

**Systematic review protocol:  
Examining the effects of introducing pay-for-performance for primary care  
physicians on diabetes outcomes in single-payer healthcare systems**

Neeru Gupta<sup>1\*</sup>  
Holly M. Ayles<sup>2</sup>

<sup>1</sup> Associate Professor and Diabetes Research Chair.

<sup>2</sup> Student of Business Administration.

\* Corresponding author: [neeru.gupta@unb.ca](mailto:neeru.gupta@unb.ca).

University of New Brunswick  
Department of Sociology  
PO Box 4400  
Fredericton, New Brunswick E3B 5A3, Canada.

Diabetes Population Health and Health Services Research Working Paper, no. 2017-01.

Fredericton, November 2017.

## **ABSTRACT**

**Background:** Although pay-for-performance (P4P) for diabetes care is increasingly common across health organizations, evidence of its effectiveness in improving population health and service delivery is deficient. This information gap is attributable in part to the heterogeneity of healthcare financing, covered medical conditions, care settings, and provider remuneration arrangements within and across countries.

**Objective:** This paper outlines a protocol for a systematic review examining the effects of introducing P4P for physicians in primary care and community settings to support guideline-based diabetes care. Our aim is to reduce the heterogeneity of evidence presented that has deterred conclusiveness of previous reviews by narrowing the focus to disease-specific P4P schemes in single-payer healthcare insurance systems. This approach enables us to minimize the risk of unintended consequences of P4P such as physicians' gaming the payment system.

**Methods:** Our review systematically searches, appraises, and synthesizes the literature concentrating on whether P4P for primary care physicians leads to better diabetes outcomes in single-payer health systems. We search 10 electronic databases and manually scan the reference lists of review articles and other global health literature. We include primary studies evaluating the effects of introducing P4P for diabetes care among primary care physicians in countries of universal health coverage. Outcomes of interest include patient morbidity, avoidable hospitalization, premature death, and healthcare costs.

**Results:** We have received funding from Diabetes Canada and the New Brunswick Health Research Foundation to conduct policy-actionable diabetes health services research. Database searches were conducted and full-texts screened by two reviewers in 2017. We aim to submit the review for publication in 2018.

**Conclusions:** We are narratively synthesizing the data. Because of the wide range of outcomes considered, we do not expect to perform a meta-analysis.

## **KEYWORDS:**

Healthcare delivery; healthcare financing; diabetes care; pay-for-performance; systematic review; research protocol.

## **INTRODUCTION**

The prevalence of diabetes mellitus has been steadily increasing worldwide in recent decades, nearly doubling from 4.7% to 8.5% among the adult population between 1980 and 2014.<sup>1</sup> The World Health Organization advocates better incentives for healthcare providers and other means of focused financing can enhance health system efficiencies.<sup>2</sup> Although pay-for-performance (P4P) for management of diabetes and other chronic non-communicable diseases has become increasingly common in many health organizations, evidence of its efficiency remains deficient. Part of this information gap is attributable to the wide range of provider incentives, which vary in scales, performance measures, and labels within and across countries for the delivery of specific services.<sup>2,3</sup> While a number of systematic reviews have examined the effects of P4P on different indicators of healthcare processes, costs, and outcomes across different contexts of health system financing, they have been marked with heterogeneity in terms of the covered medical conditions, care settings, and other contextual characteristics that deters conclusiveness of the effects of P4P on health outcomes.<sup>3-8</sup> Our objective is to narrow the field of scope to help countries and jurisdictions reach evidence-informed decisions on whether to introduce or maintain P4P for diabetes management in primary care. We are performing a systematic review to address whether incentives for physicians in primary care and community settings lead to better diabetes outcomes – compared to the absence of diabetes-related incentivising remuneration – in single-payer national health insurance systems.

## **METHODS**

We are conducting our study based on a PICOS (Population, Intervention, Comparison, Outcomes, Study type) framework, in line with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines.<sup>9</sup>

### **Research question**

We synthesize the evidence from countries with single-payer health insurance systems on effects of P4P on diabetes outcomes. Specifically, we aim to answer the question: Does the

introduction of physician incentive reimbursements for diabetes management in primary and community care led to improved diabetes-related population health and health system outcomes?

### **Eligibility**

Original quantitative research studies are eligible to be included in the systematic review if they assessed the effects of introduction of P4P for primary care physicians on diabetes outcomes through some sort of evaluative component. Evaluations can include population-based observational studies, natural experiments, and before-and-after modelling studies. Because we are interested in outcomes at the population or health system level, the generalizability and applicability of randomized control trials are limited in this case. We exclude studies that are not largely representative of the general (or adult) patient population; from countries with mixed-payment national health systems; comparing changes in incentive types or levels for pre-existing P4P schemes; or assessing incentive payments for patients, for non-medical practitioners, or for physicians operating exclusively in hospitals or other non-community settings.

We include studies where the main outcomes of interest included patient- or policymaker-important health and health system markers for chronic disease management. We exclude studies focusing exclusively on inclusion/exemption status within a P4P scheme, diagnostic screening, structural changes, or patient or provider perceptions.

### **Search strategy**

We search for relevant peer-reviewed publications in 10 electronic databases: ABI Inform, Business Source Premier, Canadian Business and Current Affairs, Cochrane Library, EconLit, PAIS, PubMed, Scopus, SocIndex, and Sociological Abstracts. The search strategy includes nomenclature related to pay-for-performance and diabetes mellitus. Free text and formal search terms and filters are translated to respect database-specific requirements, with the advice and assistance of library professionals at the University of New Brunswick's Harriett Irving Library (Fredericton, Canada). We also hand-search the reference lists of previous systematic reviews on physician remuneration and pay-for-performance,<sup>3-8,10-12</sup> as well as the World Health Organization's flagship report, *The world health report: health systems financing—the path to universal coverage*<sup>2</sup> and its relevant background papers. We include studies in English, French, Spanish, or Portuguese that were published between 2000 and 2017.

## **Data extraction**

Two reviewers independently screen a sample of the eligible full-text articles to identify and secure consensus on studies for review inclusion. We record the country and its health financing arrangement, characteristics of the incentive scheme, the study objective, the care providers and patient population in the study, the data gathering technique, the comparison groups, and the outcomes measured. We characterize the timing of outcomes as prior (baseline) or consecutive (follow-up) to the introduction of P4P.

## **Quality assessment**

We develop a standardized quality assessment tool to grade the evidence reported in each study based on two key criteria for evaluating complex social interventions: relevance for healthcare improvement; and methodological rigour.<sup>13-17</sup> We downgrade studies considering endogenous treatment of guideline-based diabetes care in the assessment (e.g., glycemic monitoring and control that might cause estimation bias). Studies are assigned a letter grade to report the outcomes relevance and methodological quality; a copy of the data extraction and quality assessment form can be found in Box 1. Any disagreements between the two reviewers in quality assessment are resolved by consensus.

## **Analysis**

A preliminary search covering the period January 2000 to April 2017 yielded 2090 records from ten databases and other peer-reviewed and grey literature sources. After applying our inclusion and exclusion criteria, we will use narrative synthesis to interpret the evidence. Because of the heterogeneity of outcomes under consideration, we will not perform a meta-analysis. Since database searches were completed prior to the publication of this protocol, it is ineligible for registration with the [PROSPERO](#) international database of prospectively registered systematic reviews in health and social care.

**Filters for inclusion in the systematic review of physician practice incentives**

- Healthcare system: single-payer
- Published between 2000 and 2017
- Published in English, French, Spanish or Portuguese
- Participants: general (adult) population, with chronic disease
- Provider type: general/family/community physicians
- Practice location: primary care (not exclusively hospital or other non-community setting)
- Intervention: introduction of pay-for-performance (not comparison of different P4P types/levels)
- Study objective: quantify/evaluate the effect of the introduction of P4P on patient/population outcomes
- Study design: experimental or observational data (not just descriptive review)

Location: \_\_\_\_\_  
 Year of publication: \_\_\_\_\_  
 Other language: \_\_\_\_\_

**Quality assessment for included studies**

Two criteria: outcomes (1 item) and methods (3 items).

Each study can get an A (high), B (medium) or C (low) quality rating on the following basis:

- A = A/B for q1 and A on all three items for q2
- B = A/B for q1 and A/B on all three items for q2
- C = Any C.

**1. Guideline for relevancy of OUTCOME measures:**

At least 2 different outcomes measured on severe morbid events (e.g. severe complications of chronic disease), hospitalization, mortality, or health care costs	A
At least 1 outcome measured on severe morbid event, hospitalization, mortality, or costs	B
Only non-critical or non-patient-important outcomes measured (e.g. physical exam or care plan without recording symptoms, glucose control, lipid profile...)	C

**2. Guideline for the METHODS quality criteria:**

**2.a. Selection bias**

Study participants very likely to represent the target (adult) population: 90 to 100% representation	A
Study participants somewhat likely to represent target population: 70 to 89% representation	B
Not likely to represent target population (<70% representation) OR Representation unreported/unclear OR Validity of data unreported/unclear OR Indication of exclusion of important population groups or major loss to follow-up of participants over time	C

**2.b. Study design**

Control group and pre- and post-intervention <i>longitudinal</i> data (i.e. same individuals pre and post)	A
No control group and pre- and post-intervention longitudinal data; OR Control group and pre- and post-intervention <i>cross-sectional</i> data (i.e. different individuals pre and post) AND no indication of major change in population	B
No control group and pre- and post-intervention <i>cross-sectional</i> data; OR Control group and pre- and post-intervention cross-sectional data AND possibility of major change in population; OR No pre-intervention data	C

**2.c. Confounders**

Control group <i>matched</i> to intervention group on key confounding variables AND supporting data/methods presented (e.g. propensity score matching, quasi experimental design, interrupted time series, difference-in-differences, difference between counterfactual and observed)	A
Stated that control group matched or was similar to the intervention group, but supporting data/methods not presented	B
No patient matching or adjustment of unobservables that may cause bias reported; OR No information on differences between intervention and control group; OR No patient control group	C

**Box 1: Inclusion/exclusion and quality assessment grid for the systematic review of effects of P4P for primary care physicians on diabetes outcomes in single-payer health systems**

## **Role of the funding sources**

Diabetes Canada and the New Brunswick Health Research Foundation provided financial support for this work. The funders have no role in the study design, data collection, data analysis, or research dissemination strategy.

## **DISCUSSION**

Given the rapidly rising rates of diabetes and other chronic non-communicable diseases worldwide, combined with global efforts to promote universal health coverage, decision makers need solid and timely evidence to inform policy and resource allocation decisions for strengthening primary care service delivery for chronic disease prevention and management. While there have been a number of systematic reviews examining the effects of P4P for improving quality of care, many focused primarily on hospitals and none screened for the health system financing arrangement. To the best of our knowledge, this is the first review aiming to focus on the evidence in single-payer health insurance systems. The restriction to single-payer systems enables us to minimize the risk of unintended consequences of P4P from physicians' gaming the payment system, in particular, to discount possible effects of physicians' moving between health organizations within a jurisdiction to benefit from an incentive or avoiding high-risk patients so as to not upset clinical performance metrics. A limitation of this approach is that relatively few (essentially high income) countries will be included in our study.

Review findings will be disseminated in peer-reviewed publications and at presentations targeting both the researcher and knowledge user communities.

## **REFERENCES**

1. World Health Organization. *Global report on diabetes*. Geneva: 2016.
2. World Health Organization. *The world health report—health systems financing: the path to universal coverage*. Geneva: 2010.
3. Scott A, Sivey P, Ait Ouakrim D, Willenberg L, Naccarella L, Furler J, Young D. The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Syst Rev* 2011; 9: CD008451.

4. de Bruin SR, Baan CA, Struijs JN. Pay-for-performance in disease management: a systematic review of the literature. *BMC Health Serv Res* 2011; 11:272.
5. Turcotte-Tremblay AM, Spagnolo J, De Allegri M, Riddle V. Does performance-based financing increase value for money in low- and middle-income countries? A systematic review. *Health Econ Rev* 2016; 6(1):30.
6. Emmert M, Eijkenaar F, Kemter H, Esslinger AS, Schöffski O. Economic evaluation of pay-for-performance in health care: a systematic review. *Eur J Health Econ* 2012; 13:755–67.
7. Chaix-Couturier C, Durand-Zaleski I, Jolly D, Durieux P. Effects of financial incentives on medical practice: results from a systematic review of the literature and methodological issues. *Int J Qual Health Care* 2000; 12:133–42.
8. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv Res* 2010; 10:247.
9. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009; 6(7): e1000100.
10. Tao W, Agerholm J, Burström B. The impact of reimbursement systems on equity in access and quality of primary care: a systematic literature review. *BMC Health Serv Res* 2016; 16:542.
11. Jia L, Yuan B, Meng Q, Scott A. Payment methods for ambulatory care health professionals. *Cochrane Database Syst Rev* 2015; 9: CD011865.
12. Flodgren G, Eccles MP, Shepperd S, Scott A, Parmelli E, Beyer FR. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Syst Rev* 2011; 7: CD009255.
13. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ, GRADE Working Group. What is ‘quality of evidence’ and why is it important to clinicians? *BMJ* 2008; 336(7651):995–98.
14. Movsisyan A, Melendez-Torres GJ, Montgomery P. Users identified challenges in applying GRADE to complex interventions and suggested an extension to GRADE. *J Clin Epidemiol* 2016; 70: 191–99.
15. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Ytter Y, Schünemann HJ. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 2011; 64(4): 395–400.
16. Berwick DM, Nolan TW, Whittington J. The Triple Aim: care, health, and cost. *Health Affairs* 2008; 27(3): 759–69.
17. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al. GRADE guidelines: 9. Rating up the quality of evidence. *J Clin Epidemiol* 2011; 64(12): 1311–16.