Journal of Pharmaceutical Care & Health Systems

Sheiman, J Pharma Care Health Sys 2016, 3:2
http://dx.doi.org/10.4172/2376-0419.1000154

Review Article Open Access

Payment Methods for Integration: Typology, Evidence and Pre-conditions of Implementation

Igor Sheiman*

National Research University, High School of Economics, Moscow, Russia

***Corresponding author:** Igor Sheiman, National Research University, High School of Economics, Moscow, Russia, Tel: +7 612 0142; E-mail: igor.sheim@g23.relcom.ru

Received date: January 14, 2016; **Accepted date:** February 20, 2016; **Published date:** February 27, 2016

Abstract

Many countries have recently started the search for new payments methods with the specific objective to encourage the integration in service delivery. This paper suggests their typology. A brief overview of these methods in the USA and Europe, including Russia, indicates that there is still no strong evidence of their effects on integration and other dimensions of service delivery performance. It is argued that relative to other integrated methods global payment is the most promising method, since it provides incentives for comprehensive organizational changes. But this method is hard to implement – mostly due to a high probability of excessive financial risks placed on providers in integrated networks. The activities to mitigate these risks are discussed based on the approaches used in the Alternative Quality Contract in Massachusetts and fundholding scheme in Russia. The major pre-conditions for global payment implementation are specified: involvement of hospitals in global payment schemes, shared savings arrangements, special set of activities to mitigate financial risks, performance transparency system. It is also argued that there is a dilemma of strong economic incentives with serious implementation problems and lower economic incentives with less substantial implementation problems.

Keywords: Medical service integration; Integrated payment methods; Pay-for-performance; Episode based bundled payment; Global payment; Fundholding

Introduction

A key instrument to encourage integration of service delivery is the adoption of health care payment system that incorporates economic incentives for better interaction between health care providers. Three major attributes of integration should be promoted through methods of payment: 1) Teamwork of various providers, 2) Coordination of their activity, 3) Continuity of care at various stages of service delivery. They most comprehensively reflect integration activities: multispecialty groups of providers are established that work on the basis of joint clinical guidelines; their members interact with each other to enhance clinical and economic outcome; every stage of patients' "route" in the health system is closely connected with the previous and the subsequent ones. Economic incentives complement organizational changes for integration, clinical information exchange, new governance structure, cultural values and other factors of integration [1].

Many countries have recently started the search for new payment methods with the specific objective to provide such incentives. The scope of activities ranges from small-scale pilots to regional schemes. This search seems to be most intensive in the countries with the dominance of fee-for-service (FFS) and the lack of regulatory mechanisms for integration. The most interesting and ambitious projects have started in the USA and then are followed by some European countries.

The starting point of new methods discussion is a widely accepted view that FFS payment system leads to the fragmentation of care and the inefficient resource allocation. For example, Brantes et al. [2] argue that FFS discourages collaboration of providers and active efforts to reduce avoidable complications of care. Halvorson [3] gives the impressive examples of disintegrating and costly provision of care due to "selling units of care" rather than "a total package of care" that can be designed to avoid hospital admissions. Comparisons of physicians paid under FFS and capitation demonstrate that rates of elective surgery, patients' consultations, diagnostic services, specialist and hospital referrals are lower under capitation [4].

The policies for integration have prompted a number of questions. Which methods of payment can encourage integration? What is the typology of these methods? What is the evidence of their implementation? What are the pre-conditions for the viable payment system for integration?

This paper attempts to answer these questions. First, the typology of payment methods for integration is discussed. Second, a brief review of the evidence of these methods is presented. It is based on the projects implemented in the USA and some European countries. The focus is made on the global payment pilots in the Russian Federation. Third, relative strengths and weaknesses are evaluated against five criteria. Finally, we discuss the pre-conditions of the global payment implementation as the most comprehensive economic approach to strengthen integration.

There is another area of integration to be incentivized by payment methods – between health care and social care providers. It is not less relevant than the integration within the health sector. Special incentives to promote this integration are discussed in the literature [5]. In this paper we discuss the incentives for health care providers only.

Typology of Payment Methods for Integration

The methods that are specifically designed to reach the above mentioned attributes of integration can be regarded as payment methods for integration or integrated payment methods. The major expectation is that they incentivize providers to focus on integrated activities rather than selling units of care.

Relative to the traditional payment methods, payment for integration methods are designed to be less directly linked to the existing service delivery organizational structures. They allow more flexible allocation of financial resources to providers that deliver each element of an extended episode of care. These methods are based on the pathway of patients' movement in the health system and designed to integrate each element of service delivery through paying for a broad (bundled) unit of service volumes.

There are attempts to classify payment for integration methods. In the review of the "Eurohealth" [6], financial incentives and payment models are distinguished. Financial incentives are based on pay-for-performance method. The bonuses are paid for specific integration activities, mostly for achieving outcomes under chronic disease management programs. Payment models are addressed as bundled payment that is defined as "an annual payment for the complete package of care required by patients with chronic diseases" [7-12].

Mechanic and Altman [13] are more specific about bundled payment in the context of the USA. They suggest two types of such payment: bundled payment for episode of care and global payment. Bundled payment in the narrow sense is related to the specific disease or episode of care, such as hospital admission. Under this financial scheme, the rate of payment covers the cost of a bundle of services, including readmission. Global payment is defined as "all-inclusive payment per enrollee for a defined scope of services, regardless of how much care is actually provided". The authors refer to capitation method that is widely used in Europe for paying PHC providers and extend this method to integrated networks of outpatient and inpatient care providers.

The latter classification seems to be more precise because it takes into account the differences between methods of bundling the unit of payment. We see three major distinctions between global and episode based bundled payment. First, global payment is related to general medical activities, while bundled payment is the reimbursement of disease-specific activities. The further examples of global payment in the USA and Russia are comprehensive schemes that cover the cost of all types of care provided by integrated networks. Second, global payment always implies enrollment of patients with a specific integration scheme, while bundled payment does not. It can be used together with a traditional FFS method, which is the case in widely cited Dutch chronic disease management programs [14]. Third, global payment is always made for a specific period of time, while bundled payment is usually (although not always) related to an episode of care that does not have clear cut time dimension.

Thus financial incentives and two integrated payment methods can be discussed – pay-for-performance (P4P), episode based bundled payment and global payment. Each of them has its own subcategories.

Pay-for-performance is not a distinct payment method and is always used together with other payment methods (FFS, capitation or bundled payment) to reward specific integration activities. Bonuses are paid to promote the following integration activities:

Management of chronic cases

The most famous example is the Quality and Outcomes Framework (QOF) in **England**. The scheme was introduced to improve the quality of primary care and to encourage improvement in chronic disease management. A set of indicators and targets are used to measure GPs performance and financial rewards are linked to the actual achievement of each target.

The use of new IT for better information exchange between providers, as well as providers and patients

In **Denmark** telephone and email patients' consultations are encouraged by bonuses to promote on-going tracking chronic patients health status, proactive managing them, coordinating care with other providers [6].

Continuity of health care

In **Russia** a high rate of emergency care calls by chronic patients after hospital admission is considered as the indicator of inability of primary care providers to ensure continuity of care. A decrease of this indicator for the specific PHC provider is used as a positive performance measure and rewarded with the bonuses [15].

Usually financial incentives are used to influence the structure and processes of integrated care. But there are examples of their focus on the outcome of care [6].

P4P can be also divided into general and disease-specific. The former rewards general integration activities, the latter – reaching targets related to specific disease-based management programs.

Episode based bundled payment

The specification of the bundle of services differs in the literature. Struijs and Baan [14] address bundled payment as the payment for integrated set of services provided by outpatient care teams - GPs and outpatient specialists under programs of chronic disease management in the Netherlands. Brantes et al. [2] see this method as a reimbursement of inpatient care cost - not only the episode of hospital admission but also "a set period of management of a chronic condition", including readmission caused by low quality of inpatient care. Thus this method is used for the reimbursement of inpatient care or both outpatient and inpatient care.

Appleby et al. [16] suggest two main variants of bundled payment that are under development in England NHS: payment based on pathway tariffs and a year of life tariffs. The aim in both cases is to set prices in relation to a bundle of services covering the episode of care involving a number of providers, and all the care episodes a patient receives over a year. Marshall et al. [17] suggest the approach that is based on the separation of two types of tariffs – for acute care phase and the 'recovery, rehabilitation and re-ablement' phase, allowing this to be commissioned in the community rather than in hospital.

A broader definition is used for the version of diagnostically related groups - Medicare Severity Diagnostic Related Groups in the USA. This is a flat payment weight for multi morbidity cases – for the principal diagnosis and up to eight secondary diagnoses. Allowing for the severity of illness provides an incentive for hospitals to improve clinical integration of various specialties [18].

Thus all definitions agree on the object of bundled payment – episode of care with a specific diagnosis. But they differ in terms of the scope of services covered (inpatient and outpatient care or only outpatient care), coverage of diagnoses (single case vs. multi-morbidity case), the time period for payment under a single rate (the episode of admission or a lengthy period of treatment after admission).

Global payment

Global payment assumes financial accountability of providers working in integrated networks. They are responsible for the deviations of actual and expected cost (for example, the cost under capitation scheme). Global payment is always designed in the way that providers can keep savings and therefore are incentivized to more integration and more control over overutilization of services. They may bear risks of overspending as well.

Global payment schemes differ according to the level of financial risk bearing and the actors that act as risk bearers – PHC providers, hospitals or the entire network of providers. The specific type of global payment is a PHC *provider-fundholder scheme* (further - fundholding). Under this scheme PCH providers become holders of funding for outpatient and inpatient care. They are paid by all-inclusive capitation method, then act as purchasers of care – commission services and pay to specialists, hospitals and other providers of care that deal with the enrolled population. Their risk bearing creates incentives for closer links with other providers to avoid overutilization of costly services and to enhance quality of care at the level of PHC. This method has first been tested in the USSR in late 80's and is currently used in some Russian regions. Similar schemes were used or piloted in various versions in the UK, Italy, Sweden, Finland, Estonia [19].

P4P, episode based bundled payment and global payment can be used in combination as mixed method. The options of combination can be different.

Summing up the above, we suggest the typology of financial incentives and payment methods for integration (Table 1). Each method, including P4P, is subdivided into a few categories based on different criteria: financial incentives – on the objectives and areas of incentives (promoting teamwork, etc.), episode–based bundled payment – on the area and design of payment systems (scope of care covered by the bundled payment, degree of service bundling specification, coverage of diagnoses, time period for a bundled payment), global payment for the general activity – on providers acting as major risk bearers. To the certain degree, the versions of integrated payment methods may be overlapping. Global payment for the general activity can have similar versions as disease-specific bundled payment, for example, global payment can be made for a set of pathways. But their major distinction remains. This is a relatively higher rate of risk bearing, because it relates to the entire or major part of integrated networks activity.

Method	Versions of the method	Country models examples
Financial incentives (pay-for-performance)	Bonus schemes rewarding: -management of chronic cases -use of new IT for better information exchange -ensuring continuity of care	Denmark Russia QOF, England NHS
Episode-based bundled payment	*1) Scope of care:* -outpatient care only -inpatient care, including readmissions -inpatient care, including rehabilitation services -outpatient and inpatient care *2) Coverage of diagnoses:* -single case -multi-morbidity case *3) Time period:* -episode of admission -lengthy period after admission *4) Degree of service bundling specification:* -pathway of service delivery is known -unknown	Geisinger Health System, the USA The US Medicare DRGs Prometheus Payment model Chronic disease management care groups, the Netherlands. Pilot projects in England NHS, including pathway-based payments for maternity cases and a "year of care" payment for life-long conditions.
Global payment for the general activity	Schemes with major risk bearers: -PHC providers (fundholding) -hospitals -all providers in integrated networks	Accountable care organizations, the USA. Alternative Quality Contract, Massachusetts, the USA. Pilots of fundholding in 10 regions of Russia.
Mixed methods	-episode-based bundled payment+P4P -global payment+P4P	Used in all episode-based and global payment schemes

	-global payment+episode-based bundled payment for specific diseases+P4P	

Table 1: Typology of financial incentives and payment methods for integration.

Design and Evidence of Payment Methods for Integration

Integrated methods are usually designed to achieve many objectives, including strengthening integration of service delivery. With the available instruments of evaluation, it is hard to disentangle the impact of these methods on integration. Therefore, the following review is focused on three major dimensions of providers' performance: a) quality of care, b) utilization of care and health spending, c) integration activities.

Regarding P4P, we rely on the evidence collected elsewhere [6] with the major conclusion that rewarding the achievement of specific targets indicators of integration, although useful, can hardly contribute substantially to integration. For example, the most famous example of P4P in Europe – the Quality and Outcomes Framework (QOF) in England – may have contributed to higher quality of care [17,20], but there are no signs that QOF has contributed to integration through promoting continuity and co-ordination of care. There has been some reported concern that the OQF can potentially result in the neglect of non-incentivized areas [21,22]. Also, it does not provide incentives to develop new ways of delivering care for people with co-morbidities and long-term conditions. There is a risk that people with co-morbidities are not treated holistically, since indicators and targets are used separately for each condition, and no single organization takes the responsibility for the total of patient's needs.

Episode-based bundled payment

Desirable outcomes of this method are twofold. The first is to encourage a shift of care from inpatient to outpatient units of hospitals. Having a bundled rate that covers the cost of admissions, physician and other clinical services, hospitals are interested to enhance the readiness of patients for admission – to have all necessary consultations and tests, plan clinical activity prior to admission. At the same time, they are interested to reduce unnecessary physician and ancillary services. The second outcome is to reduce complications and readmissions. To achieve this, the rate of payment includes expected costs for readmissions within a specified period of time. This rate is set on the basis of a targeted reduction in readmissions so that hospitals don't have an "umbrella" for the reimbursement of low quality care.

The most widely cited example of bundled payment is the **Geisinger Health System** (GHS) initiative in the USA, implemented by an integrated delivery network of physicians and hospitals. This system offers forty specific clinical processes related to managing patients after coronary artery bypass surgery. The integrated rate includes the cost of surgery, all necessary tests and post discharge follow-up of patients within 90 days. The cost of services related to possible complications and readmissions is covered by this rate and not reimbursed additionally to the regular admission. The rate assumes that the GHS will reduce the historic frequency of complications by half, that is pricing is based on growing requirements to quality of care [13].

Evaluation of the first eighteen months of the project implementation found a 44 percent readmissions reduction, shorter average length of stay, and reduced hospital charges. According to Geisinger executives, the success of the initiative is due in large part to integrated delivery network with an electronic record that allows tracking patients after discharge. The major problem is how to equitably distribute episode payments across physicians, hospitals and other providers involved in the project. Therefore, there is some resistance of physician-hospital integrated networks to bundled payment [13].

In Europe episode based bundled payments are piloted in various countries, including the Netherlands, the UK and Denmark. Contrary to the US initiatives, they are implemented mostly in the sector of outpatient care. For example, in the **Netherlands** the bundled payment was established in 2007 for diabetes care. Under the scheme, health insurers are able to purchase a bundle of services needed to manage chronic diseases through the payment of a single fee to newly created contracting entities called "care groups" These groups are clinically and financially responsible for all patients enrolled in the diabetes care program. The main objective of the bundled payment model is to encourage providers to improve coordination of care and to decrease utilization of expensive specialist care.

Evaluation of the scheme made in 2012 indicated that 25% fewer bundled payment patients utilized specialist care in comparison to care-as-usual patients (not enrolled with care groups). This contributed to some savings per patient in the cost of diabetes-specific specialist care, but when non-diabetes costs are included total specialist costs for bundled payment patients increased more than the costs for care-as-usual patients. The higher costs for bundled payment patients can be attributed to start-up costs of these groups. Also, the long-term effect is possible, since diabetes complications often take a long time to develop. It is too early to draw definitive conclusions about the long-term impact on the cost and quality of care.

Another result of this evaluation is the revealed problem of the uncertainty of cost allocation across programs of patients' management and regular set of services reimbursed through FFS. There are concerns that care groups may be double billing the insurer for the same services – through traditional FFS and bundled payment. When more chronic diseases are added to the bundled payment schemes it will be difficult to determine under which bundle certain services should be billed [13].

The recent local and national experiments with bundled payment in **England NHS** include payments for life-long conditions such as cystic fibrosis and for high-risk, multi morbidity patients, and pathway-based payments for maternity services. They combine weighted capitation with P4P and risk sharing between providers working with these cases [17].

These experiments have revealed additional problems with their implementation. Charlesworth et al. [23] suggest three reasons not to make these models a key feature of future payment systems. First, it is hard to determine a set of services to be covered by capitation payment for specific disease, and therefore to determine organizational boundaries for capitated payment. Second, special regulatory regimes are needed to protect patients from the risk of denial of care and

quality reduction (see further the Russian case of dealing with this problem). Third, this model will create winners and losers, due to the kind of geographical variation in cost and utilization, at a time of serious financial constraints for the country.

These implementation problems, in our view, highlight a broader challenge – the narrow scope of bundled payment, its limited impact on integration. The scheme is designed for the specific diseases management programs and doesn't create incentives for providers in other clinical areas. There is a potential for providers to skew their activity to the most "rewardable" schemes.

Moreover, integration efforts in one clinical area promoted by bundled payment may co-exist with fragmentation of care in other clinical area. To cope with this, more holistic approaches to payment are needed, such as global payment that is related to the entire medical activities of integrated networks.

Global payment

Two global payment models are reviewed in this section – the Alternative Quality Contract in Massachusetts, USA and fundholding scheme in the Russian Federation. Although different in their background and design, they represent the common characteristics – comprehensive coverage of services in integrated networks and high degree of risk bearing.

The comparison of these models provides insights on pre-conditions of the viable payment system to encourage integration enhance quality and avoid overutilization of care.

Alternative quality contract (AQC): Starting from 2009, this contract is negotiated between the insurance company (Blue Cross, Blue Shield) and groups of providers. A 5 year global budget is provided to these groups. It covers the entire continuum of care, including inpatient, outpatient, rehabilitation, long-term care and prescription drugs. Before the contract is signed, the rate of increase for each year is set. The goal of the AQC is to decrease the rate of growth in health spending by 50% after five years.

Providers in a global budget model not only share in savings if spending is below the budget but are also accountable for deficits when spending exceeds the budget. This is a stronger payment model than shared savings models in Accountable Care Organizations (ACOs) because it presumes higher financial risks: groups of providers may not only win (as it is the case in ACOs) but also lose their income.

Global budget is combined with quality metrics so that provider groups would have enough incentive to improve quality and cope with withholding the necessary care in order to save money. Financial incentives of up to 10 percent of total reimbursement are paid to groups of providers.

They are based on 32 measures for ambulatory care and 32 for hospital care. The bonuses are not incorporated into the budgets and must be earned each year. Savings can be kept only if P4P targets are reached. The pattern of sharing is determined by an integrated network itself.

A few instruments are used to mitigate financial risks of provider groups. First, changes in enrollees' health status are measured concurrently and capitation rate is adjusted annually using risk-adjustment model. Thus higher-than-typical proportion of sicker patients is identified with the subsequent adjustment of the budget.

Second, provider groups can choose to participate in the AQC on a risk-sharing basis rather than a full-risk arrangement. In the former case they share risks with the insurer, in the latter – accept 100 percent risk for deviations of their spending over the planned budget. The risk borne by groups ranges from 50 percent to 100 percent. Third, the groups are required to have reinsurance for covering medical spending that exceeds a specific threshold [24].

The insurance plan requires enrollees to designate a primary care physician (PCP) and assumes referrals to specialists. The PCP organization is accountable for all enrollee services, regardless of whether the enrollee receives care in this organization or any other provider. As of May 2013, a total of 17 physician-hospital groups signed AQC contracts to deliver care to 689.105 HMO members. Groups range from large physician-hospital organizations to small independent practices with common leadership.

Song et al. [25] compare the AQC system with control group for the period 2006-2009. The results of the evaluation demonstrate small savings (1.9 per cent per quarter). They were achieved through changes in referral patterns to providers with lower fees rather than decrease in utilization. The AQC system is associated with the improvement in measures of quality of management of chronic conditions in adults and pediatric care, but not for adult preventive care. All AQC groups met 2009 budget targets and earned surpluses.

The authors conclude that the AQC system demonstrates moderate positive results although not so substantial as was expected. Similar study of the same authors for the second year of the AQC implementation demonstrates more substantial positive results. Savings increased to 3.3 percent in year 2 compared to spending in nonparticipating group. Quality of care also improved with chronic care management, adult preventive care, and pediatric care within contracting groups improving more in year 2 than in year 1 [26].

There is an anecdotal evidence of a positive impact on integration. The interviews with physicians and administrators of participating groups indicate that physicians and hospitals now are more interested to work together so that to reach spending and quality targets. Sharing savings between providers is becoming the instrument to reorganize the delivery of care and ensure continuity of care. "When 50 percent of your patients are at risk, you don't stop the care when patients are discharged. We have a huge amount of outpatient work we do in the offices, in skilled nursing and rehab facilities" [27] - this is the opinion of the senior administrator the hospital-physician group participating in the AQC. Also, there are new incentives to encourage proactive actions of physicians to manage chronic cases and thereby to decrease hospital care utilization. Outpatient case managers are hired.

A special work is done to identify patients who need care but have not seen a doctor recently. With all limitations of this kind of evidence, it demonstrates the positive vector of the changes. The additional evidence is needed to see the outcome of the new scheme.

Fundholding: The version of global payment known as "polyclinic as fund holder" (further – fundholding) is implemented in 10 out of 83 regions of the Russian Federation (A region is a big area with the population ranging from one to a few million people. Most of providers are state owned and governed by regional health authorities).

The background of the scheme is substantially different if compared with the AQC in Massachusetts. First, the Russian health system is funded mostly from public sources (mandatory health insurance revenue) and every regional health system has its fixed budget.

The system is close-ended and poorly funded (public health expenditures amount to only 3.5% of GDP), therefore the major objective of the global payment is not reducing health spending but coping with the structural distortions in service delivery – under provision of PHC and overprovision of inpatient care. However, other objectives are the same – more emphasis on preventive care and chronic diseases management, enhancing quality of care, strengthening integration of service delivery. Similar to the AQC, global payment is seen as the alternative to the dominance of FFS method.

Another area of distinction is a relatively higher degree of integration in the Russian health system as a starting point for the payment reform. Outpatient care has traditionally been provided mostly by state owned multispecialty polyclinics with district physicians (Russian version of GPs) and specialists in their staff.

They have catchment area population, while residents can chose a polyclinic and district physician in it. The latter acts as a gatekeeper and traditionally has been seen as the provider responsible for the dynamic supervision of the enrolled patients. Some hospitals have polyclinics as structural units. These and some other organizational characteristics facilitate coordination of service providers and make global payment easier to implement.

The design of global payment in the Russian Federation (Description of the fundholding design is based on Sheiman [10] and unpublished data of the Russian Federation Federal Fund of mandatory health insurance) reflects the above mentioned objectives of payment reform. Polyclinics are capitated for the catchment area population, and pay for referrals to hospitals and other providers. Global payment to polyclinics includes the expected cost of outpatient care, all or part of inpatient care and care provided by special emergency care units (in Russia most of them are free-standing providers with a relatively high volume of curative responsibilities).

The range of services differs in the regional schemes from small scale global budgets (e.g. in Kaluga region - only for outpatient care) to all-inclusive payment (in Kaliningrad region). Polyclinics can keep savings, therefore are interested to increase their own volume of care and reduce referrals to hospitals. Similar to the AQC, polyclinics are accountable for deficits of finance.

Global payment to polyclinics is supplemented by the financial incentives. The number of indicators ranges across regions from 5-6 very general indicators to 20-30 detailed disease-specific quality and outcome indicators. General indicators include the coverage of check-ups, hospital admissions and emergency care rates for the list of chronic patients under supervision, home death rate, the incidence of the revealed cancer cases at terminal stages, TB incidence, etc. These indicators are designed to promote preventive care, capture "defects" of PHC providers work and cope with withholding the necessary care.

They are heavily weighted in a formal "model of performance assessment" that allows to evaluate a degree of reaching the targets for each indicator and build an "integrated coefficient of reaching performance targets", that is an integrated performance indicator.

In most of regions with fundholding scheme, there is a regulation that the savings of polyclinics can be kept by them only when most of performance targets are reached and the integrated coefficient is higher than the established minimum. The regions differ substantially in the size of this minimum.

To mitigate financial risks of fund holders, some regions (for example, Kemerovo region) use the scheme of risk sharing between health insurers and polyclinics. The latter have the limit of their financial responsibility. Particularly costly cases are reimbursed directly by insurers. This scheme requires a clear specification of the scope of risk bearing by policlinics. If they don't control utilization of certain specialized services by patients (they tend to see specialists without referrals), then these services are not included in capitation rate and reimbursed directly by insurers.

Also, like in the AQC, a degree of risk bearing by polyclinics may differ – ranging from minimum to full responsibility for inpatient care cost. But contrary to the Massachusetts scheme, a degree of risk bearing is determined by the local regulation and is universal for all participating providers.

Financial penalties are used for patients that have not been timely referred to hospital. Health insurers are responsible for revealing such cases and penalizing polyclinics. They are interested in this kind of control because can keep 10% of the financial penalty size. This is another instrument to cope with the opportunistic behavior of fund holders.

The potential of fundholding to strengthen integration is based on the assumption that polyclinics are interested in preventive care, managing chronic cases (to avoid admissions and emergency care visits), strengthening coordination function of PHC providers, cooperating with high quality hospitals and ensuring continuity of care provided by various providers. This motivation may also lead to controlling overutilization of care and enhancing quality and outcome of care.

The practical implementation of the scheme over the last 3-4 years in Russia has had mixed results. Polyclinics as fund holders tend to provide more care to chronic cases. The first innovation is to set up physician-nurses teams for home visits in case of emergency or expected aggravation of health status. For example, in Perm region the number of home visits per capita by polyclinics' personnel has doubled over the first 9 months of the new method implementation, while the number of emergency calls by chronic patients has decreased by 3 times. The second new approach is to set up "schools of diabetes" and "schools of asthma" (simplified versions of the Western programs of chronic disease management programs).

In some regions with fundholding scheme, polyclinics have attempted to be involved in planning and controlling the volumes of inpatient care. They draft the plans of their patients' needs for hospital admissions and then negotiate volumes of care with hospitals. The case of Samara city has proved that such planning can serve as a strong leverage to avoid inappropriate admissions, decrease LOS in hospitals and ensure the higher readiness of PHC providers for the treatment patients after hospital discharge.

However, hospitals oppose to such mode of planning, therefore it was implemented mostly due to the administrative pressure of the city government. Although promising, this approach has not been rolled out in other regions with fundholding.

There is the evidence of fundholding impact on utilization of services. For example, Kaliningrad region (the leader in fundholding implementation in Russia) has lower than average rates of physicians visits, volumes of inpatient care, LOS, frequency of emergency care calls – with general mortality rate and cardio mortality rate lower than the average for Russia (Table 2).

	Kaliningrad region	Average Russia
Physician visits per capita	7,1	9,6
Emergency calls per 1000	294	332
Inpatient bed-days per 1000	223	264
Average length of stay	11,4	12,4
General mortality rate	13,2	13,3
Cardio mortality rate	700	737
Source: Russian Federation Federal Fund of mandatory health insurance [1].		

Table 2: Health care utilization and outcomes in Kaliningrad region compared to the Russian average in 2013.

However, fundholding has not contributed much to integration of care yet. Polyclinics do not control patient flows due to weakening the referral system and low trust of patients in district physicians (according to the polls, only 14% of respondents trust their qualification). Interaction between outpatient and inpatient care providers remains low. Information exchange is in the initial stage [1]. Potential strengths of this payment method are mitigated by inherent drawbacks of the organization of service delivery, particularly the low role of district physicians and poorly regulated interactions between providers of all levels. Also, there are conflicting incentives for integration on the part of polyclinics and hospitals. The former are interested in more cooperation with hospitals to avoid aggravations of chronic cases, while the latter are not. They tend to increase admissions, including those which are inappropriate.

Comparison of the AQC and fundholding

There are many similarities in two schemes of global payment. First, their objectives are more or less the same, although fundholding in Russia is not directly focused on containing health spending. Second, both schemes are based on a substantial degree of risk bearing. Groups of providers are financially accountable not only for the savings but also for the deficits of revenue to cover the cost of care provision. Third, global budget is supplemented by P4P to provide additional leverage to enhance quality of care and to avoid withholding the necessary care. Fourth, both schemes face the problem of managing financial risks of providers and therefore use a set of instruments to mitigate them. Some of these instruments are similar - sharing risks with insurers and the optional level of risk bearing borne by providers.

The major distinctions are the following:

- The initial level of service delivery integration in higher in the Russian Federation, therefore some arrangements, like establishing multispecialty groups, are not relevant here.
- Risk bearing actors are different: PHC providers – in the fundholding scheme, integrated groups of providers - in the AQC.
- Formal arrangements between PHC providers and hospitals as well as risk sharing between them are the key elements of the AQC. They don't exist in fundholding. Its design is based on the conflict between polyclinics and hospitals rather than cooperation between them. Hospitals are not involved in the arrangement and therefore are not interested in decrease in overutilization and stronger links with PHC providers (Samara city case provides evidence to this). The absence of risk sharing makes the Russian scheme very vulnerable. Polyclinics face the problem of excessive financial risks.
- Contrary to the AQC, the budget for the next year in the fundholding scheme is not linked to the performance. The reason for this is that public funding in Russia is always inadequate and not totally predictable.
- Risk-adjustment of capitation rate is a regular procedure in the AQC but not in the fundholding scheme. This aggravates the problem of excessive risk bearing for the Russian polyclinics that have higher than average portion of sicker people.
- In both countries, the evidence of global payment impact is still weak. At this stage it is clear that the method is hard to implement as it requires a careful "tuning" of providers' incentives to avoid their opportunistic behaviour.

Discussion

The brief overview of newly introduced payment methods for integration indicates that there is still no strong evidence of their effects on integration activities as well as on other dimensions of service delivery system performance. However, the evidence available coupled with the experience of new methods implementation (The author has been involved in building global payment schemes in Kaluga, Kemerovo, Samara and Perm regions of Russia) provides the ground for the comparison of their relative strengths and weaknesses, as well as making generalizations on pre-conditions of their successful implementation.

We summarize P4P, specific disease bundled payment and global payment according to five criteria. Four of those were suggested by Mechanic and Altman [13] – potential to provide integration, reduce unnecessary utilization, and encourage high quality care and operational feasibility. We suggest the fifth criterion - the degree of excessive financial risks of providers. The rationale behind this is that too high financial risks may potentially lead to the unwillingness of providers to bear such risks or their opportunistic behaviour – under referring and under treatment. The results of the evaluation are presented in Table 3.

P4P method is relatively easy to implement. Providers don't bear any risks, they can only win but never lose from these scheme. The impact on integration, utilization and quality is low. Episode-based payment can promote integration in specific disease management programs with the potential of reducing unnecessary utilization and encouraging high quality care. It is relatively easy to implement. The

possibility of excessive financial risks of providers exists but it is much lower than in the case of global payment – due to a relatively small scale of services to be covered by bundled payment.

	Promoting provider integration	Controlling unnecessary utilization	Encouraging high quality care	Operational feasibility	Degree of excessive financial risks borne by providers
Pay-for-performance	*	-	*	***	-
Episode-based payment	**	**	**	**	*
Global Payment	***	***	**	*	***
Source: Author analysis					

Table 3: Evaluation of integrated payment methods based on key criteria.

Global payment method has the highest potential for integration and high performance of providers. It is more comprehensive, that is designed for the entire medical care rather than specific diseases management. Capitation rate that covers the services of big organizational entities, like physician-hospital groups, is the leverage to promote formal and virtual links between providers. As Ham puts it [28], "the more comprehensive the scope of capitated budgets, the more important this becomes".

Global payment usually implies a substantial risk bearing by providers. For example, in some integrated networks in the US hospitals agree to pay out physicians from pre-capitized accounts after procedure before receiving reimbursement from payer. Thus risks to hospitals may be excessive and this affects their willingness to be involved in the new arrangement [29]. In Russia polyclinics as fund holders are supposed to pay for hospital care even in the situation when the revenue from the payer is not enough to cover all cost or when capitation rate is not risk-adjusted enough. Such risk bearing creates strong incentives for savings and therefore for closer cooperation of providers to achieve this. The other side of the medal is that risk bearing may be excessive with the resulting opportunistic behaviour of providers, their unwillingness to be involved in networks and even growing tension between providers. For example, in the above mentioned Kaliningrad region the incentives of fundholding are so strong that bed capacity has decreased substantially, and a few hospitals have been closed. High requirements to appropriateness of admissions and LOS, as well as selective referrals to the best providers, have been opposed by many hospitals and physicians.

We may conclude with a careful generalization that more comprehensive methods of payment create stronger economic incentives for integration but at the same time make integrated networks more vulnerable. Disease specific bundled payment is less conflicting and easier to implement but its potential for integration is relatively lower.

Another area of discussion is the specification of pre-conditions for a positive impact of global payment. The experience of Russian regions with fundholding indicates that global payment per se may be neutral to any organizational changes. Moreover, it can create obstacles to changes if it is designed totally for cost containment. In the former Soviet health system in the USSR all providers had global budget that was input-based (e.g. cost of labour, utilities) irrespective of the actual volumes and quality of care, therefore did not create any incentives to improve providers performance. Integration was achieved only due to command-and control methods [30,31]. Bearing in mind this "extreme" case of providers' demotivation, we suggest six major pre-conditions for the global payment as integration instrument:

Involvement of hospitals in global payment schemes

Formal contractual arrangements between hospitals and PCH providers are needed to achieve common objectives of integrated schemes. These arrangements should specify the activities to enhance teamwork, coordination of care and continuity of care. Global payment based on conflicting interests of PHC providers and hospitals (the case of fundholding) should give way to the common incentives to ensure savings in the entire network of providers.

Shared savings schemes

This is critical to ensure a transparent pattern of savings distribution between providers in the networks. The schemes should be designed to involve all participating providers in the integrating activities. If PHC group acts as the major risk bearer, then hospitals must be encouraged to work in such networks and meet their objectives. Shared savings schemes based on formal contractual arrangements may serve as the instrument to encourage such collaboration. Global payment without shared savings is doomed to conflicting interests within groups of providers.

Performance transparency system

Each provider in a network has clear cut performance targets that comply with its general objectives. P4P is used to encourage providers to reach these targets – in addition to the incentive of shared savings. To track the performance of each provider, a monitoring system is needed, as well as expanded analytics capabilities.

Activities to mitigate financial risks

They are designed to ensure sustainability of global payment and to avoid opportunistic behaviour of providers. The experience of the AQC and the fundholding suggests a set of major activities: shared savings payment subject to reaching performance targets, sharing risks between providers and purchasers of care, optional involvement of provider in risk bearing, concurrent risk-adjustment of capitation formula. Without these activities integrated networks become vulnerable and conflicting.

Managerial control of services under global payment

Enrolment of patients, a sound referral system and careful specification of the services under global payment are the major instruments of managerial control. Global payment should be set for the services that are based on the clear cut pathways of patients' movement in the health system. The services that are not controlled by providers groups should be avoided.

Infrastructure for coordination and collaboration

Re-organization of a service delivery is seen by some commentators as the stage preceding global payment [32]. Based on the Russian experience, we can add that the global payment may not provide adequate incentives to compensate for the lack of organizational integration activities and the low coordinating role of PHC providers, let alone their shortage, which is the case in the USA and Russia [1,27]. This is a substantial obstacle to integration that can't be overcome only through new economic incentives.

These conditions of potential input of global payment are country specific. Even within one country they may differ substantially. The lack of infrastructure for coordination and collaboration may be regarded as the sign for more careful and phased out transition to integrated payment method. In countries with particularly fragmented care and the dominance of FFS as payment method, P4P and specific disease bundled payment (or their combination) may be the first step to create incentives for integration. Their major task is to foster initial integration activities in specific areas. Global payment can be used in more "mature" integrated networks – after implementing key steps to prepare payment reform.

On this way to a broader unit of payment, combination of traditional and integrated payment methods becomes inevitable. Global payment is coupled with FFS for a set of priority services, mostly preventive ones. FFS is also used for the services that are beyond control of providers as risk bearers (they are not included in capitation rate) and therefore reimbursed directly by a payer. Also, the combination of any form of bundled payment and P4P becomes critical so that to implement savings sharing schemes according to performance targets. Thus any integrated method is most likely to be transformed into the mixed payment system. Reaching a viable combination of various methods is the major problem of integrated payment implementation.

Conclusion

Three methods can encourage integration in service delivery: P4P, episode based bundled payment and global payment. Each of them has its own subcategories. A suggested typology is focused on the distinctions between episode based bundled payment and global payment.

A brief overview of these methods implementation, with the focus on the Alternative Quality Contract in Massachusetts and fundholding in Russia, indicates that global payment is the most promising method, since it provides incentives for comprehensive organizational changes. But its implementation is hard – due to the potential excessive financial risks of providers.

The major pre-conditions for the global payment as integration instrument are: involvement of hospitals in global payment schemes, shared savings arrangements, activities to mitigate financial risks, performance transparency system, managerial control of services under global payment, infrastructure for coordination and collaboration. The lack of infrastructure may be regarded as the sign for more careful and phased out transition to integrated payment method.

More comprehensive methods of payment create stronger economic incentives for integration but at the same time they are hard to implement and make integrated networks more vulnerable. There is a dilemma of strong economic incentives with serious implementation problems and low economic incentives with no or few implementation problems.

Acknowledgement

This paper is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE) in Moscow, Russia.

References

1. Sheiman I, Shevski V (2014) Evaluation of health care delivery integration: The case of the Russian federation. Health Policy 115: 128-137.
2. Brantes F, Rosenthal M, Painter M (2009) Building a bridge from fragmentation to accountability – The Prometheus payment model. New England Journal of Medicine, pp: 1033-1036.
3. Halvorson CG (2009) Health care will not reform itself. Productivity Press.
4. Dusheiko M, Gravelle H, Jacobs R, Smith P (2006) The effects of financial incentives on gatekeeping doctors: Evidence from a natural experiment. Journal of Health Economics 25: 449-478.
5. Mason A, Goddard M, Weatherly H (2014) Financial mechanisms for integrating funds for health and social care: An evidence review. Center for Health Economics, University of York, UK.
6. Hernandesz-Quevedo C (2013) Enhancing chronic care management in Denmark. Eurohealth 19: 17-18.
7. Bienkowska-Gibbs T (2013) Integrated care programs in Canada. Eurohealth 19: 13-14.
8. Hernández-Quevedo C, Llano R, Mossialos E (2013) Paying for integrated care: An overview. Eurohealth 19: 3-6.
9. Liano R (2013) Bundled payments for integrated care in the Netherlands. Eurohealth 19: 16-17.
10. Sheiman I (2011) Pervi opit realizatsii sistemi "policlnika-fondoderzchatel". Menedzcher zdravoohranenia.
11. Struijs J (2013) Payment reform and integrated care: The need for evaluation. International Journal of Integrated Care.
12. Young R (2013) Team care arrangements in Australia. Eurohealth 19: 11-12.
13. Mechanic R, Altman S (2009) Payment reform options: Episode payment is a good place to start. Health Affairs 28: 262-271.
14. Struijs J, Baan C (2011) Integrating care through bundled payments — lessons from the Netherlands. New England Journal of Medicine 364: 990-991.
15. Sheiman I (2010) Sistemi stimulirovania povishenia cachestva meditsinskoi pomozchi Upravlenie Zdravoohraneniem.
16. Appleby J, Harrison T, Hawkins L, Dixon A (2012) Payment by results: How can payment systems help to deliver better care? The King's Fund.
17. Marshall L, Charlesworth A, Hurst J (2014) The NHS payment system: Evolving policy and emerging evidence. Nuffildtrust, London.
18. Durand-Zaleski I, Obrecht F (2008) In: Nolte E, Knai C, McKee M (eds.) Managing chronic conditions: Experience in eight countries. European Observatory on Health Systems and Policies, Copenhagen.
19. McCallum B, Brommes M, Robinson R, Bergman S, Palu T (2009) The impact of primary care purchasing in Europe: A comparative case study

of primary care reforms. In: Primary care in the driver's seat? Saltman R, Rico A, Boerma W (eds.) WHO, European observatory on health systems and health policy. Open University Press, pp: 105-128.

20. Steel N, Willems S (2010) Research learning from the UK quality and outcomes framework: a review of existing research. Qual Prim Care 18: 117-125.
21. Ham C, Smith J, Eastmure E (2011) Commissioning integrated care in a Liberated NHS. Nuffield Trust, London.
22. Maresso A (2013) The quality and outcome framework in England. Eurohealth 19: 9-10.
23. Charlesworth A, Hawkins L, Marshall L (2014) NHS payment reform: Lessons from the past and directions for the future. Nuffildtrust, London, pp: 1-28.
24. Chernew ME, Mechanic RE, Landon BE, Safran DG (2011) Private-payer innovation in Massachusetts: The "alternative quality contract". Health Affairs 30: 51-61.
25. Song Z, Safran DG, Landon BE, He Y, Ellis RP, et al. (2011) Health care spending and quality in year 1 of the alternative quality contract. N Engl J Med 365: 909-918.
26. Song Z, Safran DG, Landon BE, Landrum MB, He Y, et al. (2012) The 'alternative quality contract' based on a global budget, lowered medical spending and improved quality. Health Affairs 8: 1885-1894.
27. HRSA (2013) Human resources and services administration bureau of health professions. Projecting the supply and demand for primary care practitioners trough 2020.
28. Ham C (2013) Payment systems and incentives to support integrated care. The King's Fund.
29. James R (2011) Integrated care delivery models: Insights from USA and Europe. Seminar at the National Research University High School of Economics, Moscow.
30. Davis C (1988) The organization and performance of the contemporary Soviet health system. In: Lapidus G, Swanson G (Eds.) State and welfare, USA/USSR: Contemporary policy and practice. Institute of International Studies, University of California, Berkeley.
31. Davis C (2010) Understanding the legacy: health financing systems in the USSR and central and eastern Europe prior to transition. In: Reforms and challenges implementing health financing reforms – lessons from and for countries in transition. Open University Press, Buckingham.
32. Crosson F (2009) 21st-Century health care – the case for integrated delivery systems. New England Journal of Medicine, pp: 1324-1325.