DELAYING THE IMPLEMENTATION OF PAYMENT BY RESULTS IN MENTAL HEALTH: THE FAILURE OF QUASI-MARKET

Ruoxi Wang
PhD candidate,
School of Sociology and Social Policy
University of Nottingham, UK
Email: clairewong1128@gmail.com

Abstract

Purposes: To investigate the policy-making process of introducing a prospective payment method called ‘Payment by Results in mental health’ in England and to identify the factors that caused the delay in implementation.

This study carried out a three-stage analysis using mixed-methods. Firstly, the feasibility of applying Quasi-market theory in mental health care at the fundamental level was theoretically analysed. Then, the validity of the construction of the classification system and the cost calculation mechanism at the mechanical level was theoretically evaluated. And lastly, 12 semi-structured interviews with actors from different interest groups (commissioners, hospital managers and frontline clinicians) was conducted to investigate the practical obstacles that hindered the implementation; and conducted 51 online surveys to testify the corresponding findings at the practical level. Results from the empirical data were triangulated with the literature.

The following barriers to implementation were identified: 1) Mismatches between the Quasi-market theory and mental health services; 2) Complex nature of mental disorder and inaccurate data from the classification system; and 3) Gaps between political intent and frontline including (a) inappropriately set targets (b) increasing workload (c) lack of sufficient training and (d) ‘gaming’ behaviour.

Implementing PbR policy in mental health failed to serve the purposes of improving efficiency and quality, given the fundamental problems of the commodification of mental health services, the constructional flaws in defining the mental disorder and the corresponding needs for care, along with the practical difficulties resulted from the gaps between the political intent and the frontline reality.

Key words: Payment by Results, Implementation, Delay, Commodification.

1. Introduction

Besides providing patient-centred services, improving service efficiency has attracted more attention, nowadays, given the universal common phenomenon that the healthcare industry is suffering from an intense financial pressure (Barton, 2003). Since the birth of the UK National Health Service (NHS), the English healthcare system has been funded by taxation in which the patients have same access to health care services with no need to pay for the services they receive directly (NHS Choices, 2013). However, the universal availability, in
turn, resulted in the demand side ‘moral hazard’ (Barros et al., 2008), given the empirical evidence that indicates the increase in demand or even the excessive consumption of healthcare resources (Blunt, 2014). Consequently, the English healthcare system has been under great financial pressure to improve its efficiency, although it has been rated as the most efficient system by the Commonwealth Fund (NHS Confederation, 2014).

A prospective payment method called Payment by Results (PbR) represents a new way to replace the conventional ‘cost-based’ reimbursement by relating the type of patients treated by the providers to the costs incurred by the providers and, therefore, collectively managing healthcare costs. Underpinned by the case-mix principle, it groups patients into categories according to their diagnoses and needs for care, in which patients are expected to consume the same level of hospital resources (Fetter & Freeman, 1986). The average costs for each ‘group’ are, therefore, determined as the basis for this prospective payment method. In other words, this prospective payment system pays providers a predetermined, set rate based on patient’s needs in an evidence-based way (Mayes, 2007), which is believed as able to retain cost, increase efficiency and promote quality through provider-side, non-price competition (Miraldo et al., 2006).

The casemix-based classification system and accordingly, the cost calculation system, were first developed for surgical services for which the groups with clinical meaning and economically homogeneity are easy to derive. However, due to the complex nature of mental disorder, only some countries have implemented or planned to implement casemix-based payment systems in psychiatry, among which England is in the process of extending the coverage to mental health services (Kobel et al., 2011). Despite the initial plan for it to come into effect by 2013, at the time of writing (October 2015), mental health services are still contracted under the Block Contract whereas the PbR has only been applied as a classification system with care pathways still under development. Also, the terms “dangerous” and “unintended outcomes” have been employed to describe the rush to implement PbR in this field (Lintern, 2013).

Under such circumstances, this article aims to identify the driving factors that caused the delay in implementing PbR in mental health services in England through a close look at how and to what extent political environment impacted the implementation of a particular project in practice. The main thrust of this article is on the commodification of mental health services, an area filled with complex individualised conditions that devaluate the casemix-based classification. Occupying the ‘administrative anthropology’ theory, the study carried out historical documentary analysis, semi-structured interview and an online survey to evaluate the policy making process including its initiation, development and implementation. This article, therefore, presents the findings from three levels including the fundamental theory underpinning the PbR design, the construction of the classification system and the practical obstacles that hindered the implementation process.

### 1.1 Initiation of PbR: Within the Context of Saving Costs

With the intent to meet the large demand of service and to continue providing patients quality healthcare services, the government introduced PbR in 2003 in acute services to replace the previous Block Contract. Thus, establishing a new financial framework to pay different providers on a fair and transparent basis (Department of Health, 2002). The rationale for designing PbR was to retain cost, increase efficiency and retain quality through provider-side, non-price competition in a simulated ‘market’ where commissioners purchase services amongst different providers (Miraldo et al., 2006).

The concept of PbR consists of two main components, namely, a classification system in which certain cases are categorised into specific groups that are similar in intervention design and resource consumption and a set of national fixed prices (called ‘tariff’) for
healthcare services subject to each group (O’Connor and Neumann, 2006). Underpinned by the purchaser/provider split, commissioners purchase healthcare services from a largerange of providers including public, private or voluntary sectors where the patients received treatments on national fixed tariffs. Thus, this creates a linear relationship between the amount of activity providers take and the amount of the income they receive (Farrar et al., 2009). Figure 1 shows an example of how the fixed tariff affects providers’ behaviours according to different conditions for a certain type of service.

Figure 1 Relationship between the marginal costs and the PbR price

The extent to which this financial arrangement influences a provider’s behaviour depends on the relationship between their marginal costs and the fixed national tariff. MC_A and MC_B represent the marginal costs for healthcare resource consumption of two different providers. Under the Block Contract, providers and commissioners could negotiate the price and the volume of a certain type of service locally. Thus, Provider A could provide service activity at x_A^0 amount and be paid at the price of p_A whilst Provider B could provide x_B^0 amount of activity at the price of p_B (when marginal cost equals marginal price). The introduction of national fixed price p_0 calculated from the national average unit costs under PbR means the provider could not gain extra profit until its marginal cost curve intersects the flat marginal revenue curve (MR_{PbR}) under PbR. For Provider A, the relationship between MC_A and MR_{PbR} suggests that its profit reaches to the highest at the amount of x_A^1. In this case, the financial incentive encourages providers like Provider A to increase their capacity and as such, improve their efficiency. In contrast, Provider B has a marginal curve depicted as MC_B, with higher marginal costs than MR_{PbR} in every level of activity. In other words, it costs more than this provider would receive in every case, thus, it has no choice but to reduce unnecessary costs to attain financial balance. In other words, the intent to avoid financial risks also stimulate the cost saving and efficiency enhancement of providers such as provider B (Mannion et al., 2008). Generally speaking, the standardised national tariff is believed to reduce the variation in healthcare cost and, therefore, the total expenditure. In addition, the national fixed tariff rules out price-competition. Consequently, providers can only compete on service quality to attract more patients. This, in turn, facilitates the improvement or at least the maintenance of service quality. On the other hand, the regulatory agencies exercise the authority, on behalf of the government, to guarantee the quality of service.

As Geissler et al. (2011) illustrate, the casemix-based systems have been widely employed in acute services by European countries for two major reasons. The First being the transparency of clinical services brought by the standardised classification system and secondly, the increase in efficiency led by this cost-and-volume payment mechanism. Unlike the acute services that are underpinned by a comparatively clear diagnosis system, the evaluation of its implementation in mental health services depends on the understanding of not only the
payment system itself, but also the nature of mental health services in order to facilitate evaluating the incorporation of the two. Therefore, this study evaluates PbR scheme from a top-down order: evaluate the theoretical feasibility of applying market theory in the healthcare system by looking at the fundamental basis of PbR initiation - Quasi-market; investigate the construction of the classification system and its collaboration with the sub-systems at the mechanical level and explore the practical obstacles that hindered the implementation process.

2. Research Methodology

This study employed the mixed methods (triangulation) to understand the various factors and how these factors affected PbR’s implementation in mental health services in a top-down and bottom-up order as is showed in Figure 2. The combination of theoretical analysis, semi-structured interview and online surveying enabled a multi-level and multi-angle understanding of this issue: the theoretical analysis looked at the fundamental and mechanical validity of PbR in mental health services whereas the interviews and online surveys examined its implementation in practice. For the fieldwork, the objective was to gather material that shows how the PbR scheme designed and developed and, therefore, to identify the actual outputs of the policy. The semi-structured interviews and the online surveys intended to discover frontline staff’s attitudes towards the implementation of this policy, with a particular attention paid to the gaps between the government’s priorities and the service delivery at the frontline. The fieldwork findings presented different or even conflicting perspectives between different interest groups, which increased the comprehensiveness of understanding.

Figure 2 Flowchart of research design

Qualitative analysis was mainly employed in the semi-structured interviews. Based on the top-down theoretical analysis that evaluated the fundamental and mechanical validity of PbR in mental health services, the semi-structured interviews aimed to provide up-to-date, on-the-ground information regarding PbR’s implementation. Therefore, between November 2013 and April 2014, the semi-structured interviews involved 12 participants from three interest groups including commissioners, hospital managers and frontline clinicians within Nottinghamshire. As the main method of this study, the qualitative analysis examined how
and to what extent the theoretical flaws affected everyday clinical practice and outlined the external factors that hindered the implementation progress at the current stage.

To avoid the risk of being partial (Mays and Pope, 1995), this study adopted online survey for confirmatory purpose. After gathering and analysing the qualitative data, online surveys were sent to three organisations within Nottinghamshire and one Foundation Trust in Derbyshire (N=51) between June 2014 and September 2014. Quantitative analysis was then carried out to confirm and testify the information drawn from the qualitative analysis. IBM SPSS 20 was used to analyse the quantitative data. Approaches such as descriptive statistics, cross-tabulations, chi-square tests and Mann-Whitney test were applied to conduct quantitative analysis. P<0.05 was adopted to detect the statistical significance.

3. Main findings

Figure 3 shows an example of how PbR comes to effect. Generally, it takes three major steps for an idea to translate into political outputs. Regarding the whole process, it starts with the policy makers’ general intent; then it transfers into a sequence of policies/targets that specify the initial intention; and in the last stage, the frontline staff exercise following the specific targets through which the original intention is therefore translated into everyday practice. As Shaw (1994) demonstrates, the actual collaboration between actors is not as smooth as expected in theory. This is underpinned by Lipsky’s (1980) ‘street level bureaucracy’ theory, which points out the difficulties in standardising and therefore, predicting frontline level practice. According to Lipsky (1980), in a healthcare system, patients’ individualised conditions with individualised needs for care affect the working context. Given frontline clinicians’ irreducible responsibility for providing appropriate responses to the client’s situation, which can hardly follow the authoritative agency guidelines strictly, it leads to the difficulties in rationalising or simplifying the frontline service delivery process. This points out the limitation of a forward mapping approach to implementation analysis, given the underpinning assumption that policy makers control the implementation process (Shaw, 1994). As such, it indicates the importance of backward mapping approach, which appreciates the importance of frontline level participants in the service delivery process. They are the key participants who transfer a policy into real outputs by adapting intentions into workable forms. Also, understanding their situations and perspectives will facilitate a critical judge of the outcome of the process. In this case, it is worth investigating the actors and variables at each stage of the policy implementation process. Also, according to Lipsky’s (19991) theory, a comprehensive understanding of a specific policy should consider political context, policy makers’ allocation decisions and the frontline practice. This conforms the Glennerster et al.’s (1983) theory of ‘administrative anthropology’, which includes historical documentary analysis, structured interviews and observation of practice. Occupying the ‘administrative anthropology’ theory, this study aims to have a comprehensive understanding of the policy making process of the PbR scheme by looking at the issues at the fundamental level, the mechanical level and the practical level, which will be respectively illustrated in the subsequent subsections.

Figure 3 Analytical framework
### 3.1 Fundamental problem: The failure of the Quasi-market

#### 3.1.1 Quasi-market behind PbR

As was mentioned earlier, the original aim of introducing Payment by Results was to control the cost and gradually lead to quality improvement through non-financial competition between providers. When tracing back the idea of competition, it originated from the idea of the Quasi-market, which is a modified form of the fully competitive market model and has been introduced into the NHS since the 1990s.

Given the features of public services, the government has strong responsibility on monitoring and regulating the system to meet the public interests. On the other hand, in facing the increasing demand, the increasing financial pressures calls for incentives for a higher efficiency. In order to promote cost-efficiency under such a great financial pressure while maintaining government’s regulative responsibility, Le Grand and Bartlett (1993) have developed the concept of the Quasi-market. As Kahkonen (2004) interprets, Quasi-market is a market since it aims to reap the supposed efficiency gains through replacing the monopolistic public provider and thus, embracing supply-side competition; while Quasi-market is ‘quasi’ because it differs from the conventional free market given the abandon on financial competition, as well as, the government’s regulatory power. It, therefore, serves two purposes: on the one hand, the competitive market mechanism brings competition to control cost and drive up quality for a better allocation of healthcare resources as was mentioned earlier. On the other hand, through agencies who exercise authorities on behalf of the government, the government keeps its regulatory responsibility and thus, ensures the common good of public services.

Considering the features of public services, agencies such as National institute for Clinical Excellence (NICE), Care Quality Commission and the Monitor have been established to exercise the authority on behalf of the government to regulate the service provision. National institute for Clinical Excellence (NICE) has been established to develop national service framework in order to reduce variations in performance (Gray and Harrison, 2004). Key regulatory functions are carried out by agencies such as Care Quality Commission and Monitor (Lewis et al., 2009). Acting as an economic regulator, the Monitor’s is entitled to regulate the price paid to providers and to apply competition law against anti-competitive behaviours to ensure the continuity of high-qualified service. Care Quality Commission takes responsibility to license the providers and inspect the service (Department of Health, 2010). The PbR development team is developing care pathways to standardise the frontline practices subject to each disease group. In this case, the government’s inspection and regulation are expected to be an assurance that patients from different parts of the country receive the same quality service.

### 4. Fundamental drawbacks of the Quasi-market

As a combination of market and hierarchical model, the success of the Quasi-market theory restson the preconditions of both the competitive market and the government’s power to control.

1) **Drawbacks of the market theory:**

The following four parts test four basic assumptions that enable competition principle to serve its function, including homogenous product, supply-side competition, consumers’ rational choice and zero externality. The conflicts between the assumptions and the features of health care are outlined below:
Homogeneity of product

The homogeneity of output is the basic assumption for the fully competitive market: only when outputs in the industry are identical, consumers make the choice between different providers simply based on price and thus encourage the supply-side competition. Although providers are not competing on prices in the quasi-market, the idea behind the price settlement rests on the hypothesis that services can be standardised to be identical and the adoption of average value of the costs standardises the costs. This is a process to commoditise the healthcare services. However, in the healthcare system, providers do not provide healthcare products, rather, they offer services. One prominent difference between ‘service’ and ‘product’ is that products could be produced according to certain standardised processes while services are difficult to be standardised given the individual recipients (Pratt et al., 2007). In order to distinguish services from products, Normann (2001) uses the word ‘offering’ to describe the process of which service is delivered. Offering is a process that is optimised in terms of relevant actors, rather than a physical object. Services, such as healthcare, require some collaboration between consumers and providers, such as both psychotherapists and patients work together to reveal and make use of their personal experiences for treatment. It is particularly the case in mental health services since the complicated nature of mental disorder determines key point of the treatment as knowing patients and their needs (Jones, 2004). The process involves building up relationship, trust and intimacy, which requires clinicians to interact with patients and be flexible to the specific individual patient (Jones, 2004, Perry, 2000). In other words, such kind of service is highly individualised and cannot be completely substituted by others.

Supply-side competition

Another precondition of supply-side competition is that there should be a large amount of providers evolving in the market whereas the exit of one supplier would not affect the price or the operation of market. If there are too few providers, the effects of market mechanism would be hindered by oligarch (Struyven and Steurs, 2005). However, in healthcare, due to its own nature, it is not possible to break down monopolies for a number of reasons (Lowery, 1998). Hospitals in the UK are founded on an expectation to provide a comprehensive range of services to the population within the district (Ministry of Health, 1962). Therefore, these organisations would have some extent of monopolistic power (Propper et al., 1998). However, unlike other industries in which monopoly and oligarchy are of more potential damage than the advantages they can bring, it is of great importance to have some extent of monopoly in healthcare services. According to the nature of medical treatment, one specialty cannot be delivered without the support from another (West, 1998). If a provider is to offer a comprehensive range of service, it needs a wide range of specialists to maintain a normal working order. Accordingly, it needs to have the capacity to serve a wide area in order to have a sufficient caseload to keep these specialised teams working. In this case, the size of the supplier and the range of service it provides are important to ensure the comprehensive service, which explains the necessity of developing one large general supplier rather than many small providers. Purchasing from distant providers is also seen as unattractive to both consumers and purchasers, due to the problems brought by potential transport and communication, which further reduces the intensity of provider-side competition. The lack of large amount of providers, as such, influences the contracting relationship between commissioners and providers. This causes commissioners to have to develop a dependent relationship with the existing providers given the limited number of alternatives (Petsoulas et al., 2011).
Consumers’ choices

Given the consumer/purchaser split, consumers who receive healthcare services are not the ones who pay for them. Therefore, who chooses providers and how to make rational choice affect the degree of competition in the Quasi-market. As such, it is who makes the choice that matters. The fundamental social problem is seen as the finite resources versus the infinite human demands (Bradby, 2012). Maximising the utility of resources is commissioners’ goal when considering this issue at a general level, while receiving the best individualised service is the patients’ goal when it comes to the individual level. Services that meet patients' satisfaction, may not be most effective or even necessary. Agency theory indicates that it is very difficult to align these two mutually conflicting goals (Allen, 2002). Studies undertaken by Fotaki (1999) and Dixon et al. (2011) support the idea that agents acting on behalf of patients have not been effective. Some trade-offs between efficiency and patients' preference have been done with the establishment of policies, such as encouraging patient choice (West, 1998).

Additionally, whether patients are able to make rational choices, if they are willing to, or whether commissioners are able make the right choice for patients, this is crucial to the market mechanism. Mental health is a comparatively unexplored area with complexity and high asset specificity where lacks gold standards for diagnosis and treatment. Even highly specialised professionals sometime feel difficulties in providing the most appropriate treatment to patients, not to mention patients’ understanding of what is best or efficient treatment for themselves, which in turn leads to their lack of ability to make rational choices. Commissioners, as the agents for patients, also suffer from the information asymmetry and the ability to rigorously assess such highly professionalised services (Lewis et al., 2009).

Therefore, neither consumers nor purchasers are regarded as able to obtain sufficient information and thus, make rational choice based on efficiency/quality as the economic theory predicts. Moreover, some researchers even doubt about the demand-side’s freedom to choose services between providers, due to externalities such as political considerations.

Externalities

According to the competition mechanism brought by the Quasi-market, providers with higher costs than national fixed prices will suffer from deficit, which may lead to job loss for the staff in those organisations. The entry of private providers is believed to intensify the supply-side competition, which may lead to a less stable situation. However, according to West (1998), there exists dilemma between pursuing cost-efficiency and considering national political pressure on public services. Department of Health is under the public pressure to protect jobs of the existing public organisations – the NHS hospitals when establishing any policy for reforms. Since health care remains a key general election issue, and the proposers who suppose employing market mechanism to promote competition and reduce costs will be easily labelled as ‘privatising’ or ‘dismantling’ the NHS (Greener, 2002). Under such external political pressure, substantial reforms such as ‘bringing competition’ and ‘reducing costs/budget’ are difficult to implement.

2) Government’s failure to regulate

Besides the market mechanism, the government’s regulation is the core element designed to be an assurance of service quality. The subsequent subsections investigate the appropriateness and the government’s ability to regulate service quality.
Appropriateness of service standardisation

One prominent feature that distinguishes healthcare services from other production industries is the output, which requires frontline staff to offer a continuous interaction with clients and behave according to the specific contexts (Lipsky, 1980). In health care, particularly mental health services, treating individuals according to the predetermined standards but neglecting the existing differences such as age, sex and income level is seen as unfair. Due to the personalised treatment in mental health, different therapeutic approaches and other types of treatment such as CBT and counselling can be offered to patients with similar needs (Cheshire and Pilgrim, 2004). It may be easier to standardise drug protocols and some biological approaches such as ECT but it is difficult to identify a standardised procedure for developing a relationship with patients in treatments such as CBT. Furthermore, clear evidence of efficacy amongst a range of psychiatric treatments is hard to find, and so it is common for clinicians to provide psychiatric care according to their own value systems, which may be inconsistent with their fellow members (Jones, 2004). It is these variations that disable the frontline performances from fitting into a metric of correct responses and therefore, the authoritative agency guidelines. As Pratt et al. (2007) point out, patients also want to be treated as an individual rather than a statistic, number or case. Receiving individualised treatment is patients’ demand and therefore the core code for professionals. Therefore, the centrality of co-production in healthcare services goes against the theory of care pathway, which is to standardise clinical treatments in order to provide standardised services to each patient according to their disease categories.

Government’s ability to regulate

According to Iles (2011), one is able to access information that is known or knowable by the public through searching the internet, whereas in complex situations that require professional knowledge then one needs to seek help from people with deeply held knowledge and expertise. It indicates the inevitable information asymmetry in healthcare services between the professionals and the lay people. While key regulatory functions are carried out by agencies such as Care Quality Commission and Monitor (Lewis et al., 2009). In such circumstances, these non-clinical managers’ abilities to accurately measure and regulate service quality have been questioned, which questions the external investigators’ ability to evaluate healthcare quality.

Moreover, the professionalism indicates the difficulties in specifying and quantifying every move in the individualised treatment. Therefore, only activities that can be measured are enclosed into measurement system. It is these measurable ‘facts’ and activities that form the ‘performance’ (Iles, 2011), which indicates that targets relating to measurable aspects do not encompass the whole picture of quality. The other risk brought by quality measurement is the misleading of performance. When performance is measured against objectives, the objectives are required to be specified in advance. Providers could therefore just focus solely on the targets and achieve the targets in an easier and quicker way. The outcome-oriented system may easily lead to ‘gaming’ behaviours (will be discussed later). This shows agreement with Lipsky’s ‘street-level bureaucracy’ theory (1980) indicating that the actual performance of frontline activities is virtually impossible to measure.

Professional hierarchy

In health care, such a highly professionalised area, the dominant power does not belong to the government, but to the professional bodies such as medical and surgical Royal Colleges (West, 1998). Issues, such as hospital planning at the macro level and department establishment at the micro level, are managed by the professional bodies. It is this level of professionalism that gives great autonomy power to the medical professionals. Especially the
confidential relationship between doctors and patients brings unavoidable information asymmetry between the scrutinised and the observed. Thus, this inversely influences the accuracy of performance measurement (Greener, 2002). Also, because of the information asymmetry, professionals’ actions in diagnosis, treatment, and referral become the policy of the system (Lipsky, 1980). As Exworthy et al. (1999) conclude, on the one hand the NHS is centrally financed, since the politicians set the size of the budget, whereas on the other hand it is run by professionals who also decide how to spend the budget.

According to the initial problems of market mechanism and government’s regulation discussed above, Quasi-market fails to serve its functions as theoretically predicted, given the gaps between the preconditions and the features of mental health services. This shows agreement with Kahkonen (2004) who points out that the Quasi-market is never a solution. The weak foundation in turn increases the difficulties in applying this system to managing mental health services, particularly given the constructional problems of the classification system as is discussed below.

4.1 Mechanical problems of PbR: Failure to standardise services

As was mentioned earlier, the basis for provider-side competition is the homogeneity of products. This highlights the central importance of being able to distinguish illness type and the corresponding services, a process of standardisation whereby best evidence based practice is applied to a defined set of conditions. This section illustrates the mechanism of the classification system and the relevant drawbacks.

4.2 Mental health classification system: MHCT

Figure 4 shows the theoretical working mechanism of PbR in mental health services. The clinical classification system called Mental Health Clustering Tool (MHCT) classifies patients using a decision tree. Patients are allocated into one super-cluster amongst ‘non-psychotic’, ‘psychotic’ and ‘organic’ based on its origin of disorder. Furthermore, it is narrowed down to one of seven secondary clusters attached to these super-clusters. Then, patients are rated by the rating scale called the Health of the Nation Outcome Scale (HoNOS), as shown in Figure 5. The rating scale incorporates the colour-coded rules according to which different colours indicate how likely the certain symptom is expected to be categorised as the specific score in the specific cluster. A series of standardised care pathways are attached to the corresponding clusters to guide service provision and therefore, the price settlement (Cabana et al., 1999).

Figure 4 Flowchart of disease classification and cost calculation
Acting as the basis of the payment system, the validity of care pathways is crucial for the feasibility of a cost calculation. Unlike the MHCT and HoNOS that have already been applied in everyday clinical practice, the care pathway system as a separated element is still under development. The subsequent subsections discuss the early findings of the initial problems of the classification system.

### 4.2.1 Complex nature of mental disorders

Unlike the acute services, mental disorders have been widely acknowledged as a comparatively underexplored area, due to the natural characteristic of mental illness – intangible pathology and vague aetiology, particularly since mental disorders cannot be objectively defined or diagnosed through any kind of laboratory test that is in daily use in acute services (Frances, 2010). Even in organic conditions of the neurology, in which the damages could be found on the clinical image, there exists some cases like dementia for which the brain damage is not always detectable in post-mortem (Rogers and Pilgrim, 2010).

Therefore, the lack of clear markers of diagnosing mental disorders leads to more reliance on the professional judgment on one’s mental condition. The large variations in mental disorder between individual patients are widely recognised and these variations make classification process difficult. Moreover, besides the homogeneity of patients’ needs for care, the success of standardising frontline practitioners rests on two other preconditions. On one hand, the medical skills of the frontline clinicians are at the same across the country and on the other hand, the classification system MHCT is able to standardise clinical practice by defining patients' needs for care and designing the corresponding care pathways. The lack of medical or
clinical tests makes classification difficult and leads to large variations in judgement between different clinicians even when facing the same patient (Houts, 2001). Moreover, the devolution of power, since the NHS Liberalisation, indicates an increase in geographic variation in service provision and therefore the implementation progress. According to the Mental health bulletin: Fifth report Mental Health Minimum Dataset (MHMDS) annual returns (The NHS Information Centre and Mental Health and Community Team, 2011), the provision and use of mental health services vary across England. This is also underpinned by the findings from both the interviews and the online surveys, in which participants noted the large variations in the professional skill between frontline clinicians. These variations, therefore, reduce the validity of the classification system, which has already been compromised by the ideological problems as discussed below.

4.2.2 Ideological problems of MHCT

MHCT works as a hub gathering different information and distributing patients into 21 clusters according to their severity of symptom. The re-clustering action that takes place at the end of an episode reminds clinicians to reassess the patient’s condition and to adjust interventions if necessary. It is indeed seen as helpful from an administrative perspective since this symptom-based classification system enables interventions to be provided without being firstly informed by a detailed diagnosis. Nevertheless, frontline clinicians have built up their clinical treatment system based on diagnosis, which encourages clinicians to think about the dimensions including the social dimensions understand the individual patient. More importantly, from the perspective of patient-centred care, treatment is a process of seeking the root problems of the patient, which is the cause of disease, and solving problems occurring in the trial-and-error process, which should be an integrate evolution. In other words, it makes more sense to provide integrated treatments according to the development of the patient’s condition, rather than separating a coherent treatment process into several stages. Moreover, the vague distinctions between the neighbouring clusters make this ideology more questionable, as is discussed in detail below.

4.2.3 Constructional problems of the MHCT

Without sufficient support from diagnosis or care pathways, the MHCT has been criticised for its blurring boundaries that lead to low within-group homogeneity and low between-group heterogeneity. Regarding within-group homogeneity, the vague definition of ‘stable’ raises difficulties in categorising patients into Cluster 11, where some patients stay in a high level of hallucination or other variances in mood for a long time, whereas the variations for them are normal and stable. In the interviews, the frontline clinicians questioned whether they were in the same ‘stable’ condition as expected in Cluster 11. Regarding the between-group heterogeneity, co-morbidities make cases difficult to fit precisely into a specific cluster. As is revealed from the interviews, facing a patient with two possible diagnoses as bipolar disorder and organic dementia, some frontline clinicians perceived difficulties in making decision between Cluster 12 for bipolar and Cluster 16 for organic dementia. Additionally, some grey areas have been noted between the neighbouring clusters such as Cluster 3 and Cluster 4. This may be attributed to the low validity of the HoNOS scales, which will be discussed in the following part. In particular, no prior consideration of potential services, the classification system has been criticised as a ‘labelling process’ (Callard et al., 2013, Middleton, 2013) through which patients are allocate to clusters where no one fits well.

4.3.4 Initial problems of the HoNOS

The interview findings revealed the sequence of PbR development that could partly explain the discordance between the HoNOS and the MHCT: the HoNOS was first introduced to test general health and social functioning, then it became HoNOSPbR, in preparing for the new
payment system, and the MHCT came last, as the classification system. In other words, the HoNOS was not designed for the classification system and it would inevitably generate some discordance.

A reliable classification system should be subject to a wider spectrum of symptoms in mental health and it is important to properly and accurately situate the cut-off points (Patel et al., 2014). However, mental health care is such a complicated area in which even the categorical approach with simple ‘cut-off’ groups is under the attack of blurring boundaries (Maser and Patterson, 2002). As Aboraya (2012) argues, most of the time, the boundaries between minimal and mild symptoms are not of clinical significance to result in a different clinical decision. At the moment, there lack general agreements on the number of dimensions, as well as, empirical studies on evaluating the validity and credibility of this system (Busko, 2007). Consequently, the lack of universal standards on dimensions requires more reliance upon clinicians’ ability, whereas there is a large variation in professional skill between frontline clinicians in this field, which in turn compromises the between-group heterogeneity.

Moreover, the HoNOS scales only test general health and social functioning rather than measuring the specific health care outcomes or clinical effectiveness (Harrison et al., 2004). Other specific scales such as Hamilton Rating Scale for depression work more effectively than the more general HoNOS (Wang et al., 2015). More condition-specific outcome data are necessary if patients’ needs, severity and responses to treatment are to be assessed with sufficient accuracy. In other words, solely relying on HoNOS as the standardised outcome assessment attached to MHCT clusters is not a sufficiently reliable approach.

### 4.3.5 Drawbacks of the ‘standardisation-to-the-average’ principle in price settlement

Even if the MHCT classification system is able to accurately predict resource consumption, there is still work to be done before tariffs can be established and used as the basis of remunerating NHS provider organisations. Given the national fixed prices that are majorly set based on the national average unit costs, it assumes that the costs of treating people in the same cluster follow a roughly normal distribution although incurred by different providers and the ‘deviances’ can be reduced by using average costs to standardise the costs (Department of Health, 2012). However, the study by the Price Waterhouse Coopers (2012) points out that even for the same service, the unit costs reported by providers largely differ from each other, which rejects the assumption of ‘standardisation-to-the-average’ from a statistical point of view (Wang et al., 2015).

It has also been argued that a consequence of the ‘standardisation-to-the-average’ principle is that it encourages providers to become ‘average’ rather than improving their performances (Llewellyn and Northcott, 2005). Consequently, providers are required to conduct practice in standardised pathways and receive payment at national average level, unless they can provide persuasive evidences for a new service and its impact on cost reduction, capacity enhancement or quality improvement. However, the process of proving effectiveness of a new product is often difficult, especially in this inflexible system (Appleby et al., 2012).

The above arguments regarding the fundamental theory and the construction of PbR indicate that PbR lacks theoretical validity, in terms of accurately identifying patients’ needs for care and therefore, paying different providers on a fair and transparent basis. The practical evidence that outlines the gaps between the political intent and frontline realities further confirms the lack of feasibility.
5. Practical problems of PbR: conflicts between political considerations and clinical reality

Besides the above fundamental and mechanical problems, gapshave been found between the political considerations and the clinical realities when coming down to the frontline. As is reflected from both the interviews and the online surveys, the biggest problem is the difficulty in encouraging frontline clinicians to engage in this reform. Besides the constructional problems of the classification system, issues such as not properly set targets, increasing workload and the lack of sufficient training were attributed as the reasons that cause their reluctance to engage.

5.1 Targets not properly set

Regarding the transitional period moving from the Block Contract to PbR, the interviews and online surveys have noted that a series of reform-related policies with various targets have been established simultaneously. Among these policies, on the one hand some targets were seen as confusing whereas, some policies such as ‘discharge all Cluster 11 patients back to the GPs’ and ‘have 4 successful contacts with clients everyday’ were even seen as inappropriately set. According to the interview findings, the potential damage to the service quality was noted as the main side effect brought by this target-driven policy. Also, 47% of the questionnaire participants regarded it as an either ‘important’ or ‘very important’ factor that caused delay in implementing PbR. Moreover, to prevent fraud, one policy required two clinicians to see one patient together, which was not seen with much clinical utility except the financial benefit resulting from double counting the clinical work. It has also been noted that the ‘undoable’ targets triggered the providers to lie about the outcome.

“My understanding is that the commissioners are setting targets to the trusts all over the country. However actually there is no trust which is meeting these targets. ... They (the commissioners) just pick some targets at their way not aware of that they are not achievable. They are not doable.” (Frontline clinician)

5.2 Increasing workload

According to the interview findings, the managers noted the inevitable increase in workload given the effort required to understand the new system and to facilitate quality assessment. However, this extra workload and the over focusing on checking the system raised some managers’ concern about PbR’s validity since this system was supposed to be sophisticated and efficient enough not to require constantly checking for fraud. The experiences from the frontline clinicians echo the managers’ concern: some checklists were long and redundant, which took away their time that could have been used to improve their skills, thus confused them by the dilemma between the idea of increasing efficiency and the policy of increasing paperwork. This is underpinned by the corresponding finding from questionnaires as 75% of the frontline clinicians regarded heavy workload as an either ‘important’ or ‘very important’ factor for the delayed implementation.

5.3 Lack of sufficient training

In the interviews, the managers and the commissioners attributed the inadequate training to one of the reasons that caused frontline clinicians’ lack of understanding of the PbR system. As the online surveys revealed, over 38% participants have never been to any training session. Amongst those who have attended training sessions, most of them (90.92%) have only attended once or twice.

Moreover, the interview participants further questioned the effectiveness of the training sessions. In contrast to the initial intent to facilitate a better understanding of the
system, the training sessions themselves have not turned out to be effective. According to the online surveys, 92.9% of the managers and 68.8% of the frontline clinicians doubted the effectiveness of training sessions. Reasons are partly attributed to the disjointing training sessions given the several years’ gap before the new series of training session started again. One of the consequences was the absence of opportunities to deliver up-to-date information to the frontline in a direct and accurate way: the same message delivered by different managers became different versions when it came to the frontline, which inevitably resulted in more confusion.

“In meetings with the senior managers, they would be given messages that probably go down the chain. But by the time they got to the frontline, they are probably not the accurate messages, or there might be the panicking messages.” (Manager)

5.4 ‘Gaming’

All the above problems raised frontline clinicians’ reluctance to engage, which resulted in their ‘gaming’ behaviour, particularly within a context of information asymmetry. Combining findings from the interviews and the online surveys, the study found that ‘gaming’ exists in everyday frontline practice although not in the way predicted by previous studies according to economic theories as ‘game for money’, rather it is ‘game for workload’.

Shown by the questionnaire figures, patient rated measures were seen as subjective or even misleading, given 32% of the participants who concerned about the validity of the patient rated measurements to reflect the actual service quality. In this case, the quality measurement system functions more like a means that policy makers respond to public dissatisfaction rather than an effective clinical instrument to guide clinical practice. This shows agreement with Lipsky’s (1980) theory that actual performance of frontline activities is virtually impossible to measure. The lack of accurate quality measures together with the information asymmetry created potential for providers to ‘game’ especially given the current austerity.

Frontline clinicians’ lack of willingness to engage and the ‘gaming’ behaviour further compromised the frontline data quality, which negatively affected the frontline data quality, care pathway design and therefore the whole development process of the PbR scheme.

Conclusion

This study was carried out to explore the current stage of implementing PbR in mental health care in England and to investigate the driving factors that caused its delayed implementation. This study has also sought to know whether the fundamental theories behind PbR, which is Quasi-market, had an impact on PbR’s delayed implementation and if so, to what extent it contributed to the delay. Given mental health care as an under-explored domain, together with the incomplete status of PbR development, there has been a lack of sufficient investigation in examining the current problems of implementing PbR in mental health services. Moreover, there lacked compelling evidence of PbR’s impacts on cost, efficiency and quality even in acute services where PbR fits most. It therefore raised concerns about the validity and feasibility of PbR itself, particularly when considering the complex nature of mental disorders.

This study employed mixed methods to present multi-level and multi-angle perspectives on issues surrounding PbR design and implementation. A three-stage analysis was carried out to provide multi-angle perspectives regarding issues surrounding PbR at each stage of the policy making process including its initiation, development and implementation. The theoretical analysis focused on evaluating the fundamental theories behind PbR at the macro level. The conflicts between the preconditions of the market theory and the features of the
mental health care demonstrated the Quasi-market’s failure to serve its function in this field. The constructional problems of the classification system indicated allow validity of the classification system resulting from its failure to define patients’ needs for care, which reflected the mismatches between the political intent and the proxy. The fieldwork looked at the practical problems during the development and implementation, which highlighted the problems incurred when applying the political targets into practice. Through examining the three stage of policy making process, this study hereby derived a perspective that the fundamental problems of applying the Quasi-market theory, the constructional problems of the clinical classification system together with the practical obstacles hindered the implementation of PbR in mental health domain, which indicated PbR’s failure to serve the political intents regarding saving costs, increasing efficiency and improving quality.
REFERENCE


