Social Science & Medicine 178 (2017) 157–166

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Review article

The influence of welfare systems on pay-for-performance programs for general practitioners: A critical review

Mehdi Ammi*, Grant Fortier

School of Public Policy and Administration, Carleton University, Ottawa, Ontario, Canada

ARTICLE INFO

Article history:
Received 7 June 2016
Received in revised form
10 February 2017
Accepted 12 February 2017
Available online 16 February 2017

Keywords:
Pay-for-performance
Welfare systems
General practitioners
Decommodification
Social stratification
North America
Europe

ABSTRACT

While pay-for-performance (P4P) programs are increasingly common tools used to foster quality and efficiency in primary care, the evidence concerning their effectiveness is at best mixed. In this article, we explore the influence of welfare systems on four P4P-related dimensions: the level of healthcare funders' commitment to P4Ps (by funding and length of program operation), program design (specifically target-based vs. participation-based program), physicians' acceptance of the program and program effects. Using Esping-Andersen's typology, we examine P4P for general practitioners (GPs) in thirteen European and North American countries and find that welfare systems contribute to explain variations in P4P experiences. Overall, liberal systems exhibited the most enthusiastic adoption of P4P, with significant physician acceptance, generous incentives and positive but modest program effects. Social democratic countries showed minimal interest in P4P for GPs, with the exception of Sweden. Although corporatist systems adopted performance pay, these countries experienced mixed results, with strong physician opposition. In response to this opposition, health care funders tended to favour participation-based over target-based P4P. We demonstrate how the interaction of decommodification and social stratification in each welfare regime influences these countries' experiences with P4P for GPs, directly for funders' commitment, program design and physicians' acceptance, and indirectly for program effects, hence providing a framework for analyzing P4P in other contexts or care settings.

1. Introduction

Over the last two decades, pay-for-performance (P4P) programs have become popular compensation tools for health care providers. The main objectives of P4P are to improve service quality, contain costs, foster efficiency and ultimately improve patients' health outcomes. Complementing traditional physician compensation models (e.g. fee-for-service, capitation), P4P programs use financial incentives to explicitly reward adherence to predefined standards of care and cost control (Greene and Nash, 2009). While some early programs penalized health care providers for non-compliance, the vast majority of programs reward physicians in addition to standard compensation (Town et al., 2005). Performance pay initiatives can reinforce multiple dimensions of care, which may be classified under the well-known structure (e.g. electronic health records), process (e.g. disease management for patients with chronic conditions) and outcomes (e.g. increasing the proportion of generic drugs prescribed) framework (Donabedian, 1988). In practice, some P4P programs will reward "participation" rather than "targets" – for instance, if physicians receive bonuses for every additional patient they enroll in a chronic disease management program.

Given the strong impact of general practitioners (GPs) on the utilization of health care services (namely specialists' services and prescription drugs), GPs are a natural target for cost-containment strategies including pay-for-performance (Eijkenaar, 2012; Kravet et al., 2008; Wright and Ricketts, 2010). In addition, quality primary care has been identified as a significant contributor to the overall strength of the healthcare system (Starfield et al., 2005), making P4P a strategic component of the quality improvement arsenal.

Despite the high expectations of its advocates, the literature on the effectiveness of P4P is inconclusive, with findings ranging from a modest improvement in targeted outcomes to negligible program impacts (Eijkenaar, 2012; Eijkenaar et al., 2013; Greene and Nash, 2009). This inconclusiveness is attributable to the quality of the research design and statistical analysis used, the variation in

* Corresponding author. River Building, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada.
E-mail address: mehdi.ammi@carleton.ca (M. Ammi).

http://dx.doi.org/10.1016/j.socscimed.2017.02.019
0277-9536/

individual P4P program design choices (Eijkenaar, 2012; Eijkenaar et al., 2013), as well as difference in contextual and institutional background. The outcomes of P4P also depend on the degree of awareness of the program among physicians (Li et al., 2014), the existence of a dialogue between patients, physicians and payers (Rosenthal and Dudley, 2007) and the pre-existing main payment system (Kantarevic and Kralj, 2013). Nevertheless, a more comprehensive understanding of the factors contributing to P4P effectiveness is necessary, including program acceptance by the providers of care.

Eikemo and Bambra (2008) remind us that welfare regimes are an effective way of classifying and understanding the ideological, social and institutional context of health systems in general; thus, they may prove helpful in comprehending the setting within which physician performance pay is implemented. The Esping-Andersen (EA) typology sorts the industrialized nations of Western Europe and North America into three broad welfare state regimes (Esping-Andersen, 1989, 1990): liberal (e.g. USA), corporatist (e.g. Germany) and social-democratic (e.g. Norway). States falling within each regime type are differentiated by (1) the role of the welfare state in altering market forces (decommodification), (2) whether social welfare reinforces class distinctions (social stratification), and (3) the relationship between the state, individual and family (Esping-Andersen, 1989), with membership within each welfare category heavily influenced by the first two attributes of the classification. Although the EA model's relevance has been criticized (Barrientos, 2008; Orloff, 1993; Van Der Veen and Van Der Brug, 2013), it is nevertheless a widely recognized framework that remains central to the health and social policy debates (Emmenegger et al., 2015; Powell and Barrientos, 2015).

The goal of this critical literature review is thus to identify *whether* and *how* differences between welfare regimes, as identified by Esping-Andersen, affect various P4P program-related dimensions that cover the full program lifecycle: the level of healthcare funders' (e.g. governments, insurers) commitment to P4Ps (by funding level and length of program operation), program design (specifically target-based vs. participation-based program), the physicians' acceptance of the program (attitudes of general practitioners and their medical association, and participation rates in the program) and program impacts. By examining the variation in P4P adoption between welfare systems rather than examining each program individually, we contribute to the fast growing pay-for-performance literature in a unique way.

The paper is organized as follows. We present our methodology and choice of countries in Section 2. Section 3 describes the main P4P initiatives grouped by welfare regime and explores if and how welfare regimes influence experiences with P4P. We discuss these results in Section 4 and conclude in Section 5.

2. Methodology

Using a critical literature review approach, we analyzed P4P experiences across welfare systems, with the nation state being the primary unit of analysis. We selected a large sample of countries from each regime for analysis to reduce the risk of selecting outliers in each system category while using the consistency of the results across countries within a typology to ensure the validity of our conclusions. For consistency, we used Esping-Andersen's (1990) original categories of welfare regimes and their constituent nations (see Table 1). While not all countries fit neatly within the original classification, we retain these groups to ensure a reasonably representative selection of cases across all three categories, and to assess the predictive power of EA's original typology in terms of P4P experiences.

We conducted a survey of available peer-reviewed and grey English-language P4P literature using PubMed and Google Scholar, taking an approach similar to a scoping review to map out what is currently known. Firstly, we conducted searches for each country in the study using the following keywords in varying combinations: “pay-for-performance,” “P4P,” “programs,” “healthcare,” “incentives,” “physicians” and “general practitioners.” Secondly, we adopted a snowball strategy by investigating individual sources that were cited in the documents turned up through the initial search, and keeping the referenced document when it provided more detailed information about the experience with P4P in a given country. Finally, the review of international P4P programs by Eijkenaar (2012) identified a number of reliable sources on specific programs and was used for triangulation.

Table 1
Countries selected by welfare system type.

Liberal	Corporatist	Social democratic
Australia	Austria	Denmark
Canada	France	Finland
New Zealand	Germany	Norway
United Kingdom	Italy	Sweden
United States		

The goal of these searches was to find evidence of at least one pay-for-performance program active in each country being studied. The program had to be an official initiative in effect at a national or regional/state/county level that provides financial incentives to general practitioners for the meeting of specific “performance” objectives. In the event a nation had multiple P4P programs in place, we selected only one program study based on each program's significance, how widespread it was amongst physicians and whether sufficient information about the program's history and effects was available. Note that in some countries, our search failed to turn up any physician pay-for-performance programs. Nevertheless, the absence of P4P in a country is a significant result in itself.

Having gathered our information, we examined the P4P program literature for four key areas of interest: (1) the level of commitment to the P4P program, as reflected in the program's start date and the mean proportion of annual physician income derived from P4P participation; (2) physician acceptance of the P4P programs. Indicators include general attitudes of physicians and their representatives (i.e. medical associations) towards proposed P4P programs, and the participation rates of eligible physicians, in cases where participation is voluntary; (3) the program design, particularly as whether the P4P is target-oriented or participation-based; (4) the effects of the programs on physician practice, in terms of measured performance.

These outcomes are significant because they are indicative of how the key targets of P4P programs (physicians) will affect (program design) and react to a proposed program before (general attitude and acceptance), during (participation rates and income from performance pay) and after (effects) its implementation, covering the full lifecycle of a given program.

3. Results

The different outcomes for the P4P programs in each welfare system type are synthetized in Table 2 and discussed hereafter.

3.1. Liberal system

Physicians and medical associations in liberal regimes are broadly supportive of the concept of P4P and view it as a legitimate policy tool for public and private insurers to improve quality of care