations in the literature regarding other pay for performance schemes, rather than the one this paper proposes. It should include ways to circumnavigate these matters identified, such as the small numbers of indicators that were identified from other projects.

Discretionary Revisions

The authors state there is little evidence to judge the effect of pay for performance programmes on quality. However in primary care this is not the case, there are several research articles from the UK on this matter, some examples are:


Walker S et al. Value for money and the quality and outcomes framework in primary care in the NHS. British Journal of General Practice 2010;60(574):352-57

Major Compulsory Revisions

The abstract also contains limitations and states- " In addition, before implementing P4P programs in Italy, several other factors should be clarified: which clinical conditions should be included, the threshold for each quality measure, the amount of financial incentives and the modes for their delivery." In a way this has identified major limitations that need to be addressed before proposing a pay for performance programme, which seems to be the aim of the
paper-so perhaps these questions need addressing to make the paper more robust. Indeed in the background section, para. 3 the paper states "Two types of information are required for a correct functioning of a P4P program. These include measures of health processes/outcomes that can be positively affected by clinical management, and information about specific treatments and clinical services that can transform individuals' initially poor health status into better outcomes". However the paper does not then go on to answer these questions fully.

There are further limitations that need fuller discussion. The aims need further clarification- then the limitations need to include ‘no user input’ either from health professionals or patients, the small number of areas for which quality measures were found, the generalisability of a survey of 3 hospitals and an approximation of costs involved in delivering a pay for performance programme.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?

Major Compulsory Revision

The authors state that they have performed a literature review which has identified important international literature, though there appear to be some important gaps. A major omission is a systematic review of pay for performance referred to previously by Van Herck et al.

8. Do the title and abstract accurately convey what has been found?

Because it is not entirely clear what the precise aim of the paper is it is hard to comment on this.

9. Is the writing acceptable?

yes

"Minor issues not for publication"

The list of contraindications in the British National Formulary for beta blockers are somewhat different to those quoted.

There are a few minor corrections, ‘s’ missing from the end of hospitals in several places in the abstract and paper.

I didn’t quite understand ‘DRG’ (3rd para, background) and this could be defined more clearly.

The abstract word count is incorrect.

Level of interest: An article of limited interest

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a
statistician.

**Declaration of competing interests:**

I have received funding from the Policy Research Programme in the Department of Health [England] for research into the health gain and cost effectiveness of the UK pay for performance programme.