Introduction

On April 1, 1983, the first patient classification system (PSC) to be used for paying hospitals for the services they provided was adopted by the US Congress. For the first time, a payer – in this case Medicare – had a way of comparing the outputs of one hospital with those of another and a basis for paying hospitals in a standardized fashion for the "products" they produced.

This system, known as Diagnosis-Related Groups, (DRGs) was developed by a team of researchers at Yale University under the direction of Robert Fetter and John Thompson and sparked a revolution in the health care sector in the United States. At a moment in time when there was increasing concern in Congress and elsewhere about the rapid rise of costs in health care, hospitals could no longer justify higher costs simply by asserting their patients were sicker than anyone else's. By classifying patients according to the resource consumption patterns that were typically associated with particular diagnoses, the DRG case-based system promised to introduce both transparency and operational efficiency into a production process that had previously been largely opaque.

The US, however, was not the only country struggling with increasing costs in health care in the 1980s. A number of other countries, particularly in Western Europe, were experiencing similar increases and were in the hunt for solutions. When Congress adopted the Prospective Payment System with DRGs as the underlying patient classification system, other countries took notice, and soon a number of them began experimenting with various kinds of PCSs, most of which were modelled, directly or indirectly, on the DRG system. France, the UK, Portugal, and Belgium, among others, were "fast followers" and began to explore the possible use of PCSs in their own health systems.

In 1993, John Kimberly and Gérard de Pouvourville published a book titled *The Migration of Managerial Innovation: Diagnosis-Related Groups and Health Care Administration in Western Europe*. The book detailed the

experiences of nine different countries in Western Europe with DRGs through 1991, including that of Germany, a country that at the time had decided not to go down that path. Viewing the spread of DRGs through the lens of innovation diffusion, the authors concluded that in addition to its potential to help control costs, a principal reason for the adoption of DRGs was the system's flexibility, flexibility that allowed it to be adapted to fit a variety of national priorities and policy contexts.

By 2005, fourteen years later, a number of other countries in other parts of the world had begun to use patient classification systems. What had their experiences been? How similar to or different from the original nine were they? What had happened in some of those original nine countries in the meanwhile? And what might be learned about introducing change into national health systems by comparing their experiences? To try to answer these questions, Kimberly and Pouvourville enlisted the help of Tom D'Aunno at INSEAD, and the three of us resolved to identify knowledgeable individuals in a number of countries who could write chapters for their country around a common set of themes. We turned to Jean Marie Rodrigues of France and Céu Mateus from Portugal, both of whom have been centrally involved in an organization called Patient Classification Systems International (PCS-I), for suggestions. After much discussion, we identified individuals in France, Belgium, the UK, Sweden, Switzerland, Portugal, Denmark, and Germany, eight of the nine countries in the original book, and in Hungary, Italy, Australia, Singapore, Japan, and Canada, new additions, to write chapters. We have also included a chapter on the US, feeling that it would be useful for the DRG story to be told as context for the rest of the accounts in the book.

An authors' meeting was held at INSEAD in December of 2004 to orient each chapter author to the principal themes in the book and to permit some sharing of experiences. We asked each author to provide a brief overview of the health system in their country to provide the context into which patient classification systems were being introduced and then to address a common set of questions: when was PCS introduced, what motivated its introduction, who were the key actors in its introduction, how did the implementation process unfold, and, finally, what has the impact been and what debates and controversies have emerged around its introduction and implementation?

First drafts were produced in 2005, and a second authors' meeting was held, again at INSEAD, in December of 2005 to discuss, review and identify strengths and weaknesses in each chapter that had been written up to that

point. Final drafts and updates were produced in 2006, as was the concluding overview chapter.

Audience

The result is a book that provides rich descriptions of the fate of PCSs in each of the fifteen countries represented as well as an analysis of the commonalities and differences among them. As such, the book is intended to appeal to three principal audiences: *health policy makers and managers* concerned with designing and implementing new initiatives that will have broad impact and will engage many different sets of actors in complex and often highly contentious ways; *students*, both undergraduate and graduate, taking courses in health administration, comparative health systems, and/or the management of change; and *researchers* working on problems of innovation and change in general or, more specifically, on international and comparative health policy and management.

In an era when health reform is high on the political agendas of most countries around the globe, the book provides an overview of how health care is organized and financed in fifteen of these countries and how patient classification systems are being used in these efforts. In so doing, it illustrates a range of alternative solutions, and is thus likely to be of interest to those concerned with the problem of health reform in general as well. The story of health reform, of course, continues to unfold, and Patient Classification Systems and their continued evolution are only one piece of a much larger puzzle.

1 Origins of DRGs in the United States: A technical, political and cultural story

Jon Chilingerian

Introduction

In 1983, DRGs became the price-setting system for the Medicare program in the United States. Why did the United States choose DRGs?[1] The idea of setting 518 diagnostic payment rates for 4,800 hospitals seemed unimaginably complicated, too technical and an exercise in formula-driven cost control to some observers – an ambitious endeavor unlikely to succeed.[2] Nevertheless, since its inception, the DRG system has been called the single most significant post-war innovation in medical financing in the history of the United States (Mayes 2006), and may be the most influential health care management research project ever developed. As the chapters in this volume attest, worldwide adoption of DRGs followed in the wake of this American experiment.

Other competing patient classification systems could have been selected (Pettingill and Vertrees 1982). The range of policy options included flat rates per discharge, capitation, expenditure caps, negotiated rates, and competitive bidding (Smith 1992). Although researchers continue to experiment with alternative patient classification systems, a critical mass has formed around DRGs as the *dominant design* for measuring a hospital's casemix. A dominant policy design not only obtains legitimacy from the relevant community, future innovations must adhere to its basic features (Utterback 1996). A dominant design does not have to outperform other innovations; it merely has to balance the stakeholder interests.

Though the control of rising health costs is a major policy issue, American hospitals had come to expect "pass-throughs, bail-outs, and hold-harmless clauses" from the political system (Smith 1992, p 44). The implementation of DRGs, however, challenged this assumption and created

a financial incentive for hospitals to manage Medicare patients more efficiently or lose money, contradicting the fundamental interests of hospital providers and professional monopolies. After Medicare shifted from retrospective reimbursement to prospective payment under DRGs, nine years later in 1992, Congress passed an equivalent prospective payment system for physicians. In 1997, outpatient services, skilled nursing care, long-term care, home health, and rehabilitation services also went to prospective payment. Hence, the implementation of DRGs stimulated a massive transformation of payment and financing for health care in the United States.

Transformational innovations like DRGs don't just happen; the inertia that characterizes health care organizations makes them remarkably resistant to policy changes. In fact, DRGs were never designed to be a payment mechanism; they were designed for managing hospitals. Support for DRGs did not come because the approach offered a perfect technical policy solution. This novel patient classification system was selected because it became closely aligned with the social–cultural system and the political system. To understand why the United States adopted a DRG-based payment system, the events, actors, and incidents within the historical context must be understood.

The remainder of this chapter is organized into five sections. The first section discusses the basic ideas behind DRGs. The second introduces a framework for understanding the policy environment in the United States. The third is a brief history of the introduction of DRGs in terms of the technical, political, and social–cultural systems. The fourth examines some of the impacts and controversies around DRGs. The final section summarizes and discusses the lessons learned.

ABCs of DRGs

In the United States, DRGs are a patient classification technique that defines and measures a hospital's casemix (ProPAC 1985).[3] Designed to segment clinically similar groups of patients by their hospital resource requirements, DRGs pay a flat amount per diagnosis. Each year the United States federal government uses DRGs to set 518[4] diagnostic payment rates for 4,800 short-term acute hospitals treating Medicare eligible patients.

Classification of patients depends on several partitioning variables such as: principal discharge diagnosis, a patient's age, and up to eight

Table 1.1 US DRGs with highest volume FY2003

DRG	DRG Name	Discharges %	No. (000)
127	Heart failure and shock	6%	693
89	Simple pneumonia and pleurisy > 17 w/cc	4	519
29	Major joint and limb reattachment	4	427
88	Chronic obstructive pulmonary disease	3	397
182	Esophagitis, gastroenteritis, and misc. digest. > 17	2	292
296	Nutritional metabolic disorder > 17 w/cc	2	261
174	GI hemorrhage w/cc	2	259
143	Chest pain	2	246
14	Intracranial hemorrhage / cerebral infarction	2	242
320	Kidney and urinary	2	211

Source: Federal Register, May 19, 2004, p 28195–28818

co-morbidities and complications. DRGs have a decision-tree structure, where each patient is categorized into one of twenty-five major diagnostic categories, and then into either a surgical or medical treatment strategy. Each DRG represents a class of patients with similar clinical work processes and medical service bundles. The most common DRGs are shown in Table 1.1. Ten DRGs account for nearly 30 percent of acute hospital admissions.

The number of DRGs has remained manageable, evolving from 468 when federal DRGs began in 1984 to 518 in 2005. However, as diseases and treatments change, so must DRGs. Each year they are reviewed and sometimes amended, or new DRGs are created. For example, there were no ICD–9 codes for HIV infections until 1986. MDC 25 was created for three new DRGs: HIV with major operating room procedure, HIV with major related condition, HIV without other related conditions. The Centers for Medicare and Medicaid Services (CMS) is the US federal agency that has been granted the power to establish new DRGs when a group of patients require more costly procedures. For example, DRG 482 was developed for tracheostomy and patients who require ninety-six hours or more of mechanical ventilation were put into DRG 541 or 542.

Each DRG is assigned a relative weight (RW) that represents national average costs (i.e. the expected resource consumption for a *typical* patient at an *average* hospital). So a DRG with a RW of 1 is expected to consume half the resources as a RW of 2. Each year, every hospital's diagnoses are aggregated and summarized into an overall RW. Pettengill and Vertrees

(1982) found that the casemix index range went from a low of .51 to a high of 1.83. Hospitals with higher RW receive more money.

Recalibration means a DRG's weight is increased or decreased. Each year the 518[5] different DRG weights are reanalyzed and recalibrated. For example, in 1985 a fracture of the femur (DRG 235) had a casemix index (i.e. relative weight) of 1.08; in 2005 the index was reduced to .7512.[6] To set prices with DRGs, Medicare calculates a national average unit price for every unit of service of care received during a hospital stay. In general, a hospital's DRG payment is: *Payment per discharge = (DRG relative weight × Standardized Base Payment)*. Medicare creates base payment rates that include operating (i.e. room and ancillary services) and capital costs (i.e. interest and depreciation). Adjustments are made for market conditions, medical education, care for low-income populations, casemix complexity, and new technology (see appendix).

The American health policy environment

Figure 1.1 displays a general framework for understanding the American policy environment. The health policy environment is composed of a shifting constellation of organizations – political institutions, associations, government agencies, research groups, NGOs, and so on. These organizations can be grouped into three loosely coupled systems, which influence behavior and must be managed: a technical medical care system, a

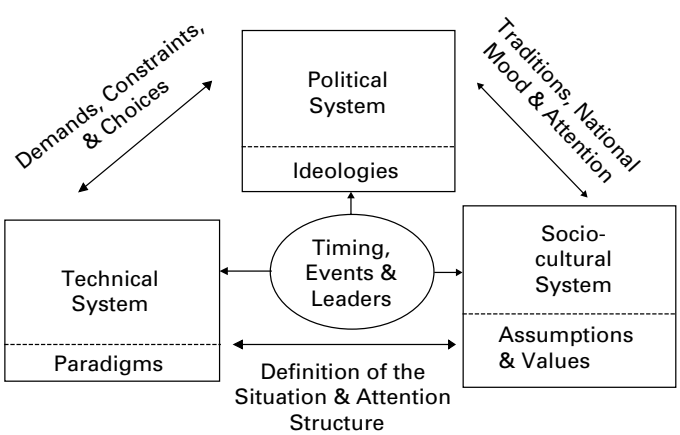

Figure 1.1 Three loosely coupled systems in the health policy environment

sociocultural system, and a political system (see Parsons, 1960; Tichy 1981; Tichy 1983; Tichy 1990; Chilingerian 2004).

The *technical system* is an invisible college of practitioners, delivery organizations, research scientists, planning staff, research and evaluation bureaus, consultants, and of course, academic professors working on ideas. Although personal and organizational values often influence how they frame problems, focus attention, and formulate paradigms, whenever new problems emerge the political system negotiates the demands, constraints, and choices for a problem (Pfeffer and Salancik 1978), and the sociocultural system helps to interpret the situation (Kingdon 1984). The technical system goes right to work obtaining research funding, conducting research, writing reports, publishing and circulating papers, attending hearings, offering testimony, and testing new ideas (Kingdon 1984). Although technical policy proposals that are politically unappealing have little chance of surviving, timing the release of a technical solution can make a difference in its reception.

The *sociocultural system* reduces the likelihood that an idea's merit alone will cause it to rise to the top of the agenda, because solutions are closely embedded in personal relationships, attention structures, and networks (Granovetter 1985). The sociocultural system creates a context for understanding, building commitment and offering justification for the superiority of a technical solution (Parsons 1960). Brickman (1987) has argued that decisions are important because we infer values from decisions; consequently, policy must fit with the cultural values of the mainstream (Kingdon 1984).

In the United States, health policy does not rest with a few powerful individuals; there is a *political system* that allocates power, influence, and attention to a political agenda, and resolves uncertainties about decision rights, relationships that affect resource allocation, internal status and career paths, and conflicts affecting the technical and sociocultural systems (Bacharach and Lawler 1982). Power is shared between the federal government and the fifty state governments. States experiment locally with health policy innovations, and if perceived as successful, new ideas may be adopted at the national level. This shared power is fragmented among a constellation of law-making politicians, executive branch bureaucrats, as well as providers and professional organizations. The fragmentation of power and influence was built into the foundation of American government with the phrase "checks and balances." Morone (1994) describes how this plays out:

Proposals must run an awesome political gauntlet: first, the Office of the President; then, the competing, overlapping committees in each branch of Congress (five separate committees have major jurisdiction over health care bills); after that, the Washington bureaucracy; and, finally, the multiple layers of American federalism – all divided by function and bedeviled by an extensive (much called upon) judiciary. It is not a policy-making apparatus designed for swift or concerted action. On the contrary, American government is designed to be maladroit at securing broad, coordinated policy changes – like national health care reform. (Morone 1994, p 154)

The health policy environment can be thought of as a weather system occasionally subject to severe conditions such as mini-tornados. Small events and incidences, and leadership inside the political system give rise to coalitions of powerful, influential people. Dominant coalitions determine how decisions are made.

Novel or complex problems and uncertainties can cause each of the systems to overload. Events in the situation (i.e., emergent leaders, chance meetings, values at stake, availability of information, struggle for control, etc.) will determine whether the political and/or cultural systems resolve the uncertainty. ''Great ideas'' are powerful, but their truth value is not enough to gain acceptance. Research has found that policy decisions represent a solution acceptable to the coalition (March and Simon 1958). So adoption and implementation require executive and political leaders who organize attention and advocate effectively, and build commitment to new ideas (Yergin and Stanislaw 1999). American policy leaders must convince bureaucrats, congress people, journalists, interest groups, and, indirectly, public opinion that an innovation makes sense. From time to time, leaders mobilize a *dominant coalition* around a great idea that fits with the *dominant ideology* of specialists, and the *dominant cultural values* of the mainstream. Under these conditions, an innovative idea can become a dominant paradigm. Urgency, timing and the depth of leadership matter more than well-designed solutions, competence, and commitment.

Why DRGs were selected: A brief history

Throughout the 1970s the United States experienced unbridled rising hospital costs. The technical experts argued that without a valid and reliable measure for hospital casemix, a fair prospective payment system would be difficult if not impossible. What virtually no one knew was that a small

research group at Yale University had begun to develop a workable casemix measure. However, even the Yale researchers did not understand what they would eventually design, because of the way they had originally framed the problem. They asked – *In order for the hospital to manage and control cost-per-case, how can a hospital product be defined and measured?*

Research & Development at Yale University: "We have this new thing"

DRGs were the brainchild of two Yale professors: Bob Fetter, and John Thompson. Richard Averill, who later became director of health-related research at Yale, was a graduate student working with both Fetter and Thompson. Averill remembers that when they started to talk about these problems, Bob Fetter would ask "But John, what is the product?" John Thompson, whose background was in nursing, would say, "We treat patients." Then Bob Fetter would argue that to design a control process, you would have to be able to differentiate among a hospital's acute products.

While being developed at Yale, this method of describing the hospital's products was not called Diagnostic Related Groups (DRGs). In fact, when Thompson would encourage people to consider using DRGs he merely said "You have to do this new thing." This new thing that could describe and measure the entire range of patients treated in an acute hospital, from newborns and children to adults.

DRGs began as a "pure-research endeavor." In 1967, a group of local physicians asked for help with utilization review (Fetter 1991).[7] Could industrial engineering techniques be adapted to hospitals? Subsequent research in the early 1970s looked at why maternity and newborn care and costs varied among accredited, not-for-profit Connecticut hospitals. To address those questions, the Yale team worked iteratively among the conceptual, empirical, and policy domains. Fetter (1991) recalled in 1969 bringing together a panel of physicians to describe clinical work processes. As in the case of manufacturing, tens of thousands of "unique" hospital patient types existed. For example, there were thirty-nine ways to describe a cataract care process. Researchers were searching for an underlying structure focused on similarities to "identify the ordinary, the usual, the routine, and applying the techniques of statistical process control, to filter out the aberrant cases in order to understand the causes of aberrations" (Fetter 1991, p 6).

Previous work found significant correlations between length of stay (LOS), total charges and casemix complexity (see Lave and Leinhardt 1976;