Summary

The design and effects of payment systems in health care is an area that has received a great deal of attention among health economists. One explanation is that a payment system is perceived as an important part of a wider governance structure that influences the incentives and priorities across health care providers. From a historical perspective, there is a clear link between the design of payment systems and key problems discussed in different time periods. In order to solve problems of escalating expenditures within Swedish health care in the 1980s, fixed global budgets was introduced. To promote increased productivity and cost-effectiveness in the 1990s, compensation to providers was instead based on the volume of services.

In view of current developments towards transparent comparisons and "report-cards" on provider quality, and discussions about shortcomings in the quality of health care services and the gap between evidence-based methods and clinical practice, it should not come as a surprise that payment systems are being developed yet again. The new aspirations are to give hospitals and primary care providers economic compensation based on their performance and accomplishments towards defined objectives, so called pay-for-performance (P4P). In the past, the economic compensation has mainly been linked to resource inputs or outputs in terms of number of visits and/or number of cases. With P4P, the attention is instead on outcomes or even results or value for the patients. P4P may also be linked to the growing interest in evidence-based medicine. By giving providers extra payment if defined quality measures or targets are reached, or to threaten the same providers with economic penalties if targets are not reached, a further stimulus that supports a positive development can be implemented with the aid of economic incentives.

P4P is already used within Swedish health care, especially within primary care but to some extent also for hospital services and when the national government allocates grants to the 21 county councils responsible for health care services. A widespread interest to develop P4P further can be noted, although uncertainties about how to design the P4P-programs are relatively high. Moreover, many stakeholders want to participate and influence the development. This combination of a common interest, uncertainty and many stakeholders whowant to make a contribution, makes it especially important to develop a knowledge base that can support the future development of P4P in Swedish health care.

The systematic literature review presented in this report is based on searches in four relevant databases (Pubmed, HEED, NHS-HEED and Econlit) that resulted in just over 1000 abstracts and a review of 75 complete articles. P4P in health services is used in several countries. The majority of empirical studies about the effects of P4P, however, come from primary care in the UK or from health care services in the US. In spite of the fact that financing and delivery of health care services varies significantly between the UK and the US, it is interesting to note that the experiences from P4P are quite similar. So far, P4P have only contributed to quite modest benefits in terms of improved quality. In several studies, notably in the US, it has been difficult to separate the effects of P4P from other parallel changes, such as transparent comparison across providers based on report-cards.

In the UK, almost every primary care practice has received additional payments based on their performance and the indicators used in the so called Quality and Outcomes Framework. From one perspective, this is a clear indication that providers do care about payment systems and economic incentives. However, improved documentation has been an important factor to reach defined targets, rather than improved quality of care. The most significant effect is that the income of General Practitioners has increased substantially. There is some evidence of improvements in process measures of care for patients with a chronic disease, e.g. controls of patients with diabetes. Interviews with physicians and nurses also points towards changes in the activities carried out during a patient visit. In order to meet the requirements that follow from the P4P-program, the agenda during a patient visit has been pushed towards a standardized protocol and “box-ticking”, leaving more limited time for the patient to describe their individual problems.

Although a positive development has been documented for process measures of care, benefits from the patients' perspective and when focusing on outcome measures are more uncertain.

Also in the US, the main experience is that the benefits of P4P have been modest. In several programs, the lack of benefits should not come as surprise as the additional P4P-payment has been modest and only valid for a few of each provider's patients. In spite of the lack of clear benefits, the P4P-program is being developed further. Similar to developments in the UK, performance targets based on outcomes are slowly replacing targets based on process measures, which have dominated most P4P-programs so far. Developments in the US have also initiated a demand for changes at a larger scale. Several payers have argued for more fundamental changes in the payment systems, to the favor of comprehensive and capitation payment rather that the traditional payment based on the volume of services.

Although the initial high expectations have not been met, the interest in P4P remains. There is a general belief across first of all practitioners but also researchers that the lack of significant effects so far can be explained by poorly designed programs. Based on the analysis presented in this report there are indeed several questions to consider when designing a P4P-program:

- What is the problem?
- Which performance targets and indicators should be used?
- How difficult should it be to reach defined targets?
- Should individual physicians, teams or organizational units receive the payment?
- Are there alternative ways to reach objectives and solve defined problems?

The design of P4P-programs requires an explicit strategy where the answer to these questions are developed based on the overall vision of using P4P. A necessary condition in order to succeed is that the problems to be solved have been identified. Experiences show that the extra payment in P4P-programs often is allocated to providers that already have a high quality and performance level, and thereby are able to reach defined targets with little extra effort. If the purpose of P4P-programs is to give incentives to substantial improvements across a broader range of providers, knowledge about the base-line as well as variation across providers in terms of

performance is necessary. What problems are there?? Do problems exist among all providers or only among a few? Without such knowledge it is not possible to determine the appropriate performance target. Moreover, it is not possible to decide the appropriate size of the payment or if payments should be based on relative or absolute performance targets (or both), in order to give incentives to the appropriate providers.

Development of P4P-programs should also be looked upon as a process, where performance targets and payments included in the program changes in parallel to developments of clinical practice. This also means that in developing P4P-programs it is necessary to take the specific context into account which might vary across regions. At a more aggregate level, however, there are significant benefits if the 21 county councils can collaborate to develop evidence that can support the choice of valid performance targets and indicators and, not least, the development and use of medical quality registers and IT-systems.

As described in the report, a P4P-program can be based on several different types of performance targets. The choice of relevant targets and indicators can be discussed from several perspectives:

- Should clinical or system-oriented performance targets be used?
- Should targets focus on a specific diagnose or broader patient groups?
- Should targets focus on process or outcome measures?
- Should many targets or only a few targets be used?

Depending on the type and number of performance targets in any P4P-program, requirement for preparations as well as economic incentives will vary. The same decision will also have an impact on the probability that the appropriate providers get rewarded. On a more general level, the choice between different types of targets can also be linked to the purpose of the P4P-program. A program that builds on system-oriented targets, such as waiting-times for patients, may become a tool for implementation of political priorities. But a P4P-program may also rest on scientific evidence if the purpose is to support implementation of processes and methods that have a firm evidence base based on clinical research. If the choice of appropriate methods or processes is more

uncertain, a P4P-program can also be designed to support innovations and new ways of addressing patient needs by the use of outcome measures or targets addressing value for patients.

Since the choice between different types of performance targets is linked to different purposes, it is difficult - if not impossible - to provide a general answer to questions about which targets to choose. In the long run, however, the design of P4P-programs in Swedish health care should be able to focus on outcome measures and patient benefits, rather than process measures. Even if the use of outcome measures also can be associated with disadvantages, there is one clear advantage as providers are incentivized to develop innovations based on their own professional competence. Objectives based on process measures, which rely on certain evidence-based methods or procedures, can be supported in other ways, for example through direct demands on publicly owned providers or as part of the accreditation of private providers funded by public sources. Evidence also suggests that process measures are better suited to identify providers with poor quality, rather than to separate between the many providers with relatively good quality.

Even if P4P-programs are based on outcome measures and patient benefits they can be designed in multiple ways. It is therefore not possible to say anything about costs and consequences on a more general level. Consequences and benefits also depend on parallel measures to support a positive development. If the use of outcome measures is combined with a general support for a organizational culture that puts a high value on innovations and continuous improvements, the effects will most likely be greater compared to a P4P-program based on outcome measures implemented in isolation. The benefit of any specific program, however, can only be assessed in formal evaluations. By investment in more formal evaluations of Swedish programs it will also be possible to build an evidence-base that can support a stronger design of future P4P-programs.

A failure when it comes to the design of P4P-programs may in the worst case mean that actual benefits for patients develops to the worse, while the documented quality at the same time indicate improvements. Health care providers are rewarded in the belief that improvements are made, in spite of the fact clinical practice develops in the wrong direction. Examples are when providers avoid complex patients for whom it is more difficult to reach defined performance targets or if services to patients become

dominated by protocols and "box-ticking" rather than an interest for the patient's individual needs.

Even if a further development of P4P-programs in Swedish health care can be associated with risks of unintended effects there is also disadvantages with the alternative of leaving current payment systems as they are. In Swedish health care, there is a traditional and strong focus on fixed payment to both hospitals and primary care providers. Different and sometimes competing providers usually get the same payment, irrespective of existing differences in the quality of services. With this in mind, P4P may be seen as a much needed contribution to prevent too much focus on cost containment and the risk of under-provision of care and services that follows from fixed payment. There will always be a risk that the wrong providers are rewarded in a P4P-program or that it results in unintended effects and crowding out of objectives that are not part of the program. The programs will never be perfect. However, without using P4P as part of the payment systems to providers the wrong providers will to some extent be rewarded with certainty, as every provider get the same payment, irrespective of the quality of services..