Guest editorial: Payment for performance: comparing policy making, design and implementation in health, social care and education

Over the last four decades payment for performance (P4P), also known as performance-based financing (PBF) or results-based financing (RBF), has been increasingly applied in a variety of contexts (and sectors) with the expectation that it can improve the performance of health systems in low- and middle-income countries (LMICs) and high-income countries (HICs). Initially adopted as a method for improving health worker and health organisation performance, P4P/PBF has been increasingly seen as a mechanism for more strategic reform of health systems (especially in LMICs).

P4P has its roots in private sector management and is based on the use of financial incentives to extrinsically motivate workers to improve their performance and thus improve organisational performance. In the 1970s and 1980s, such initiatives were introduced in public sector organisations as New Public Management (NPM) concepts and started to become embedded in public administrations (Pollitt, 1993). However, it is important, especially in the context of the papers presented in this special edition, to note the difference between NPM as a concept or ideology and PBF modalities as management practices that are associated with NPM but are not necessarily predicated on the adoption of NPM as an ideological approach.

As reported by other papers published in this journal over the past decades (Dixon, 2021; Hur, 2018), the application of pay for performance staff reward schemes within public sector became widespread internationally. However, the practice also spread into team and organisational performance related payments. Aoki (2019) demonstrated the widespread use of performance related pay in education services across 60 countries world-wide finding that such policies were, in 2012, most widely used in less accountable and less wealthy economies. In global health financing, it has been suggested that PBF experienced a similar phenomenon, where generalised programmes acted as "travelling models" with major international organizations promoting similarly designed schemes across different country contexts (Olivier de Sardan *et al.*, 2017; Paul and Ridde, 2022). This was particularly the case in LMICs, with a limited evidence-base nor space for adaptation (Barnes *et al.*, 2015). The research by Aoki further highlighted an important "wave effect" and that results often demonstrated a situation where early adopters had already ended schemes, suggesting that there have not been many "lessons learned" from previous programmes. For example, by the 1990s, there were already concerns being expressed about both the efficacy of such programmes in the private sector and their potential for adoption in the public sector (Ingraham, 1993). By the 2000s, several scholars were also suggesting the end of NPM type approaches in the public sector with the emergence of new concepts of public governance (Dunleavy *et al.*, 2006; Osborne, 2006).

In public health, P4P programmes can be broadly defined as a mechanism through which health providers can be partially financed and/or health professionals receive financial incentives based on their performance when measured against predetermined indicators/targets (Saddi *et al.*, 2018). Over the last two decades, not only did P4P types of programmes

International Journal of Public Sector Management
Vol. 36 No. 6/7, 2023
pp. 481-491

0951-3558
DOI 10.1108/IJPSM-10-2023-355

continue to be used in a variety of public sector services, but there has also been an expansion of the use of such schemes in health systems. These programmes, however, differ from country to country in design, while operating idiosyncratically, moderated by various contextual and socioeconomic factors (country income level, centralized vs decentralized systems, cultural practices, etc.).

In HIC, P4P more commonly reflects direct payment incentives to health workers for select services and not always linked to wider health systems or service reforms. In the USA, in 2005, over half of all health maintenance organisations adopted some form of P4P (Rosenthal *et al.*, 2006) and in the UK in 2003 a universal P4P scheme, the Quality and Outcomes Framework (QOF), was introduced for all primary care providers (Lester and Campbell, 2010). The QOF was linked to the work of the general practitioner and their contracted services. In a review of performance measurement programmes in public health systems in North America, Europe, Australia and New Zealand, Schwartz and Deber (2016) identified 55 different schemes suggesting widespread implementation.

In low-income countries, and especially in African countries, PBF is the more common concept employed and has proliferated from "diffusion entrepreneurs" who played an important role in spreading PBF schemes (Gautier *et al.*, 2018). The World Bank, international non-governmental organizations (INGOs) and other global health funding agencies had a major role in the training, implementation and monitoring of PBF (Saddi *et al.*, 2018, p. 2). By 2017, 32 out of 46 sub-Saharan African countries implemented PBF (Gautier *et al.*, 2018).

Despite the significant proliferation of P4P and performance-based schemes in health systems and their utilisation for wider system initiatives such as promoting universal health care, raising service quality or increasing investment in specific areas of care; the evidence of health outcome effect remains mixed (Renmans *et al.*, 2016; Paul *et al.*, 2018; Singh *et al.*, 2021; Diaconu *et al.*, 2022; Saddi *et al.*, 2023). Moreover, the paper by Kristensen and colleagues in this special edition highlight the vast range in system design while arguing that research and literature on performance schemes often remains siloed in LMICs, MICs and HICs with little cross-over or use of comparative analysis, as also emphasised as well by two earlier special issues (Saddi *et al.*, 2018, 2021) and a satellite session at Health Systems Global (Anselmi *et al.*, 2020) – something we start to address in this special edition. Nevertheless, while comparing schemes will help in the design of P4P/PBF schemes and their implementation, it says little about wider system impacts and the need to situate these programmes in the wider health policy and health systems literature.

Yet, concerns about the effectiveness of P4P/PBF and its underwhelming evidence-base as a strategic reform tool stem from the 1980s in relation to the private sector, the point at which many of the practices were being adopted in public service sectors. In particular, critical attention has been focused on the importance of considering both extrinsic (which covers schemes like PBF and other policy drivers and the wider context of implementation) and intrinsic drivers which have been shown to be wider than simply monetary reward (Saddi *et al.*, 2023).

Although there is a mixed and contested evidence-base in support of P4P/PBF, there has been a continued implementation of PBF schemes and their promotion by donor organisations and consultants (Paul *et al.*, 2017; Diaconu *et al.*, 2022). As one example, the RBF scheme in Zimbabwe has recently undergone a major expansion of its performance indicators, which has occurred just a few years after the scheme underwent a full national rollout (Kadungure *et al.*, 2021). Moreover, despite a recent World Bank report downplaying the significance of PBF in producing positive health outcomes in comparison to other forms of financing (de Walque *et al.*, 2022), the World Bank's Global Financial Facility has continued to promote its PBF model in its negotiations with "front runner countries" like Mozambique, while a results-based rationale has further been introduced as an important component for achieving pandemic preparedness by the new Pandemic Fund (World Bank, 2023).

In Brazil, although there have been challenges and failures in the implementation of PMAQ, especially its health education component (Saddi *et al.*, 2021a, b; Coelho *et al.*, 2023), P4P was not removed from the policy agenda. In 2019, Brazil adopted a new P4P scheme in primary care, partially inspired in the UK QOF. The new Prevent Brazil Programme is the new primary care funding model in Brazil, including P4P as one of its main components. This expansion of P4P was done without considering inequality in health services and suggests that the logics of P4P remain entrenched within many health reform strategies.

That said, there are some signs of an initial shift away from P4P/PBF to thinking about better methods of resource allocation rather than solely focusing on drivers of behavioural economic incentivization based on payment. As suggested above, a recent World Bank report on financial incentivization concluded that it "produced gains in health outcomes compared to business-as-usual, but these gains did not necessarily result from the specific financial incentives and associated monitoring components of projects", suggesting that "impactful health financing reform may mean pivoting from performance pay while retaining other important aspects of PBF projects—like transparency, accountability, and decentralized frontline financing" (de Walque *et al.*, 2022). What this signals, to some degree, is a creeping acceptance that despite over 30 years of implementation in global health, the overall benefits of PBF have remained indeterminate. This indeterminacy exists despite pockets of success and recognition that certain aspects of programme design are worth maintaining. Additionally, it remains an open question whether these pockets of success outweigh the costs, especially when considering growing concerns in LMICs about overall P4P/PBF financial sustainability, equitable delivery and adverse knock-on effects.

Although the recent World Bank report suggests that P4P/PBF might be waning within certain policy circles, the logic of PBF and its core principles seemingly retain much of its ideational hegemony within the development aid for health (DAH) lexicon. For example, there remains an interlocking policy relationship between strategic purchasing and PBF, with several World Health Organization and World Bank documents suggesting that PBF is a key "entry point for strategic purchasing" and for driving reforms toward universal health coverage among other health goals (Kutzin, 2016; WHO, 2017; Suvcat *et al.*, 2017; Mathauer *et al.*, 2017, 2019; McIsaac *et al.*, 2018). This overlap is also prevalent in the PBF literature, where PBF is often understood as a crucial mechanism for strategic purchasing (Witter *et al.*, 2019). This has led some to suggest that despite P4P/PBF's lack of delivery, it is being "repackaged" into health reform strategies by its champions (Gautier *et al.*, 2019) and that there is need for a more conceptual and empirical re-examination of strategic purchasing's priorities and its link with PBF (Paul *et al.*, 2021).

Given that key principles of P4P/PBF continue to be embedded in new policy designs, it is crucial to generate a clearer picture of what works, for whom and how. In response, a key aim of this special edition is to examine comparable mechanisms of P4P/PBF across settings to provide a more comprehensive understanding of the benefits and burdens of P4P/PBF, particularly as they manifest in different contexts.

The need for a new comparative research agenda

A recent review of the literature on PBF in low- and middle -income countries (Singh *et al.*, 2021) showed that too little is known about how contextual conditions influence PBF programmes, moderate their effectiveness and affect its relevance, reinforcing a similar conclusion made 5 years earlier (Renmans *et al.*, 2016). Moreover, this lack of reliable evidence-base can be tracked back to an effectiveness-oriented research agenda and a reliance on methodological approaches that have traditionally focused on results rather than processes and on the intervention rather than the context.

In trying to show P4P/PBF effectiveness or lack of it, many studies have resorted to quasi -experimental designs (Appel *et al.*, 2023; Sieleunou *et al.*, 2020; Tawfiq *et al.*, 2019). Fuelled by the

belief that these are the gold standard of evaluation and research methods, the World Bank invested millions of dollars in randomized controlled trials through the Health Results Innovation Trust Fund, later called the Global Financing Facility (Fernandes and Sridhar, 2017). This claim about the gold standard status of quasi-experimental designs has long been disputed (Bhaskar, 2008; Krauss, 2018; Victora *et al.*, 2004), while it has become equally clear that it does not fare well in elucidating the influence of context (Pawson and Tilley, 1997). A simple explanation is that considering context within statistical approaches means you need to introduce interaction terms which lowers the statistical power and increases the number of observations needed, which is not always feasible. A more profound explanation is that contextual factors are often approached as confounders, while it is assumed that P4P/PBF has an intrinsically causal power where net-effects can be determined (Marchal *et al.*, 2013; Ragin, 2014 [1987]). This leads to uncritically pooling data from very different contexts and P4P/PBF schemes, while aiming to analyse the intrinsic effectiveness of PBF. For example, Gage and Bauhoff (2021) performed an analysis of the pooled data of Burundi, Lesotho, Senegal, Zambia and Zimbabwe to study the impact of P4P/PBF on neonatal mortality, without truly taking into account the many differences between countries and there P4P/PBF schemes.

To be sure, not all studies have been focusing on impact and several studies have instead focused on the processes and mechanisms triggered by P4P/PBF (De Allegri *et al.*, 2018; Lohmann *et al.*, 2018; Renmans *et al.*, 2020; Singh *et al.*, 2021; Saddi *et al.*, 2021a, b, 2023). This certainly helps us to better understand how these programmes work and even to some extent how context may impact its effectiveness. However, individual case studies make it difficult to investigate the effect of broader system-level contextual factors, like cultural norms, work ethics, institutional arrangements, existing lines of accountability, governmental trust, etc. Indeed, the influence of such factors may only be thoroughly studied using cross-case comparisons in which variation exists across these factors.

The lack of such studies is in part testimony to the many research challenges encountered by them. First, while P4P/PBF schemes often follow a similar set-up, they also differ in significant ways (Renmans *et al.*, 2017). Second, the context in which these projects are being implemented may either differ too much (at the micro/meso-level) or too little (at the macro-level). Third, data sources may not be comparable. All three obstacles strongly complicate and limit our ability to compare schemes across countries and sectors. Moreover, when we try to compare studies from different countries, via systematic reviews, we are hampered by the different methodologies and data sources used (Diaconu *et al.*, 2022). As a result, when comparative studies are being implemented, they often end up being two or more individual case studies in parallel, failing to truly benefit from the comparison.

While the studies in this special edition go some way to fill this research gap, a different philosophical perspective may offer some solution to these challenges. Comparative studies based on the idea that the intervention has intrinsic causal power are most affected by the obstacles. A configurational view on causation may partly overcome them or even benefit from them (Ragin, 2009, 2014 [1987]). This view asserts that a causal relationship is context-dependent and that not one factor or intervention, but a combination of factors (intervention and contextual conditions) leads to a certain phenomenon/outcome. This idea has gained considerable traction in the social sciences given the growing influence of Bhaskar's (2008) critical realism, Ragin's (2014 [1987]) qualitative comparative analysis, and Pawson and Tilley's (1997) realist evaluation. When looking at comparative studies through this lens, differences between contexts and interventions become valuable sources of information about the influence of contextual factors.

Combining such a configurational approach with a theory-based approach can help to further our understanding of how P4P/PBF works and interacts with contextual conditions. Indeed, through the process of abstraction, analytical generalization (Yin, 1994), and the development of middle-range theories (Merton, 1949), we can connect the findings from one

context to those from another context, enabling comparative research designs and furthering our knowledge on PBF and its interaction with context.

One increasingly relevant approach that does just that is the realist approach as developed by Pawson and Tilley (1997). Instead of asking the question "does it work?", it focuses on the much broader question "what works for whom, when, where, under what circumstances, how and why?". It is based on a context-sensitive generative view of causation (see also Bhaskar (2008)), which means that in certain contexts certain mechanisms generate specific outcomes. The evaluator/researcher tries to develop an understanding of how an intervention works by developing a middle-range program theory. As said, it is this connection with theory that may be the missing link to make comparative research on PBF possible, relevant and effective.

This special issue aims to start a renewed critical debate on how to produce contextual evidence in P4P/PBF from the perspective of comparative analysis of policy processes and/or effects, to better understand how management and policy similarities and differences take place across different countries, across administrations or health teams (or selected units of analysis) in a single country and from a global viewpoint.

Papers published in this special issue

This special issue gathers three types of comparative studies: cross-country comparisons (Zitti *et al.*, 2023; Andersen, 2023), intra country comparisons (Aktas *et al.*, 2023; Srinivasan and Sarin, 2023; Coelho *et al.*, 2023; Dos Anjos *et al.*, 2023) and world-wide meta-analysis from existing literature (Kristensen *et al.*, 2023; Briganti *et al., 2023*). Authors carry out policy process analyses, employing qualitative, mixed methods and review techniques, combined with comparative analytical descriptive accounts.

Cross-country comparisons of PBF programmes

Two papers compare some specific units of analysis across two countries, capturing differences and similarities in respect to the variables analysed. One of them bring new data collection and analysis on Mali and Burkina Faso (Zitti *et al.*, 2023), while one carries out a review comparing the US and the UK (Andersen, 2023).

Zitti *et al.* (2023) develop a qualitative analysis comparing cases of community verification (CV) processes during the implementation of PBF in Mali and Burkina Faso. Authors carried an inductive type of qualitative analysis based on participatory observation and semi-structured interviews. Results stress that implementation of CV strategies was characterised by challenging issues such as the difficulty in finding certain users in the community due to incorrect spelling of names, lack of telephone details and incomplete information. Moreover, both countries revealed that investigators were used to data fabrication, attributed to having inadequate time or difficulties in reaching certain users. Authors conclude that CV should be simplified and be better adapted to context, and state that CV results need to be employed more carefully for service improvement purposes.

Andersen's (2023) article examines the process of implementing Social Impact Bond (SIB) instruments in the field of healthcare, social care and employment/education, based on a literature review of empirical articles and evaluations of SIB projects completed in the UK and USA. At the same time, the article proposes a conceptual framework for SIB implementation studies, considering different types of barriers that may hinder its implementation, such as adopted SIB models, organizational competencies, data infrastructure, stakeholder engagement and institutional context. The author highlights the importance of investigating this type of instrument, which has increasingly gained attention in developed countries, as a way of incorporating private initiative into public policies.

Intra-country PBF program comparisons

Another block of four papers develops in depth single country studies, breaking the analysis into smaller analytical units in different locations, examining sub-national or sub-regional aspects or administrative variations within a country. They develop cases analysis of original or new data focusing on Turkey (Aktas *et al.*, 2023), India (Srinivasan and Sarin, 2023) and Brazil (Coelho *et al.*, 2023; Dos Anjos *et al.*, 2023).

The paper by Aktas *et al.* (2023) explores how physicians use adaptation and forms of resistance during the implementation of P4P in hospitals in Turkey. The findings highlight that doctors had considerable autonomy to deal with policy restrictions, thus the implementation displayed challenging characteristics such as cherry-picking (high volume, low risk procedures) and pro-social rule breaking. Moreover, policymakers were not aware of the main challenges take place at the front-line level. The article offers knowledge on conflicts between policy and professional roles conflicts, which can emerge from top-down P4P policies.

Srinivasan and Sarin (2023) analyse Community Health Workers' (CHW) prosocial effort within policy implementation of a health/nutrition program in West India using randomized lab-in-the-field inquiry with 344 Anganwadi workers. The study shows that there is no simplistic causation of incentives and effort related to prosocial tasks. For instance, CHWs not receiving PBF compensation tended to invest the greatest number of efforts, but they compromised task quality. Authors conclude that health/nutrition programmes should consider detailed policy issues when offering P4P compensations, such as distribution of products (like health goods) – that can rely on volunteer effort – while volunteer's engagement can be counterproductive for quality focus.

The study by Coelho *et al.* (2023) shed new light on how P4P mechanisms alter frontline staff participation in performance-oriented health programmes in Brazil. The study compares how diverse contexts and policy mechanisms generated different approaches to frontline staff participation during the implementation of a Brazilian P4P programme. Results highlight that sustained participation resulted from better management and integration between frontline health workers in the unit/team. Moreover, critical working conditions led to low levels of motivation.

The paper by dos Anjos *et al.* (2023) analyses the relationship between performance payments for professionals and the improvement of primary health services, based on the Brazilian experience. Based on a cross-sectional analysis, with data from 27.5 thousand Family Health teams, collected at two moments (2013/2019), the authors classified contexts in which there was or was not the use of P4P mechanisms and compared its effects on the quality of services offered to the population. The results demonstrated that, between the two periods studied, there was an improvement in the quality of what the teams developed, but the results of those who received bonuses were not much higher than those of those who did not.

World-wide meta-analysis of PBF

Two papers develop what might be called world-wide meta-analysis, which seeks to synthesise understandings from a global body of literature on P4P. One is based on recognised published cross-country reviews, bringing a global perspective to key issues in PBF between HIC and LMIC settings (Kristensen *et al.*, 2023), while the other performs a type of meta-synthesis qualitative review (Briganti *et al.*, 2023).

Kristensen *et al.* (2023) highlight the heuristic value of generating shared learning by comparing PBF initiatives in different income contexts. Conducting a literature review on the use of PBF in healthcare in high-, middle- and low-income countries, the authors found that the evidence-base produced has grown in quantity and quality, which can help demonstrate

the effects of different PBF designs in different contexts. Although there is still a need for more comparative studies, the authors argue that there is existing evidence to suggest that different income level countries adapt their schemes, allowing for better pooled learning across different contexts.

The final paper by Briganti *et al.* (2023) investigates the effects of P4P schemes on the satisfaction of healthcare professionals. The findings indicate that P4P systems have an impact on professional satisfaction based on the institutional, organizational, geographic and cultural context. In favourable contexts, PBF schemes impact motivation, occupational well-being and work engagement. At the same time, the results indicate the existence of a virtuous/vicious circle in which P4P leads to more satisfaction in contexts where professionals are already satisfied and leads to more dissatisfaction in environments where there is already a lot of dissatisfaction.

Future challenges

Although this special issue starts filling a relevant gap in comparative knowledge related to P4P/PBF, there is still a need to develop comparative analysis between countries across different regions and from different income groups. There is also a lack of studies either employing and testing theories or extracting theoretical lessons related to P4P/PBF, which can then lead to more strategic thinking about PBF's role as a health reform tool (if at all). As argued above, these more theoretical efforts might better capture for whom and under what circumstances PBF works (or does not work). This shift in methodology combined with a wider comparative element would allow for the generation of more nuanced studies of P4P/PBF that could develop a more sophisticated evidence-base, from which to inform policy and practitioners. This improved knowledge base would not only shed light on how to implement a more successful P4P/PBF programme or scheme, but also provide insights on what types of successful management strategies and initiatives exist, particularly in relation to different contextual constraints. Getting a more comprehensive and sophisticated understanding of when P4P/PBF works and for whom is crucial, since reforms associated with P4P/PBF require high levels of motivation and engagement by managers and front-line health professionals (not to mention how programmes translate into health outcomes). Without better understanding of their specific needs, contextual conditions and appropriate measures of success in relation to the "real" performance of PBF, the future of P4P/PBF will remain indeterminant.

Fabiana da Cunha Saddi
Programme of Post-Graduation in Development, Society and International Cooperation, Center for Advanced Multidisciplinary Studies, University of Brasília, Brasilia, Brazil

Stephen Peckham
Centre for Health Services Studies, University of Kent Centre for Health Services Studies, Canterbury, UK

Garrett Wallace Brown
POLIS, University of Leeds, Leeds, UK

Dimitri Renmans
Ecole de Santé Publique, Université Libre de Bruxelles – Campus Erasme, Bruxelles, Belgium, and

Allan Nuno Alves de Sousa
Ministério da Saúde, Brasilia, Brazil

References

Aktas, P., Hammond, J. and Richardson, L. (2023), "Pay-for-performance in healthcare provision: the role of discretion in policy implementation in Turkey", *International Journal of Public Sector Management*, Vol. 36 Nos 6-7, pp. 530-545, doi: 10.1108/IJPSM-12-2022-0282.

Andersen, M.M. (2023), "Barriers to social impact bond implementation: a review of evidence from the UK and US", *International Journal of Public Sector Management*, Vol. 36 Nos 6-7, pp. 512-529, doi: 10.1108/IJPSM-05-2022-0134.

Aoki, N. (2019), "After all these years, what has happened to the international prevalence of NPM-inspired managerial practices?", *International Journal of Public Sector Management*, Vol. 32 No. 4, pp. 403-417.

Appel, I., Lohmann, J., De Allegri, M., Koulidiati, J.-L., Somda, S., Robyn, P.J. and Brenner, S. (2023), "Improving the readiness and clinical quality of antenatal care - findings from a quasi-experimental evaluation of a performance-based financing scheme in Burkina Faso", *BMC Pregnancy Childbirth*, Vol. 23 No. 1, pp. 1-9, (352), doi: 10.1186/s12884-023-05573-x.

Barnes, A., Brown, G.W. and Harman, S. (2015), *Global Politics of Health Reform in Africa: Performance, Participation, and Policy*, Palgrave Macmillan, London.

Bhaskar, R. (2008), *A Realist Theory of Science*, Verso Books, London.

Briganti, P., De Genaro, D., Buonocore, F., Varriale, L. (2023), "Job satisfaction in pay-for-performance healthcare systems: a meta-synthesis of qualitative literature", *International Journal of Public Sector Management*, Vol. 36 Nos 6-7, pp. 606-631.

Coelho, G., Saddi, F.C., Peckham, S., Silva, M.A., Silva, J.D., Barretos, M.L.P., Rocha, G.A.S., Silva, A.E.N. (2023), "Policy mechanisms and types of participation in the implementation of the Brazilian pay-for-performance programme PMAQ", *International Journal of Public Sector Management*, Vol. 36 Nos 6-7, pp. 563-577.

De Allegri, M., Bertone, M.P., McMahon, S., Mounpe Chare, I. and Robyn, P.J. (2018), "Unraveling PBF effects beyond impact evaluation: results from a qualitative study in Cameroon", *BMJ Global Health*, Vol. 3, pp. 1-8, e000693, doi: 10.1136/bmjgh-2017-000693.

de Walque, Damien Kandpal, E., Wagstaff, A., Friedman, J., Neelsen, S., Piatti-Fünfkirchen, M., Sautmann, A., Shapira, G. and Van de Poel, E. (2022), *Improving Effective Coverage in Health: Do Financial Incentives Work?*, Policy Research Report series, World Bank, Washington, doi: 10.1596/978-1-4648-1825-7.

Diaconu, K., Witter, S., Binyaruka, P., Borghi, J., Brown, G.W., Singh, N. and Herrera, C.A. (2022), "Appraising pay-for-performance in healthcare in low- and middle-income countries through systematic reviews: reflections from two teams", *Cochrane Database of Systematic Reviews*, Vol. 5 No. 5, pp. 1-3, ED000157, doi: 10.1002/14651858.ED000157.

Dixon, R. (2021), "Performance management in social impact bonds: how an outcomes-based approach shapes hybrid partnerships", *International Journal of Public Sector Management*, Vol. 34 No. 3, pp. 356-376.

dos Anjos, D.S.O., Scherer, M.D.d.A., Cantalino, J.L.R. and da Silva, E.N. (2023), "Pay for performance in Brazilian primary health care: quality of work processes, actions and services", *International Journal of Public Sector Management*, Vol. 36 Nos 6-7, pp. 578-591, doi: 10.1108/IJPSM-01-2023-0022.

Dunleavy, P., Margetts, H., Bastow, S. and Tinkler, J. (2006), "New public management is dead–long live digital-era governance", *Journal of Public Administration Research and Theory*, Vol. 16 No. 3, pp. 467-494, doi: 10.1093/jopart/mui057.

Fernandes, G. and Sridhar, D. (2017), "World Bank and the global financing facility", *British Medical Journal*, Vol. 358, pp. 1-4, j3395 doi: 10.1136/bmj.j3395.

Gage, A. and Bauhoff, S. (2021), "The effects of performance-based financing on neonatal health outcomes in Burundi, Lesotho, Senegal, Zambia and Zimbabwe", *Health Policy and Planning*, Vol. 36 No. 3, pp. 332-340, doi: 10.1093/heapol/czaa191.

Gautier, L., Tosun, J., De Allegri, M. and Ridde, V. (2018), "How do diffusion entrepreneurs spread policies? Insights from performance-based financing in Sub-Saharan Africa", *World Development*, Vol. 110, pp. 160-175, doi: 10.1016/j.worlddev.2018.05.032.

Gautier, L., De Allegri, M. and Ridde, V. (2019), "How is the discourse of performance-based financing shaped at the global level? A poststructural analysis", *Globalization and Health*, Vol. 15 No. 6, pp. 1-21, doi: 10.1186/s12992-018-0443-9.

Hur, H. (2018), "The unintended consequences of a pay-for-performance rule change: US Department of Defense's National Security Personnel System", *International Journal of Public Sector Management*, Vol. 31 No. 6, pp. 654-671.

Ingraham, P.W. (1993), "Of pigs in pokes and policy diffusion: another look at pay-for-performance", *Public Administration Review*, Vol. 53 No. 4, pp. 348-356.

Kadungure, A., Brown, G.W., Loewenson, R. and Gwati, G. (2021), "Adapting results-based financing to respond to endogenous and exogenous moderators in Zimbabwe", *Journal of Health Organization and Management*, Vol. 35 No. 3, pp. 287-305, doi: 10.1108/JHOM-06-2020-0215.

Krauss, A. (2018), "Why all randomised controlled trials produce biased results", *Annals of Medicine*, Vol. 50 No. 4, pp. 312-322, doi: 10.1080/07853890.2018.1453233.

Kristensen, S.R., Anselmi, L., Brown, G.W., Fichera, E., Kovacs, R., Loewenson, R., Singh, N., Midzi, N., Mustapha, F., White, L. and Borghi, J. (2023), "Pay for performance at a crossroads: lessons from taking a global perspective", *International Journal of Public Sector Management*, Vol. 35 Nos 6-7, pp. 592-605, doi: 10.1108/IJPSM-03-2023-0084.

Kutzin, J. (2016), "RBF: from program to entry point for strategic purchasing", *Paper Presented at the: Sharing and Debating Country Experiences on Health Financing Reform, Health Financing Technical Network Meeting*, Geneva, 14 December.

Lester, H. and Campbell, S. (2010), "Developing quality and outcomes framework (QOF) indicators and the concept of 'QOFability'", *Quality in Primary Care*, Vol. 18 No. 2, pp. 103-109.

Lohmann, J., Wilhelm, D., Kambala, C., Brenner, S., Muula, A.S. and De Allegri, M. (2018), "The money can be a motivator, to me a little, but mostly PBF just helps me to do better in my job.' an exploration of the motivational mechanisms of performance-based financing for health workers in Malawi", *Health Policy and Planning*, Vol. 33 No. 2, pp. 183-191, doi: 10.1093/heapol/czx156.

Marchal, B., Westhorp, G., Wong, G., Van Belle, S., Greenhalgh, T., Kegels, G. and Pawson, R. (2013), "Realist RCTs of complex interventions – an oxymoron", *Social Science & Medicine*, Vol. 94, pp. 124-128, doi: 10.1016/j.socscimed.2013.06.025.

Mathauer, I., Dale, E., Jowett, M. and Kutzin, J. (2019), *Purchasing Health Services for Universal Health Coverage: How to Make it More Strategic?*, World Health Organization, Health Financing Policy Brief, No. 6, pp. 1-21.

McIsaac, M., Kutzin, J., Dale, E. and Soucat, A. (2018), "Results-based financing in health: from evidence to implementation", *Bulletin of the World Health Organization*, Vol. 11, pp. 730-730A, doi: 10.2471/BLT.18.222968.

Merton, R.K. (1949), "On sociological theories of the middle range", in Merton, R.K. (Ed.), *Social Theory and Social Structure*, Simon & Schuster, the Free Press, New York, pp. 39-53.

Olivier de Sardan, J.P., Diarra, A. and Moha, M. (2017), "Travelling models and the challenge of pragmatic contexts and practical norms: the case of maternal health", *Health Research Policy and Systems*, Vol. 15 No. Suppl 1, No. 60, doi: 10.1186/s12961-017-0213-9.

Osborne, S. (2006), "New public governance?", *Public Management Review*, Vol. 8 No. 3, pp. 377-387.

Paul, E. and Ridde, V. (2022), "Results-based financing (RBF) in Africa. A discussion of a travelling model", *Revue internationale des études du développement*, Vol. 248, pp. 169-204, doi: 10.4000/ried.334.

Pawson, R. and Tilley, N. (1997), *Realistic Evaluation*, Sage, London.

Ragin, C.C. (2009), *Redesigning Social Inquiry: Fuzzy Sets and beyond*, University of Chicago Press, Chicago.

Ragin, C.C. (2014), *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*, University of California Press, Berkeley, CA.

Renmans, D., Holvoet, N., Orach, C.G. and Criel, B. (2016), "Opening the 'black box' of performance-based financing in low- and lower middle-income countries: a review of the literature", *Health Policy and Planning*, Vol. 31 No. 9, pp. 1297-1309, doi: 10.1093/heapol/czw045.

Renmans, D., Holvoet, N., Criel, B. and Meessen, B. (2017), "Performance-Based Financing: the same is different", *Health Policy and Planning*, Vol. 32 No. 6, pp. 860-868, doi: 10.1093/heapol/czx030.

Renmans, D., Holvoet, N. and Criel, B. (2020), "No mechanism without context: strengthening the analysis of context in realist evaluations using causal loop diagramming", *New Directions for Evaluation*, Vol. 167, pp. 101-114, doi: 10.1002/ev.20424.

Rosenthal, M.B., Landon, B.E., Normand, S.L.T., Frank, R.G. and Epstein, A.M. (2006), "Pay for performance in commercial HMOs", *New England Journal of Medicine*, Vol. 355 No. 18, pp. 1895-1902.

Saddi, F.C., Peckham, S. and Santos Mundim, P. (2018), "Presentation: pay for performance, formulation, implementation and policy success: in low, middle and high income countries", *Sociedade E Cultura*, Vol. 21 No. 2, doi: 10.5216/sec.v21i2.56304.

Saddi, F.C., Forbes, L.J.L. and Peckham, S. (2021a), "Guest editorial", *Journal of Health Organization and Management*, Vol. 35 No. 3, pp. 245-248, doi: 10.1108/JHOM-05-2021-476.

Saddi, F.C., Harris, M., Parreira, F.R., Pêgo, R.A., Coelho, G.A., Lozano, R.B., Mundim, P.d.S. and Peckham, S. (2021b), "Exploring frontliners' knowledge, participation and evaluation in the implementation of a pay-for-performance program (PMAQ) in primary health care in Brazil", *Journal of Health Organization and Management*, Vol. 35 No. 3, pp. 327-343, doi: 10.1108/JHOM-04-2020-0154.

Saddi, F.C., Peckham, S., Coelho, G.A. and Lloyd-Sherlock, P. (2023), "A framework for measuring the effects of policy processes on health system strengthening", in Lewis, J.M. and Triantafillou, P. (Eds), *Handbook on Measuring Governance*, [Book submitted for publication. Edward Elgar Publishing] Forthcoming in December 2023.

Schwartz, R. and Deber, R. (2016), "The performance measurement–management divide in public health", *Health Policy*, Vol. 120 No. 3, pp. 273-280.

Sieleunou, I., De Allegri, M., Roland Enok Bonong, P., Ouédraogo, S. and Ridde, V. (2020), "Does performance-based financing curb stock-outs of essential medicines? Results from a randomised controlled trial in Cameroon", *Tropical Medicine & International Health*, Vol. 25 No. 8, pp. 944-961, doi: 10.1111/tmi.13447.

Singh, N.S., Kovacs, R.J., Cassidy, R., Kristensen, S.R., Borghi, J. and Brown, G.W. (2021), "A realist review to assess for whom, under what conditions and how pay for performance programmes work in low- and middle-income countries", *Social Science & Medicine*, Vol. 270, pp. 1-17, 113624, doi: 10.1016/j.socscimed.2020.113624.

Srinivasan, S. and Sarin, A. (2023), "Frontline workers' performance in prosocial tasks: evidence from a lab-in-the-field in western India", *International Journal of Public Sector Management*, Vol. 36 Nos 6-7, pp. 546-562.

Tawfiq, E., Desai, J. and Hyslop, D. (2019), "Effects of results-based financing of maternal and child health services on patient satisfaction in Afghanistan", *Journal of Health Services Research & Policy*, Vol. 24 No. 1, pp. 4-10, doi: 10.1177/1355819618802944.

Victora, C.G., Habicht, J.P. and Bryce, J. (2004), "Evidence-based public health: moving beyond randomized trials", *American Journal of Public Health*, Vol. 94 No. 3, pp. 400-405, doi: 10.2105/AJPH.94.3.400.

Witter, S., Bertone, M.P., Namakula, J., Chandiwana, P., Chirwa, Y., Ssennyonjo, A. and Ssengooba, F. (2019), "(How) does RBF strengthen strategic purchasing of health care? Comparing the

experience of Uganda, Zimbabwe and the Democratic Republic of the Congo", *Global Health Research and Policy*, Vol. 4 No. 3, doi: 10.1186/s41256-019-0094-2.

World Bank (2023), "The pandemic fund results framework", World Bank, WA, available at: https://thedocs.worldbank.org/en/doc/eac1acfe37285a29942e9bb513a4fb43-0200022022/related/PF-Results-Framework-with-Indicator-Reference-Sheets-Feb-09-2023.pdf

World Health Organization (2017), "Strategic purchasing for universal health coverage: unlocking the potential", *World Health Organization Global Meeting Summary and Key Messages*, Geneva, 25-27 April, available at: https://cdn.who.int/media/docs/default-source/health-financing/strategic-purchasing/strategic-purchasing-meeting-2017-report.pdf?sfvrsn=54868676_4 (accessed 26 September 2023).

Yin, R.K. (1994), *Case Study Research: Design and Methods*, Sage Publications, Thousand Oaks, CA, United States.

Zitti, T., Coulibaly, A., Gali-Gali, I.A.Z., Ridde, V. and Turcotte-Tremblay, A.-M. (2023), "A comparative study of community verification processes in the context of performance-based financing in Mali and Burkina Faso", *International Journal of Public Sector Management*, Vol. 36 Nos 6-7, pp. 492-511, doi: 10.1108/IJPSM-02-2023-0063.

Further reading

Archer, M.S. (1995), *Realist Social Theory: the Morphogenetic Approach*, Cambridge University Press, Cambridge.

Diaconu, K., Falconer, J., Verbel, A., Fretheim, A. and Witter, S. (2021), "Paying for performance to improve the delivery of health interventions in low- and middle-income countries", *Cochrane Database of Systematic Reviews*, Vol. 5 No. 5, pp. 1-266, Cd007899, doi: 10.1002/14651858.CD007899.pub3.

Saddi, F.C, Peckham, S., Coelho, G., Silva, J.D., Barretos, M., Rocha, G. and Novais, A. (2019), "Exploring how streams-advocacy and learning-feedback relations can affect policy capacity and re-design: a framework applied to uncover the re-formulation process of a Brazilian program (PMAQ, 3rd round)". *Paper presented at the International Conference of Public Policy*, 26-28 June, Montreal, available at: https://www.ippapublicpolicy.org//file/paper/5d0b8e14d8539.pdf (accessed 20 September 2023).

Skiles, M.P., Curtis, S.L., Basinga, P. and Angeles, G. (2013), "An equity analysis of performance-based financing in Rwanda: are services reaching the poorest women?", *Health Policy and Planning*, Vol. 28 No. 8, pp. 825-837, doi: 10.1093/heapol/czs122.

Soucat, A., Dale, E., Mathauer, I. and Kutzin, J. (2017), "Pay-for-Performance debate: not seeing the forest for the trees", *Health System & Reform*, Vol. 3 No. 2, pp. 74-79, doi: 10.1080/23288604.2017.1302902.