Université de Montréal

The Unintended Consequences of a Complex Intervention Combining Performance-Based Financing with Health Equity Measures in Burkina Faso

Par Anne-Marie Turcotte-Tremblay

Département de médecine sociale et préventive
École de santé publique de l'Université de Montréal (ESPUM)

Thèse présentée en vue de l'obtention du grade de Philosophiae doctor –
Doctorat (Ph.D.) en santé publique, option organisation des soins

Mars, 2020

Université de Montréal
Département de médecine sociale et préventive,
École de santé publique de l'Université de Montréal (ESPUM)

Cette thèse intitulée

The Unintended Consequences of a Complex Intervention Combining Performance-Based Financing with Health Equity Measures in Burkina Faso

Présentée par

Anne-Marie Turcotte-Tremblay

A été évaluée par un jury composé des personnes suivantes

Pierre Fournier
Président-rapporteur

Valéry Ridde
Directeur de recherche

Manuela De Allegri
Codirectrice

Roxane Borgès Da Silva
Membre du jury

Maman Joyce Dogba
Examinatrice externe

Résumé

Contexte : La mauvaise qualité et la faible utilisation des services de santé contribuent aux taux élevés de morbidité et de mortalité dans plusieurs pays à faible et à moyen revenu. Face à cette situation, le gouvernement du Burkina Faso a testé une intervention novatrice qui combine le financement basé sur la performance (FBP) à des mesures d'équité en santé. Les formations sanitaires ont reçu des prix unitaires pour des services de santé fournis ainsi que des bonus conditionnels à la qualité des soins. Des comités communautaires ont sélectionné les indigents pour leur octroyer des exemptions de paiements des soins. Malgré le peu d'études sur le sujet, des acteurs en santé mondiale craignent que l'intervention puisse avoir des conséquences non intentionnelles importantes.

Objectif : Cette thèse vise à accroître les connaissances scientifiques sur les conséquences non intentionnelles du FBP combiné à des mesures d'équité en santé dans un environnement à faible revenu.

Méthodes : Nous avons développé un cadre conceptuel basé sur la théorie de la diffusion des innovations. Une étude de cas multiples a été réalisée avec neuf formations sanitaires au Burkina Faso. Cinq mois sur le terrain ont permis d'effectuer 104 entrevues semi-structurées, 266 séances d'observation et des conversations informelles avec un large éventail d'acteurs incluant les prestataires de soins, les patients et les vérificateurs. Les données qualitatives ont été codées avec QDA miner pour faciliter l'analyse thématique. Nous avons également utilisé des données quantitatives du système de gestion pour décrire l'évolution des services et trianguler les résultats.

Résultats : La nature et la mise en œuvre de l'intervention ont interagi avec le système social et les caractéristiques de ses membres pour engendrer des conséquences non intentionnelles importantes, dont la plupart étaient indésirables. Les prestataires de soins ont démontré une fixation sur les mesures de rendement, ont falsifié les registres médicaux et ont enseigné de mauvaises pratiques aux stagiaires pour augmenter leurs subsides et bonus. Comme conséquence non intentionnelle désirable, certaines formations sanitaires ont limité la vente de médicaments sans prescriptions pour encourager les consultations. Les vérifications communautaires, durant lesquelles les patients sont retrouvés pour vérifier les services déclarés, ont entraîné la falsification des données de vérification, la perte de la confidentialité des patients et certaines craintes chez les patients, bien que certains étaient heureux de partager leurs opinions. Enfin, les prestataires de soins ont limité les services offerts gratuitement aux indigents, ce qui a déclenché des conflits.

Discussion : Cette thèse contribue au développement des connaissances scientifiques sur la façon dont le FBP, combiné à des mesures d'équité, peut engendrer des conséquences non intentionnelles. Les résultats sont utiles pour affiner ce type d’intervention et éclairer une mise en œuvre efficace dans le secteur du financement de la santé. Plus largement, cette thèse démontre la faisabilité et la valeur ajoutée d'utiliser un cadre conceptuel pour étudier les conséquences non intentionnelles. Elle pourra guider les chercheurs à élargir leur angle d’analyse afin de rendre compte des conséquences intentionnelles et non intentionnelles des interventions complexes en santé.

Mots-clés : Financement basé sur la performance, vérification, satisfaction des patients, exemption du paiement des soins, indigents, conséquences non intentionnelles, théorie de la diffusion des innovations, étude de cas multiple, Burkina Faso

Abstract

Background: Poor quality and low utilization of healthcare services contribute to high levels of morbidity and mortality in many low- and middle-income countries (LMICs). In response, the government of Burkina Faso tested an innovative intervention that combines performance-based financing (PBF) with health equity measures. Healthcare facilities received unit fees for targeted services and bonuses conditional upon the quality of care. To reduce inequities in access to care, community-based committees selected indigents, i.e., the poorest segment of the population, to offer them user fee exemptions. Facilities were also paid more for services delivered to indigents. Despite the potential of this type of intervention, many global health actors argue that it could lead to important unintended consequences that influence its overall impact. Yet, little attention has been given to studying the unintended consequences of this complex intervention.

Objective: This thesis aims to increase the scientific knowledge on the unintended consequences of PBF combined with health equity measures in a low-income setting.

Methods: We developed a conceptual framework based on the diffusion of innovations theory. Using a multiple case study design, we selected nine healthcare facilities in Burkina Faso. Over five months of fieldwork, we collected multiple sources of qualitative data including 104 semi-structured interviews, 266 recorded observation sessions, informal conversations and documentation. Participants included a wide range of stakeholders, such as providers, patients, and PBF verifiers. Data were coded using QDA miner to conduct a thematic analysis. We also used secondary data from the PBF routine management system to describe the evolution of services and triangulate results.

Results: Interactions between the nature and implementation of the intervention, the nature of the social system, and its members' characteristics led to important unintended consequences, most of which were undesirable. Providers were fixated on performance measures rather than on underlying objectives, falsified medical registers, and taught trainees improper practices to increase subsidies and bonuses. As a desirable unintended consequence, we found that some facilities limited the sale of non-prescribed medication to encourage patients to consult. Community verifications, in which patients are traced to verify the authenticity of reported services and patient satisfaction, also led to unintended consequences, such as the falsification of verification data, the loss of patient confidentiality, and fears among patients, although some were pleased to share their views. Lastly, health equity measures also triggered changes that were not intended by program planners. For example, providers limited the free services and medication delivered to indigents, which led to conflicts between parties.

Discussion: This thesis contributes to the development of scientific knowledge on how PBF interventions, combined with equity measures, can trigger unintended consequences in a low-income setting. The results are useful to inform effective implementation and refine interventions, particularly in the health financing sector. More broadly, this thesis demonstrates the feasibility and added value of using a conceptual framework to study the unintended consequences of complex health interventions. This thesis can inspire and guide future researchers to broaden their analytical horizons to capture both intended and unintended consequences of health interventions.

Keywords: Performance-based financing, Verification, Patient satisfaction, User fee exemption, Indigent, Unintended consequences, Diffusion of innovations theory, Multiple-case study, Burkina Faso

Table of content

List of tables

List of figures

List of acronyms and abbreviations

AEDES: *Agence Européenne pour le Développement et la Santé* / European Agency for Development and Health

ANC: ante-natal care

CBHI: community-based health insurance

CHE: catastrophic health expenditure

CHW: community health worker

CMA: *centre médical avec antenne chirurgicale* / medical centre with surgical satellite

CVA: contractualization and verification agency

COGES: *comités de gestion* / healthcare facilities management committees

CSPS: *centre de santé et de promotion social* / health and social promotion centre

DAC: Development Assistance Committee (of the OECD)

DEReC: DAC Evaluation Resource Centre

DRC: Democratic Republic of the Congo

FBP : *Financement basé sur la performance*

F CFA: CFA franc

GHR-CAPS: Global Health Research Capacity Strengthening Program

HIC: high-income country

HRITF: Health Results Innovation Trust Fund

HSPEP: Health Sector Performance Enhancement Project

IMCI: integrated management of childhood illnesses

LIC: low-income country

LMIC: low- and middle-income country

MRC: Medical Research Council

NGO: non-governmental organization

NPM: New Public Management

OECD: Organisation for Economic Co-operation and Development

OPP: out-of-pocket payment

P4P: pay for performance

PADS: *Programme d’appui au développement en santé* / program to support health development

PAHO: Pan American Health Organization

PAP: *plan d’amélioration de la performance*

PBF: performance-based financing

PBI: performance-based incentive

PFMR: *pays à faible et à moyen revenu*

QED: quasi-experimental design

QOF: Quality and Outcomes Framework

RBF: results-based financing

RCH: reproductive and child health

RCT: randomized control trial

REALISME : *REcheche AppLiquée Interventionnelles en Santé Mondiale et Équité*

SSA: sub-Saharan Africa

SSDI-PBI: Support for Service Delivery Integration Performance-Based Incentives

ST-FBR: *Service technique-FBR* / results-based financing technical service

UHC: universal healthcare coverage

UNDP: United Nations Development Programme

USAID: United States Agency for International Development

WHO: World Health Organization

I dedicate this thesis to the people who participated in the study in Burkina Faso,
some of whom passed away too young.
May the findings help us "leave no one behind".

Acknowledgements

If “it takes a village to raise a child”, then it must be fair to say it takes a society to produce a doctoral thesis. It is simply remarkable how many people, organizations, and places came together to produce this long-term achievement. I am grateful and forever indebted to the people and organizations below for their continuous support throughout this life-changing journey.

- Thanks to Dr. Valéry Ridde, my supervisor and mentor, who broadened my horizons by exposing me to a new research environment and new methods. His productivity, commitment, and passion make him a force of nature. Among his positive influences on my career was his encouragement to publish on the different angles of my doctoral study (e.g., theories, methods, ethics, reflexivity), demystifying the publication process.

- Thanks to Dr. Manuela De Allegri, my co-supervisor and mentor, who always encouraged me to improve the quality of this research. Her constant availability, rigour, and humanity made her irreplaceable throughout this doctoral study.

- Thanks to my collaborators, particularly Idriss Ali Gali Gali, and all the participants in Burkina Faso, without whom this study would simply not have been possible. They welcomed me with tremendous generosity, shared their experiences, and touched my soul forever.

- Thanks to all my family members, especially my parents, Lucie Turcotte and André Tremblay, and my spouse, Jason Pitre, for their continuous support of my passion for global health research.

- Thanks to all my colleagues and friends from *Université de Montréal*, *Thèsez-vous?, Thèsez-vous ENSEMBLE,* and the *Global Health Research Capacity Strengthening Program* for making this journey so enriching on both the personal and professional levels.

- Thanks to the co-authors of the publications included in this thesis. It was a pleasure to develop ideas, analyse and reflect together.

- Thanks to all the professors and researchers who encouraged me and inspired me to pursue a career in research.

- Thanks to the following organizations for their financial support: Canadian Institutes of Health Research (CIHR), *Fonds de recherche du Québec – Santé (FRQS)*, International Development Research Centre (IDRC), Global Health Research Capacity Strengthening Program (GHR-CAPS), Population Health Intervention Research Network (PHIRNET), *Institut de recherche en santé publique de l'Université de Montréal*, *Centre de recherche du Centre hospitalier de l'Université de Montréal*, *Faculté des études supérieures et postdoctorales,* and *Fédération des associations étudiantes du campus de l'Université de Montréal*.

- Finally, thanks to the jury members, reviewers and editors who generously contributed their time and expertise.

Together, you all enriched the content of this thesis.

Chapter 1 – Introduction

Health is a fundamental right of every human being (United Nations, 1948). This right entails that healthcare services must be available, accessible, acceptable and of good quality for everyone (The Right to the Highest Attainable Standard of Health, 2000). Every state in the world has ratified at least one international human rights treaty recognizing the right to health (United Nations, 2008). More recently, the practical expression of this right has been conveyed through universal health care coverage (UHC) (Ooms et al., 2014). In 2005, member states of the World Health Assembly took an oath to promote UHC by providing all people with access to quality health services and ensuring that the use of these services does not expose users to financial hardship (World Health Organization, 2005, 2010). This momentum led to the 2030 Agenda for Sustainable Development. Its goal 3.8 is to "*achieve universal health coverage, including financial risk protection, access to quality essential health-care services and access to safe, effective, quality and affordable essential medicines and vaccines for all*" (United Nations, 2015, p. 16).

Despite concerted efforts of the global health community, universal health coverage remains a major challenge, especially in low- and middle-income countries (LMICs). At least half of the global population do not have full coverage for essential health services (WHO, 2018b). According to the World Health Statistics Report (2018b), more than 40% of all pregnant women are not receiving early antenatal care globally. In many LMICs, more than half of all births are not assisted by skilled health personnel, compared to less than 10% in most high-income countries (WHO, 2018b). These performance gaps within health systems contribute to preventable deaths. In 2015, 303 000 women died due to complications of pregnancy or childbirth (WHO, 2018b). Almost all of these deaths (99%) were in LMICs, with almost two-thirds (64%) occurring in the WHO African Region.

Of those who manage to access healthcare services, many endure financial hardship. In 2010, 11.7% of the world's population spent at least 10% of their household budget on healthcare services (WHO, 2018b). The incidence of catastrophic health expenditures (CHE) is higher in low-income countries that require out-of-pocket payments (OPP), as it is often the case in sub-Saharan African (SSA) countries (Njagi et al., 2018). Studies show that socio-economic factors drive CHEs (Njagi et al., 2018). For example, a household's income level is a consistent determinant of CHEs, with higher-income groups being less likely to incur CHE than are middle- or lower-income groups (Njagi et al., 2018).

1.1 The expansion of performance-based financing in LMICs

Global health actors are experimenting with a wide array of innovations, that is, new practices, to accelerate progress towards UHC. In the field of healthcare financing, new purchasing mechanisms are being designed and implemented to transfer funds in ways that can encourage the production and distribution of high-quality healthcare. In recent years, results-based financing (RBF) is one high-profile intervention that has been widely implemented in LMICs to shift from input-based financing to output-based financing. RBF is an umbrella term defined as "*a cash payment or non-monetary transfer made to a national or sub-national government, manager, provider, payer or consumer of health services after predefined results have been attained and verified.*" (Musgrove, 2011, p. 1). The incentives can target both the supply and demand sides. A subset of RBF is called performance-based financing (PBF). With this approach, "*health care providers are paid for delivering specific services, provided the services follow explicit protocols, with a system of inspection and auditing to assure compliance and to raise quality where necessary. Performance-based payments are also provided for the teams that carry out these inspections, to motivate them to be thorough and accurate*" (Musgrove, 2011, p. 4). According to WHO (2020b), PBF has the following distinguishing features: 1) incentives are directed only to providers, not beneficiaries; 2) awards are purely financial; and 3) payments are conditional on performance, often defined in terms of process or output indicators, adjusted by some measure of quality. Underlying PBF is the assumption that financial incentives at the individual and organizational levels, combined with increased accountability, managerial autonomy, separation of functions, planning, monitoring, and feedback, will provide actors of the healthcare system with the motivation and resources necessary to improve the quantity and quality of services provided (Fritsche et al., 2014). Different PBF models bearing various labels (e.g., pay for performance [P4P], performance-based incentives) have been implemented across LMICs to strengthen healthcare systems (Eldridge & Palmer, 2009; Musgrove, 2011). Musgrove (2011) provides a detailed review of the different concepts and terms, although many are used interchangeably in practice.

1.2 Combining equity measures with PBF in Burkina Faso

To date, however, few PBF models have integrated and tested specific health equity measures that target vulnerable groups (Ridde et al., 2018). In fact, a review of 23 PBF implementation manuals found that only six documents addressed issues of accessibility, while seven addressed either equity or non-discrimination (Cole et al., 2019). Moreover, research shows that PBF is not inherently pro-poor or well-suited to tackling health inequities (Lannes et al., 2015; Priedeman Skiles et al., 2013; Ridde et al., 2018). Thus, to achieve universal health coverage by 2030, additional measures may be necessary to adopt comprehensive approaches that simultaneously tackle the supply and demand sides of care.

Weighing these considerations, the government of Burkina Faso, supported by the World Bank, tested a new intervention that combined PBF with health equity measures. This complex intervention comprised four complementary components targeting both the demand and supply sides: 1) pay healthcare facilities according to the quantity and quality of healthcare services provided; 2) pay healthcare facilities more for services delivered to the poor; 3) exempt the poorest 20% of the population from paying user fees; and 4) provide community-based health insurance (CBHI). Due to its innovative nature, this intervention provides a unique opportunity to produce knowledge on the implementation processes and consequences of combining PBF with health equity measures in a real-life setting.

1.3 Could this complex intervention trigger unintended consequences?

Despite its potential to spark change within health systems, many global health actors are concerned that the processes and outcomes of PBF may not be as simple and linear as envisioned by promoters and program planners (Paul et al., 2018; Renmans et al., 2017). Some hypothesize that PBF may lead to important unintended consequences that go well beyond the targeted objectives of the intervention, thereby influencing its overall effectiveness (Fretheim et al., 2012; Kalk et al., 2010). The concept of unintended consequences, which will be further defined and operationalized in the next chapters, refers to changes brought by an intervention other than those it aimed to achieve (Jabeen, 2016). Those can be desirable or undesirable changes, depending on the perspective of the stakeholders. They can affect various actors, such as patients, providers, community members, healthcare managers, and government officials. Such concerns continue to fuel policy debates within the global health community (Ma-Nitu et al., 2018; Paul et al., 2018).

Yet few empirical studies have been conducted to investigate whether and how PBF interventions combined with equity measures can trigger unintended consequences in LMICs. For example, the World Bank funded impact evaluations of PBF interventions in countries like Burkina Faso, Benin and Cameroon without explicitly aiming to assess the unintended consequences that emerged (Lemière et al., 2015; Souares et al., 2013; The World Bank, 2012). More broadly, this reflects a general shortcoming in evaluation and research practices. Impact evaluations tend to focus on effectiveness, defined as "*the extent to which a given development intervention's objectives were achieved, or are expected to be achieved…*" (OECD, 2002, p. 20). Reflecting on current evaluation practices, Heider (2017) explains that, "*the way effectiveness has been defined has kept attention focused on intended results. Most evaluations grapple with getting evidence to determine whether objectives were achieved and to measure an intervention's contributions. Fewer evaluations are able to collect evidence on effects outside the immediate results chain and identify unintended consequences*."

Numerous explanations have been put forward to explain the lack of attention paid to unintended consequences. These include: the common assumption that innovations or development efforts produce beneficial results (i.e., pro-innovation bias, paternalistic bias); conflicts of interests of funders; time and budget constraints; the training of most practitioners and researchers being aimed mainly at assessing effectiveness; and difficulties in predicting, measuring and responding to unintended consequences (Bamberger et al., 2016; Jabeen, 2016; McQueen, 2014; Morell, 2018; Oliver et al., 2019; Rogers, 2003).

There are numerous reasons for studying the unintended consequences of PBF in LMICs. First, the likelihood that a complex intervention like PBF would trigger desirable or undesirable unintended consequences going well beyond the objectives of the intervention is high. There is much uncertainty about how interventions with multiple interacting components that target multiple groups and organizational levels will actually unfold in complex systems such as the healthcare system (Greenhalgh & Papoutsi, 2018; Rogers, 2003; Woolcock, 2013). According to Woolcock (2013), the uncertainty surrounding complex interventions is related to the numerous pathways and feedback loops connecting inputs, actions, and outcomes, the intervention's exposure to exogenous influences, and the capacity of the people involved to exercise discretion (i.e., to act independently of or in accordance with rules, expectations, precedent, passions, professional norms, or self-interest). Greenhalgh and Papoutsi (2018) argue that engaging with complexity requires a new paradigm that carefully considers unintended consequences.

Another reason for studying unintended consequences is that the findings can inform decision-makers. Although they may be less discernible, unintended consequences may be far-reaching and just as important as the intended consequences. To be able to judge the overall value of an innovative intervention, stakeholders must have a comprehensive understanding of both its intended and unintended consequences. In the past, fear of potential unintended consequences has impeded the diffusion of promising innovations (Ash, Sittig, Poon, et al., 2007). Moreover, funding agencies that ignore unintended consequences may continue to support programs that are not achieving their objectives, or are doing so in a less efficient way (Bamberger et al., 2016). With full knowledge of the evidence, program developers and implementers may be able to plan more effective interventions by capitalizing on desirable unintended consequences or by altering their strategy to mitigate undesirable ones (Jabeen, 2016).

There are also ethical reasons to study the unintended consequences of an intervention like PBF (McQueen, 2014). According to Mittelmark (2014), large-scale interventions that inject a new agenda, money, and people into a setting might disturb it in unplanned ways. Some of the effects may be seriously untoward. Jabeen (2016) shows that well-intended development programs or policies (e.g., neo-liberal policies in Latin America) sometimes worsen the conditions of beneficiaries, which directly contradicts the core purpose of development aid, which is to improve the situation of the poor in LMICs. Thus, monitoring the environment for unplanned effects is the outsider's minimal ethical obligation, to ensure that interventions do not cause harm to populations (Jabeen, 2016; Mittelmark, 2014).

1.4 Objectives of the thesis

There is an urgent need to draw attention to the study of unintended consequences of complex health interventions. PBF’s rapid expansion across LMICs provides a unique opportunity to evaluate whether this complex intervention triggers unintended consequences.

The overall objective of this thesis is to increase scientific knowledge on the unintended consequences of an innovative intervention that combines PBF with health equity measures in a low-income setting. To this end, we adopted a multiple case study design to examine the PBF intervention implemented in healthcare facilities in Burkina Faso. More specifically, we pursued three aims, each of which is the focus of an article presented in the results section:

1. to analyze the unintended consequences of the overall PBF program in primary healthcare facilities in Burkina Faso;
2. to decipher the unintended consequences specific to community verifications and client satisfaction assessments conducted for the PBF program; and
3. to examine the unintended consequences of selecting the poorest 20% of the population through a community-based process and offering them user-fee exemptions while paying healthcare facilities more for services delivered to the poor.

1.5 Structure of the thesis

This thesis is organized as follows: This first chapter reviews the general background and objectives of the thesis. The second chapter defines the concepts used in this thesis and reviews the state of knowledge: 1) the unintended consequences of complex interventions; 2) the role that global health actors played in the rapid expansion of PBF in LMICs (Article 1); and 3) the unintended consequences of PBF combined with equity measures. The third chapter lays out our theoretical model and research questions, and the fourth presents our scientific endeavour and methods employed. The fifth chapter reports the findings in the form of three scientific articles published in peer-reviewed journals (referred to as Articles 2, 3 and 4). Each article addresses one of the three aims of this thesis by focusing on different dimensions of the intervention (i.e., PBF in primary healthcare facilities, community verifications, and equity measures). The sixth chapter discusses: 1) the advancement of empirical, theoretical and methodological knowledge, including a reflection on the oscillating role of researchers during observation sessions (Article 5); 2) cross-cutting themes of the thesis; 3) strengths and limitations of the study; 4) knowledge translation strategies, including reflections on ethical challenges encountered during the dissemination of the findings locally (Article 6); and 5) the practical implications of the thesis for research and practice. The seventh chapter offers concluding remarks. The thesis is partially based on the publications presented in Table I.

Table I. List of publications included in the thesis

#	Section	Title, Authors, Journal, Year
1	**Literature review**	**The role of global health actors in the expansion of performance-based financing in low- and middle-income countries** [translation] A.-M. Turcotte-Tremblay, L. Gautier, O. Bodson, N. E. Sambieni, and V. Ridde *Journal de Gestion et d'Économie Médicales, 2018*
2	**Results**	**An exploration of the unintended consequences of performance-based financing in 6 primary healthcare facilities in Burkina Faso** A.-M. Turcotte-Tremblay, I. A. Gali Gali, and V. Ridde *International Journal of Health Policy and Management, 2020*
3	**Results**	**The unintended consequences of community verifications for performance-based financing in Burkina Faso** A.-M. Turcotte-Tremblay, I. A. Gali Gali, M. De Allegri and V. Ridde *Social Science & Medicine, 2017*
4	**Results**	**The unintended consequences of combining equity measures with performance-based financing in Burkina Faso** A.-M. Turcotte-Tremblay, M. De Allegri, I. A. Gali Gali and V. Ridde *International Journal for Equity in Health, 2018*
5	**Discussion**	**Oscillating between passive and active roles during non-participant observation in global health research** A.-M. Turcotte-Tremblay *Cahiers REALISME, 2018*
6	**Discussion**	**A reflection on the challenge of protecting confidentiality of participants while disseminating research results locally** A.-M. Turcotte-Tremblay and E. Mc Sween-Cadieux *BMC Medical Ethics, 2018*
7	**Appendix**	**Does performance-based financing increase value for money in low- and middle-income countries? A systematic review** Turcotte-Tremblay, A.-M., J. Spagnolo, M. De Allegri and V. Ridde *Health Economics Review*, *2016*

Chapter 2 – Literature Review

2.1 Unintended consequences

2.1.1 Defining the terminology

Robert K. Merton (1936) is recognized as the first social scientist to have systematically analyzed the issue of “unanticipated consequences” (Norton, 2008). In his seminal article, “The Unanticipated Consequences of Purposive Social Action”, Merton deplored the diversity of terms used to address the issue, which “*tended to obscure the definite continuity of its consideration*” (Merton, 1936, p. 2). More than 80 years later, the terminology used to represent this research area continues to vary greatly in the scientific literature, indicating a lack of conceptual clarity among researchers (Jabeen, 2018). Terms commonly used interchangeably or with fuzzy definitions include “unintended consequences” (Bardach & Cabana, 2009), “unintended outcomes” (van de Ruit, 2019), “unintended effects” (Mittelmark, 2014), “unanticipated effects” (de Zwart, 2015), "unexpected effect", "unpredictable effects" (Morell, 2018), “unforeseen effects”, "side effects" (Jabeen, 2016), "spillover effects" (Benjamin-Chung et al., 2017), "non-targeted effects", "non-specific effects", "externalities", "surprises" (Morell, 2010), "secondary effects", “iatrogenic effects” (Bonell et al., 2015; Desclaux & Boye, 2014), “differential effects” (Jabeen, 2016), “unintended harm” (Allen-Scott et al., 2014), “adverse effects” , "undesirable effects" (Desclaux & Boye, 2014), “negative effects” (Biesma et al., 2009), “regressive effects” (Jabeen, 2016), “unwanted consequences” (Cherkaoui, 2004) and "perverse consequences" (Chang et al., 2012). Teasing apart the connotations and research implications of each of these terms can be daunting for researchers or program evaluators. Yet, the selection of a specific term can change the analytical lens, creating blind spots that overlook some areas of interest. Establishing clearer definitions would be useful to collectively pursue a systematic and scientific treatment of the issue (Jabeen, 2016).

For the purpose of this doctoral research, we adopted the term "unintended consequences"[1], defined as *changes for which there is a lack of purposeful action or causation that occur to a social system as a result of the adoption of an innovation* (Ash, Sittig, Poon, et al., 2007; Rogers, 2003)**.**

The use of the term “unintended consequences” has the following advantages:

- The neutrality of the word “unintended” enables researchers to include desirable (positive), undesirable (negative), and neutral changes in their analysis;
- The word “unintended” enables researchers to include changes that are both anticipated and unanticipated by program planners. Effects that are not intended might very well be anticipated by program planners or researchers (de Zwart, 2015; Koch & Schulpen, 2018);
- The word “consequences” enables researchers to focus on both the processes and effects of innovations. Changes can emerge at any point in time during the cycle of the intervention.

2.1.2 Underlying factors

Researchers have identified various underlying factors influencing the emergence of unintended consequences. Merton (1936) offered five causes for what he called “unanticipated consequences”: 1) ignorance because no actor has perfect information; 2) erroneous ideas about the situation or the actions to be executed; 3) imperious immediacy of interest (a certain myopic focus on the objectives to be obtained); 4) value-driven decisions that work in the short term but not the long term; and 5) self-defeating prophecies. According to this approach, many unintended consequences could be avoided if planners would only think in an objective and well-informed manner (Koch & Schulpen, 2018).

More recently, Allen-Scott et al. (2014) conducted a literature review showing that there are five underlying factors associated with unintended harm of public health interventions: 1)

---

[1] To avoid confusion, please note that the French translation of the term "unintended consequences" is "conséquences involontaires" or “conséquences non intentionnelles”, while the French translation of "unanticipated consequences" is "conséquences inattendues".

limited or poor quality evidence; 2) the boomerang effect (prevention of one extreme leading to another); 3) lack of community engagement; 4) ignoring of root causes; and 5) the one-size-fits-all fallacy, referring to how a successful intervention developed in one context/culture may cause harm in another. The authors focus on the phenomenon of interventions developed in high income countries subsequently causing unexpected harm in low-income countries.

Other authors attribute unintended consequences to the complexity of the world. Systems thinkers argue that there are a "*multitude of interconnections, non-linearities, multi-dimensionalities and unpredictabilities that interact with external interventions. These complexities render it difficult to understand the potential ambiguous effects of external interventions in integrated systems*" (Koch & Schulpen, 2018, p. 2). In line with this approach, Morell (2010) presents seven risks factors contributing to the prevalence of unintended effects resulting from external interventions: 1) tightness and richness of linkages between the program and the external environment; 2) size of the program relative to the boundaries of the system in which it lives; 3) stability of the environment between implementation and results of the program; 4) time elapsed between the program and intended results; 5) robustness of an external intervention across time and place; 6) level of innovation in a program; and 7) level of existing knowledge about the context.

2.1.3 Proposed typologies

Some authors have proposed typologies to classify unintended consequences. In pediatric care, Bardach and Cabana (2009) grouped unintended consequences of quality improvement programs according to direct and indirect effects on resource utilization, provider behaviours, or patients. In public health, Bloomrosen et al. (2011) identified six types of unintended consequences: cognitive, care process, organizational, social/legal, fiscal, technology. These unintended consequences can affect different types of stakeholders, such as patients, providers, organizations, vendors, payers, and governments. In contrast, Lorenc and Oliver (2013) use the following categories to classify the adverse effects of public health interventions: direct harms; psychological harms; equity harms; group and social harms and opportunity harms. More recently, Allen-Scott et al. (2014) conducted a literature review of the unintended harm of public health interventions, omitting beneficial unintended consequences, which led them to develop the following typology of unintended harm: physical; psychosocial; economic; cultural; and environmental. Instead of using a thematic typology, Koch and Schulpen (2018) proposed that unintended consequences be classified according to whether they are: really unintended; anticipated; positive; spill-over effects; unavoidable; and bigger than and relevant to the initial effect. Other authors have chosen to further classify unanticipated consequences as foreseeable versus unforeseeable, depending on whether the changes of the intervention *could* have been predicted beforehand, had adequate efforts been made (Jabeen, 2018; Morell, 2010). Taken together, these different typologies can be useful to encourage researchers to consider the broad specifications of unintended consequences that can emerge during data collection and different ways of classifying unintended consequences during the analyses.

In summary, these typologies highlight that researchers should adopt a broad perspective in order to understand the different types of consequences that emerge. They also highlight that there is a lack of consensus regarding the best way to categorize unintended consequences. Thus far, there does not appear to be a single, best typology.

2.1.4 Approaches and methods

The sub-section below presents the approaches used as well as the benefits of adopting qualitative methods to assess unintended consequences of interventions.

Evaluation approaches: Numerous approaches in program evaluation refer, directly or indirectly, to the concept of unintended consequences. In 2000, the Medical Research Council (MRC) published its *Framework for the Development and Evaluation of RCTs for Complex Interventions to Improve Health*, aimed at helping researchers evaluate complex health interventions. This guide states that "*Consideration also has to be given to assessment of outcomes not targeted by an intervention, adverse and unanticipated outcomes"* (Medical Research Council, 2000, p. 17). Unfortunately, this section on unintended consequences was excluded from the accompanying BMJ paper that was widely cited (Campbell et al., 2000). In 2008, the MRC revised and updated its guide. This time, both the guide and the accompanying article stated that "*A single primary outcome may not make best use of the data; a range of measures will be needed and unintended consequences picked up where possible.*" (Craig et al., 2013, p. 588; Medical Research Council, 2008, p. 7). They also stated that process evaluations, which explore the way in which interventions are implemented, can provide valuable insight into why an intervention "*fails or has unexpected consequences*". However, no further technical guidance on how to assess unintended consequences was provided, and none of the 14 case studies presented focused on that topic.

Jabeen (2016) reviews seven approaches in evaluation that hold potential to uncover unintended program effects: 1) social impact assessment; 2) goal-free evaluation; 3) Sieber's framework to evaluate unintended effects; 4) Sherrill's two-step approach; 5) theory-based evaluation, 6) developmental evaluation; and 7) agile evaluation. **Table II**, based on Jabeen's review, presents a brief description and the main limitations of each of these approaches. This review highlights the numerous shortfalls of current evaluation approaches to assess unintended consequences, including lack of clarity regarding the various types of unintended effects, insufficient elucidation of methodological guidelines, and lack of feasibility with regard to the evaluation of established programs. The author concludes that "*evaluation theory is clearly under-developed regarding examination of unintended effects…previous approaches do not provide sufficient theoretical and empirical guidance for practising evaluators*" (Jabeen, 2016, p. 6).

Table II. Summary of seven approaches that contributed to the study of unintended effects according to Jabeen (2016)

Approach	Main description	Main limitations
Social impact assessment	• Aims to maximise benefits and minimize the negative impacts of development interventions by estimating all the likely effects of an intervention related to all groups • Participatory approach involving stakeholders	• Ex-ante evaluation approach only dealing with foreseeable consequences and excluding unforeseen or unforeseeable ones • Consequences emerging during or after implementation are not considered
Goal-free evaluation	• Aims to assess the "actual" effects of an intervention, including good and bad "side effects" • Evaluator remains intentionally unaware of program goals to avoid a narrow view	• Lack clarity on various types of unintended effects • Lack of explicit guidelines on how to undertake a goal-free inquiry
Sieber's framework to evaluate unintended effects	• Aims to predict and test various unintended effects • Focuses on unintended effects which worsen the original situation (regressive effects) • Similarities with theory-based evaluations	• Excludes positive unintended outcomes • Unforeseeable effects can not be identified • Some technical and sensitive points are left to the readers' judgement
Sherrill's two-step approach	• Aims to predict unintended outcomes • Uses longitudinal experimental designs to measure them	• Inability to assess the unforeseeable unintended outcomes
Theory-based evaluation	• Refers to a variety of ways of developing a causal model linking an intervention's inputs and activities to a chain of intended or observed outcomes, and using this model to guide evaluation	• Guidance on examination of unintended effects remain limited • Does not differentiate between various types of unintended effects
Developmental evaluation	• Aims to help those engaged in an intervention to understand the emergent effects of their action in order to make ongoing changes • Flexible in terms of methods to detect both intended and unintended outcomes • Evaluator is part of the project team from the start	• Lack of feasibility with regards to evaluation of already established programs • Requires evaluator's long-term engagement which is not always feasible when funds are limited
Agile evaluation	• Aims to evaluate surprises by using evaluation designs that are agile and flexible • Distinguishes between foreseeable and unforeseeable unintended outcomes	• Some concepts are not explained sufficiently • Methodological guidelines to capture all types of unintended effects are not elucidated sufficiently

The benefits of qualitative methods: The literature suggests that the selection of appropriate methods is important to study unintended consequences. Bamberger et al. (2016) explain that randomized control trial (RCTs) and quasi-experimental designs (QED) are limited in their ability to identify unintended consequences. One reason is that RCTs and QEDs normally test a small number of null hypotheses for narrowly-defined anticipated project outcomes, which does not allow for the exploration of other questions, such as unintended consequences. Another reason is that contextual factors, which are sometimes ignored in RCTs in order to model laboratory settings, tend to be related to unintended consequences (e.g., the political dynamics of local populations).

As such, Bamberger et al. (2016) recommend that studies incorporate qualitative methods to help identify unintended consequences. Qualitative components permit the use of flexible designs that can capture the evolution of programs and their contexts. In the same vein, Morell (2018) argues that methods such as open-ended interviewing, focus groups, observational techniques, and document reviews are ideally suited to reveal unintended consequences. The significant contribution of anthropologists to the development of empirical knowledge on the unintended consequences of innovations supports this perspective (Rogers, 2003). For example, classic anthropological studies have examined the unintended consequences of introducing snowmobiles or steel axes in isolated communities (Pelto, 1973; Sharp, 1952). Their focus on contexts and their longitudinal data gathering techniques enable them to capture the unintended consequences that emerge over time (Patton, 2011; Rogers, 2003).

2.1.5 A neglected topic in health evaluations and studies

Empirical evidence shows that public health[2] and global health[3] interventions can trigger a wide range of desirable and undesirable unintended consequences (Allen-Scott et al., 2014; Koch & Schulpen, 2018). In regard to governance, for example, Biesma et al. (2009) found that initiatives for HIV/AIDS in LICs led to the distortion of the recipient countries' national policies, as they distracted governments from making coordinated efforts to strengthen health systems and led to the re-verticalization of management. More recently, however, Dijkstra (2018) conducted a literature review showing that the sum of all foreign aid can also have desirable unintended effect, such as the reduction of domestic conflicts in recipient countries. In regard to community health, van de Ruit's (2019) study in South Africa showed how structural factors and dynamic local processes caused a community health worker (CHW) program to produce undesirable unintended consequences. CHWs freelanced in communities without regulation, and patients experienced adverse outcomes due to uncoordinated care**.**

Despite the magnitude of unintended consequences of health interventions, their systematic assessment continues to be neglected in both research and evaluation (Bonell et al., 2015; Merton, 1936; Mittelmark, 2014; Morell, 2005). To examine this, Jabeen (2016) searched the database of the DAC Evaluation Resource Centre (DEReC), which contains an extensive collection of documentation from 30 member countries and multilateral development institutions. This search retrieved only 24 evaluation reports, undertaken or commissioned by 12 agencies, referring to either unintended or unanticipated consequences. While some agencies acknowledged the importance of incorporating strategies to evaluate unintended outcomes in their design, this rhetoric was not translated into evaluation practices. For example, the World Bank and Asian Development Bank evaluation reports did not include the study of unintended

---

[2] Based on the literature in public health, we propose the following definition of public health: an area for study, research, and practice that aims to maintain and improve population health through four essential functions: ongoing surveillance of population health status and its determinants; promotion of population health; prevention of illness, psychosocial problems, and injuries; and protection of population health (APHA, 2020; Ministère de la Santé et des Services sociaux, 2015; Pommier & Grimaud, 2007).

[3] Following Koplan et al. (2009), we define global health as "an area for study, research, and practice that places a priority on improving health and achieving health equity for all people worldwide."

consequences in the evaluation's objectives, terms of reference, or questions, which is in contrast with the emphasis placed on this issue in these organizations' evaluation guidelines. In the case of the UNDP and the Aga Khan Foundation, no systematic effort was made to capture unintended consequences, even if this was included in the evaluation objectives or terms of reference.

Similarly, a review of program evaluations for the United States Agency for International Development (USAID) showed only 15% of cases considered "unplanned/unanticipated results" (Hageboeck et al., 2013), a decrease from previous years. This neglect of unintended consequences is also apparent in PBF program evaluations. In Mali, for example, when asked whether a PBF program was associated with "unintended impacts (positive or negative)", evaluators simply indicated "*None reported*", while other sections of the report were filled out more thoroughly (Independent Evaluation Group, 2017).

2.2 Performance-based financing

2.2.1 Article 1: The role of global health actors in the expansion of PBF in LMICs

The sub-section below presents an article on the role that global health actors played in the rapid expansion of PBF in LMICs. While the article does not focus on unintended consequences, it is useful to introduce the broader context within which the intervention emerged and spread in LMICs. Moreover, the article alludes to some of the underlying factors that can contribute to unintended harm, according to the existing literature. In their review, Allen-Scott et al. (2014) suggest that the implementation of a public health intervention in an LMIC by a high-income country is one of the main underlying factors of unintended harm associated with public health interventions. This is linked to a lack of adaptation to local context and often-limited health resources. In this vein, past research has presented PBF as a policy tool transferred from high-income countries to LMICs (Barnes et al., 2015; Eldridge & Palmer, 2009; Olivier de Sardan et al., 2017). Thus, examining the role that global health actors played in the expansion of PBF in LMICs may be useful to understand the context in which unintended consequences may emerge. The article also briefly alludes to the lack of community engagement and the limited quality of evidence on PBF in LMICs, which Allen-Scott and colleagues (Allen-Scott et al., 2014) have also identified as underlying factors of unintended harm associated with public health interventions.

Article 1: The role of global health actors in the expansion of performance-based financing in low- and middle-income countries

English adaptation of the article: *Le rôle des acteurs de la santé mondiale dans l'expansion du financement basé sur la performance dans les pays à faible et à moyen revenu*

Anne-Marie Turcotte-Tremblay, Lara Gautier, Oriane Bodson,
N'koué Emmanuel Sambieni and Valery Ridde

Journal de gestion et d'économie médicales 2018, 36

DOI 10.3917/jgem.185.0261

Available online:
https://www.cairn.info/revue-journal-de-gestion-et-d-economie-medicales-2018-5-page-261.htm

Title: The role of global health actors in the expansion of performance-based financing in low- and middle-income countries

Anne-Marie Turcotte-Tremblay[a,b*]

Lara Gautier[a,b,c]

Oriane Bodson[d]

N'koué Emmanuel Sambieni[e]

Valéry Ridde[a,f]

[a] University of Montreal Public Health Research Institute

[b] School of Public Health, University of Montreal

[c] Centre d'Études en Sciences Sociales sur les Mondes Africains, Américains et Asiatiques, IRD, Université Sorbonne Paris Cité, Paris, France

[d] ARC Effi-Santé, Faculté des sciences sociales, Université de Liège, Liège, Belgique

[e] Université de Parakou, Parakou, Bénin

[f] IRD, CEPED, Université Sorbonne Paris Cité, ERL INSERM SAGESUD, Paris, France

Corresponding author: Anne-Marie Turcotte-Tremblay

Contributions of authors:

Anne-Marie Turcotte-Tremblay designed and wrote parts of the first draft, contributed to analyses, and managed the revision process.

Lara Gautier contributed to the analyses and critically revised the manuscript.

Oriane Bodson collected data, drafted sections on Burkina Faso and on transferring policies, contributed to analyses, and revised the manuscript.

N'koué Emmanuel Sambieni collected data, drafted the section on Benin, contributed to analyses, and revised the manuscript.

Valéry Ridde conceived the original idea, contributed to analyses, and critically revised the manuscript.

ABSTRACT

Performance-based financing (PBF) has spread rapidly in low- and middle-income countries (LMICs) despite the lack of evidence on its implementation process and effectiveness. The objective of this analysis is to understand the role of global health actors in the expansion of PBF in LMICs. We present two concrete cases in Burkina Faso and Benin to illustrate the importance of the role that global health actors played in the conception, promotion, modelling and evaluation of PBF in LMICs.

Keywords: Global health actors, Performance-based financing, Development aid, Reform, Burkina Faso, Benin.

RÉSUMÉ

Le financement basé sur la performance (FBP) s'est rapidement répandu dans les pays à faible et à moyen revenu (PFMR), malgré le manque de données probantes sur les processus de mise en oeuvre et son efficacité. L'objectif de cette analyse est de comprendre le rôle des acteurs de la santé mondiale dans l'expansion du FBP dans les PFMR. Nous abordons, entre autres, deux cas concrets au Burkina Faso et au Bénin afin d'illustrer l'importance des acteurs de la santé mondiale dans la conception, la promotion, la modélisation et l'évaluation du FBP.

Mots-clés : Acteurs de la santé mondiale, Financement basé sur la performance, Aide au développement, Réforme, Burkina Faso, Bénin.

1. INTRODUCTION

Health is a "*a fundamental human right indispensable for the exercise of other human rights*" [1]. Its realization is a collective responsibility. Global health actors, such as international organizations, international financial institutions, government agencies, and multilateral cooperation agencies, must use their competencies to ensure the realization of the right to health. Yet, in low- and middle-income countries (LMICs), access to quality healthcare remains inadequate and mortality rates worrisome.

To achieve universal health coverage (UHC)—partly aimed at increasing the use and quality of health services [2]—and to improve the efficiency of health financing strategies, many global health actors have supported the diffusion of a new financing model, inspired by the New Public Management approach. Commonly referred to as performance-based financing (PBF), this strategy provides financial incentives to motivate actors in the healthcare system to improve the quantity and quality of care provided. Following an experience in Rwanda, PBF spread across LMICs despite the lack of evidence on its implementation processes and effectiveness [3-5]. How did this intervention emerge and spread so quickly? What was the role of global health actors in this expansion? In this paper, we analyze the role of global health actors in the conception, promotion, modelling, financing, and evaluation of PBF in LMICs. We present two concrete cases in Burkina Faso and Benin to illustrate, using empirical data, the role that global health actors played in PBF.

2. RESULTS AS THE BASIS FOR FUNDING

The term PBF encompasses a wide range of terminology and approaches aimed at providing incentives to improve the performance of health systems. Some define it as "*the transfer of money or material goods conditional on taking a measurable action or achieving a predetermined performance target*" [6]. PBF is in line with the New Public Management approach, in that it focuses on incentives, performance evaluation, accountability, competition, and contractualization. In many cases, the theory of intervention for PBF assumes that financial incentives, combined with increased accountability, responsibility, and autonomy in the allocation of resources, motivate and enable healthcare providers to improve the quantity and quality of services. As with New Public Management, performance agreements are established between a purchasing agency (e.g., Ministry of Health, NGO) and health facilities. Performance indicators and compensation modalities vary according to the promoter's intervention model [7]. To facilitate accountability, primary data collection tools (e.g., registers) with information to track patients (e.g., address and telephone number) can be used [8]. Independent auditors[4] conduct verifications. Depending on the intervention model, institutions obtain different degrees of autonomy to distribute PBF payments. For example, in a guide published by the World Bank, it is recommended that about 50% of what health facilities earn based on their performance be allocated to staff payments and that the remaining funds be allocated to recurrent costs unrelated to salaries (e.g. purchase of equipment) [9].

2.1 THE EXPANSION OF PBF IN LMICs

2.2 A BURGEONING INTERVENTION MODEL

In LMICs, the implementation of PBF appears to have begun in the 2000s. The oldest studies on PBF included in a systematic review date back to 2003 [5]. The contractual approach in LMICs, which can be considered the ancestor of PBF, was first implemented in Cambodia and Haiti in the late 1990s. In sub-Saharan Africa, one of the first LMICs to adopt this approach was Rwanda in 2002. Following pilot projects in three districts, the government adopted a national policy in 2005. Subsequently, global health organizations used the Rwandan experience as an example to demonstrate the effectiveness of PBF and encourage other African decision-makers to undertake this intervention model, including through the organization of multiple study tours [10]. The following decade was characterized by a significant expansion of PBF in LMICs.

According to Fritsche and colleagues [9], the number of World Bank intervention sites related to PBF in sub-Saharan Africa increased from about 3 in 2006 to 34 in 2013. Some of these intervention sites were in discussion at the time but have since implemented PBF. When considering all LMICs, the number of PBF intervention sites of this organization reached about 46 in 2015 [11].

---

[4] The terms “auditors” and “verifiers” are used interchangeably in this document.

2.3 AT THE CROSSROADS OF NUMEROUS ORGANIZATIONS

We examined, using a convenience sample, 47 scientific articles or documents from the grey literature on PBF in LMICs to identify examples of global health actors involved in the expansion of PBF (see Supplementary file A). These actors were classified according to the categories used by Robert and Ridde [12]: intergovernmental organizations, international non-governmental organizations, government agencies, and transnational networks and working groups. To this we added international public-private partnerships and private organizations, to provide a more complete and precise picture of relevant actors. While far from being exhaustive, this review nevertheless provides a portrait of the types of global health actors who have played a role in the expansion of PBF in LMICs.

Table III shows that a wide variety of global health actors are involved. A rapid review of the 47 documents allowed us to identify the involvement of more than 30 organizations, not just the World Bank, as is generally expected. In Canada, government agencies such as Global Affairs Canada are currently involved in the development of PBF projects. In Haiti, for example, a new project of C$39.8 million, conducted by Global Affairs Canada, the Pan American Health Organization (PAHO), and the World Health Organization (WHO), will use performance contracts in hospitals to improve maternal and child health over a four-year period [13].

Table III. Examples of global health actors involved in the expansion of PBF in LMICs

Intergovernmental organizations	• World Bank • United Nations Children's Fund • World Health Organization • European Union
Governments and their agencies	German cooperation • Gesellschaft für Internationale Zusammenarbeit / German Technical Cooperation American cooperation • Millennium Challenge Corporation • President's Emergency Plan for AIDS Relief • United States Agency for International Development (inc. Health Systems 20/20) Belgian cooperation • Belgian Development Agency (Enable) Canadian cooperation • Global Affairs Canada Dutch cooperation • Dutch Development Organisation Japanese cooperation • Japan International Cooperation Agency Norwegian cooperation • Government of Norway United Kingdom cooperation • Department for International Development Swiss cooperation • Swiss Tropical Institute
International non-governmental organizations	• Bureau diocésain des œuvres médicales • Catholic Organization for Relief and Development Aid and Cordaid-Memisa • Doctors for Global Health • International Rescue Committee • Save the Children

Transnational networks and working groups	• Performance-Based Financing Community of Practice • Multi-country PBF network
International public-private partnerships	• Global Alliance for Vaccines and Immunization • The Global Fund to Fight AIDS, Tuberculosis and Malaria
Private organizations	• European Agency for Development and Health • Albert Einstein College of Medicine • Center for Global Development • Clinton Health Access Initiative • HealthNet/TPO • Health Partners International • Royal Tropical Institute • Summa Foundation • Management Sciences for Health

2.4 GLOBAL HEALTH ACTORS AT THE HEART AND BODY OF REFORMS

This section analyzes the role of global health actors, inspired by Rocher's sociological analysis of reforms [14].

Conception: According to Rocher's framework [14], conceptors are those who initially had the idea of conducting a reform. They are at the origin of the process. In this respect, several global health organizations conceived the idea of PBF, initially developed by a few individuals. They demonstrated an explicit intention to initiate a shift in the financing of healthcare systems. The World Bank's 1993 report “Investing in Health” emphasized the need to achieve concrete results and improve transparency within health systems and encouraged other organizations to adopt PBF [15]. This report inspired other global health actors (e.g., Cordaid and HealthNet TPO) to propose a shift from input-based to output-based financing for health facilities [15]. However, Barnes and colleagues [16] point out that there is a discrepancy between the discourses of current and former employees of organizations involved in PBF expansion. Current staff contend that the conception of PBF was mainly driven by South-South initiatives, whereas former staff assert that public international organizations were behind PBF in LMICs.

Promotion*:* According to Rocher's analytical framework [14], the promoters of a reform are those who make the official announcement, commit to support its realization, and encourage their members to support it. Several global health actors were directly involved in the promotion of PBF. They supported PBF in their discourse and encouraged governments to subscribe to it. A study in Cameroon, for example, showed that a group of political entrepreneurs associated with the World Bank used several forms of influence (e.g., financial, ideational, networking) to put PBF on the national agenda [10]. At a press conference in 2014, the President of the World Bank said, "*Evidence shows that results-based financing has a significant impact – saving lives and expanding access to quality, essential health services for the poorest women and children in developing countries*" [17]. More recently, WHO experts stated that, "*there should be no PBF or P4P debate. Country experience has demonstrated that moving from more passive to more strategic purchasing of health services contributes to progress toward policy objectives embedded in universal health coverage…So the question is not whether to introduce a reform with the label P4P but rather how to introduce, extend, and institutionalize mechanisms for strategic purchasing of health services*." Thus, senior WHO representatives joined other global health actors actively involved in promoting PBF in LMICs.

In addition, some global health actors supported and organized study tours to Rwanda for high-level political leaders (e.g., heads of state, ministers of health) to help them understand the functioning of PBF and disseminate information in their respective countries. Several courses, short internships, and exchange workshops were also organized. In December 2018, a consulting firm (SINA Health) regularly hired by the World Bank for contracts offered its 75th course on PBF. In 2014, the World Bank organized a one-week PBF exchange workshop in Argentina that brought together more than 200 people from 50 countries. In addition, global health organizations (e.g., Norwegian Agency for Development Cooperation) support a Community of Practice (CoP) on PBF. It enables its 200 members to share experiences and thoughts, create a sense of cohesion, defend "certain principles" that are central to them, and even..."*trigger a revolution*" [19].

Currently[1], global health actors are promoting the implementation of PBF through significant financial support. One of the main funds is the Health Results Innovation Trust Fund (HRITF) created in 2007 with support from the governments of Norway and the United Kingdom [20], which together contributed US$537 million. The International Development Association has contributed US$2.4 billion. Currently coordinated by the World Bank, the HRITF aims to support the development, implementation, monitoring, and evaluation of PBF interventions. It also aims to strengthen institutional capacity and to generate and disseminate evidence on PBF. Barnes, Brown, and Harman [21] found that the content and webinars on the fund's website (see www.rbfhealth.org) deliver discourses that are favourable to PBF. In 2013, the HRITF had already enabled the rapid expansion of PBF by funding three national programs, 17 pilot projects, six advanced planning phases, and discussions on the adoption of PBF in nine countries [9].

In 2013, the World Bank and the Global Fund to Fight AIDS, Tuberculosis and Malaria announced a new partnership to support the expansion of PBF in the field of maternal and child healthcare [17]. This partnership has made it possible for health services related to HIV/AIDS, malaria, and tuberculosis to be included in PBF programs.

Financial support for PBF has been a key factor in its expansion, particularly in Africa: "*We found evidence of non-PBF success, yet the Ministry did not want to hear this because it would affect future funding from the World Bank for national rollout…*" [16]. Similarly, in Burkina Faso, some PBF activities, such as community verifications and patient satisfaction surveys, have been implemented to meet the donor's funding conditions, despite the fact that their utility is questioned by local actors [8].

While some global health actors currently share a certain enthusiasm for PBF, their role in financing this intervention in the medium and long terms is uncertain. There is little empirical evidence on the financial sustainability and efficiency of PBF in LMICs [22]. One of the main threats facing PBF in Burundi, for example, is related to its financial sustainability [15]. Actually, the State contributes only 52% of the funding. In Burkina Faso, PBF was suspended in healthcare centres for an indefinite period in July 2017, when World Bank funding ended. In November 2018, the government of Burkina Faso and the World Bank signed a new agreement under the Global Financing Facility, but the future of PBF will depend, among other things, on the renewal of external funding and the government's willingness/capacity to meet the World Bank's requirements [2].

Modelling: According to Rocher's analytical framework [14], modellers are those who are responsible for translating the project's intentions into administrative structures or changes in practice. In the case of PBF, global health actors were involved in "drawing" the intervention models. They helped write the discourse surrounding PBF. They also developed and recommended practical modalities for PBF implementation. For example, World Bank staff and external collaborators wrote and distributed a free implementation guide on PBF, the PBF Toolbox, available in several languages [9]. Surprisingly, this guide contains several statements with little tangible scientific evidence, regarding such topics as the impact of PBF on equity.

In addition, some global health actors have funded the creation of local technical teams trained to adapt the standard intervention model (such as the one presented in the PBF Toolbox) and organize its implementation in LMICs. In Burundi, for example, independent agencies funded by international NGOs were in charge of the contracting process with healthcare centres and the verification of reported results [23]. In South Africa, several authors have suggested that, in discussions around the possible implementation of PBF, consultants from global health organizations (e.g., Clinton Foundation, USAID) greatly influenced the selection of targets and the means to reach them [16]. In addition, local stakeholders interviewed during our research reported that employees of global health organizations exerted pressure to implement the right type of PBF, referring to the standard PBF model, as we will explain below.

Evaluation and research: Global health actors supporting PBF are sometimes involved in evaluation and research on these interventions [24]. Pilot projects have been implemented to test intervention models. However, in some cases, PBF interventions have been scaled up before pilot projects are evaluated, or without those evaluations being considered, which is a shared responsibility (e.g., government, global health actors). For evaluations, calls for proposals are sometimes launched to select teams of consultants. To date, global health organizations involved have tended to focus more on assessing PBF effectiveness than on understanding implementation processes. Measurement methods and tools are often standardized to be able to compare results with intervention sites in other countries, which could constrain the researchers' or consultants' methodological choices, beyond the obligation to comply with specifications.

The involvement of promoters in the evaluation of PBF interventions could lead to conflicts of interest. It is not uncommon for members of global health organizations to have professional or personal knowledge of consultants selected to evaluate programs. The fact that global health organizations directly contribute to a consultant's salary could limit the latter's scientific independence. In Africa, there is a risk that consultants may self-censor to avoid displeasing sponsors [25]. According to Barnes et al. [16], the positive perception of PBF influences how evidence is generated and interpreted. Yet potential conflicts of interest are not always declared in scientific publications [22]. For example, a scientific article examining the effects of PBF on the use and quality of care in Burundi does not mention the involvement of some co-authors in promoting PBF [23].

3. EXAMPLES FROM TWO EXPERIENCES

In this section, we present the cases of Burkina Faso and Benin to illustrate, using original empirical data, the role that global health actors played in the expansion of PBF. In both cases, data collection included documentary research and fieldwork, conducted between February 2014 and February 2015. During the fieldwork, the third and fourth authors conducted participant observation and interviews with several types of key stakeholders, including individuals who held senior positions during the PBF study tours to Rwanda. The snowball technique was sometimes used, in which one participant was asked to identify others. This empirical data, still exploratory, graphically illustrate the points presented above.

3.1 THE BURKINABE EXPERIENCE

Burkina Faso, receiving technical and financial support from the World Bank, launched a PBF pre-pilot project in 2011 [3] by testing its national strategy in three health districts for a period of nine months. The main objective was to "test tools and materials in the field and enable actors to build their capacity" [26] in order to revise the intervention and gradually expand it on a larger scale until it reached national coverage (which has not been accomplished to date).

The World Bank's involvement in the promotion and modelling of the intervention was apparent. In effect, the introduction of PBF in Burkina Faso followed a study tour to Rwanda, driven by the World Bank, in which four officials of the Ministry of Health participated in 2009. "*Upon our return, we decided to implement PBF in the health sector*," said one participant, attesting to the impact of such a field trip on the agenda-setting process. Subsequently, the World Bank continued to act as a promoter by supporting the participation of officials from the Ministry of Health in two other study tours to Mali and Rwanda in May and July 2010, as well as by funding the organization of a PBF training workshop for decision-makers in the health sector in June 2014. These workshops promoted PBF and provided an opportunity to share some practical ways of implementing PBF. For example, the introduction of PBF in Burkina Faso was accompanied by the creation of a technical unit (ST-FBR) in charge of implementing PBF in the health sector, as had been recommended during a training workshop held with support from the World Bank that brought together 120 officials from the Ministry of Health. Some members of the technical unit subsequently worked as consultants for the World Bank to help other African countries implement PBF.

Based on some qualitative and quantitative results presented as encouraging by the Ministry of Health, despite the evaluation's limitations, the country launched, in 2014, the pilot phase of its national program in 15 districts for a period of three years. In addition to this territorial expansion, the government developed a new and original component in Africa that combined PBF with both a community-based indigent selection process and community-based health insurance.

However, a review of official documents suggested that the decision to expand PBF on a larger scale (15 districts) was made well before the results of the evaluation were available. The first document describing the content of the pilot project was published in November 2011, while the evaluation report of the pre-pilot project, produced by a consulting firm that was ultimately awarded a contract to coordinate the indigent selection process, was published in 2013. We compared the PBF pilot project developed in the 2011 document with the intervention model planned for 2014. This revealed several changes resulting from the Bank's conditionalities, highlighting its role in the intervention modelling process.

The World Bank funded an impact evaluation of the PBF pilot project. A search for scientific evidence appeared to have led to two important modifications between 2011 and 2014: (1) the use of a randomized controlled trial design with four intervention arms instead of a simpler design using "treatment" and "control" districts[4], and (2) the addition of a sixth region (Centre-East) to the five regions previously identified. These changes in the PBF modelling appeared have been made, not so much in response to a national demand, but rather to methodological considerations on the part of the World Bank to produce evidence on the effectiveness of the intervention.

Related to the randomization, one of the most remarkable changes between 2011 and 2014 was undoubtedly the emergence of a concern for equity in access to health services for the poorest. This concern was accompanied, in some intervention arms, by a user-fee exemption for the poor. Curative consultations for poor patients paid by the intervention could represent up to 10% of all consultations. One participant pointed out the external origin of this cap: "*It's the external consultant's idea, likely inspired by the literature and his experience*."

Finally, the same external consultant from the World Bank, who had been providing training on PBF for many years, tried to influence the terminology used. He supported the use of the term "subsidies"—to replace the lengthy notion of "subgrant packages for primary healthcare"—to refer to performance-related payments, although this term had never been used before in health policies in Burkina Faso. According to one participant, "*the World Bank did not want to see the notion of 'premiums' appear*" and therefore suggested this change. Ultimately, the terms "subsidies" and "premiums" were both used locally.

3.2 THE BENINESE EXPERIENCE

In Benin, persistent deficiencies in healthcare and low utilization of services, especially in the field of maternal and child health, led the country to organize a General Assembly on the health sector, mainly with the support of the World Bank in 2007. One of the healthcare system's main problems was the poor management, demotivation, and ineffectiveness of human resources. This observation led to the search for effective solutions. Policy documents were developed, but as one participant stated, "*these documents are often written to please donors*".

That same year (2007), one of the first initiatives to encourage health workers to improve their performance, developed with support from USAID and the HRITF, was implemented. This unprecedented experiment offered financial rewards to midwives who had "worked well". These rewards were linked to zones with difficult access and low levels of healthcare services penetration in the population. The intervention had a rudimentary conceptual organization, mobilized a small number of international and local actors, and was allotted a relatively small financial envelope in relation to the challenges that needed to be addressed.

Subsequently, the World Bank became involved in the promotion and modelling of PBF in Benin. A key step for the adoption of PBF was the World Bank's initiative to organize a study tour to Rwanda for executives from the Ministry of Health of Benin. Visitors were seduced: "*Sometimes you see your neighbour wearing a dress and you like it. It may be convenient for you, just as it may not be convenient for you*," said one participant. Upon their return, these managers began to produce the technical documents for the Health Sector Performance Enhancement Project (HSPEP - *Projet de renforcement de la performance du secteur de la santé*), which was entirely funded by the World Bank.

The pilot phase of this project started in 2010 in two regions, each supported by a different donor, the World Bank (with the HSPEP) and Enable, the Belgian cooperation agency, which already had another project supporting health districts in its intervention areas (Mono-Couffo and Atacora-Donga). These two organizations (World Bank and Enable) applied different intervention models for PBF: the first drawing on the models used in the other funded countries and the second drawing on local approaches such as the one presented at the beginning of this section.

In 2014, Benin decided to extend PBF to all health districts, although no evaluative study supported the decision-making process, as pointed out by a senior World Bank official during an interview. The expansion was financially supported by new partners such as the Global Fund and the GAVI Alliance. UNICEF also intervened to some extent to support the inclusion of community relays for child health. The budget was mobilized and managed by the HSPEP in a common fund. This national scale-up was based on the results of implementation monitoring in pilot districts, particularly in some health districts such as Banikoara, where the results were considered very satisfactory according to HSPEP officials. Some international organizations asserted that expanding PBF without evaluating the implementation of the pilot phase was not recommended. However, these organizations agreed to contribute to the financing of PBF.

In general, the World Bank, through the HSPEP, designed the different strategies to scale up PBF, according to a manager interviewed. Between June and July 2013, the World Bank supported the revision of the PBF's framework strategies prior to the scaling-up phase. In May 2014, it supported the development of a guide to implement and manage the indigent selection process. It also supported the participation of several officials from ministries and other state institutions in various PBF training sessions held in Cotonou, Benin. Every year since 2010, the Cabinet BEST-SD, in collaboration with SINA Health and BASP96, has been organizing training courses on PBF for people from French-speaking African countries, according to HSPEP officials.

The involvement of multiple donors (two and then four) led the two main financial partners to set up a discussion platform in 2012 with the intention of adopting a consensual model for PBF, especially regarding performance indicators, verification, and technical development of various matrices. Several joint missions were organized by the Ministry for this, but no consensus could be reached, due to resistance on both sides [28] and the weakness of the State (through the Ministry of Health) in arbitrating ideas and decisions.

In July 2017, the HSPEP was without funding, as the World Bank and other partners had ceased their support. The Belgian cooperation continued to finance the project, but only in the areas for which it was responsible. The single model could not be adopted. There was no progress on issues of technical, political, and financial sustainability. In 2018, PBF was no longer being implemented in Benin. The government did not take up the intervention.

4. AVENUES FOR ACTION AND RESEARCH

4.1. REVIEWING THE POLICY TRANSFER PROCESS

The policy transfer process for PBF raises the issue of how global health actors encourage the diffusion of a single, standardized model, devised by experts who are sometimes out of touch with local norms and contexts. To date, few studies have examined the policy transfer process for PBF. Yet the case of PBF represents the dissemination of elements that are both hard (i.e., public policy instruments, structures and practices) and soft (i.e., standards, lessons) [29], with the World Bank acting beyond the role of a knowledge bank by promoting PBF on the basis of key examples set up as reference models (notably in Rwanda).

In development and humanitarian aid, it is not unusual for health interventions to be transferred without sufficiently taking local institutions and contexts into account, sometimes contradicting social norms and even ignoring them “superbly” [30]. In the case of PBF, this was illustrated by the lack of consideration for community authorities, namely the management committees of healthcare facilities (COGES) in Burkina Faso. The fact that the COGES were not better integrated into PBF highlights a certain lack of alignment between the intervention and existing health policies, a situation also encountered in Benin. In Burkina Faso, PBF was also dissociated from other health reforms or policies covering emergency obstetric and neonatal care, services for poor pregnant women, and prenatal consultations. These activities were purchased through PBF while being subsidized by the State through other mechanisms. The integration of PBF with these reforms was carried out very late in its implementation. It therefore seems important to reconsider the degree to which PBF is integrated into different health contexts where it is introduced while taking into account previous reforms.

New norms that are introduced are known to be often incomprehensible for a portion of actors. In the case of PBF in Burkina Faso, this was illustrated by the random assignment of health facilities to different PBF modalities (intervention arms): "*For T1, T2, T3 [the intervention arms], I did not understand anything,*" stated a nurse. When asked about the origin of the randomization idea, another participant said, "*I don't know. (...) It seems that it's a condition at the World Bank level. It had to be done: they're looking for evidence.*" This lack of knowledge indicates the lack of participation of local actors in the design of its content, as shown elsewhere [16]. The nature of the participation clearly varied depending on context. It could be formal or informal. However, it appeared to remain low, far from the image of South-South learning where international organizations only act as channels of exchange. In Tanzania and Zambia, for example, national actors preferred to co-lead the implementation, leaving donors in charge of developing the PBF policy. The authors of a study in Benin also highlight a lack of local ownership: "*in the two districts visited, field actors regret not having been consulted during the elaboration of PBF policies*" [31]. In this respect, we can only recommend that actors remain attentive to the importance of improving the participatory nature of the intervention model and give due consideration to the many—sometimes discrete—actors manoeuvring in the health landscape and whose contribution could be beneficial for the intervention and therefore for the health of populations.

Ethiopia's hesitation regarding the implementation of PBF and Benin's decision to stop it paved the way for a global reflection on the conditions for the institutional transfer of PBF that was expected to unfold over the coming years [32] and attracted more attention. Following a pre-implementation audit of PBF in Ethiopia, the Office of the Inspector General for the Global Fund Secretariat recommended the adoption of a model that was more appropriate for the country's current context and capacity. Since 2013, however, Ethiopia has received funding from the HRITF [33].

4.2 A CALL FOR PRUDENCE, BUT NOT INACTION

Faced with the needs of vulnerable populations, several global health actors have been mandated to take action to improve the quality and use of, and access to, healthcare in LMICs. The Sustainable Development Goals and, before that, the Millennium Development Goals and the Alma Ata Charter, invite the international community to take action. Thus, global health actors have a certain legitimacy in supporting healthcare organization models in LMICs. From an ethical standpoint, some argue that global health actors have a responsibility to promote interventions that are evaluated as being the most effective at a minimum cost. For example, Kass [34] argues that data must substantiate a program's positive impact to justify its social costs (e.g., restrictions on freedom).

However, the decisions of international organizations in public health often fall within a "zone of relative uncertainty" [35]. Indeed, global health actors, driven by their values and ideologies, rarely have evidence leading to certainty in the analysis of health systems in various settings. In recent years, global health actors seem to promote PBF in LMICs on the basis of theoretical efficacy, derived from underlying theories, rather than on empirical evaluations conducted in intervention settings.

According to the uncertainty principle [35], global health actors driving such reforms should take into account the margin of uncertainty that characterizes the analysis of health system issues and PBF effectiveness. The complexity of interactions among health system factors, the lack of scientific evidence on the effectiveness of PBF, the limitations of evaluation methods that are still too often focused on impacts without understanding the role of context, the potential conflicts of interest, the particularities of the various intervention contexts, the potentially harmful consequences of PBF, and lessons from the past regarding intervention models that have now been refuted should induce these global health actors to exercise great caution. We recall, for example, their promotion of out-of-pocket user fees in the 1980s, completely reversed today. At the time, many researchers and non-governmental organizations were concerned that these recommendations went against the scientific evidence that was available. Will history repeat itself? The uncertainty principle does not confine global health actors to inaction, but rather to prudence in determining and implementing actions. In addition, it calls for global health actors to be sensitive to the particularities of intervention contexts while avoiding the mechanical implementation of interventions. This principle also suggests that global health actors have a moral responsibility to recognize and communicate the degree of uncertainty about the effectiveness of interventions and the adverse consequences that can result from them.

4.3 NEEDS TO BE ADRESSED IN RESEARCH

Several questions remain about the role of global health actors in the expansion of PBF in LMICs. We can identify several themes for future research. It would be useful to study the emergence of PBF by looking at the reasons and factors underlying the decision of global health actors to promote that intervention over others. To this end, political science offers various approaches—related, for instance, to the role of ideas, the diffusion of policies [36], and Kingdon's framework on multiple currents, [10, 37]—that could help understand the factors that contributed to political decisions in favour of PBF. We are currently undertaking such studies in Benin, Burkina Faso, and Mali. Future results will be useful to better understand the situation, given that few reforms in Africa have been studied using such frameworks.

It would also be interesting to further study the fidelity and implementation processes, particularly to understand the dynamics and power games between international, national, and regional actors. How will the role of global health actors evolve over the coming years, and what will the consequences be for interventions? In addition, comparative studies could examine how the implementation of various PBF models adopted by global health actors has led to different levels of performance over time.

Future research should also analyze the sustainability process to better understand how global health actors influence the scaling-up of PBF, the different degrees of sustainability, and the capacity of African states to pursue PBF [38].

Likewise, a comparison remains to be made with other health reforms in LMICs, as it may be that the role of global health actors is similar, as was shown regarding direct payment in the 1980s. If this were the case, their involvement in PBF would be the norm rather than the exception.

5. CONCLUSION

PBF experienced a strong expansion in LMICs over the past decade, despite insufficient and mixed evidence on its relevance, implementation process, and effectiveness. We identified more than 30 global health actors that have been closely involved in the expansion of PBF in LMICs. Our exploratory data on the cases of Burkina Faso and Benin illustrate their influence in the conception, promotion, modelling, and evaluation of the intervention. Due to the lack of evidence on PBF, we call on international organizations, as well as public health authorities, to exercise caution in deciding on actions to improve the health of populations and reduce social health inequalities.

CONFLICTS OF INTEREST:

VR was a co-researcher for the baseline study of the impact evaluation in Burkina Faso, which was supported by the World Bank, but was not remunerated for this activity. The other authors have no conflicts of interest related to this article.

ACKNOWLEDGEMENTS

The authors thank the people who collaborated in Burkina Faso and Benin for their hospitality and availability.

FUNDING

AMTT received a doctoral research award from the Canadian Institutes of Health Research (CIHR). LG received a doctoral research award from the Fonds de Recherche du Québec – Société et Culture (FRQSC). A CIHR-funded research program (ROH-115213) supported the data collection for the two case studies.

REFERENCES

[1] United Nations Committee on Economic, Social and Cultural Rights. The Right to the Highest Attainable Standard of Health. 2000.

[2] Musango L, Basenya O, Elovainio R. Une approche novatrice pour contribuer au progrès vers la couverture universelle au Burundi : Intégration du financement basé sur la performance et de la gratuité des soins. Afr Health Monit. 2013;17.

[3] Basinga P, Mayaka S, Condo J. Performance-based financing: the need for more research. Bull World Health Organ. 2011;89:698–9.

[4] Renmans D, Holvoet N, Orach CG, Criel B. Opening the "black box" of performance-based financing in low- and lower middle-income countries: a review of the literature. Health Policy Plan. 2016;31:1297–1309.

[5] Witter S, Fretheim A, Kessy FL, Lindahl AK. Paying for performance to improve the delivery of health interventions in low- and middle- income countries. Cochrane Database Syst Rev. 2013;2:1–82.

[6] Eichler R. Performance Incentives for Global Health: Potential and Pitfalls. CGD Books; 2009.

[7] Gergen J, Josephson E, Coe M, Ski S, Madhavan S, Bauhoff S. Quality of care in performance-based financing: How it is incorporated in 32 programs across 28 countries. Glob Health Sci Pract. 2017;5:90–107.

[8] Turcotte-Tremblay A-M, Gali Gali IA, De Allegri M, Ridde V. The unintended consequences of community verifications for performance-based financing in Burkina Faso. Soc Sci Med. 2017;191:226–36.

[9] Fritsche G, Soeters R, Meessen B. Boîte à outils : financement basé sur la performance. Banque mondiale. Washington, DC; 2014.

[10] Sieleunou I, Turcotte-Tremblay A-M, Fotso J-CT, Tamga DM, Yumo HA, Kouokam E, *et al.* Setting performance-based financing in the health sector agenda: a case study in Cameroon. Glob Health. 2017;13:52.

[11] World Bank. Achieving Results for Women's and Children's Health. Progress Report. 2015. http://www.rbfhealth.org/sites/rbf/files/2015%20Progress%20Report_0.pdf.

[12] Robert E, Ridde V. Global health actors no longer in favor of user fees: a documentary study. Glob Health. 2013;9:29.

[13] PAHO, WHO. Canada and PAHO sign an important agreement for Haiti: Essential Health Care for Mothers and Children (SESAME) project. 2017. http://www.paho.org/hq/index.php?option=com_content&view=article&id=13100%3Acanada

-paho-sign-important-agreement-haiti-essential-health-care-mothers-children-sesame-project&catid=1443%3Aweb-bulletins&Itemid=135&lang=en.

[14] Rocher G. Les réformes: une perspective sociologique. Montréal; 2004.

[15] Falisse J-B, Ndayishimiye J, Kamenyero V, Bossuyt M. Performance-based financing in the context of selective free health-care: an evaluation of its effects on the use of primary health-care services in Burundi using routine data. Health Policy Plan. 2014.

[16] Barnes A, Brown G, Harman S, Papamichail A. African participation and partnership in performance-based financing: A case study in global health policy. 2014. https://www.rbfhealth.org/resource/african-participation-and-partnership-performance-based-financing-case-study-global-health.

[17] Reynolds C. World Bank, Global Fund, Partner to Expand Results-Based Financing for Maternal and Child Health. The World Bank. 2013.

[18] Soucat A, Dale E, Mathauer I, Kutzin J. Pay-for-Performance Debate: Not Seeing the Forest for the Trees. Health Syst Reform. 2017;3:74–9.

[19] Meessen B. Communauté de Pratique “Financement Basé sur la Performance”: cinq ans déjà! Health Financing in Africa. 2015. http://www.healthfinancingafrica.org/home/communaute-de-pratique-financement-basesur-la-performance-cinq-ans-deja.

[20] World Bank. Results-Based financing for Health. 2014. https://www.rbfhealth.org/project/our-projects.

[21] Barnes A, Brown GW, Harman S. Global Politics of Health Reform in Africa. Performance, Participation, and Policy. Palgrave Macmillan. Basingstoke; 2015.

[22] Turcotte-Tremblay A-M, Spagnolo J, De Allegri M, Ridde V. Does performance-based financing increase value for money in low- and middleincome countries ? A systematic review. Health Econ Rev. 2016;6:30.

[23] Bonfrer I, Soeters R, Poel EV de, Basenya O, Longin G, Looij F van de, *et al.* Introduction of Performance-Based Financing in Burundi Was Associated with Improvements in Care And Quality. Health Aff (Millwood). 2014;33:2179–87.

[24] Gautier L, Ridde V. Health financing policies in Sub-Saharan Africa: government ownership or donors' influence ? A scoping review of policymaking processes. Glob Health Res Policy. 2017;2:23.

[25] Olivier de Sardan J-P. Promouvoir la recherche face à la consultance. Cah Détudes Afr. 2011;202–203:511–28.

[26] Ministère de la Santé. Evaluation finale de la phase-test du financement basé sur les résultats dans les districts sanitaires de Boulsa, Léo et Titao. Ouagadougou, Burkina Faso; 2013.

[27] Ridde V, Turcotte-Tremblay A-M, Souares A, Lohmann J, Zombré D, Koulidiati JL, *et al.* Protocol for the process evaluation of interventions combining performance-based financing with health equity in Burkina Faso. Implement Sci. 2014;9:149.

[28] Paul E, Lamine Dramé M, Kashala J-P, Ekambi Ndema A, Kounnou M, Aïssan JC, *et al.* Performance-Based Financing to Strengthen the Health System in Benin: Challenging the Mainstream Approach. Int J Health Policy Manag. 2017;0. http://www.ijhpm.com/article_3352.html. Accessed 25 May 2017.

[29] Stone D. Transfer agents and global networks in the "transnationalization" of policy. J Eur Public Policy. 2004;11:545–66.

[30] Olivier de Sardan J-P. La manne, les normes et les soupçons. Les contradictions de l'aide vue d'en bas. Revue Tiers Monde. 2014;:1–22.

[31] Paul E, Sossouhounto N, Eclou DS. Local stakeholders' perceptions about the introduction of performance-based financing in Benin: a case study in two health districts. Int J Health Policy Manag. 2014;3:207–14.

[32] The Global Fund. Global Fund Grants to the Federal Democratic Republic of Ethiopia Pre-Implementation Review. Audit Report. Geneva, Switzerland; 2015.

[33] The World Bank Group. RBF Health - Ethiopia. 2014. http://www.rbfhealth.org/rbfhealth/country/ethiopia.

[34] Kass N. An Ethics Framework for Public Health. Am J Public Health. 2001;91:1776–82.

[35] Massé R. Éthique et santé publique: enjeux, valeurs et normativité. Presses Université Laval; 2003.

[36] Gautier L, Tosun J, De Allegri M, Ridde V. How do diffusion entrepreneurs spread policies? Insights from performance-based financing in Sub-Saharan Africa. World Development. 2018;110(2018):160–175

[37] Kiendrébéogo JA, Shroff ZC, Berthé A, Yonli L, Béchir M, Meessen B. Why Performance-Based Financing in Chad Failed to Emerge on the National Policy Agenda. Health Syst Reform. 2017. http://www.tandfonline.com/doi/abs/10.1080/23288604.2017.1280115. Accessed 19 Sep 2017.

[38] Seppey M. La pérennisation de programmes de financement basé sur les résultats : Étude de cas au Mali. Mémoire. 2016.

SUPPLEMENTARY FILE A

List of selected documents mentioning global health actors involved in PBF in LMICs[5]

1. Basinga P, Gertler PJ, Binagwaho A, Soucat AL, Sturdy J, Vermeersch CM. Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation. The Lancet. 2011; 377(9775): 1421–8.
2. Basinga P, Mayaka S, Condo J. Performance-based financing: the need for more research. Bull World Health Organ. 2011 Sep 1; 89(9): 698–9.
3. Noltze M, Gaisbauer F, Schwedersky T, Krapp S. Contribution analysis as an evaluation strategy in the context of a sector-wide approach: Performance-based health financing in Rwanda. Afr Eval J [Internet]. 2014 Dec 18 [cited 2015 Apr 29]; 2(1). Available from: http://www.aejonline.org/index.php/aej/article/view/81.
4. Soeters R, Peerenboom PB, Mushagalusa P, Kimanuka C. Performance-Based Financing Experiment Improved Health Care In The Democratic Republic Of Congo. Health Aff (Millwood). 2011 Aug 1; 30(8): 1518–27.
5. Zeng W, Rwiyereka AK, Amico PR, Avila-Figueroa C, Shepard DS. Efficiency of HIV/AIDS health centers and effect of community-based health insurance and performance-based financing on HIV/AIDS service delivery in Rwanda. Am J Trop Med Hyg. 2014 Apr; 90(4): 740–6.
6. Zeng W, Cros M, Wright KD, Shepard DS. Impact of performance-based financing on primary health care services in Haiti. Health Policy Plan. 2013 Sep; 28(6): 596–605.
7. Fox S, Witter S, Wylde E, Mafuta E, Lievens T. Paying health workers for performance in a fragmented, fragile state: reflections from Katanga Province, Democratic Republic of Congo. Health Policy Plan. 2014 Jan; 29(1): 96–105.
8. Kantengwa K, De Naeyer L, Ndizeye C, Uwayitu A, Pollock J, Bryant M. PBF in Rwanda: what happened after the BTC-experience ? Trop Med Int Health. 2010 Jan 1; 15(1): 148–9.

9. Ssengooba F, McPake B, Palmer N. Why Performance-Based Contracting Failed in Uganda-An “Open-Box” Evaluation of a Complex Health System Intervention. Soc Sci Med. 2012; 75(2): 377–83.

10. Witter S, Zulfiqur T, Javeed S, Khan A, Bari A. Paying health workers for performance in Battagram district, Pakistan. Hum Resour Health. 2011 Oct 7; 9: 23.

11. Ridde V, Yaogo M, Zongo S, Somé P-A, Turcotte-Tremblay A-M. Twelve months of implementation of health care performance-based financing in Burkina Faso: A qualitative multiple case study. Int J Health Plann Mgmt. 2017; 1–15.

12. Paul E, Sossouhounto N, Eclou DS. Local stakeholders’ perceptions about the introduction of performance-based financing in Benin: a case study in two health districts. Int J Health Policy Manag. 2014; 3(4): 207–14.

13. Oxman A, Fretheim A. Can paying for results help to achieve the Millennium Development Goals ? A critical review of selected evaluations of results-based financing. Journal of Evidence-Based Medicine. 2009; 184–95.

14. Falisse J-B, Meessen B, Ndayishimiye J, Bossuyt M. Community participation and voice mechanisms under performance-based financing schemes in Burundi. Trop Med Int Health. 2012; 674–82.

15. Beane CR, Hobbs SH, Thirumurthy H. Exploring the potential for using results-based financing to address non-communicable diseases in low- and middle-income countries. BMC Public Health. 2013 Feb 1; 13(1): 92.

16. Meessen B, Soucat A, Sekabaraga C. Performance-based financing: just a donor fad or a catalyst towards comprehensive health-care reform ? Bull World Health Organ. 2011 Feb 1; 89(2): 153–6.

17. Chee G. Using financing to motivate a for-profit health care provider to deliver family planning services: Is it a cost-effective intervention ? A study of AAR health services in Kenya. The International Journal of Health Planning and Management. 2003; 205–20.

18. McNamara P. Quality-based payment: six case examples. Int J Qual Health Care J Int Soc Qual Health Care ISQua. 2005 Aug; 17(4): 357–62.

19. Rusa L, Ngirabega J de D, Janssen W, Van Bastelaere S, Porignon D, Vandenbulcke W. Performance-based financing for better quality of services in Rwandan health centres: 3-year experience. Trop Med Int Health TM IH. 2009 Jul; 14(7): 830–7.

20. PAHO, WHO. Canada and PAHO sign an important agreement for Haiti: Essential Health Care for Mothers and Children (SESAME) project [Internet]. 2017.

21. Seppey M. La pérennisation de programmes de financement basé sur les résultats : Étude de cas au Mali. 2016.

22. Ashir GM, Doctor HV, Afenyadu GY. Performance Based Financing and Uptake of Maternal and Child Health Services in Yobe Sate, Northern Nigeria. Glob J Health Sci. 2013 Jan 29; 5(3): p34.

23. Chimhutu V, Lindkvist I, Lange S. When incentives work too well: locally implemented pay for performance (P4P) and adverse sanctions towards home birth in Tanzania - a qualitative study. BMC Health Services Research. 2014;

24. Fretheim A, Witter S, Lindahl AK, Olsen IT. Performance-based financing in low- and middle-income countries: still more questions than answers. Bull World Health Organ. 2012; 90(8): 559–559A.

25. Borghi J, Mayumana I, Mashasi I, Binyaruka P, Patouillard E, Njau I, *et al.* Protocol for the evaluation of a pay for performance programme in Pwani region in Tanzania: a controlled before and after study. Implement Sci. 2013; 8:80.

26. Baral G. An assessment of the safe delivery incentive program at a tertiary level hospital in Nepal. J Nepal Health Res Counc. 2012 May;10(21): 118–24.

27. Mayaka S. Le financement basé sur la performance dans un système de santé complexe : cas de la République Démocratique du Congo. Université Catholique de Louvain; 2015.

28. Soeters R, Habineza C, Peerenboom PB. Performance-based financing and changing the district health system: experience from Rwanda. Bull World Health Organ. 2006 Nov; 84(11): 884–9.

29. Miller JS, Musominali S, Baganizi M, Paccione GA. A process evaluation of performance-based incentives for village health workers in Kisoro district, Uganda. Hum Resour Health. 2014 Apr 8; 12(1): 19.

30. Meessen B. Communauté de Pratique “Financement Basé sur la Performance”: cinq ans déjà! Financing Health in Africa: Le Blog.

31. Multi Country PBF Network. Multi Country PBF Network [Internet]. 2017. Available from: http://www.multicountrypbfnetwork.org/index.html.

32. Sieleunou I, Turcotte-Tremblay A-M, Habakkuk YA, Kouokam E, Taptué Fotso J-C, Magne Tamga D, *et al.* Transferring the Purchasing Role from International to National Organizations During the Scale Up Phase of Performance-Based Financing in Cameroon. Health Syst Reform. 2017; 2.

33. Honda A. 10 best resources on pay for performance in low- and middle-income countries. Health Policy Plan. 2013 Aug; 28(5): 454–7.

NOTES

[1] In the 2000s, the Inter-American Development Bank also strongly encouraged some Latin American and Caribbean States to embark on this type of process oriented towards a contractual approach. This was not a great success in Haiti, for example, although the idea now seems to be coming back with the launch of a new PBF project in 2015...which took a long time to begin.

[2] Sanfo, A. (2018). Financement du GFF : le Burkina Faso matérialise son engagement avec la Banque mondiale. Radiodiffusion Télévision du Burkina.

[3] This intervention is commonly referred to as results-based financing (RBF) in French-speaking African countries.

[4] See Ridde *et al.* [27] for a description of the four intervention arms tested during the randomized control trial.

[5] Only 33 of the 47 articles selected mentioned global health actors involved in PBF.

2.2.2 Unintended consequences of PBF

Many global health actors argue that PBF could, at least theoretically, trigger serious unintended consequences (Barnes et al., 2015; Ireland et al., 2011; Kalk, 2011; Oxman & Fretheim, 2008, 2009). Some of these concerns stem from the literature in high-income countries (HICs). A systematic literature review of reviews in HICs provides some evidence that PBF had unintended consequences (Eijkenaar et al., 2013). Undesirable unintended consequences include risk selection, gaming behaviour, effects on providers' intrinsic motivation, and less time for holistic care, although the findings on some of these outcomes are mixed (Eijkenaar et al., 2013). In terms of desirable unintended consequence, Eijkenaar et al.'s review only found a few studies showing that PBF had positive effects on unrewarded aspects of care. More recently, Hysong (2017) found that incentives had positive spillover effects, whereby providers educated themselves better and were more professional. However, many questions remain about the unintended consequences of PBF in high-income countries. Cashin et al. (2014) examined 12 pay-for-performance programs in HICs and concluded that none of the programs carefully assessed unintended consequences. Petross et al. (2020) found that the existing literature from high-income settings defines unintended consequences very specifically in relation to effects on non-incentivized clinical indicators and relies exclusively on quantitative methods.

It is also important to note that the results of the PBF studies in HICs cannot automatically be generalized to LMICs, because the contexts and resources differ significantly. Research has shown that the specific design features and contextual factors influence the adoption, adaptation, implementation, and effects of PBF (Bertone et al., 2018; De Allegri, Bertone, et al., 2018; Eijkenaar et al., 2013). Stakeholders in LMICs may adopt different intervention models depending on the funding agency and may have to deal with distinctive barriers and facilitators that influence the emergence of unintended consequences.

To date, however, multiple literature reviews have demonstrated the astonishing lack of empirical evidence documenting the unintended consequences of PBF in LMICs (Chalkley et al., 2016; Gorter et al., 2013; Witter et al., 2012). Of these, a Cochrane systematic review found only two non-published studies reporting on the unintended consequences of PBF in Zambia and Tanzania and concluded that the quality of the evidence was low (Witter et al., 2012). In 2019, an update of this Cochrane review highlighted that the knowledge gap persists (Witter et al., unpublished data). The authors extracted 1,980 data points corresponding to a diverse set of outcomes of PBF in LMICs. Of these, only four data points were categorized as being related to "unintended effects". It should be noted, however, that the stringent selection criteria and the conceptualization of unintended effects likely influenced the quantity of studies included.

To broaden our understanding of the unintended consequences of PBF in LMICs and better delineate the knowledge gap, the empirical evidence related to this issue is compiled and synthesized in the following section. Although it is not exhaustive[5], this review maps the various types of unintended consequences that have emerged in the scientific literature following the implementation of PBF interventions in LMICs. **Box 1** lists the types of unintended consequences of PBF identified in empirical studies conducted in LMICs.

Box 1. Types of unintended consequences identified in empirical studies in LMICs

- Focusing on targeted services or patients at the expense of others
- Measure fixation
- Gaming
- Misrepresentation of performance data
- Prioritizing quantity over quality of care
- Overburdened healthcare staff and verifiers
- Ossification
- Reduction in staff's intrinsic motivation
- Conflicts among actors of the healthcare system
- Provision of unnecessary services
- Adoption of coercive strategies to attract patients to healthcare facilities
- Inequities between populations, staff members or facilities
- Suboptimal value for money
- Pride

---

[5] Our initial search strategy, developed with two university librarians, using MESH terms and keywords in six databases, retrieved almost 10,000 scientific articles, highlighting the vast amount of literature published on performance-based initiatives across the world. The elevated quantity was also related to the fact that we had not restrained our search to specific outcomes, in order that all unintended consequences might emerge, not only those labelled as such by the authors. A subsequent screening of the titles and abstracts reduced the number of potential articles to about 1,000. For reasons of feasibility, this review targets the most relevant scientific literature on performance-based financing in low- and middle-income countries.

Focusing on targeted services or patients at the expense of others

One potential unintended consequence of PBF is the risk that providers would focus on PBF indicators while neglecting activities that are not remunerated. This is sometimes referred to as "multitasking", "distortions", or "substitution effects" (Chalkley et al., 2016; Chimhutu et al., 2014; Huillery & Seban, 2014; Ireland et al., 2011; Kalk et al., 2010). Similarly, Smith (1995) explains that the public reporting of performance scores can lead to "tunnel vision", defined as an emphasis *by management* on phenomena that are quantified in the performance measurement scheme, at the expense of unquantified aspects of performance.

To date, research on this topic has been limited in LMICs (Chalkley et al., 2016). In Tanzania, a controlled before-and-after study found that P4P was associated with a significant reduction in non-targeted outpatient visits at dispensaries (Binyaruka et al., 2015). In Rwanda, Basinga et al.'s impact evaluation (2011) showed that PBF had the greatest effect on services with the highest payment and requiring the least effort from providers (e.g., institutional deliveries). However, the study does not inform about the effects of PBF on the provision of non-incentivized services or a substitution effect. In Kalk et al.'s qualitative study (2010), participants regularly reported the neglect of essential activities as a result of additional workload created by the P4P system. Providers chose between activities seen as necessary and those required for the rewards (e.g., filling out forms). Potentially life-preserving activities in intensive care units were neglected.

In contrast, other studies have found that providers under a PBF scheme did not focus on the incentivized services at the expense of others. Huillery and Seban's (2014) field experiment in the Democratic Republic of Congo showed that increased efforts on targeted services did not come at the expense of non-targeted services. There was no substitution of non-rewarded activities for rewarded ones. In Burkina Faso, Kuunibe et al. (2020) examined the changes in the service mix following the implementation of PBF. The authors conclude that PBF can produce modest increases in service provision without altering the overall service mix, possibly due to the wide range of indicators that were incentivized by the program. The results do show, however, that there was a decrease in child complete immunization. It is not clear whether this was due to a shortage in vaccines in the region or the fact that this service was not strongly incentivized relative to the amount of work required for providers (unpublished data).

In addition to privileging incentivized *services or tasks*, healthcare providers may also have vested interests in selecting or avoiding certain types of *patients*. Cherry-picking or cream-skimming refers to the fact that providers can chose patients who help them score well and with whom they gain more benefits (e.g., income and/or time) (Borgès Da Silva et al., 2015; Oxman & Fretheim, 2008). In contrast, dumping refers to the fact that providers can explicitly avoid patients who cause them to score poorly and cost more in terms of income and/or work time (Borgès Da Silva et al., 2015). Ellis (1998) also refers to the concept of skimping, defined as the under-provision of services to high cost patients. In high income settings, economic models have shown that changes in financial incentives associated with different types of patients can encourage patient selection (Borgès Da Silva et al., 2015). In Taiwan, a pay-for-performance program improved the care for some diabetes patients but doctors appear to have cherry-picked certain patients and excluded sicker ones from the program (Chang et al., 2012). Studies on this P4P program showed that lower severity cases were selectively enrolled (Chang et al., 2012). Older patients and patients with more comorbidities or more severe conditions were prone to be excluded (Chen et al., 2011).

In LMICs, the risk of cherry-picking patients has been discussed but little empirical evidence exists on this topic (Chalkley et al., 2016; Gorter et al., 2013; Ireland et al., 2011; Kalk et al., 2010; Witter et al., 2012). In the Democratic Republic of Congo, an impact evaluation report funded by USAID states that, on probing, none of the respondents alluded to cherry-picking but the authors concluded that it cannot be ruled out (Sadaphal & Bongiovanni, 2016). The concepts of dumping and skimping patients have received even less attention in the literature on PBF in LMICs. Future research should examine whether the distinctive features of the PBF models promoted in LMICs encourage cherry-picking, skimping or dumping. For example, studies could assess whether providers are more likely to cherry-pick children or indigent patients if the fee for treating these patients is higher and if they do not require too much work. Alternatively, busy providers may be more likely to skimp or dump children or indigent patients if these types of cases are too severe and they take away from the providers' leisure time.

Measure fixation

Smith (1995) defines measure fixation as an emphasis on measures of success rather than on the underlying objective. Unfortunately, we only found one study reporting data on this unintended consequence in Rwanda (Kalk et al., 2010). To understand the strengths and weaknesses of the PBF program, 69 interviews were conducted with Ministry of Health staff, hospital management and administrative staff, doctors and nurses in a district hospital, and patients. Participants explained that indicators are supposed to inform about "a bigger issue behind", which is more difficult to measure. However, by offering incentives for very precise indicators, these indicators become "dissociated" from their meaning and lose their rationale. For example, participants indicated that the existence of a correctly filled-out partograph can be completely disconnected from a successful delivery. Future research in LMICs should examine whether PBF leads providers to fixate on performance measures rather than on the underlying objectives, especially since this was found to be a key unintended consequence of the Quality and Outcomes Framework (QOF) in the United Kingdom (Lester et al., 2011).

Gaming

In economics, gaming refers to using a system's policies, regulations, and procedures to increase one's financial gain in a manner contrary to the intended purpose of those policies, regulations, and procedures (Hurley, 2010). Smith (1995) defines gaming as the deliberate manipulation of behaviour to secure a strategic advantage. For example, participants in Rwanda reported not distributing the last drug box of the pharmacy to avoid a stock-out (Kalk et al., 2010). In his thesis, Abomo Kele (2018) reviews reports from the *Agence Européenne pour le Développement et la Santé* (AEDES), a consultation agency in charge of conducting external verifications for PBF in Benin. According to Abomo Kele, the reports reveal the adoption of numerous gaming strategies to increase PBF scores. For example, providers staged an emergency service and maintained a human presence in the room to show PBF evaluators and get the 100 points assigned to this item in the evaluation grid. Moreover, the staff members' schedules presented during the PBF verification did not correspond to reality. A nurse's aid confirmed that she worked in a laboratory and had received instructions to occupy another department on the day of the verification only. Similarly, a study on PBF in Cameroon showed that providers sometimes developed strategies to improve their personal gains (Sieleunou et al., 2019). Some facilities used separate drug cupboards for the tracer drugs evaluated during PBF verifications to ensure they met evaluation criteria, causing a fragmentation of the drug management system.

Misrepresentation of performance data

The misrepresentation of performance data refers to the deliberate manipulation of information, such that reported behaviour differs from actual behaviour (Smith, 1995). Some authors classify this as a type of gaming strategy (Chalkley et al., 2016). In Rwanda, Kalk et al. (2010) found that providers regularly distorted information to improve reported results. This included the arbitrary and retrospective filling-out of forms. Most staff felt justified in distorting information because of the perceived inappropriateness of some indicators and because of lack of time to do the job properly. In another study, providers in Rwanda reported that P4P incited them to forge data in health facilities in order to meet targets (Chimhutu et al., 2014). A year later, however, most participants argued that forging data was impossible for practical reasons. The authors did not provide an explanation for this change in discourse.

In his thesis, Abomo Kele (2018) presents AEDES reports revealing that providers in Benin falsified medical registers. For example, a false prenatal consultation was added for a woman who had given birth in order to make her maternal form meet the PBF's criteria of four prenatal consultations. In his study, local stakeholders also questioned the effectiveness of community verifications for identifying cases of cheating and applying penalties. One participant explained that after 10 years of field experience, he had come to the conclusion that few decisions were taken on the basis of these community verifications (Abomo Kele, 2018).

In Burkina Faso, Kuunibe et al. (2019) found that healthcare providers overreported and underreported the quantity of healthcare services in the PBF forms. Misreporting varied by service indicator and health district. However, the authors cautioned that there is a need to conceptually distinguish fraud from misreporting, as the latter can also be due to calculation errors. Moreover, it is important to note that 100% concordance between the quantity that providers declared and the quantity that the PBF verifier manually counted in the medical registers does not guarantee those services were truly provided to patients, especially if falsification occurs directly in the medical registers.

To our knowledge, there are no studies documenting "upcoding" in LMICs. This occurs when providers claim a reimbursement for a service that pays more than the service that was actually provided (Hurley, 2010). The term "upcoding" is quite popular in the literature on the unintended consequences of PBF in high-income countries. The lack of studies documenting this type of behaviour in LMICs may be due to the specific design features of the PBF programs in LMICs, the difficulty of studying this unintended consequence, or simply the lack of attention paid to this issue. However, the PBF toolkit does explain that paying for consultations among children under five poses verification difficulties in actual settings. This service is easy to manipulate because exact ages can be difficult to verify (Fritsche et al., 2014).

Prioritizing quantity over quality of care

An important concern regarding PBF is that paying for volume of services can encourage providers to reduce the quality of care when it is not noticeable by patients or difficult to measure. In Rwanda, healthcare providers felt that P4P was forcing them to prioritize quantity over quality of care (Chimhutu et al., 2014). The study revealed concerns that PBF led to “turnstile medicine", wherein consultations are rushed to fit more in, so as to reach a specific targeted quantity. To minimize this risk, the World Bank's PBF model recommends that the purchasing of healthcare services be conditional upon their quality (Fritsche et al., 2014). A quantified checklist can be used to measure and reward technical quality (e.g., physical infrastructure). However, some dimensions of quality are difficult to capture during verifications due to non-verifiability, lack of time, or financial constraints. Moreover, the thresholds for quality indicators are low.

Overburdened healthcare staff and verifiers

One potential adverse unintended consequence of PBF is the work overload that it may place on verifiers and healthcare workers. This issue was reported in various contexts. In Kalk et al.’s (2010) study in Rwanda, participants frequently described PBF as putting additional stress on a system already overstretched. Seventy-two per cent of medical staff reported to regularly work supplementary hours and to feel constantly tired because of the workload under the PBF program. Participants argued that clinical work, meetings introduced for P4P supervision and the filling of all forms would require about 12 hours daily for each nurse and medical doctor. Similarly, participants in Benin complained that PBF increased the workload and time worked without being sufficiently compensated by performance premiums (Paul et al., 2014). In Tanzania, staff spent 17 percent of their time on data generation and verification for PBF

(Binyaruka et al., 2015). In Uganda, a study on why performance-based contracting failed showed that the overwhelming workload of extracting performance data and the limited ability of auditors to identify and tally services caused difficulties (Ssengooba et al., 2012). It should be noted, however, that the auditing team in this last study was concurrently tasked with collecting data for the impact evaluation, thereby contributing to their workload.

Ossification

Smith (1995) defines ossification as an organizational paralysis brought about by an excessively rigid system of performance evaluation. In Uganda, Ssengooba et al. (2012) conducted a study aimed at understanding the performance-based contracting (PBC) intervention, its mechanisms, and its potential effects. Results showed that inadequate time was allowed for the selection of service targets by health centres. Targets were selected hastily during the district meetings organized to generate pilot awareness. Without prior warning, the invited managers from each health centre were asked to select the targets and sign a contractual agreement for the pilot. Participating centres were then locked-in to the poor choices they had made in haste and whose contexts subsequently changed significantly. For example, some of the selected targets became less relevant following the implementation of parallel health interventions in the country.

Reduction in intrinsic motivation

A potential unintended consequence often debated in the literature is the possibility that PBF would reduce or “crowd out” intrinsic motivation (Kalk et al., 2010). While presenting the taxonomy of human motivation is beyond the scope of this review, it is important to note that humans have different levels and kinds of motivation, which can be influenced by performance-contingent rewards. Deci and Ryan (2000), the authors of the Cognitive Evaluation Theory (a subtheory of the Self-Determination Theory), distinguish between intrinsic motivation, which is related to the interest and enjoyment of conducting a task itself, and extrinsic motivation, which refers to doing something because it leads to a separable outcome. The extent to which basic human needs for autonomy, competence, and relatedness are satisfied either facilitates or undermines extrinsic or intrinsic motivation. The concern that PBF can undermine intrinsic motivation stems from Deci and Ryan’s work:

"*The research began with the demonstration that extrinsic rewards can undermine intrinsic motivation…, which we interpret in terms of the reward shifting people from a more internal to external perceived locus of causality. Although the issue of rewards has been hotly debated, a recent meta-analysis…confirms that virtually every type of expected tangible reward made contingent on task performance does, in fact, undermine intrinsic motivation.*" (Ryan & Deci, 2000, p. 59)

Ryan and Deci (2000) argue that the quality of performance can be very different when one is behaving for intrinsic rather than extrinsic reasons. It should be noted, however, that there are different types of extrinsic motivation, some of which represent impoverished forms of motivation and some of which represent active, agentic states. These depend on the extent to which the social contextual conditions support one's feelings of competence, autonomy, and relatedness. Lohman et al. (2016) explain that recent literature in psychology differentiates motivation according to whether it is congruent with the individuals' goals, values, and preferences. Externally stimulated behaviour may have an internal locus of causality if it is aligned with one's goals, values, and preferences.

Within this debate, PBF proponents claim that the motivation arguments regarding PBF are "ambiguous and cannot be directly applied to LMIC", where salaries are very low (Fritsche et al., 2014, p. 309). They also point out that PBF entails elements that can foster intrinsic motivation, such as autonomy in the use of funds and strong management support (Lohmann et al., 2016). Moreover, some proponents have adopted an intensity approach, arguing that it is the overall amount of motivation that matters rather than its composition (e.g., intrinsic versus extrinsic) (Lohmann et al., 2016).

A few empirical studies have examined whether PBF reduces or crowds out intrinsic motivation in LMICs. In Malawi, Lohman et al. (2018) conducted a mixed methods study showing that PBF did not affect health workers' overall intrinsic motivation levels. The intervention had both positive and negative effects on the health workers' basic psychological needs for autonomy, competence, and relatedness, which are central to intrinsic motivation. There was substantial individual variation in the extent to which positive or negative effects appeared to predominate.

In Ghana, Aninanya et al. (2016) conducted a quasi-experimental study design with pre- and post-intervention measurement to determine the impact of performance-based incentives on maternal health workers' motivation. The results of a difference in difference analysis comparing pre- and post-intervention differences in intrinsic motivation scores between intervention and comparison arms were not statistically significant. All healthcare workers revealed that extrinsic rather than intrinsic motivation inspired them to work harder.

In Burundi, Bertone and Meessen (2013) found no evidence of crowding-out of intrinsic motivation in two PBF schemes. In one of the schemes, however, the capped bonus became a "fixed" extra remuneration for most facilities, possibly leading to a reduction of intrinsic motivation. Health workers began to perceive this bonus as a right and expected it every month, even when performance did not justify it. Providers spent time on endless discussions to obtain it rather than on providing services to patients.

In Uganda, Lutwama, Roos, and Dolamo (2013) assessed the implementation of PBF through questionnaires and semi-structured interviews with providers and managers, respectively. The findings suggest that the loopholes in the implementation of PBF likely influenced the providers' motivation negatively. For example, there were inadequacies in setting performance targets, irregular performance assessments, limited prospects for career progression, inadequate performance feedback, and poor reward mechanisms. Eighty percent of the workers were not satisfied with the fringe benefits received. Healthcare workers were promised financial incentives that never fully came through.

In Burkina Faso, Fillol et al. (2019) conducted a multiple case study to examine how the organizational context within which PBF was implemented shaped health workers' motivational reactions. The results showed that the extent to which PBF contributed to positive, sustainable forms of motivation depended on partly on: 1) the extent to which existing hierarchies fostered participation and transparency; 2) the managers' handling of the performance feedback; and 3) the facility's pre-PBF levels in regard to infrastructure, equipment, and human resources.

More recently, Maini et al. (2019) conducted a cross-sectional study in the Democratic Republic of Congo to examine the differences in motivation between health workers who recently had PBF withdrawn and workers who had never received PBF. Results showed that workers in facilities where PBF had been removed scored significantly lower on all dimensions of motivation except “satisfaction with tasks” (no significant differences) and “level of extrinsic motivation” (marginally significantly higher scores) than those in non-PBF facilities. The removal of the PBF scheme was blamed for an exodus of staff due to the reduction in income and had a negative impact on relationships between staff and the local community. Some nurses noted that colleagues had become less welcoming and even rude to patients since donor payments to workers had ceased. The authors concluded that donors and governments unable to sustain PBF payments should have clear exit strategies and institute measures to mitigate any adverse effects on motivation following withdrawal.

Also in the DRC, Huillery and Seban’s (2014) study found that provider attendance was higher in the incentivized health facilities than in the comparison facilities when the incentives were in place. However, attendance was found to be lower a few months after the incentives were withdrawn. Also, the previously incentivized health workers were found to attach more importance to job material benefits than to non-material benefits, in contrast to non-incentivized health workers.

In summary, the unintended consequences of PBF on motivation has fueled some debate in the literature. A few studies using different quantitative and qualitative methods found mixed results regarding this complex issue. These studies have shown that the organizational context, individual variations, implementation loopholes as well as the type and sustainability of incentives can influence how PBF influences intrinsic and extrinsic motivation.

Conflicts among actors of the healthcare system

The PBF model aims to promote team spirit and collaboration among actors of the healthcare system (Fritsche et al., 2014). However, some research suggests that PBF can cause conflicts among actors of the healthcare system. Lohman et al. (2018) found that the PBF intervention in Malawi led to interpersonal challenges by allowing healthcare facilities to freely decide how to share subsidies among staff members. Almost all participants reported that the tension and fighting that ensued were a source of frustration. On the other hand, some participants described how the intervention stimulated improved team work as staff members worked towards the common goal of achieving PBF targets.

In Tanzania, Chimhutu et al. (2016) found that the distribution of P4P bonuses caused frustrations. Providers who delivered services related to reproductive and child health (RCH) received more bonuses than did others, due to the centrality of their role in meeting targets. Non-RCH staff objected to getting less bonuses than RCH staff, while RCH staff, who were providing the targeted services, complained about not getting more bonuses. The frustrations that ensued negatively affected work motivation, undermined teamwork across departments, and created tensions among workers in health facilities.

In Ghana, Aninanya et al.'s study (2016) a minority of participants reported that PBIs fostered competition between maternal health workers that stifled teamwork and that PBIs were not equitable because only some benefited.

Together, these various studies suggest that inequitable distribution of subsidies may be a common source of conflict and frustration in PBF interventions. A comparison of findings across studies also suggest that PBF may improve teamwork when healthcare workers are working towards a common goal but that the competition introduced by PBF interventions may also stifle teamwork in some contexts. These studies, however, did not sufficiently elaborate on the possibility that the effects of PBF may be different for intra-facility versus inter-facility collaboration.

Provision of unnecessary services

To improve the quality of care and services, PBF aims to rationalize the prescription of medication and ensure that patients receive the appropriate services they need. Appropriateness of care has been defined as the right care provided by the right providers, to the right patient, in the right place, at the right time, resulting in optimal quality care (Canadian Medical Association, 2015). In some cases, however, fee-for-service payments may actually hinder appropriate care. Healthcare workers who want to increase their profit may provide wrong, unnecessary or excessive tests, treatments and procedures. Avoidable hospitalizations constitute a common example of inappropriate care. Participants in Kalk, Paul, and Grabosch's study in Rwanda (2010) were aware of the risk of inappropriate care. Some participants argued that unnecessary clinical interventions, whether it be a simple vaccination or a caesarean section, represented a physical injury inflicted rather than a medical act rendered. Thus, it is essential that the additional healthcare delivery be supplied to the right patients. In this literature review, however, we did not find empirical studies that explored this issue in LMICs, potentially due to the rarity of research on unintended consequences and the challenge of documenting this issue.

Adoption of coercive strategies to attract patients to healthcare facilities

Chimhutu, Lindkvist, and Lange (2014) conducted a qualitative study in Tanzania to explore the strategies that health workers employed to reach PBF targets. Results showed that providers developed negative strategies to attract and/or force patients to utilize their services. Women were regularly told that they would be fined if they delivered at home or that they would be denied a live birth card or vaccination for their newborn. Although fines were just a threat, some sanctions were carried out. Providers had few scruples about sanctioning women who delivered at home.

Inequities between populations, staff members, or facilities

PBF is often presented as a way to increase equity in service delivery (Eldridge & Palmer, 2009). Research suggests that PBF could increase or decrease equity at three levels: 1) population; 2) staff; and 3) facilities.

At the population level, PBF could potentially influence equity by providing staff incentives to lower formal user fees or informal payments in order to attract patients (Fritsche et al., 2014). In Tanzania, for example, Binyaruka et al. (2015) found that P4P was associated with a 5% reduction in those paying out-of-pocket for deliveries, but there was no effect on the average amount paid. In the Democratic Republic of Congo, Huillery and Seban (2014) found consistent evidence that user fees for targeted services were lower in PBF facilities than in fixed-payment facilities. In Cameroon, the impact evaluation suggested that there was a reduction in formal and informal user fees in the PBF group compared to other control groups (de Walque et al., 2017).

Other researchers, however, have argued that providers may cherry-pick patients who are the healthiest (adverse selection) or are the easiest to reach, rather than those who are most in need of healthcare services (Ireland et al., 2011). Mixed evidence has been published on the pro-rich nature of PBF schemes. Using data from an impact evaluation in Rwanda, Lannes et al. (2015) showed that PBF improved access to healthcare services for patients who were easier to reach and generally more affluent. Regarding institutional deliveries, for example, PBF favoured those who did not have a financial barrier to access the service, that is, women with higher economic statuses and those with lower economic statuses who had health insurance (Lannes et al., 2015). These results contradicted a prior study in Rwanda based on Demographic and Health Survey data, which found no consistent pattern in facility deliveries relative to household wealth status (Priedeman Skiles et al., 2013). Research conducted in Burundi also found some heterogeneity of effects across poor and non-poor populations. The PBF effect on the probability of using care when ill was found to be smaller for the poor (Bonfrer, Soeters, et al., 2014). Similarly, institutional deliveries increased significantly among the non-poor only (Bonfrer, Van de Poel, et al., 2014). In Afghanistan, however, P4P did not have an effect on the equity of institutional deliveries or on children's utilization of outpatient services (Engineer et al., 2016).

For staff, PBF can influence equity if the distribution of premiums is unfair. A study examining PBF programs in Benin revealed that this was a major issue (Paul et al., 2014). It was not always the staff who produced results in the field who obtained compensations. Staff suspected the hierarchy of monopolizing the benefits of PBF. This was confirmed by the level of premiums reported by participants. In Rwanda, participants also perceived the distribution of rewards within the service unit as unfair (Kalk et al., 2010). Medical doctors were viewed as more privileged than others. Alternatively, Huillery and Seban's (2014) study conducted in the Democratic Republic of Congo found that the autonomy of payment allocation among facility staff in the PBF group led to a more egalitarian distribution of payments among workers compared to a fixed-payment group, in which each worker was entitled to a given amount of government payment depending on his/her grade and experience. In the fixed-payment group, 77% of health workers received a share of the payment, compared to 93% of workers in the PBF group. PBF benefited non-technical workers (pharmacists, managers, secretaries, receptionists, and maintenance workers) who were not on the government payroll and therefore did not receive a share of the fixed payment.

At the facility level, PBF may influence equity by rewarding those who are in a better position to meet targets or who are in areas with more resources (Eldridge & Palmer, 2009; Ireland et al., 2011). To avoid this, some PBF interventions offer unit fees that are adjusted based on the location of the facility (Bodson et al., 2018).

In Benin, however, participants perceived the PBF system as unfair due to the fact that the same indicators and quality checklists were applied to all facilities regardless of their size or initial situation (Paul et al., 2014). Thus, providers felt penalized when working in disadvantaged facilities that were unable to respect norms because they lacked material, equipment, or staff.

A study on how PBF influenced the availability of essential medication also highlighted concerns regarding inequities at the facility level (Sieleunou et al., 2019). With PBF's principle of management autonomy, facilities received less support from regional entities. Thus, small facilities with low service utilization had to buy small quantities of drugs on a regular basis. They did not have sufficient funds to acquire large quantities. As a result, fixed costs exploded and funds to restock the pharmacies dwindled, triggering frequent stock-outs.

In Tanzania, Canavan and Swai (2008) found that a Cordaid P4P program produced inequities in the allocation of staff bonuses between the different types of facilities (e.g., dispensary, health centre, hospital). The bonuses paid to providers ranged between \$2 and \$7 per month, depending on the facility level, although it is not clear whether this reflected different workloads or performances. In the same vein, Binyaruka et al. (2018) examined a P4P program introduced by Tanzania's Ministry of Health and Social Welfare with support from the Government of Norway. The study found that performance payouts were initially higher in higher-level facilities (hospitals and health centres) than in dispensaries, and higher in facilities with more medical commodities and those serving wealthier populations, but these inequalities declined over time. Moreover, P4P had greater effects on coverage of institutional deliveries among facilities with low baseline performance, serving middle-wealth populations, and located in rural areas. These results suggest that the design of incentives and facility characteristics affects providers' response.

Suboptimal value for money

Unnecessary increases in costs have been identified as a type of unintended consequences of health programs (Bardach & Cabana, 2009). In regard to PBF, some advocates uphold that PBF programs in LIMCs can increase technical efficiency[6] by motivating healthcare providers to make better use of existing resources at the health facility level (Fritsche et al., 2014; Witter et al., 2012). According to the PBF toolkit (2014, p. 308), "*well-designed research in Rwanda indicates that PBF leads to more and better quality health services as compared to just providing more money*", although the complete reference is not cited. In contrast, PBF critics have denounced the high costs of developing and strengthening the necessary structures for PBF (e.g., independent purchasing authorities, entities in charge of verification, civil society organizations charged with community oversight, health information systems) (Ireland et al., 2011; Kalk, 2011).

In 2016, our research team published a review of the literature on whether PBF influenced value-for-money in LMICs (Turcotte-Tremblay et al., 2016). For this article, which is available in Appendix A, we considered PBF to be efficient when improved care quality or quantity was achieved with equal or lower costs, or alternatively, when the same quality of care was achieved using fewer financial resources. The results showed that the overall strength of the evidence was weak. None of the articles selected were full economic evaluations making clear connections between the costs and effects of PBF. Our research team concluded that stronger empirical evidence is needed on whether PBF represents good (or bad) value for money in LMICs.

More recently, a study in Benin found that PBF verification processes are quite costly and time-consuming (Antony et al., 2017). For every $1 paid to providers, about $0.50 was used for verifications. The performance data collected during verifications were often not analyzed or used to provide feedback to stakeholders, limiting the potential to improve services. Similarly, in Cameroon, a key stakeholder reported that for every $1 paid to providers, about $0.40 was used for verifications (unpublished data). The latter rose to about $0.60 once the PBF program reached national coverage. The verification model was described as unsustainable and much too intensive. In 2020, attempts to review the verification model were underway but it was perceived to be a difficult task (unpublished data). In Tanzania, the many costs of pay-for-performance left open to debate whether the strategy was cost-effective (Borghi et al., 2015). The financial

---

[6] Technical efficiency refers to the production of the maximum possible amount of output from the inputs used, given the chosen production method (Hurley, 2010).

cost of the pay for performance pilot in Tanzania in 2012 was $1.2 million, and the economic cost was $2.3 million. The incremental cost per additional facility-based birth ranged from $540 to $907. In this low-income setting, the costs of the program were deemed substantial. The budget for the pilot intervention was distributed as follows: 28% for management, 37% for data production, 13% for verification processes, and 22% for payouts. In Malawi, De Allegri et al. (2019) found that the overall economic costs of the Support for Service Delivery Integration Performance-Based Incentives (SSDI-PBI) program, implemented over four years, amounted to $3,402,187, equivalent to $6.46 per targeted beneficiary. Costs for personnel were higher than the incentive payments. Together, these studies highlight the important opportunity cost for health systems and call for more evidence on the costs of PBF in LMICs.

Pride

In Ghana, Aninanya et al.'s (2016) quantitative data revealed that PBI did not significantly change workers' level of pride. However, in the qualitative interviews, 32 of 33 intervention health workers reported that PBIs helped them feel recognized as important stakeholders in the delivery of maternal care, which enhanced their pride. Participants positively reported meeting colleagues from different facilities and from the district level during awards ceremonies. Obtaining an award made some participants feel recognized. However, one health worker reported no increase, saying she was already proud of her job as a midwife.

2.2.3 Unintended consequences of combining health equity measures with PBF

Few studies have been published on the integration of equity measures in PBF, and particularly its unintended consequences (Ridde et al., 2018). Although program planners believe that these approaches can have a high synergistic potential, there is a possibility that combining them would not work as planned and that they would undermine rather than support one another. The interactions between the different rationales, goals, and operating procedures could lead to unintended consequences.

In Cameroon, Flink et al. (2016) examined a system targeting the poorest of society (i.e., indigents) in a PBF program. Positive effects for indigents included fewer financial worries, an improvement in economic status due to better health, and exposure to other services, such health education and access to food and water. The study also revealed that community health workers focused on easily identifiable groups (e.g., elderly, orphans) in the selection of indigents. Some non-indigent patients reported that the healthcare fees had increased and that the facility was less flexible about giving loans. There were also higher workloads for providers and conflicts about the distribution of performance bonuses (Flink et al., 2016).

In Burkina Faso, Kuunibe et al. (2020) examined what happens to the quantity of healthcare services when performance-based financing meets free healthcare for maternal and child services. Using an interrupted time-series analysis with independent controls, the authors showed that, in the period before the free healthcare policy, PBF produced modest increases for a wide range of maternal and child services. However, in the period after the introduction of the free healthcare policy, PBF did not affect service provision in intervention compared to control facilities. This suggests that the joint implementation of PBF and health equity measures can lead to unexpected results.

Beyond the PBF literature, some studies have shown that user-fee exemption policies for the poor can lead to unintended consequences. A literature review of the disruptive effects of user fee exemption policies on health systems found that they can lead to: 1) an increase in healthcare workers' feelings of being exploited; 2) the unavailability of drugs and delays in the distribution of consumables; and 3) revenue losses for health centres and reimbursement (Ridde et al., 2012).

2.2.4 Main conclusions of the literature review

A few conclusions can be drawn from the literature review presented above. First, little research set out to explore the unintended consequences of PBF from the outset, although some mixed evidence has begun to emerge in LMICs. Second, community surveys and client satisfaction surveys are neglected research topics, despite the importance of these verification mechanisms within PBF intervention models. Third, studies on combining PBF with health equity measures, and especially its unintended consequences, are rare. Fourth, certain useful data collection methods, such as observation on site, have not sufficiently been used. Few studies combined both qualitative and quantitative data. Pertinent studies in HICs rely on quantitative methods only. Sixth, the vast majority of unintended consequences reported in the literature were undesirable. The rest of this thesis will address some of these issues.

Chapter 3 – Theoretical Framework and Research Question

3.1 Applying the theory of diffusion of innovation

We developed a theoretical framework to answer the following research question: **What are the unintended consequences, and their contributing factors, of an intervention combining PBF with health equity measures in Burkina Faso?** This theoretical framework was based on the theory of diffusion of innovations. Everett M. Rogers, a professor of communication studies, popularized this theory in his award-winning book *Diffusion of Innovations,* first published in 1962 (Rogers, 2003). Although it originated in rural sociology, the theory has been influential in numerous scientific fields, including medical sociology, health promotion, development studies, health communications, organization studies and complexity studies (Greenhalgh et al., 2004; Rogers, 2003).

According to Rogers (2003), an innovation is defined as an idea, practice or object that is perceived as new by the adopters. It can be a cluster composed of distinguishable elements that are perceived as being closely interrelated, such as the different components of the complex PBF intervention combined with health equity measures in Burkina Faso. During the diffusion process, the "*innovation is communicated through certain channels over time among the members of a social system*" (Rogers, 2003, p. 11). The meaning of the innovation is gradually worked out through a process of social construction. Because innovations are novel, there is always some level of uncertainty involved in the diffusion process, which implies "*lack of predictability, of structure and of information*" (Rogers, 2003, p. 6).

In some cases, individual members of the system, such as healthcare workers, have to implement the decision made by an authority, such as the ministry of health (Rogers, 2003). Such "authority innovation-decisions" are made by relatively few people who possess power, status, and technical expertise. These actors can initiate the innovation process when they perceive a performance gap, such as the low quantity and quality of healthcare services in Burkina Faso.

Rogers explains that one branch of diffusion research looks into the consequences of innovations. According to the author's literature review, studies in this area tend to examine four categories of independent variables, which can interact to influence consequences: 1) the nature of the innovation, 2) the characteristics of members, 3) the nature of the social system, and 4) the use of the innovation. Recognizing their comprehensiveness and parsimony, we selected these four categories of independent variables as a lens to examine the contributing factors of consequences. Each of these will be discussed below.

First, the nature of the innovation refers to its attributes. One important attribute is the "relative advantage" of the innovation, that is, the degree to which an innovation is better than the idea it supersedes. As in the context of PBF, change agencies sometimes increase the relative advantage of a new practice by offering financial incentives to adopters. Although incentives increase the quantity of adopters of an innovation, the theory stipulates that the "quality" of such adoption decisions may be low, thus limiting the intended consequences of adoption. Another important attribute is "compatibility", that is, the degree to which an innovation is consistent with existing values, experiences, and local needs. According to Rogers, program planners often fail to consider cultural values of adopters, resulting in unanticipated consequences. Three other attributes that have been found to influence the diffusion process of innovations are: the level of "complexity", which refers to the degree to which an innovation is difficult to use; the level of "observability", which refers to the degree to which the results of an innovation are visible; and the level of "triability", which is the degree to which an innovation may be experimented with (Rogers, 2003).

Second, the characteristics of members of the social system are useful for understanding the diffusion process and its consequences. Examples of such characteristics are members' socioeconomic status, their attitude towards change, their perceptions of the innovation, and their past experiences. According to the theory, change agents are rarely able to predict people's subjective perceptions of the innovation or of its consequences, which leads to unforeseen consequences.

Third, it is important to examine the nature of the social system within which the innovation is introduced in order to understand its consequences. This includes local norms as well as the inner and outer environment. Greenhalgh et al. (2004) highlight that organizations assimilate innovations more readily if they are mature, functionally differentiated, and specialized, if they have slack resources for new projects, and if they have decentralized decision-making structures. The organizations' absorptive capacity for new knowledge and the presence of a receptive context for change are also important features that may influence an innovation's consequences.

Fourth, the use of the innovation can influence what consequences emerge. In organizations, the innovation process is divided into two main subprocesses (Rogers, 2003). The initiation subprocess consists of information gathering, conceptualizing, and planning for the adoption of an innovation. The implementation subprocess consists of all of the events, actions, and decisions involved in putting the innovation into use. Once an organization has made a decision to adopt, implementation does not always follow directly. During what Rogers calls the restructuring stage, the innovation is often modified and re-invented to fit more closely with the organization's needs and structure, and the organizational structures are altered to fit with the innovation. The innovation process is facilitated when the innovation is initially developed within the user organization and inhibited when adopters are not provided opportunities to re-invent it. Moreover, Greenhalgh et al. (2004) suggest that adoption is more likely to be successful if adequate feedback is provided to the intended adopters about the consequences of adoption.

Rogers (2003) defines consequences as the changes that occur to a social system as a result of the adoption or rejection of an innovation. The unpredictability of an innovation's consequences is a type of uncertainty in the diffusion process. Consequences can take many forms and are expressed in various ways. Rogers (2003) established three categories for classifying consequences, which we adapted based on recent literature and the need to operationalize these concepts.

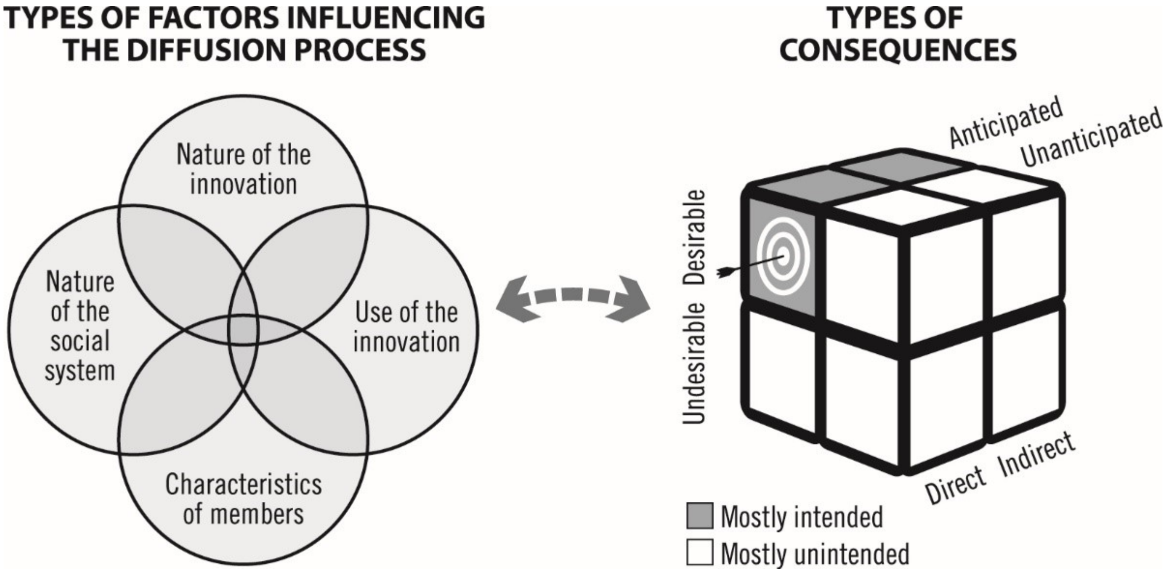
First, consequences can be classified as desirable or undesirable, depending on whether the effects of an innovation are functional for the healthcare system (i.e., positive, producing additional benefits, helping the system work properly) or dysfunctional (i.e., negative, causing harm, not helping the system work properly). Most innovations cause both desirable and undesirable consequences. For example, a desirable consequence can be an increase in effectiveness, while a common undesirable consequence can be the widening of socioeconomic gaps (Rogers, 2003). It should also be noted that a consequence could potentially be both desirable and undesirable, depending on the point of reference, or even neutral, if it does not affect the system in a desirable or undesirable way (Ash, Sittig, Dykstra, et al., 2007; Jabeen, 2016; Koch & Schulpen, 2018).

Second, consequences can be direct or indirect, depending on whether the changes to a social system occur as an immediate response to an innovation or as a second-order result of the direct consequences. To operationalize these concepts in the context of this study, we integrated Ash's (2007) approach, by considering direct consequences to be related to processes and indirect consequences to outcomes. Processes referred to what is actually done in giving and receiving healthcare, while outcomes referred to the consequences of services on patients, including changes to behaviour, knowledge, patient satisfaction, and health status (Donabedian, 1988).

Third, consequences can be anticipated or unanticipated, depending on whether the changes are recognized by the members of a social system. The literature provides little guidance to operationalize this classification. The level of anticipation can vary depending on the stakeholders (e.g. researchers, decision-makers, providers, patients) so we found it necessary to define our perspective. For this study, we considered consequences to be anticipated if they were addressed in the implementation guides or if they were in line with the "spirit of the intervention" or its "ideas" (i.e., beliefs, assumptions or perceptions)[19], according to PBF experts.

Rogers did not actually define or use the labels "intended" or "unintended" to categorize consequences, although the concept of intentionality is a common theme permeating his work. Thus, we refined Rogers' classification by considering that the following types of consequences tend to be unintended by program planners: undesirable/anticipated, undesirable/unanticipated, and desirable/unanticipated. Our rationale for classifying these consequences as unintended was that program planners are not likely to purposefully target changes they consider undesirable or that they have not anticipated. Like Bloomrosen et al., (2011) we expected that consequences that are desirable/anticipated would tend to be intended by program planners. As Jabeen[20] argued, program planners trying to promote a new intervention are likely to have listed and exhausted all the desirable outcomes that they foresee in the program's objectives. These intended consequences were beyond the focus of this study and were examined through a separate impact evaluation (De Allegri, Lohmann, et al., 2018). However, we did remain open to the possibility that some desirable/anticipated consequences could be unintended if they were, for example, positive spillover effects that were foreseen but not initially targeted by program planners. This conceptualization is consistent with recent literature suggesting that unintended consequences can be either anticipated or unanticipated as well as desirable or undesirable (de Zwart, 2015; Koch & Schulpen, 2018). **Figure 1** illustrates the framework presented above.

Figure 1. Theoretical framework



3.2 Justification for the selection of this theory

Rogers' theory on the diffusion of innovations is useful to investigate the unintended consequences of complex interventions (Rogers, 2003). The theory has been used to guide process evaluation in the past (Scott et al., 2019). Yet it remains an original conceptual framework that has not been applied to PBF in LMICs. It is one of the rare theories that provides a detailed typology of consequences. It is also comprehensive and insightful in taking into account the entire diffusion process of innovations as they course through the structure of a social system. It has helped us identify the four pillars that will guide this research: the actors, the social systems, the innovation, and the consequences. It also integrates useful concepts for studying healthcare systems, such as complexity, uncertainty, the interrelations between parts of a system, temporality, social inequalities, etc. Moreover, the theory's constructivist approach will be pertinent to answer our research question. This approach will enable us to examine how the perceptions of local actors and the social processes influence the emergence of unintended consequences. The theory's multidisciplinary nature will also allow us to enrich our analysis by drawing knowledge and methods from various scientific fields (e.g. anthropology).

Chapter 4 – Methodology

4.1 Research paradigm

Paradigms refer to how researchers view the world and go about conducting research (Creswell & Plano Clark, 2007). This research was conducted from a constructionist[7] approach, which suggests that multiple realities are perceived, constructed, and interpreted (E G Guba & Lincoln, 2005). The understandings or meanings of phenomena are formed through the participants' subjective views in interaction with the researcher. What people perceive as real has real consequences (Patton, 2015). Thus, it is important for researchers to visit participants at their sites to collect data representing their different perspectives (Creswell & Plano Clark, 2007).

Key features of constructionism are pertinent for studying the emergence of unintended consequences resulting from health interventions. First, constructionism emphasizes people's actions and roles in creating and transforming social structures (Rouleau, 2007). Structures and modes of governance exist through the people that produce, adapt, and transform them during their daily practice, interactions, and conversations. Thus, to study unintended consequences of health interventions, it is crucial to understand how people perceive, adapt, and transform the intervention model in the field, triggering real consequences that are not targeted by program planners.

Second, constructionism embraces the complexity of social phenomena (Gaudet & Robert, 2018). For example, it recognizes the necessity of taking into consideration the contexts of actions that result from the unpredictable movement of alliances among different interest groups or different concomitant events (Rouleau, 2007). The results of the interactions between such

---

[7] The terms constructionism and constructivism are often used interchangeably, but some authors have identified distinctions (Gaudet & Robert, 2018). For this thesis, I opted for the term constructionism, which tends to be more oriented towards interactions between individuals or groups, whereas constructivism tends to be more focused on individuals and their capacity for action (Rouleau, 2007). Moreover, the former is associated with a more moderate perspective than the latter (Gaudet & Robert, 2018).

complex phenomena are difficult to control. Thus, constructivism posits that intentional action rarely results in expected outcomes (Rouleau, 2007). This premise is perfectly coherent with the current research question.

Within the overarching approach of constructionism, this study was more closely aligned with a moderate position, referred to as “realism”, which recognizes that the nature of scientific reality and of social reality does contain an objective dimension (Rouleau, 2007). The objective world imposes real constraints on individual actions, but the interpretation of this world is diverse (Rouleau, 2007).

It is also possible to distinguish between different approaches within constructionism by considering the way one conceives the relation between individuals and social structures (Rouleau, 2007). In this regard, the position adopted in this study was more closely aligned with structurationists, who consider the weight of both actions and structures (e.g., rules, resources) in the construction of social and organizational facts. This perspective, often adopted to study organizations, is coherent with the need to examine interactions between people and structures to understand the emergence of unintended consequences of PBF.

Constructionism is often associated with qualitative methods, but quantitative methods are sometimes used to conduct exploratory statistical analyses (Gaudet & Robert, 2018). For the choice of methods, I advocate for pragmatism, wherein researchers are encouraged to collect data according to “what works” to address the broad scope of research questions that we face in public health, combining both qualitative and quantitative data when deemed useful (Creswell & Plano Clark, 2007). This approach resulted in the use of both qualitative and quantitative data in this study.

4.2 Research strategy

This doctoral research was nested within a larger longitudinal process evaluation of the PBF intervention in Burkina Faso. The protocol and some results on the fidelity of implementation and processes have been published (Bodson, 2014; Fillol et al., 2019; Ridde et al., 2014, 2017a).

For the current doctoral research project, we developed and conducted a multiple case study with several embedded levels of analyses to examine the unintended consequences of the PBF intervention in Burkina Faso as well as the contributing factors (Yin, 2009). We selected a multiple case study design because it provides high internal validity due to the depth of analyses (Contandriopoulous et al., 2005; Yin, 2009).

4.3 Study setting

4.3.1 The context in Burkina Faso

General description of the country

This multiple case study took place in Burkina Faso, commonly referred to as "the country of honest men". While the government's official language is French, there are over 70 languages spoken across the country, highlighting its cultural diversity (Eberhard et al., 2019). This landlocked country shares borders with Mali, Niger, Ghana, Benin, Togo, and Ivory Coast, which leads some ethnic groups to migrate for seasonal work. The nation's tropical climate, characterized by a rainy season between May and September and a dry season between October and April, exposes local populations to risks including extreme heat, flooding, drought, and harmattan (dry wind). Burkina Faso's rapidly growing population is estimated at 20.32 million, making it the 59th most populous country in the world (World Population Review, 2019). The majority of the population (77.3%) lives in rural regions and is engaged in agriculture and livestock farming (Ministère de la Santé, 2011). Yet 43.9% of the population lives below the poverty line (Ministère de la Santé, 2011). The country ranks 183rd among 189 countries according to the Human Development Index, a summary measure based on life expectancy, years of schooling, and gross national income per capita (United Nations Development Programme, 2018).

Description of the health situation

Burkina Faso obtained one of the lowest scores across the globe (167/183 countries) for the UHC service coverage index, with considerable wealth-related inequality (Hogan et al., 2018). According to the Atlas of African Health Statistics (2018a), more than 50% of pregnant women did not make at least four ANC visits, which can be life-saving. More than 34% of births were not attended by a skilled healthcare provider.

The dire state of healthcare services has important repercussions on Burkina Faso's morbidity levels. The maternal mortality ratio is 371 per 100,000 live births (WHO, 2018a), while the child mortality rates is 89 per 1,000 live births (WHO, 2018a). The pace of progress for these indicators will not be sufficient to meet SDG targets 3.1 and 3.2, which aim to reduce the maternal mortality ratio to less than 70 per 100,000 live births and the under-five mortality rate to 25 per 1,000 live births by 2030 (WHO, 2018a).

Moreover, Burkina Faso's population urgently needs better financial protection. Out-of-pocket-user fees for consultations and medication severely limit access to healthcare services (Ministère de la Santé, 2011). According to government reports, households directly contribute more than 34% of the sources of health financing, despite various strategies to subsidize some services for targeted groups (Ministère de la Santé, 2017). Wagstaff and colleagues (2018) recently estimated that the incidence of catastrophic spending at the 10% threshold was 4%.

Description of the healthcare system

Since the 1960s, Burkina Faso's national healthcare system has undergone several reforms, which continue to shape its organization (Ministère de la Santé, 2017). In the 1960s and 1970s, the healthcare system was strongly centralized, with weak infrastructure and low involvement of local populations. Following the Alma Ata Declaration, the primary healthcare approach was introduced, along with a National Health Program, which led to a more pyramidal structure. In the 1990s, following the Bamako Initiative, the Ministry of Health introduced user fees and further decentralized the healthcare system through the creation of community management committees in facilities, locally called *Comités de gestion (COGES*). In June 2016, the government introduced a nationwide free healthcare policy for maternal and child healthcare services, which overlapped with the implementation of the PBF intervention in some districts (Kuunibe et al., 2020).

The healthcare system is organized according to two dimensions: administration and delivery of services (Ministère de la Santé, 2017, p. 2). The administrative dimension comprises three levels: central (Ministry's office), intermediate (regional health departments), and peripheral (health districts). The 70 health districts, spread throughout the country, coordinate the activities of the primary healthcare facilities. More specifically, they are responsible for the operational planning and implementation of health programs.

The delivery of primary healthcare is handled at two levels. First, the 1,698 *Centres de santé et de promotion sociale* (CSPS – health and social promotion centres) generally provide a standard minimum activity package that includes promotion, preventive, and curative services in exchange for user fees (Ministère de la Santé, 2017). Second, the *Centres médicaux avec antenne chirurgicale* (CMA – medical centres with surgical satellite) act as reference structures for other health facilities within their district. Healthcare providers (e.g., nurses, midwives, itinerary health workers) in these facilities receive monthly salaries based on scales for civil servants (Ministère de la Santé, 2017). Other motivational mechanisms for providers exist, such as commissions (*ristournes)*, indemnities, letters of congratulation, continuous training, and appointments to leadership positions. To promote population involvement, representatives from the community are involved in the facilities' management committees, serving as president, accountant, secretary, etc. Also, community-based healthcare workers, selected by their community, volunteer in each village to assist providers with outreach activities (e.g., vaccination campaigns, screenings) (Ministère de la Santé, 2017).

Numerous government reports show that improving the performance of the healthcare system poses major challenges (Ministère de la Santé, 2017; Société d'études et de recherche en santé publique, 2014). Service delivery is characterized by a shortage of high-quality services, especially for vulnerable populations. Reported challenges regarding governance and leadership include corruption, the poor management capacity of district teams, and low community participation. In regard to human resources, challenges include shortages of qualified staff, difficulty retaining staff in some regions, the staff's inhospitality, and lack of staff motivation. Concerns regarding medication and technology include shortages of some essential medications, providers' irrational prescription practices, insufficient maintenance of infrastructure and equipment, and complex procedures for acquiring equipment. With regard to the national health information system, current challenges relate to the provision of quality data to inform decision-making and to the filling-out of necessary medical registers. Health financing issues include the inadequacy of funding, the lack of risk-sharing mechanisms, and the high dependence on external funding (Ministère de la Santé, 2017; Société d'études et de recherche en santé publique, 2014).

Some statistics are useful to understand the health financing sector in Burkina Faso. According to government reports, allocations to the health sector, which increased progressively from 2001 to 2010, represent an average of 12% of the total state budget per year (Ministère de la Santé, 2011). Since 2008, the country has achieved its goal to devote at least 15% of annual budgets to improving the health sector (Ministère de la Santé, 2011). In 2014, the total expenditure on health represented 5% of the gross domestic product (WHO, 2020a). Per capita, the total health expenditure amounts to USD 82 (WHO, 2020a). In 2013, government expenditure represented 58% of total health expenditure which includes contributions by development partners (De Allegri, Lohmann, et al., 2018). Implementation of the 2011–2020 national health development plan was expected to cost USD 12.68 billion, with funds coming from the State's budget, households, local authorities, the private sector, national and international non-government organizations, and bilateral and multilateral partners (Ministère de la Santé, 2011).

4.3.2 The intervention model implemented

Between 2011 and 2018, the Government of Burkina Faso was awarded USD 41.60 million from the World Bank's International Development Association and from the Health Results Innovation Trust Fund, supported by the governments of Norway and the United Kingdom, to fund the design, implementation, and evaluation of PBF (The World Bank, 2019; World Bank, 2014). In 2011, the government of Burkina Faso conducted a pre-pilot PBF test in three districts to "*gain field experience in order to refine the national strategy*" (Ministère de la Santé, 2013b, p. 11). In 2014, the intervention model was modified and scaled up to an additional 12 districts. The intervention covered 4.5 million people and involved over 576 healthcare facilities. Four intervention arms combining different PBF and equity measures were tested:

- PBF 1: *Performance-based financing only*. Health facilities were paid according to the quantity and quality of healthcare services delivered. Patients were required to pay user fees.
- PBF 2: *Performance-based financing + systematic targeting and subsidization for indigents*. PBF was combined with a community-based selection of indigents (initially

targeting 15–20% of the population), who were exempted from paying user fees. A specific PBF indicator was introduced for purchasing services for the poor, using a higher fixed-unit price than for non-targeted patients.

- PBF 3: *Performance-based financing + systematic targeting and subsidization for indigents + additional provider motivation*. PBF was combined with a community-based selection of the poor, who were exempted from paying user fees (initially targeting 15–20% of the population). Healthcare providers were paid more[8] for services than in PBF 2. The higher reimbursement rate was supposed to act as a financial incentive for healthcare providers.
- PBF 4: *Performance-based financing + community-based health insurance, including targeting of indigents*. PBF was combined with CBHI provided to the general population. The insurance premiums of the poor (targeting 15–20% of the population) were subsidized. Healthcare services were purchased at varying payment levels, as defined in PBF 3 (with higher payment levels for services provided to the poor). Thus, subsidization of services provided to the poor, who did not pay the premiums, passed via the PBF mechanism at budget-neutral levels, while capitation payments replaced out-of-pocket payments for the general population.

Figure 2 presents the intervention's logic model to illustrate the program planners' hypothesized intended consequences. A description of the intervention, inspired by a toolkit published by the World Bank, was presented in a 350-page implementation guide (Fritsche et al., 2014; Ministère de la Santé, 2013a).

This guide presents the separation of functions among the various actors involved in the intervention in order to increase transparency and minimize conflicts of interests (e.g., purchasing, regulating, providing services, paying for services). At the national level, a technical unit called the *Service technique FBR* [ST-FBR – results-based financing - technical service] was integrated into the Ministry of Health to manage the implementation of PBF. Its mission was to develop guidelines, establish PBF quantitative indicators and quality criteria, provide

---

[8] To avoid repetition, the unit fees of the different intervention models will be presented in the appendices of the articles in the Results section.

training to contractualization and verification agencies (CVA), and conduct technical monitoring. In turn, the CVAs were expected to establish contracts with facilities to purchase their services following verifications.

For the monthly quantitative verifications, providers were expected to manually count how many services were recorded in medical registers and report the "declared quantities" on a PBF form. Then, an independent PBF auditor from the local CVA was to travel to each facility to manually recount the services recorded in the medical registers and report what was locally called the "validated quantities" (i.e., verified quantities) on the same PBF form. If the discrepancy between a provider's declared quantities and the PBF auditor's validated quantities for a category of service was less than 10%, the facility was supposed to receive unit fees for each subsidized service provided. If the discrepancy was greater than 10%, that category of service was to be disqualified and unpaid, regardless of the explanation (e.g., calculation error, attempted fraud). In total, 23 indicators were supposed to be subsidized for the quantity of care at the primary care level. The unit fees for some services (e.g., maternal and child health) were higher, in an attempt to incentivize them and cover expenditures (e.g., transportation fees). The unit fees evolved over the course of the study due to what local actors called the "dynamic" nature of the intervention model as well as the implementation of the free healthcare policy for maternal and child services two years after PBF was expanded. To reduce inequities, an equity bonus was to be applied by varying the unit fees depending on the facilities' location and characteristics. After the data were processed and synthesized at higher levels, a "consolidated invoice" was to be released with the amounts of subsidies earned by each facility within the district, to stimulate competition between facilities. The payer, i.e., a government body called the *Programme d'appui au développement en santé*[9] (PADS – program to support health development) was expected to transfer quantity-related payments, locally called "subsidies", within 23 working days following the quantity verifications.

For quarterly quality verifications, a team composed of district management team members (e.g., doctor, midwife, nurse and pharmacist) was to assess the facilities' technical quality of

---

[9] The PADS is Burkina Faso's initiative to better coordinate and manage resources for the implementation of the National Health Development Plan (Beaugé et al., 2018).

care by sampling records from various medical registers and observing the facilities' environment. Scores were to be reported in a 113-item grid. The quality scores determined part of the quality-related payments, locally called "quality bonuses". At the end of verifications, PBF auditors formulated recommendations and discussed points to be improved with healthcare staff.

To strengthen the voice of the population, the level of patient satisfaction also influenced quality bonuses. Thus, every three months, auditors were supposed to select a sample of patients who visited the facility. Then, local associations were to trace these patients in their community to conduct patient satisfaction surveys. The results of the survey determined the facility's "patient satisfaction score", which represented 15% of the overall quality score. During the time period covered by this study, only facilities with an overall quality score of 50% or more could receive "quality bonuses" from the PADS which, as the payer, was expected to transfer funds to facilities within 31 working days following quality verifications. The minimal score to receive quality bonuses was increased in later years. The formula to calculate the quality bonus was:

Quality bonus = total amount of subsidies received for quantity of care during the previous three months x 25% x overall quality score

Patient satisfaction surveys were coupled with community verifications to check whether the medical data recorded in registers had been falsified. In theory, the results of community verifications could lead to sanctions if fraud was detected.

Facilities were also supposed to receive "quality improvement bonuses" to cover important expenditures planned by facilities. The amounts for quality improvement bonuses were supposed to be based on the budget available, the needs expressed in the performance improvement plans, and the relevance of expenditures.

According to the implementation guide (Ministère de la Santé, 2013a), PBF also aimed to introduce the use of new management tools. Every month, for example, providers were expected to fill out an "index tool" to manage budgets and distribute PBF-related payments. The tool synthesized the facilities' total revenues (e.g., user fees, sales of medication, PBF subsidies and bonuses), planned expenditures (e.g., salaries of staff, medication purchases, operating costs, investments), and funds to be placed in bank reserves. The tool also specified the amounts that could be distributed to motivate staff members (in addition to their monthly salaries), promote their "spirit of initiative", improve their responsiveness to patient needs, and, according to the local discourse of program planners, discourage the sales of medication "under the table". The distribution of incentives, locally called "premiums", to staff members was to be based on their qualification, seniority, responsibilities, days of absence, and results on an individual performance assessment. The index tool was supposed to involve the participation of the entire staff to make the funds management more transparent and create a positive group dynamic.

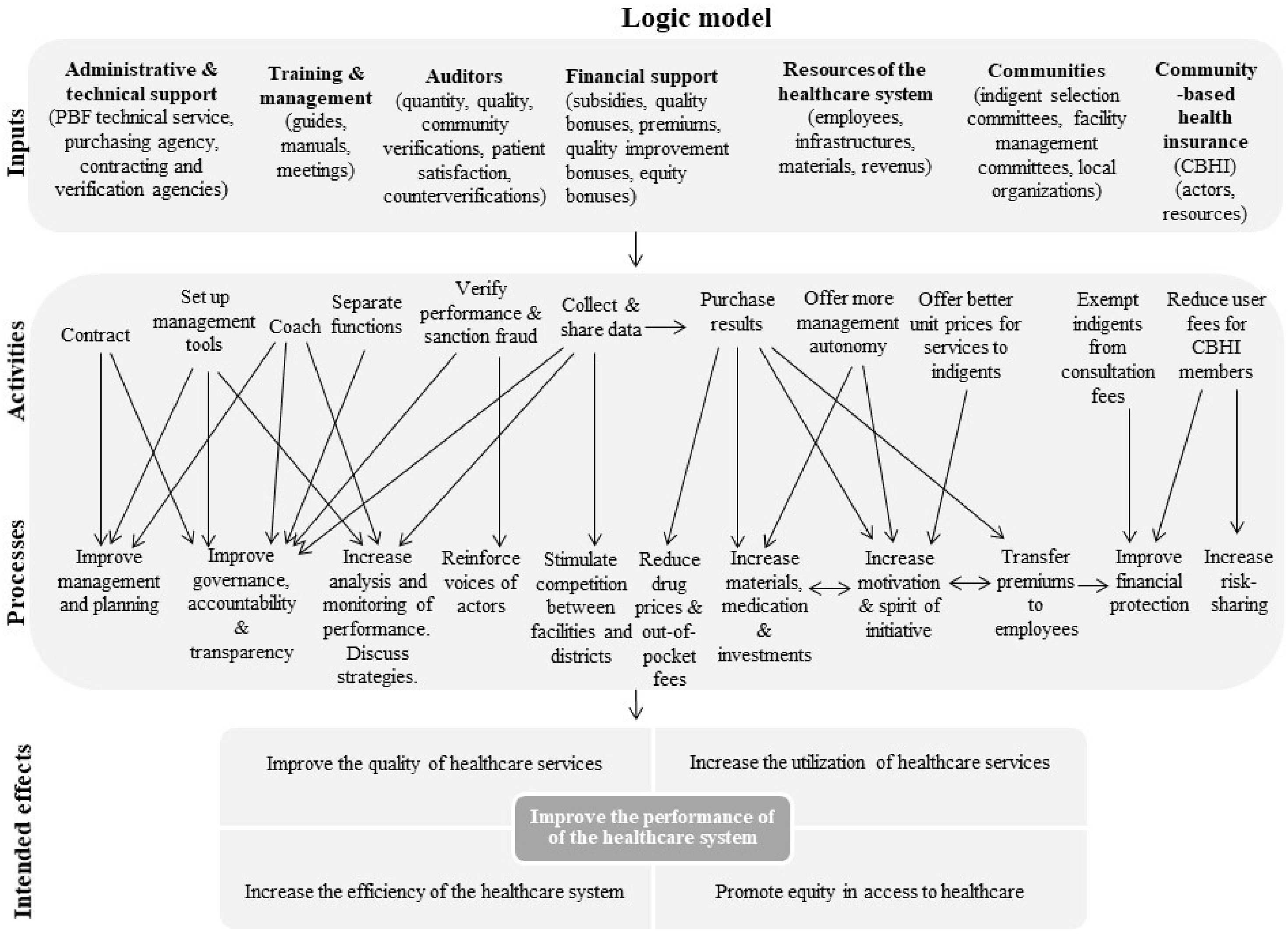
PBF program planners also introduced a management tool called "*Plan d'amélioration de la performance*" (PAP – Performance improvement plan), considered to be a condition for PBF contracts. Providers and management committees were expected to complete these plans together, on a quarterly basis, to set objectives, specify strategies to improve the quality and quantity of care, and estimate associated budgets. The development of the PAP was expected to take 1-2 days, with the assistance of contracting and verification agents, who were expected to coach facility managers and providers in the use of new management tools.

Increasing the management autonomy of facilities was an important principle in the intervention theory. As such, facilities managers were supposed to have the choice to purchase their inputs from independent distributors that operate in competition. The facilities were also supposed to be free to hire and fire staff and decide how to use their own funds.

The intervention model implemented in three intervention arms (PBF 2, PBF 3 and PBF 4) was intended to promote equity by providing indigents access to free healthcare services (Ministère de la Santé, 2014). According to the intervention guide, an indigent was defined as "an extremely socially and economically deprived person unable to support themselves and without internal or external support" (Ministère de la Santé, 2013a). In each village, selection committees were supposed to be set up to select indigents based on their knowledge of the population, living conditions, and the local understanding of the concept of indigence. These committees were supposed to be composed mostly of community leaders (e.g., community-based healthcare workers, counsellors) and at least seven people, three of whom should have been women. To minimize targeting errors, the lists of indigents selected were supposed to be validated by Local Validation Groups and the Local Implementation Teams. Then, the enumerators were supposed to collect data and photos of indigents using electronic tablets, in order to make indigent cards. When the cards were distributed, indigents were supposed to be informed about their rights and the benefits of the intervention. Indigents with cards were not supposed to pay for the following services nor the necessary medication in their primary healthcare facilities: 1) curative consultations, 2) inpatient stays, 3) eutocic deliveries, and 4) family planning consultations (Ministère de la Santé, 2013a, 2014). To compensate providers for the loss of revenues and to encourage them to provide free care, the purchase prices for services delivered to indigents were higher than for non-indigent patients. However, to avoid a moral hazard, a maximum of 10% of the services mentioned above could be delivered to indigents. With a total of 102,609 indigents identified, the financial cost and the economic cost per ultra-poor person were respectively USD 5.73 and USD 11.83 (Beaugé et al., 2018).

The overall PBF intervention ended in 2018 with the end of the World Bank program that provided funding (The World Bank, 2018). However, a successor program called "strategic purchasing" is underway. Deliberations are taking place to integrate PBF with user-fee exemptions measures.

Figure 2. Logic model of the PBF intervention combined with health equity measures in Burkina Faso



4.4 Sampling of facilities

The cases for this doctoral study on unintended consequences were eight CSPSs and one CMA located in two districts in Burkina Faso. Our assessment focused on a subset of cases that were selected for the overarching process evaluation (Ridde et al., 2014). The selection of cases for the process evaluation followed a multi-stage screening procedure that relied on mixed and participatory methods described below (Zombré et al., 2016).

First, the research team developed selection criteria that would stratify facilities into categories reflecting the diversity of contexts (i.e., districts, types of facility, intervention arms). During this stage, we selected three of the 12 districts involved in PBF in Burkina Faso. We selected these three districts because they are located in diverse regions and they represent the normal context of the healthcare system. Also, within each district, we decided to select the regional hospital and six healthcare centres representing the various intervention arms available (PBF 1, 2, 3 or 4).

Second, we used an *etic* approach (i.e., based on the perspective of the scientific observer) to classify facilities according to their initial level of performance. More specifically, each facility's initial performance was assessed by examining the evolution of key activity indicators for maternal and child health (i.e., assisted deliveries, consultations for children under the age of five, prenatal consultations, and children fully vaccinated) over the two years that preceded the intervention. For each facility, a composite score was created to estimate the facilities' initial levels of performance. Within each intervention arm in a given district, performance scores were ranked into quintiles to allow us to select facilities with contrasting levels of performance.

Third, we used an *emic* approach (i.e., from the perspective of insiders) to finalize the selection of facilities. As such, in October 2014, members of our research team visited key informants (i.e., district management teams) in each district to enlist their participation in the selection of healthcare centres that they considered to be representative of their category of performance and that they believed would offer opportunities to gain significant insight (Patton, 2015; Stake, 2005; Yin, 2009). To facilitate discussions with key informants, the research team members briefly shared the results of the quantitative analyses (i.e., Excel® files of performance scores by district/FBR arm as well as graphs on the evolution of indicators). These quantitative performance scores were contrasted with the key informants' *emic* perceptions of the facilities' initial performance. Facilities in the lowest and highest quintiles in each intervention arm were selected through discussions. This dialogue with local informants helped us to avoid selecting facilities that would be considered outliers or unrepresentative, which would have undermined the utility of research results.

For this doctoral study, we further narrowed the selection of cases due to limited time and resources for each phase of the study (i.e., pilot phase, phase 1, and phase 2). For the pilot phase, we selected two pilot cases that would enable us to test the data collection methods and gain a better understanding of the intervention within the local healthcare system. For this, we selected both a high- and a low-performing facility within one of the two districts included in the process evaluation that implemented intervention arms with user-fee exemption measures for indigents selected through a community-based process. More specifically, we selected facilities within PBF 2, as it represented a middle way between the traditional PBF model and an innovative PBF model with equity measures. For phase 1, however, we had to move the data collection to another district due to security issues for foreigners conducting prolonged fieldwork in the region. Instead, we opted for the other district selected in the process evaluation with user-fee exemption measures for indigents. Within that district, we selected a total of four high- and low-performing facilities, which were assigned to two intervention arms (PBF 1 and PBF 3). They enabled us to examine the unintended consequences of the traditional PBF, which is widely implemented in LMICs (PBF 1) and to compare them to the unintended consequences of a more innovative PBF model with user-fee exemption measures for indigents (PBF 3). For this phase, we opted for PBF 3, under the rationale that higher financial incentives for services provided to indigents might increase the likelihood that unintended consequences would emerge and be more visible to researchers.

For phase 2, on community verifications, we were able to expand the data collection to the seven cases that had been included for the process evaluation within this district. This was possible due to the manageable quantity of data derived from this more focused topic. Ultimately, we completely excluded the fourth intervention arm (PBF 4) from this analysis because it represented a more complex intervention with community-based health insurance, and additional resources would have been required to analyze this different context. Figure 3 illustrates the cases selected for each phase of the study. Table IV summarizes the characteristics of these nine cases.

Figure 3. Cases selected for each phase of the study

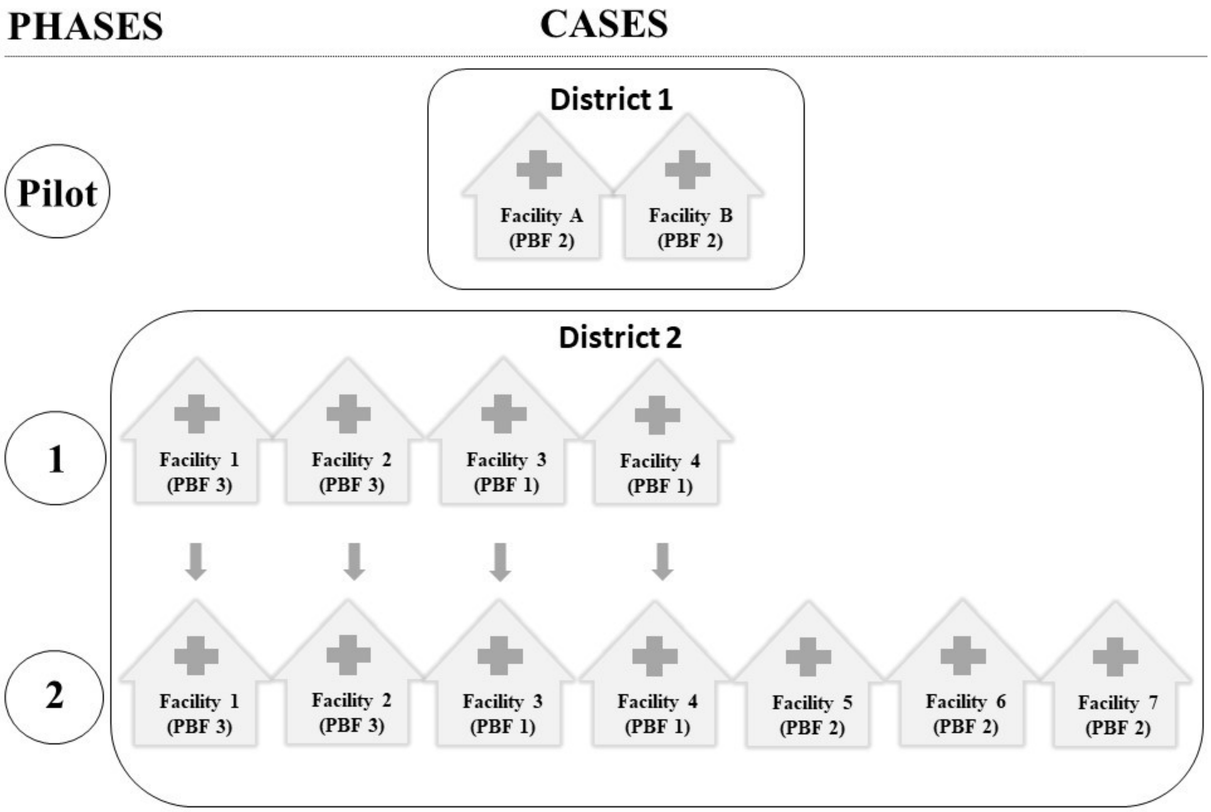


Table IV. Description of nine cases selected in the study

	Pilot cases		Primary cases				Secondary cases		
Descriptors	**Facility A**	**Facility B**	**Facility 1**	**Facility 2**	**Facility 3**	**Facility 4**	**Facility 5**	**Facility 6**	**Facility 7**
Intervention arm	PBF 2	PBF 2	PBF 3	PBF 3	PBF 1	PBF 1	PBF 2	PBF 2	FBR 2
Facility type	CSPS	CSPS	CSPS	CSPS	CSPS	CSPS	CSPS	CSPS	CMA
Zone	Rural	Rural	Rural	Rural	Rural	Rural	Rural	Rural	Urban
Initial performance	Low	High	Low	High	Low	High	Low	High	-
PBF payments owed or transferred between Trimester 1 2014 & Trimester 1 2016*	10 117 781 F CFA	12 909 022 F CFA	9 389 071 F CFA	6 450 040 F CFA	12 610 680 F CFA	6 412 805 F CFA	19 261 460 F CFA	10 380 719 F CFA	72 652 283 F CFA
Average scores for quality verifications between T1 2014 & T1 2016*	74%	84%	71%	70%	86%	67%	75%	80%	64%
Staff members	2 nurses, 1 auxiliary midwife, 1 drug manager, 2 janitors, 1 guard, 2 trainees (temporary)	1 nurse, 2 auxiliary midwives, 1 itinerary health worker (IHW), 1 drug manager, 2 janitors, 1 guard, 5 trainees	1 nurse, 1 auxiliary midwife, 2 IHWs, 1 drug manager, 1 janitor, 1 guard, 1 IHW volunteer	1 nurse, 1 auxiliary midwife, 1 IHW, 1 drug manager, 1 janitor, 1 guard	2 nurses, 1 midwife, 2 IHWs, 1 auxiliary midwife, 1 drug manager, 1 guard, 1 janitor, 3 trainees	1 nurse, 1 auxiliary midwife, 1 IHW, 4 trainees	2 nurses, 2 auxiliary midwives 2 IHWs, 1 drug manager, 1 guard, 1 janitor	2 nurses, 1 auxiliary midwife, 1 IHW 1 drug manager, 1 guard, 1 janitor	> 74 staff members
# of villages & population in catchment area	8 villages, 8 900 people	10 villages, 7 700 ppl.	5 villages, 8 000 ppl.	8 villages, 3 600 ppl.	22 villages, 11 000 ppl.	6 villages, 3 700 ppl.	8 villages, 14 100 ppl.	5 villages, 4 500 ppl.	23 villages, 53 700 ppl.
Ethnic majority	Mosse	Mosse	Dagara	Lobi	Lobi	Birifor, Djan	Dagara	Djan	Djan, Lobi, Dagara, Dioula, Birifor

*** Data available online: http://www.fbrburkina.org/data**

4.5 Sampling for interviews

Participants included a wide range of stakeholders, such as healthcare workers, service users, COGES representatives, community-based health workers (CHW), members of the district management team, and representatives of the national level. For the analysis focusing on equity measures, we included indigents and members of the indigent selection committee. For the analysis focusing on community verifications, we included community verifiers and service users selected for the community verifications. Overall, participants were purposefully selected during observation sessions based on their ability to provide relevant information and their accessibility. Then, following the snowball approach, some key informants referred us to other potential participants who could shed light on the intervention. These approaches enabled us to apply the diversification principle to select participants with a variety of intrinsic characteristics, such as different occupations, socio-economic status, and gender (Patton, 2015).

4.6 Dimensions studied

An exploratory qualitative approach was useful to understand how factors such as the social system, the characteristics of its members, and the nature and use of the innovation interacted to influence unintended consequences over time. This section summarizes the different dimensions that were analyzed to answer the research questions.

Social system: We closely examined the relationship between the social system and the unintended consequences. The social system was defined as a set of interrelated units that are engaged in joint problem-solving to accomplish a common goal. These units were individuals, informal groups, or organizations (e.g., healthcare centres). We focused on the social system's formal structure, which provides regularity and stability to behaviour, as well as on the informal structure, for example, by tracing who interacted with whom and under what circumstances. We considered informal practices that emerged over time. We also examined local norms, as they are important characteristics of the social system. They define a range of tolerable behaviour and serve as a guide or standard for the behaviour of members of a social system (Rogers, 2003). Belief systems and past experiences were also important parts of the social system that could influence the emergence of unintended consequences.

Characteristics of the members: Individuals who use an innovation shape it by giving it meaning. Thus, it was important to examine the perceptions and subjective appraisals of the members of the social system. We also examined how the characteristics of these members were related to unintended consequences. For example, attitudes towards change, the socioeconomic status, the social status in the hierarchy were important factors to examine.

Nature of the innovation: The complex intervention implemented in Burkina Faso is an innovation cluster composed of PBF and health equity measures. We examined the interactions between these different components to analyze the unintended consequences that emerged. We also examined how the intervention's hardware (e.g., tools or physical objects) and software (e.g., instructions) led to unintended consequences. We focused on the characteristics of innovations, such as their relative advantage, complexity, and compatibility with local values, beliefs, and past experiences.

Use of the innovation: Redefining and restructuring occurs when the innovation is modified and re-invented to fit the organization's needs and structure (Rogers, 2003). We examined how this process leads to unintended consequences.

Unintended consequences: Consequences were considered to be: 1) desirable or undesirable, 2) direct or indirect, and 3) anticipated or unanticipated. As explained in Chapter 3, we excluded from our analyses consequences that were simultaneously desirable and anticipated because they were considered to be intended by program planners. Instead, we focused on the unintended consequences that were undesirable and/or unanticipated.

4.7 Qualitative data

4.7.1 Qualitative data collection method

We collected qualitative data during three sequential phases, with each informing the methods for the next. In the pilot phase (April 2015), a research team member (AMTT) conducted fieldwork in two facilities in the same district (facilities A and B). Serving as pilot case studies, each facility was visited for three days to validate the feasibility of the methods and gain a better understanding of the healthcare system. In phase 1 (January-April 2016), AMTT conducted three months of fieldwork, examining four facilities in another district with longer field visits and more participants, for greater depth (facilities 1–4). Each facility was visited for two weeks. AMTT also attended a 6-day annual national PBF review meeting in 2015. In phase 2 (May 2016), another research team member (IAG) conducted 20 days of fieldwork in seven cases (facilities 1–7) to deepen the assessment of community verifications. For each phase, we collected data through observation, interviews, and informal discussions, as described below.

Observation: The researchers' immersion in the milieu provided a better understanding of the context, helped create a sense of trust with stakeholders, and established rapport so that participants felt more comfortable disclosing information (Creswell & Miller, 2000). Observation sites included facilities (e.g., maternity wards, dispensaries, drug depots), residences of providers, villages for outreach activities, and social settings where people like to interact (e.g., *maquis*, markets, funerals, initiations). In the observations, we consulted and collected a wide range of intervention documents (e.g., guidelines, forms and checklists for quality and quantity verifications, medical registers, performance reports) to fuel our analyses. Field notes on observations and reflexive thoughts were systematically recorded in research diaries within a few hours maximum. In total, 266 entries were recorded in the research diaries.

Interviews: We constructed semi-structured interview guides that drew upon previous questionnaires used for innovation diffusion research (Spicer, 1952; Warford, 2005) but were tailored to this study's needs and adapted to the different types of stakeholders. The guides assessed how social system factors, characteristics of the members, and the nature and use of the intervention interacted to produce consequences over time. Interview guides used in phase 1 and phase 2 are presented in the supplementary files of the articles in the results section. The flexible nature of semi-structured interviews allowed pertinent data to emerge according to the participants' knowledge, perspective, and experiences. Interviews lasted between 30 minutes and 2 hours depending on the type of participant. Overall, we conducted 104 semi-structured interviews: 11 in the pilot phase, 76 in phase 1, and 17 in phase 2. Local community members not directly involved with the intervention served as interpreters during 15 interviews. Interviews were recorded and transcribed.

Informal discussions: Research team members resided within the vicinity of healthcare centres, which created many opportunities for informal discussions with local stakeholders regarding the implementation and consequences of the intervention over time. Field notes on informal discussions were systematically recorded in research diaries within a few hours to avoid memory bias.

Applying the principle of saturation, we stopped collecting data when interviews and observations no longer provided information that was sufficiently different to justify continuing. **Table V** summarizes the qualitative data collected during observation sessions and interviews for the overall study.

Table V. Summary of data collected during observation sessions and interviews for overall study

Data	Quantity
Observation & informal discussions	
Entries reported in field notes	266
Interviews	
Facility level	
Healthcare providers	21
Other support staff (drug manager, janitor, security guard)	15
Volunteers & trainees	7
Community leaders (e.g., COGES, community health worker, counselor)	26
Service users	18
District level	
Administrative staff (e.g. manager, accountant)	4
CVA members	4
Members of local associations conducting community verifications	7
National level	
Representative from the *Programme d'appui au développement en santé* (PADS – program to support health development)	1
Representative from the *Service technique – financement base sur les résultats* (ST-FBR – results-based financing - technical service)	1
Total interviews	**104**

4.7.2 Qualitative data analyses

The primary unit of analysis was each healthcare facility. We conducted thematic analysis on the data. Attention was also given to subunits such as healthcare providers and community members (Yin, 2009). Data were triangulated by comparing different information sources (Olivier de Sardan & Tidjani Alou, 2015). Using a hybrid deductive-inductive approach, we assigned data to predefined themes and derived new themes while reading through the data (Pluye & Hong, 2014). We used QDA Miner 4 to code and retrieve text segments. We integrated the results from all data collection phases and used a cross-case synthesis to draw general conclusions. Following a replication logic, we considered results arising independently from more than one case to be more powerful than those from a single case, and gave the former more importance in the results section (Yin, 2009).

We conducted member checks to establish the credibility of findings, confirm the researchers' interpretation of data, triangulate results, and validate conclusions (Creswell & Miller, 2000; Fereday & Muir-Cochrane, 2006). This consisted of presenting data and interpretations to local members intimately involved in the intervention so that they could provide their feedback. One extensive member check was organized during a face-to-face meeting in Burkina Faso in April 2016. Other smaller member checks were conducted with local stakeholders via email exchanges between May 2016 and March 2019. Their reactions and comments during the presentation of main findings were useful for triangulating results (Fereday & Muir-Cochrane, 2006). In turn, we incorporated the members' comments into the final narrative (Creswell & Miller, 2000).

Following Yin (2009), the empirical findings of this case study were compared and generalized to broader theoretical propositions. This mode of generalization is called analytic generalization.[10] As explained by Gilson, "*analytic generalisation entails the development of general conclusions that, although derived from a limited number of particular experiences, provide theoretical insights that can be put forward for consideration, and testing, in other, similar situations. This includes […] theory that offers ideas about the causal mechanisms likely to underpin interventions...*" (Gilson et al., 2011, p. 3). We used the framework based on Rogers' theory of diffusion of innovation as a template with which to compare the empirical results. We considered that, if two or more

---

[10] Yin (2009) explains that analytic generalization contrasts with another way of generalizing results known as statistical generalization.

cases were shown to support the theory, replication could be claimed (Yin, 2009). Moreover, the multiple cases decreased the risk that the results were unique to a case, thereby increasing the study's external validity (Contandriopoulous et al., 2005).

4.8 Quantitative data

4.8.1 Quantitative data collection method

We used secondary data on healthcare services delivery that are publicly available on the Ministry of Health's PBF portal (www.fbrburkina.org). These longitudinal data are collected monthly in each healthcare centre for PBF verifications. Healthcare workers report the quantity of healthcare services delivered to patients, based on the medical registers. Then PBF officers verify the reported data by manually recounting the quantity of services. They enter the data into an electronic platform.

For the analyses on the unintended consequences of PBF (Article 2), we used the data collected between January 2014 and December 2016 on the number of integrated household visits as well as the number of people who underwent voluntary HIV screening reported per trimester. Our sample focused on the four primary cases (facilities 1–4) included in the qualitative phase 1. We also used quantitative data presented at the annual national PBF review meeting for 2015 on the evolution of the proportion of children under 5 years of age treated using the integrated management of childhood illnesses (IMCI) strategy, as well as on the proportion of births conducted using a partograph in PBF districts compared to control districts.

For the analyses on the unintended consequences of combining PBF with health equity measures (Article Y), we used the data collected between October 2015 and September 2016, that is, before and after fee-exemption cards were distributed to indigents starting in November 2015. The main sample for the quantitative component consisted of the two facilities with equity measures (PBF 3) included in the qualitative phase 1. To assess the transferability of the findings across the study district, however, we examined all seven facilities within the district that were assigned to the same intervention arm as the two selected for inclusion in the qualitative component (PBF 3) and for which data were available. To assess the transferability of findings more widely, we also examined all 196 facilities in the intervention districts that belonged to intervention arms with similar measures for indigents (PBF 2 and PBF 3) and for which data were available.

4.8.2 Quantitative data analyses

We used descriptive statistics to examine how the quantity of services provided to different types of patients evolved over time. We used Excel® to create graphs and conducted visual analyses to highlight patterns that emerged over time (Shadish et al., 2002). The quantitative findings were interpreted in light of the qualitative findings. This process allowed us to triangulate and assess the transferability of some of the qualitative findings.

4.9 Ethical considerations

Due to the international nature of the study, the protocol was approved by the research ethics committees in Burkina Faso (deliberation N_ 2015-12-07) and at the University of Montreal Hospital Research Centre (CE 13.358) (see Appendices B and C). It was determined that the study's benefits for the advancement of knowledge outweighed the minimal risks and inconveniences posed to participants. The study respected ethical principles put forward in both the *Tri-council Policy Statement: Ethical Conduct for Research Involving Humans* (2014) and the University of Montreal's *Guide d'information sur le consentement libre, éclairé et continu* (2018).

We obtained informed consent by explaining the objectives of this study to participants and answering any questions they had. We also informed subjects that participation was voluntary and that they could withdraw at any time. Financial compensations were not provided in exchange for their involvement to avoid creating excessive incentives or expectations regarding future data collections. Participants provided written or verbal consent, depending on their level of literacy. The consent form presented to participants is displayed in Appendix D. Throughout the study, only two people in facility 2 (an assistant midwife and a trainee) explicitly or implicitly refused to be formally interviewed. They did not provide a justification for this decision.

The potential loss of data confidentiality was probably the main risk for participants. Considerable measures were taken to protect the confidentiality of participants, including:

1) conducting interviews in hidden locations when possible; 2) keeping participants' names separate from raw data; 3) keeping raw data (e.g., interview recordings, field notes) out of reach and in locked locations; 4) keeping electronic data (e.g., transcriptions) on a password protected computer; 5) modifying and blurring pictures used to illustrate results; and 6) concealing the names of the health districts to protect the confidentiality of communities. During our fieldwork, however, the difficulty of protecting the confidentiality of participants while disseminating results locally became apparent, inspiring the publication of an article presented in the discussion section.

Chapter 5 - Results

5.1 Article 2: An exploration of the unintended consequences of performance-based financing in six primary healthcare facilities in Burkina Faso

Anne-Marie Turcotte-Tremblay, Idriss Ali Gali-Gali and Valéry Ridde

Manuscript submitted to the *International Journal of Health Policy and Management*

See Appendix E for reviewer's comments

Final version published online in open access in 2020:

https://www.ijhpm.com/article_3836.html

Title: An exploration of the unintended consequences of performance-based financing in six primary healthcare facilities in Burkina Faso

Anne-Marie Turcotte-Tremblay[a,b*], Idriss Ali Gali-Gali[c,d], Valéry Ridde[e]

a. University of Montreal Public Health Research Institute, Montreal, Canada

b. School of Public Health, University of Montreal, Montreal, Canada

c. Association Action Gouvernance Intégration Renforcement (AGIR), Ouagadougou, Burkina Faso

d. Association Zama Forum pour la Diffusion des Connaissances et des Expériences Novatrices en Afrique (Zama Forum / ADCE-Afrique), Bobo-Dioulasso, Burkina Faso

e. IRD (French Institute for Research on Sustainable Development), CEPED (IRD-Université Paris Descartes), Universités Paris Sorbonne Cités, ERL INSERM SAGESUD, Paris, France.

* Corresponding author: Anne-Marie Turcotte-Tremblay, University of Montreal Public Health Research Institute, 7101 Avenue du Parc, Montreal, QC H3N 1X9, Canada

Contributions of authors:

Anne-Marie Turcotte-Tremblay conceived the original idea, developed the protocol, collected and analyzed data, and wrote the first draft of the manuscript.

Idriss Ali Gali-Gali helped interpret results and revised the manuscript.

Valéry Ridde reviewed the protocol, helped interpret results, and critically revised the manuscript.

Abstract:

Background: Performance-based financing (PBF) is promoted to improve the quality and quantity of healthcare services in low-income countries. Despite the complexity of the intervention, little attention has been given to studying its unintended consequences. Our objective is to increase the scientific knowledge on the unintended consequences of PBF in Burkina Faso.

Methods: Using the diffusion of innovations theory, we conducted a multiple case study. The cases were six healthcare facilities in two districts. We collected data through 101 interviews, discussions, observations, and documents. We conducted thematic analysis using a hybrid deductive-inductive approach. We used secondary data to illustrate the evolution of reported services.

Results: A desirable unintended consequence of performance-based financing was that some facilities limited the sale of non-prescribed medication to encourage patients to consult. Undesirable unintended consequences included fixation on measures rather than on underlying objectives, the pursuit of narrow and less relevant performance indicators, gaming, falsification of medical registers, and teaching trainees improper practices. Dissatisfaction emerged in relation to the distribution of premiums, the lateness of subsidies and bonuses, and the non-attribution of quality points for services delivered by certain staff considered “unqualified” in guidelines. Results revealed suboptimal planning and a perception of the intervention as “budgetivorous”, as well as tensions related to the principle of managerial autonomy.

Conclusions: Performance-based financing led to numerous unintended consequences that could undermine the intervention’s overall effectiveness. The findings contribute to a more comprehensive picture of the consequences of implementing performance-based financing. Policy makers can use the results of this study to devise effective strategies that can minimize the undesirable unintended consequences.

Keywords: Performance-Based Financing, Unintended Consequences, Burkina Faso, Multiple Case Study

Key Messages:

1. Implications for policy makers

- Policy makers should carefully consider the breadth and scope of unintended consequences before pursuing or scaling up PBF interventions.
- Reducing undesirable consequences of PBF may require some adjustments to transfer subsidies on time, ensure that their distribution is perceived as equitable amongst actors involved, improve the staff members' internalization of quality standards, adapt performance indicators to the local context or seek truly independent PBF auditors.
- Policy makers should be wary of incentives and performance pressure that can encourage the deliberate and systematic falsification of medical registers.
- Program planners should increasingly monitor desirable and undesirable unintended consequences of PBF to gain a more comprehensive understanding of its impact on health systems.

2. Implications for public

There is an urgent need to improve healthcare systems in low-income countries. Governments and funding agencies are increasingly implementing an intervention called performance-based financing (PBF) to improve the quantity and quality of care. Healthcare facilities receive a unit fee for each service provided and bonuses based on the quality of care. We conducted a study in Burkina Faso to examine whether PBF led to unintended consequences. We found that PBF led to numerous unintended consequences such as gaming and fixation on performance indicators rather than on underlying objectives. Providers spend considerable amounts of time falsifying registers to improve their performance scores. Dissatisfaction grew amongst staff due to the lateness of bonuses and the distribution modalities of premiums. Many people viewed the intervention as too costly. These findings will be useful to develop strategies that help prevent or minimize unintended consequences in order to successfully improve the healthcare systems' performance.

Background

Improving healthcare systems performance is key to achieving universal health coverage by 2030. The World Health Organization is encouraging low- and middle-income countries (LMICs) to move from passive to strategic purchasing of health services[1]. Performance-based financing (PBF) is one means of introducing elements of strategic purchasing[2]. With PBF, facilities can receive a unit fee for each targeted service provided, as well as bonuses conditional on quality of care.

While PBF is rapidly expanding in LMICs, many public health actors hypothesize it can have important unintended consequences that influence its overall effectiveness. These unintended consequences are defined as "changes for which there is a lack of purposeful action or causation that occur to a social system as a result of an innovation"[3]. These changes can be desirable or undesirable, as well as anticipated or unanticipated, depending on the stakeholders' perspectives. They can affect various actors such as patients, providers and managers.

Although it often sparks debates, research on the unintended consequences of PBF remains scarce. In high-income countries (HICs), a synthesis of reviews found some evidence that PBF was associated with risk selection, spillover effects, gaming behavior, and changes in the providers' intrinsic motivation[4]. Yet, after examining 12 pay for performance programs in HICs, Cashin et al.[5] concluded that unintended consequences were never carefully assessed.

In LMICs, at least two literature reviews have demonstrated the lack of evidence documenting the unintended consequences of PBF[6,7]. However, some empirical evidence is beginning to emerge. For example, Basinga et al.[8] suggested PBF in Rwanda had the greatest impact on services requiring less effort. Other studies in the country found that information was regularly distorted, that providers used gaming strategies and that consultations were rushed to reach targets[9,10]. A study in Cameroon found that PBF raised concerns regarding drug quality and inequities between facilities[11]. In Uganda, Benin, and Burkina Faso, studies highlighted that audits generated overwhelming workloads[12-14]. A study on PBF community verifications revealed falsification of data, loss of patient confidentiality, and fears among patients[14]. In Malawi, PBF had both positive and negative effects on the health workers' basic psychological needs for autonomy, competence, and relatedness, which are central to intrinsic motivation[15]. More recently, in the Democratic Republic of Congo, Maini et al.,[16] showed that the removal of PBF negatively affected many dimensions of motivation for staff members. However, an important gap in the literature remains because none of these studies specifically aimed to explore all the potential unintended consequences that could emerge. Thus, they did not use frameworks and methods that enabled them to fully capture these phenomena.

There are several reasons for studying the unintended consequences of PBF in LMICs. The likelihood that PBF triggers unintended consequences going well beyond the objectives of the intervention is high. There is a lot of uncertainty about how new practices will function in complex systems such as healthcare systems[17]. Although they may be less discernible, unintended consequences may be far-reaching and as equally important as intended consequences. Stakeholders must have a comprehensive understanding of both the intended and unintended consequences in order to judge the overall value of an intervention.

This paper is intended to fill a knowledge gap on the neglected topic of unintended consequences of PBF in LMICs. A pilot PBF test implemented in Burkina Faso to improve the healthcare system’s performance provided a unique opportunity to develop scientific knowledge on the unintended consequences of PBF in a real-life setting. More specifically, we posed the following research question: What are the unintended consequences of PBF, and their contributing factors, in primary healthcare facilities in Burkina Faso?

Theoretical framework

We based our theoretical framework on Rogers’ diffusion of innovations theory[17] for several reasons. While it has proven its utility to analyze the consequences of health innovations[3,18] it constitutes an original approach to study PBF in LMICs. It is also one of the rare theories that provides a detailed typology of consequences (see below) while remaining sufficiently flexible to be applicable to any innovation. Moreover, it is comprehensive by taking into account the entire diffusion process of innovations as they course through the structure of a social system.

To summarize, the theory postulates that the implementation of innovations such as PBF does not always conform to plan[17]. Adopters (e.g. healthcare providers) often modify the innovation to suit the organization's needs and structure, just as the organization's structures are altered to fit the innovation. Change agencies (e.g. funders, ministries of health), which promote innovations when they perceive a performance gap, can offer financial incentives to hasten their adoption. According to the theory, the main purpose of incentives is to give the new practice a relative advantage over previous ones. Although incentives increase the quantity of innovation adopters, the quality of adoption may be low, thus limiting the intended consequences. Key variables influencing the diffusion process pertain to: 1) the nature of the social system; 2) its members' characteristics; 3) the nature of the innovation; and 4) the use of the innovation.

Inspired by Rogers[17], we classified consequences into three categories: 1) desirable or undesirable; 2) direct or indirect, depending on whether the changes related to processes or outcomes; and 3) anticipated or unanticipated. Inspired by Rogers[17], we classified consequences into three categories: 1) desirable or undesirable; 2) direct or indirect, depending on whether the changes related to processes or outcomes; and 3) anticipated or unanticipated. We considered a consequence to be anticipated if it was addressed in the implementation guides or if it was in line with the "spirit of the intervention" or its "ideas" (i.e., beliefs, assumptions or perceptions)[19], according to PBF experts. We further refined Rogers' classification by considering that the following types of consequences tend to be unintended by program planners: undesirable/anticipated, undesirable/unanticipated and desirable/unanticipated. Our rationale for classifying these consequences as unintended was that program planners are not likely to purposefully target changes they consider undesirable or have not anticipated. Like Bloomrosen et al.[18], we expected that consequences that are desirable/anticipated would tend to be intended by program planners. Similar to what Jabeen[20] previously argued, program planners trying to promote a new intervention are likely to have listed and exhausted all the desirable outcomes that they foresee in their rhetoric. Nevertheless, we did consider that some desirable/anticipated consequences could be unintended if they were, for example, positive spillover effects that were foreseen but not initially targeted by program planners. This conceptualization is consistent with recent literature suggesting that unintended consequences can be either anticipated or unanticipated as well as desirable or undesirable[21-23]. **Figure 4** illustrates our framework. Its applicability has been presented elsewhere (reference blinded for review).

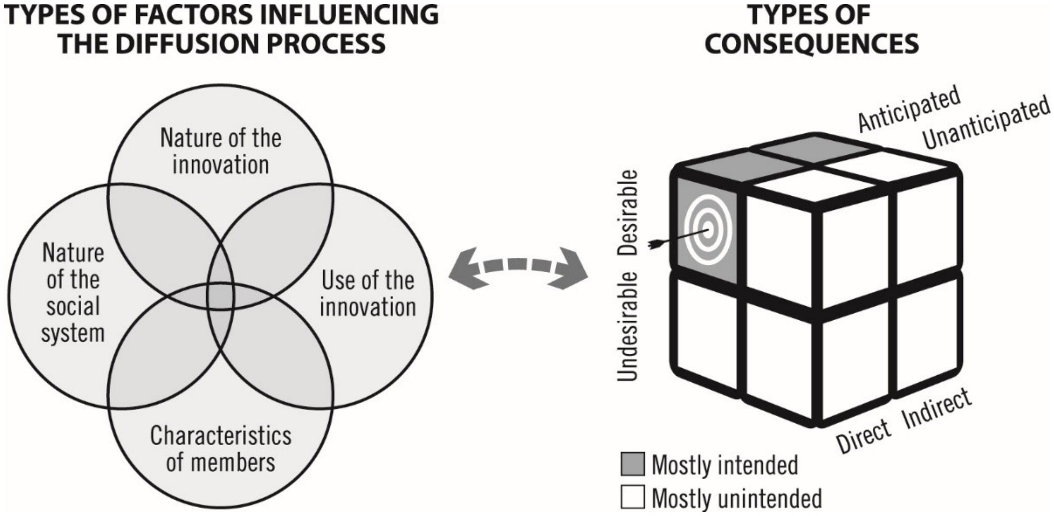


Figure 4. Theoretical framework

Methods

Study setting

The study took place in two rural districts of Burkina Faso where improving the healthcare system's performance remains a challenge. The low quality of healthcare is often characterised by the staff's inhospitality, insufficient equipment/medication and lack of training[24].

In 2011, the government, with World Bank support, conducted a pre-pilot PBF test in three districts to address generalized quality deficiencies and improve healthcare system performance[25]. In 2014, this intervention was modified and expanded to 12 more districts. Its specific objectives were to 1) increase the utilisation of healthcare services ; 2) improve the quality of healthcare services; 3) improve the efficiency of the healthcare system; 4) ensure equity in access to healthcare services ; 5) reinforce the motivation of personnel ; 5) improve community participation ; 6) reinforce the health information system ; 7) consolidate public – private partnership and 8) reinforce the governance of the healthcare system. To determine the best approach, four intervention arms were implemented combining PBF modalities with different unit fees for service and equity measures (Appendix 1). The intervention model is available online[26]. The quantity of services provided was audited monthly to establish the amount of subsidies. Every trimester, a quality audit team assessed the care to determine bonuses (Appendix 2). Community verifications were also supposed to be conducted every trimester. Subsidies and bonuses were used to pay for facility expenditures and premiums to motivate staff. Providers were required to use an index tool every month to update the facilities' revenues, plan expenditures, and determine the distribution of premiums (Appendix 3). Each trimester, providers were also required to produce a performance improvement plan to set objectives and plan activities.

Study design

We conducted a multiple case study with several embedded levels of analyses, using both qualitative and quantitative data[27]. The cases were six primary healthcare facilities, called Centres de santé et de promotion sociale (CSPS), in two districts. Case selection followed a multistage screening procedure described elsewhere (reference blinded for review). **Table VI** describes each primary healthcare facility included.

Table VI. Description of 6 cases included

	Pilot cases		**Primary cases**			
Descriptors	**Facility A**	**Facility B**	**Facility 1**	**Facility 2**	**Facility 3**	**Facility 4**
Intervention arm	PBF 2	PBF 2	PBF 3	PBF 3	PBF 1	PBF 1
Initial performance	Low	High	Low	High	Low	High
PBF payments owed or transferred between Trimester 1 2014 & Trimester 1 2016*	10 117 781 F CFA	12 909 022 F CFA	9 389 071 F CFA	6 450 040 F CFA	12 610 680 F CFA	6 412 805 F CFA
Average scores for quality verifications between Trimester 1 2014 & Trimester 1 2016*	74%	84%	71%	70%	86%	67%
Staff members	2 nurses, 1 auxiliary midwife, 1 drug manager, 2 janitors, 1 guard, 2 trainees (temporary)	1 nurse, 2 auxiliary midwives, 1 itinerary health worker (IHW), 1 drug manager, 2 janitors, 1 guard, 5 trainees	1 nurse, 1 auxiliary midwife, 2 IHWs, 1 drug manager, 1 janitor, 1 guard, 1 IHW volunteer	1 nurse, 1 auxiliary midwife, 1 IHW, 1 drug manager, 1 janitor, 1 guard	2 nurses, 1 midwife, 2 IHWs, 1 auxiliary midwife, 1 drug manager, 1 guard, 1 janitor, 3 trainees	1 nurse, 1 auxiliary midwife, 1 IHW, 4 trainees
# of villages &	8 villages	10 villages	5 villages	8 villages	22 villages	6 villages
Population in catchment area	8 900 people	7 700 people	8 000 people	3 600 people	11 000 people	3 700 people

* Data available online: http://www.fbrburkina.org/data

Abbreviations: PBF, performance-based financing

Sampling for interviews

Participants included a wide range of stakeholders, including providers, patients, and community representatives (e.g. members of the facility management committee). Participants were purposefully selected based on their ability to provide relevant information and their accessibility. In each facility, we selected all the providers, support staff and volunteers for interviews. Then, following the snowball approach, some key informants referred us to other participants who could shed light on the intervention[28]. This strategy was used to identify potential participants who were especially knowledgeable about or had a particular experience with the intervention (e.g., auditors and administrative staff at the district level). Overall, we conducted 101 semi-structured interviews.

Data collection

We adopted a broad, exploratory approach in order to capture all changes that were not initially targeted by program planners. Through observation, semi-structured interviews and informal discussions, we collected data on various dimensions of the healthcare system including service delivery, governance, human resources, medication, health information system and financial management.

Data were collected during three sequential phases, with each informing methods for the next. In the pilot phase (April 2015), the first author conducted one week of fieldwork in two facilities in the same district (facilities A and B). These served as pilot case studies to validate the feasibility of the methods. In phase 1 (January–April 2016), the first author conducted three months of fieldwork, examining four facilities in another district with longer field visits and more participants, for greater depth (facilities 1–4). Each facility was visited for two weeks. The first week primarily served to conduct observation within the facilities and the second week served to conduct semi-structured interviews with participants. The first author lived in the facilities which enabled her to conduct observation as well as informal discussions around the clock. The first author also attended a 6-day annual national PBF review meeting for 2015. In phase 2 (May 2016), the second author conducted 20 days of fieldwork in those facilities and neighbouring ones, to deepen the assessment of community verifications.

Of the 101 semi-structured interviews conducted, 11 were in the pilot phase; 76 in phase 1; and 14 in phase 2. Our interview guides (see Appendix 4) built upon previous questionnaires used for research on diffusion of innovations[29,30] but were tailored to our objectives and participants. They enabled us to assess how factors related to the social system, characteristics of the members, and the nature and use of the intervention interacted to produce consequences over time. Interviews were recorded and transcribed.

In total, 258 observation sessions were recorded in research diaries. Observations sites included health facilities and villages. During observation, we collected a wide range of intervention documents (e.g., quantity and quality verification reports, index tools) to fuel our analyses.

We also used publicly available secondary quantitative data on service delivery (www.fbrburkina.org). These longitudinal data were collected monthly in each facility for PBF audits.

We used several strategies to increase the trustworthiness of findings including: 1) prolonged engagement on the field, 2) peer debriefing with members of the research team, 3) collection of audio recordings and photographs that can be used to test findings, 4) triangulation between sources of information and methods and 5) member checks with stakeholders to confirm results[31].

Data analyses

The primary unit of analysis was each healthcare facility. We combined deductive and inductive thematic analysis[32]. We began by developing a template of themes based on our theoretical framework. Then, we carefully read the transcripts and field notes to assign the raw data to the predefined themes. At the same time, we derived new themes that were not included in the initial template but that emerged from the data and were judged relevant to our research topic. Data were triangulated by comparing different information sources[33].QDA Miner, a qualitative data analysis software, was used to code and retrieve text segments. We integrated the results from all data collection phases and used a cross-case synthesis to draw general conclusions. Following a replication logic for multiple case studies, we considered results arising independently from more than one case to be more powerful than those from a single case, and gave the former more importance in the results[27]. To avoid cherry-picking results within the rich material, we only present unintended consequences that emerged in multiple healthcare facilities. We organized a member check in Burkina Faso to confirm the researchers' data interpretation, triangulate results, and validate conclusions[34]. Further member checks were conducted subsequently on specific elements.

Ethics

The protocol was approved by the research ethics committees in Burkina Faso (deliberation N_ 2015-12-07) and at the University of Montreal Hospital Research Center (CE 13.358). Participants provided consent to participate, as required by the ethics committees.

Results

We conducted semi-structured interviews with 101 participants: Thirty-one were women and 70 were men. This discrepancy is due to gender inequalities in the workplace. **Table VII** shows that over a third of interviewees (n=36) worked in facilities either as providers or support staff. Almost a quarter of interviewees (n=25) were community leaders involved in the healthcare system.

Table VII. Demographic characteristics of interview participants

Characteristics	N=101 n
Gender	
Females	31
Males	70
Place of residence	
Rural	84
Semi-urban	15
Urban	2
Education	
Primary or less	29
Secondary	29
Post-secondary	43
Age group	
18-30	16
31-50	80
51-70	5
Status	
Facility level	
Providers	21
Support staff (drug manager, janitor, guard)	15
Volunteers & trainees	7
Community leaders (e.g., COGES, CHWs, counselors)	25
Patients	16
District level	
Administration (e.g. manager, accountant)	4
Contracting and verification agents (auditors)	4
Members of associations conducting community verifications	7
National level	
Representative from the program to support health development	1
Representative from the results-based financing - technical service)	1

Abbreviations: COGES, facilities' management committees; CHWs, community health workers.

PBF led to important unintended consequences, classified according to our model in **Table VIII**, and illustrated hereafter with verbatim citations and examples from our data

Table VIII. Classification of unintended consequences

	Anticipated		Unanticipated	
	Direct (Process)	Indirect (Outcome)	Direct (Process)	Indirect (Outcome)
Desirable	∅ No unintended consequence detected in this category	∅ No unintended consequence detected in this category	• Limits on medication sales without consultations	∅ No unintended consequence detected in this category
Undesirable	• Gaming • Fixation on indicators and subsidies • Falsification of medical registers and documents • Complacency, collusion and complicity	∅ No unintended consequence detected in this category	• Teaching trainees improper practices • Overwhelming paperwork • Pursuit of narrow performance indicators • Manipulation of index tools • Tensions and conflicts related to index tools • Staff's dissatisfaction and demotivation due to payment delays • Suboptimal planning due to payment delays • Financial issues • Frustrations for providers not eligible for quality points • Tensions between managerial autonomy and top-down control • Activities delayed and reduced due to gradual withdrawal of other funding • A "budgetivorous" intervention	∅ No unintended consequence detected in this category

Appendix 5 specifies how the anticipated consequences were addressed in the intervention guides.
Note that intended consequences are not included in this analysis.
According to our framework, the dark and light grey segments indicate "mostly intended" and "mostly unintended" consequences, respectively.

Desirable and unanticipated

Limits on medication sales without consultations: Three facilities adopted medication-related strategies to increase the number of consultations recorded in registers. In two facilities, staff refused to sell medications to people who did not first consult a provider. In another, the head nurse doubled the cost of medication for people who did not consult. These were perceived as desirable changes that would reduce self-medication. However, a drug manager reported that a small number of patients left without consulting because they could not afford additional costs.

> "*Before PBF, many people came to buy drugs but few went for consultations so the head nurse requested that everyone gets a consultation. That way, we can record them in the registers which increases the quantity score when the PBF verifier comes.*" (Facility3_ observation)

Undesirable and Anticipated

Gaming: Providers adopted gaming strategies, defined as deliberate manipulation of behaviour to secure strategic advantage. One common strategy involved staging facilities when PBF audits were announced. Medication managers in two facilities reported keeping medication boxes on the floor and placing them on shelves just before PBF auditors arrived to get quality points. Janitors reported working more when informed that auditors were coming. Another example of staging were the extra lab coats with identity badges that providers made to meet PBF criteria. Although providers often received high scores for their attire, our observations showed they usually did not wear these coats in their daily practice due to heat. These gaming strategies, adopted in multiple facilities, were instrumental in obtaining PBF points and bonuses.

> "*My coat is heavy when it's hot. The day they [PBF auditors] come, though, I wear it so as not to lose points.*" (Facility1_drugmanager22_interview)

Fixation on indicators and subsidies: PBF sometimes encouraged a narrow emphasis on indicators rather than underlying objectives. For example, some facilities installed curtains to meet PBF confidentiality criteria. Providers often received excellent scores for visual privacy. Yet our observations showed patient confidentiality was regularly compromised, with multiple patients examined simultaneously. Moreover, untrained individuals (e.g. guards, friends) freely entered consultation rooms, breaching confidentiality.

Staff members appeared fixated more on paperwork than on care provision. In facility 3, for example, a nurse falsified the register for integrated management of childhood illnesses (IMCI) while unqualified staff treated a child. In facility 4, the medication manager falsified records before a PBF audit while providers sold medication directly to patients (a prohibited practice).

Another example of fixation was that providers filled out many medical registers and new PBF management documents retrospectively (sometimes weeks later) with arbitrary or approximate information to satisfy PBF criteria. In all facilities, growth curves were systematically filled out retrospectively, limiting their utility in clinical practice. Incomplete registers automatically received a score of zero during PBF audits, so providers often invented information to fill in blanks. While some retrospective filling of registers had occurred before PBF, it was now done more systematically to avoid leaving any blanks. Providers openly admitted doing more retrospective filling with PBF so they would not be shamed or outperformed by other facilities in the reporting of scores.

> "*The manager, at the end of each month, tells me he has to update his papers to be compliant so that the other CSPSs don't outdo him…. he explains to me that they received such-and-such a resource.*" (Facility3_COGES79_interview)

Providers also displayed fixation on subsidies. For example, providers in all facilities were fixated on the number of paid integrated household visits (IHV), perceived as a "quota" not to be exceeded. In facility 3, the head nurse expressed discontent when IHVs were disqualified during a PBF audit because he "lost money", but displayed no concern regarding their poor quality, the reason for disqualification. Fixation on subsidies also motivated providers' threats to stop certain activities if PBF payment delays continued.

> "*Before PBF, people just worked, no one complained; now, with PBF, all people talk about is subsidies, subsidies, and it's become a kind of obsession that's a constant hassle.*" (National_manager105_interview)

Falsification of medical registers and documents: Providers across facilities deliberately manipulated medical registers and documents, such that the reported quantity and quality of care differed from what was actually delivered (Appendix 6). Providers routinely modified documents ahead of audits to meet PBF criteria. This falsification was time-consuming and conducted openly. We were able to infer causal relationships between PBF and falsification of registers by combining complementary evidence: 1) providers explicitly referred to PBF while falsifying registers; 2) some routinely falsified registers were created specifically for PBF; 3) some of the falsification was conducted in preparation for PBF quantity and quality audits; and 4) PBF audit reports showed providers were initially criticized for not filling out registers, which were then falsified during later stages. Ultimately, falsification of registers and documentation was instrumentalized to obtain higher scores and subsidies.

Participants reported numerous factors that explained this practice, such as pressure to perform, competition between facilities, implementation challenges (e.g. shortage of qualified staff, time required to complete registers), strict PBF criteria unadapted to local realities (e.g. lengthy forms with no leeway for omissions), and desire for premiums. Moreover, some registers had not been part of the providers' daily practice before PBF. Providers sometimes dismissed the registers as "papers" (i.e., externally imposed bureaucracy) and explained that they did not, in fact, subscribe to their importance. They also reported that some registers did not serve their needs.

Audits did not always detect falsification. Providers entered false consultations directly into medical registers, then manually counted the numbers of monthly services (real and false). They declared these numbers during audits. PBF auditors checked these numbers by manually recounting the services reported in the same medical registers. Because the original source of information (i.e., registers) had been tampered with, auditors often could not distinguish between real and false consultations. Occasional differences detected between numbers declared by providers and auditors' validated numbers usually reflected calculation errors related to manual counting rather than falsification attempts.

Some participants at the district and national levels reported being aware of the falsification of registers. PBF auditors were trained to look for signs that data had been falsified, such as use of the same pen or corrector fluid. Providers adapted their falsification strategies accordingly to avoid detection. Auditors explained they were unable to determine whether patients truly received services reported in the registers because they did not observe care in real time.

Table IX presents the types of services routinely falsified, including registers for IMCIs, IHVs, and maternity ward consultations. The section below presents examples of the various types of falsifications for incentivized services.

Table IX. Examples of falsified healthcare services or information to qualify for PBF subsidies or bonuses

Services falsified	Examples of citations
Integrated Management of Childhood Illness (IMCI)	"*On seeing the drug manager filling out the IMCI register at his home, with no patients, a midwife from a neighbouring facility asked, "You are filling those out because of PBF, [aren't you]?" The drug manager mumbled a response. The midwife quickly said, "I'm not a PBF auditor!" and changed the subject*." (Facility4_observation) "*Providers systematically enter children in the register and consider those children to have been managed with the IMCI approach, even if the IMCI procedure was not used. Some districts even have 100% of consultations using the IMCI approach, which is false… there's money to be made with PBF, so there are risks of fraud*." (NationalManager_discussion)
Partographs	"*Yes, I do deliveries and when I do, I don't use the partograph. I put the time of arrival, I do the delivery, and when the birth attendant comes, she does her partograph.… Because if I do the partograph, the [PBF auditors] will invalidate it because I'm an AIS [itinerant health worker] and I'm not supposed to do deliveries.*" (Facility4_provider9_interview) "*On Sunday, March 13… the birth attendant sat on her mat with the partograph register. She filled out partographs for March 8 and March 11 from A to Z.… For the delivery on March 11, the birth attendant was not working at the facility. She was at a wedding in another city*." (Facility2_observation)
Integrated household visits	"*The drug manager and the AIS trainee sat down to finalize the household visit forms because the PBF auditor was supposed to be there at 3 p.m. for the quantity audit. They were stressed! "Give me a date!" the manager said to the AIS trainee. He randomly added about a dozen dates for visits and another dozen for follow-up appointments. Then he signed for the community health workers and even for heads of households. The other trainee arrived and asked, "What, lying again!?"… The drug manager counted the forms and realized that the strong and weak points and the analysis had not been filled out.… He asked me to fill out the forms, even though I wasn't present during the visit and am not trained as a health provider*." (Facility4_observation)

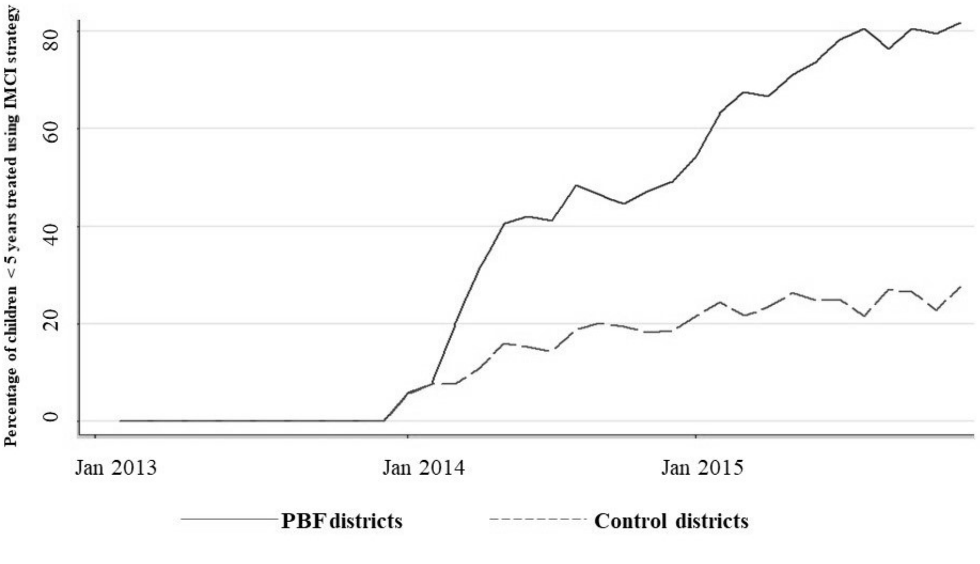
Consultations in maternity ward	"*Three providers met at the itinerant health worker's home to count the number of consultations for the PBF audit that was to take place in two days. They started at 6:13 p.m. and ended at 10:16 p.m.… "It's low, low", the head nurse said to the birth attendant on seeing the number of children between 12-23 months seen in consultation… The itinerant health worker said, "We just have to add in the register for those who didn't come. We'll fix it." The head nurse replied, "We'll count first and fix afterward, if need be." At one point, the head nurse added a consultation. He filled in an entire column, even though we had no patients.*" (Facility2_observation) "*The midwife filled out the postnatal consultation register using the birth register. She filled in several consultations with no patients there… She left the maternity ward saying, "PBF gets on my nerves! Just hearing the name gives me a headache!*" (Facility3_observation) "*The birth attendant was filling out the prenatal consultations register in the maternity ward. She added at least 10 consultations, even though there were no patients or pregnant women near her.*" (Facility3_observation)
Appointment dates	"*Some women are illiterate. We try to tell them the appointment is in four Thursdays, but sometimes they come a few days early. Some villages are more than 10 km away round-trip. Those women come by bicycle or even on foot if their husbands aren't there. We can't tell those women it's the wrong day, come back in 5 days, because they won't come back. But if we enter the real date, we'll be penalized by PBF because it's not one month later. So we don't write that date… There are reasons for low attendance that aren't due to the providers' motivation, such as illiterate women or the distances of villages. PBF should have more flexible criteria for that.*" (Facility3_provider_67_observation)
Providers' identity and qualification	"*The head nurse recopied all the consultations into the real register and signed as if he had delivered the services. But he was in another city…. He made corrections as he went along.*" (Case2_observation)
HIV screening	"*There were also overdeclarations...*" (NationalManager_discussion) "*We can't even figure out where the reagents the providers use come from. They're not from the healthcare system.*" (NationalManager_106_discussion)
Prescriptions	"*What we've seen is that sometimes [providers] report in registers having prescribed what the [diagnostic and treatment] Guide recommends, but really they've prescribed something else.*" (NationalManager_106_interview)
Other health data	"*The midwife came and added at least 3 prenatal consultations, even though there were no pregnant women here. One of the*

	additions was for March 25, but it was March 26. Then, she counted the total number of prenatal consultations for their monthly report to be submitted to the district management team.” (Facility3_observation)
Absences	“*the book is there, but absences are not noted.*” (Facility1_staff_23_interview)

Abbreviations: PBF, performance-based financing; IMCI, integrated management of childhood illness

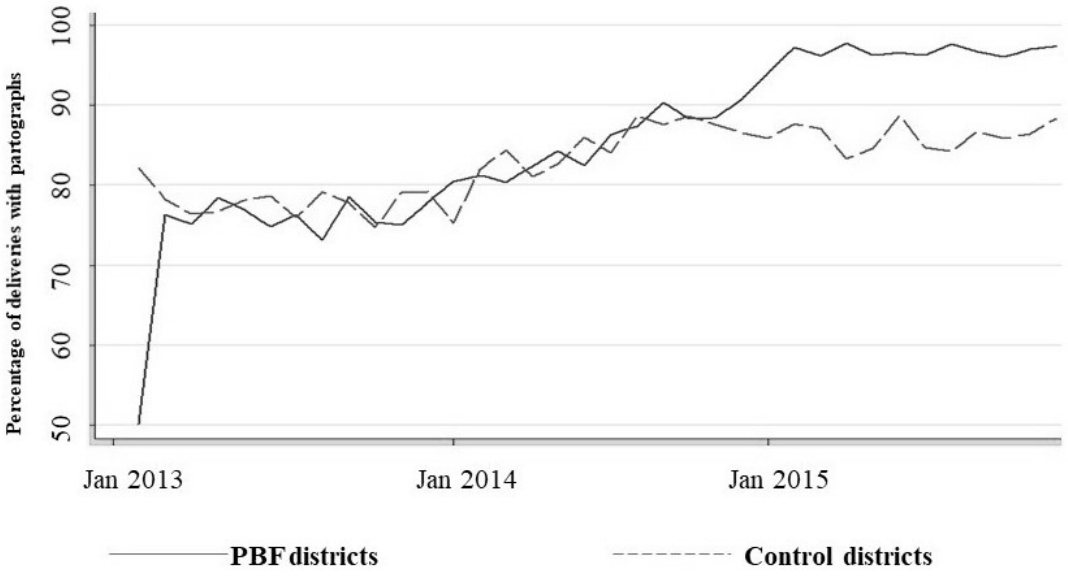
IMCI: In an annual PBF meeting, **Figure 5** was presented to show that, in facilities in PBF districts, the percentage of children treated using the IMCI strategy increased after PBF was implemented (January 2014) compared to facilities in other districts. Participants at the local and national levels argued that the increase in the utilization of the IMCI strategy was one of the main strengths of PBF. Our observation, however, showed that IMCI registers, which determined about 10% of PBF quality scores, were consistently falsified and filled out retrospectively in at least three facilities; they were never used or filled out during consultations, despite the fact that some questions required patients’ input. Those registers were often filled out by another provider than the one who provided care and subsequently signed by a provider who met the qualifications for PBF audits. In facility 4, for example, the drug manager filled out the IMCI registers at his house during his free time, even though he was neither qualified nor present during consultations.

Figure 5. Percentage of children treated using the IMCI approach in PBF and control districts[35]



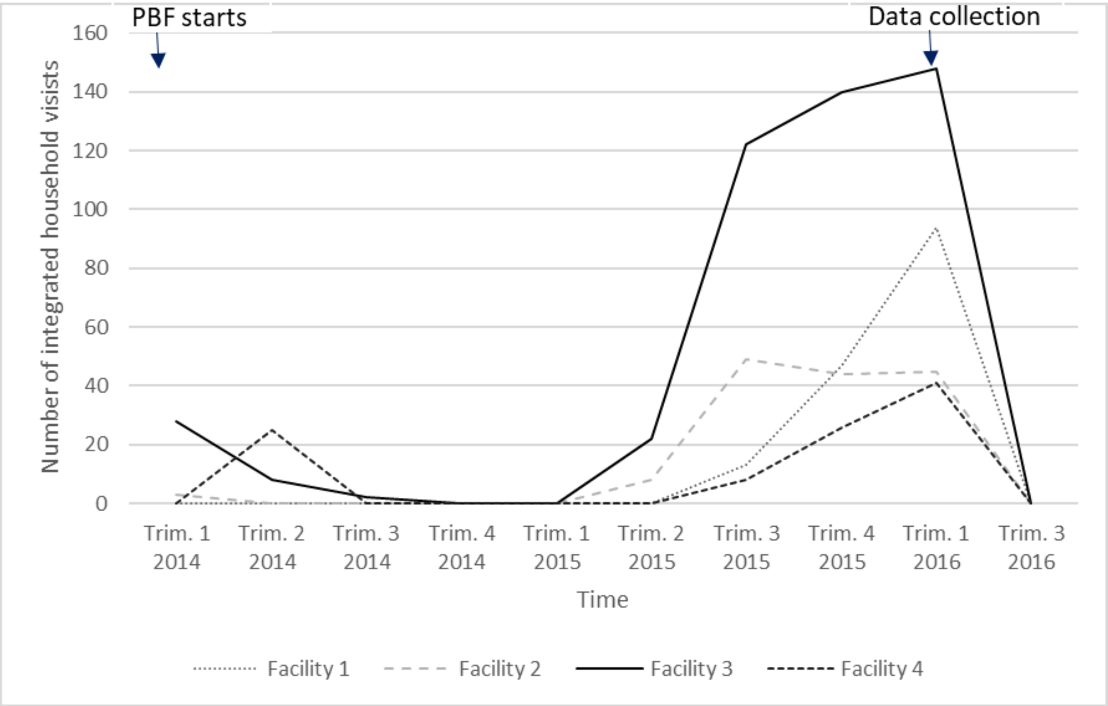
Partograph: According to the official discourse of local actors, the percentage of births conducted using a partograph increased considerably due to PBF (see **Figure 6** presented in an annual PBF meeting). Facilities often received high quality scores for the "proportion of births followed with the help of a partograph" and for "quality of the surveillance for labour and delivery". Observation, however, showed that partograph registers were not routinely used during the childbirth process, despite the fact that some information needed to be reported in a timely fashion to guide clinical decisions. Data reported in partographs were estimated or invented to meet quality evaluation criteria. Partographs were commonly filled out by a qualified provider who did not necessarily attend the birth, sometimes days after the delivery, just before PBF audits. In case 2, for example, the birth attendant calmly filled out multiple partographs in her home, while sitting on a mat, drinking tea. She created a false partograph for a delivery conducted by the itinerant health worker, in her absence, to get PBF quality points. In case 4, the itinerant health worker explicitly explained how the partographs were falsified for PBF verifications.

Figure 6. Proportion of births conducted using a partograph in PBF districts compared to control districts[35]



Integrated household visits (IHV): Under PBF, providers were required to adopt new procedures and forms for conducting IHVs. Participants pointed to the increased number of IHVs conducted over time as one of the main positive effects of PBF (**Figure 7**). Providers across facilities often received high quality scores for IHVs. During the data collection period, however, we did not witness a single provider conduct a proper IHV, despite the fact that providers reported the maximum number of paid IHVs. Observation showed that providers used various strategies to manipulate the actual conduct of IHVs, which were considered "well paid". Providers in at least three facilities falsified IHV forms and conducted IHVs of low quality. For example, providers filled out IHV forms minutes before PBF audits and falsified their content, including signatures of individuals purported to have been present, dates of appointments, and household analyses. The forms were sometimes filled out by individuals who were either not present during the reported visits nor even qualified to conduct them. In case 3, a midwife used women in the maternity ward to complete the forms rapidly without actually visiting their households. Another nurse conducted eight household visits in three hours, despite reports that each visit takes one hour. The content of the lengthy IHVs forms was often superficial, citing the same strengths, weaknesses, causes, and plans across all households. PBF managers confirmed that they observed "fraud" and "major abuses" regarding IHVs and attempted various strategies to resolve this issue (e.g. putting a cap on the number paid, suspending the purchase of IHVs).

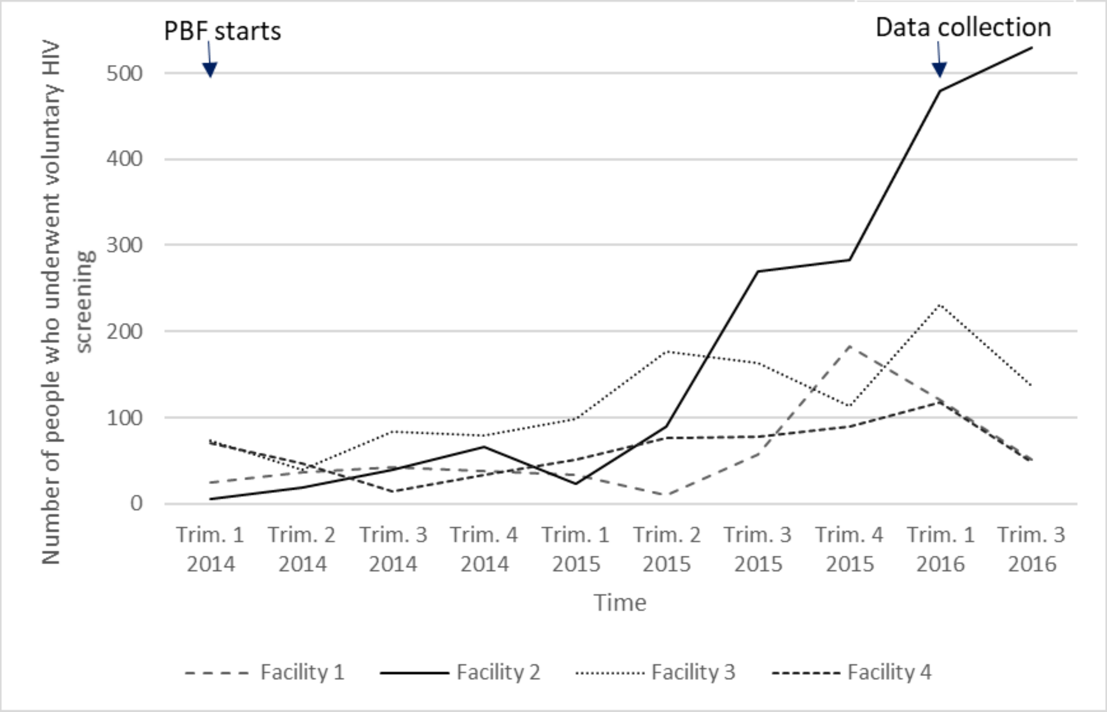
Figure 7. Number of integrated household visits per trimester for four main facilities



Source: www.fbrburkina.org/data/element/16

HIV screening: Quantitative data suggests an increase in the number of voluntary screenings in 2015 and early 2016 (**Figure 8**). However, qualitative data suggested that providers falsely reported HIV screenings. Observation revealed few HIV screenings during consultations across facilities. In case 2, the facility with the highest number of screenings, observation revealed that providers prepared for a PBF audit by creating a new HIV register to report voluntary screenings for past patients and ensuring dates were concordant to avoid looking suspicious. PBF managers explained that "over declarations" were indicated by non-concordance between the stock of reagents available in the country and the number of reported screenings. To discourage abuses, the unit fee for this activity was lowered. PBF auditors became stricter, verifying concordance between the quantity of reagents used and the number of people reported as screened. This partly explains the reduction in reported screenings seen in **Figure 8** during the third trimester of 2016.

Figure 8. Number of people who underwent voluntary HIV screening over time



Source: www.fbrburkina.org/data/element/17

Consultations: False consultations were added for patients followed in the maternity ward. For example, when preparing for a PBF audit, providers in case 2 realized they had conducted few consultations for healthy children between 12-23 months, so they simply added consultations for former patients. Similarly, in case 3, providers added false prenatal consultations for: 1) pregnant women who missed their appointments; 2) postnatal consultations for women who gave birth in the facility; and 3) children treated for malnutrition. Expected consultation dates were filled out with a pen in advance and patient information was filled in even if they did not come.

"*The midwife was sitting on the bench with the registers for moderately and severely acute malnourished children. She recorded data for about ten additional malnourished children who weren't there.... She had no health booklets in her hands and no children around her*." (Case3_observation)

Complacency, collusion, and complicity: There was complacency, collusion, and complicity between providers and managers around manipulating data to improve PBF scores. In all facilities, providers regularly witnessed each other falsifying registers without intervening. Supervisors were sometimes directly or indirectly involved in data falsification and manipulation. In facility B, before a quality audit, a district-level manager asked providers to place a trash container in front of the facility and to wear their lab coats. One PBF auditor also told the medication manager how to prepare receipts that would meet PBF criteria without giving these receipts to patients. In facility 1, a district supervisor advised providers to report a single absence so they could meet PBF criteria without having to report real absences. In facility 3, the head nurse encouraged the midwife to treat children as severely malnourished (an incentivized service), regardless of her clinical assessment. In all these examples, participants explicitly referred to PBF to justify their behaviour.

Under PBF, the different healthcare system levels received performance-based payments. The performance of one level (e.g. facilities) influenced that of others (e.g. districts). Thus, managers, some of whom were PBF auditors, had vested interests in protecting facilities. In 28 months of implementation, no district management team reported any fraud. No sanctions were given for data falsification in any intervention district. This is consistent with the broader social system, wherein providers have a history of mutual protection.

"*The supervisor said, 'It's not possible that no one was absent during the month! Chief, you have to take the hit yourself and put yourself down for one day absent, just one day, so we [district management team] can get our 65 points! We got zero in the last trimester because of that. It sent shivers down our spine!*" (Facility1_observation)

Unanticipated and undesirable

Teaching trainees improper practices: Trainees doing internships in PBF facilities often witnessed or participated in the falsification of registers to increase PBF scores. Providers showed them tricks to avoid detection, such as filling out partographs in reverse, and ensuring consistency in information, and even handwriting, across registers.

"*The midwife arbitrarily changed the register filled out by the trainee in an earlier consultation, telling her, 'Everything must be filled. Everything! Otherwise, it's zero! They don't care about you. PBF makes us write a lot. Too much! All information needs to be consistent. Otherwise, they know you want to cheat, but that you can't.*'" (Facility4_observation)

Overwhelming paperwork : PBF considerably increased the paperwork load. While many registers existed prior to the PBF implementation, they were generally neglected in day-to-day practice. With PBF, some registers were modified to collect more information (e.g. providers' signatures and qualifications). Other documents were added for PBF purposes (e.g. performance improvement plans, index tools, household visit forms). Consequently, participants across facilities complained that PBF required them to write "*too much*", considering the staff shortages. To illustrate this, one head nurse exclaimed, "*PBF is ink!*".

Pursuit of narrow performance indicators: PBF auditors and managers sometimes focused on narrow/specific performance indicators that were perceived locally as irrelevant in the context, unrealistic, or too costly. For example, one PBF indicator referred to having a fence around the facility. No facilities were fenced, so auditors repeatedly recommended fence-building. At a national meeting, the district management team even presented the lack of fences as the primary difficulty relating to quality of care, exhibiting a "tunnel vision" focused on phenomena that were quantifiable in the performance measurement scheme. Under pressure, providers often included "building a fence" or "documenting facility boundaries" as objectives in performance improvement plans. However, most providers interviewed explicitly expressed low buy-in or disagreement with these objectives.

Manipulation of index tools : With PBF, facilities were required to fill out a financial planning instrument, the index tool, each month to determine the premiums each worker should receive (Appendix 3). The amount available for staff premiums depended on the characteristics of the facility (e.g. revenues, expenses, savings) and the healthcare staff (e.g. qualifications, years of experience, absences, individual evaluation scores). Some head nurses and staff manipulated data in the index tool for their own financial gain by: 1) reducing the number of years of experience of other staff members (sometimes illiterate); 2) reporting the wrong staff qualification category of other staff members; 3) lowering a staff member's individual evaluation score; 4) artificially inflating planned expenses to keep money for themselves; 5) under-reporting real expenditures to have more funds available for staff premiums; and 6) failing to report actual absences.

"*PBF is tactical. If we buy another [childbirth] delivery table, there will be nothing left for the workers.*" (Facility3_provider_observation)

"*They deliberately decided not to buy drugs to increase their profit margin and thereby increase their premium.*" (National_manager106_interview)

Tensions and conflicts related to index tools: The index tool caused tensions and conflicts among stakeholders. First, providers were frustrated to learn they were receiving a considerably lower percentage of revenues as premiums than in the pre-pilot study. At the time of data collection, the index tools explicitly stipulated that the percentage of premiums given to providers should not exceed 30% of a facility' revenues (Appendix 3). Second, some providers were dissatisfied with the points attributed for different levels of responsibility. The tool automatically gave head nurses 20 points, which did not necessarily reflect the workload distribution. Third, the lack of transparency of some head nurses when filling out the index tool often provoked internal conflicts. Lastly, participants reported that the lack of formal inclusion of community representatives in index tools caused dissatisfaction, conflicts, and even demotivation. In PBF, community representatives involved in the management committee were required to update documentation, purchase medication, maintain the outside premises, withdraw subsidies from banks, etc. Participants revealed how this devalued their work, stating that they were "*excluded*", "*not part of the team*", "*not important among these people*", and "*doing nothing to increase subsidies*".

Frustrations for providers not eligible for quality points: With PBF, medical registers were modified so that providers specified their names and qualifications. Some services provided by certain categories of providers were not eligible for quality points. However, this criterion clashed with local practices. For example, prior to PBF, itinerant health workers, trained to conduct health promotion activities, and birth attendants, trained to provide support in maternity wards, routinely treated patients by themselves. During quality audits, however, some of their consultations automatically received a score of zero under the justification that they were "unqualified". Itinerant health workers and birth attendants expressed frustration at this perceived injustice. Providers strongly contested this evaluation criterion, arguing that it was not adapted to the local context, given: 1) staff shortages and the difficulty of hiring staff; 2) these workers' life-saving work; 3) head nurses' mobility; and 4) the fact that all providers followed the same diagnostic and treatment guide.

In five facilities, providers developed strategies to systematically falsify the identities of unqualified providers who treated patients. Itinerant health workers, birth attendants, and trainees delivered services but left the signature/qualification columns blank. Later, qualified providers signed their names and qualifications despite their absence during these consultations.

"*What hurts me with PBF is that our actions are not considered quality. This morning, I did a delivery and it went very well, but it's not considered quality. Yet I do the same acts. So, the midwife or head nurse will sign the register. It's not fair.*" (Facility2_itineraryhealthagent_observation)

Staff's dissatisfaction and demotivation due to payment delays: While most participants indicated that PBF was more advantageous than previous practices, the long payment delays were a source of dissatisfaction and demotivation. During the study period, payment delays for quantity-related subsidies were over six months, while those for quality-related bonuses were over 16 months. The quality improvement bonus was cancelled altogether. Many participants reported that delays were getting longer and PBF was losing its dynamism. Their dissatisfaction was exacerbated by the lack of communication regarding the causes of delays and their increased workload.

Suboptimal planning due to payment delays: Every month, providers were required to plan expenditures using PBF management tools (e.g. performance improvement plan, index tool). Providers were required to fill out the index tool as if monthly subsidies were already available. However, long payment delays limited the practical application of PBF management tools across facilities. Participants described this as having “virtual money” that could not actually be spent.

“*The problem with the performance improvement plan is that you plan activities but then don’t have the means to do them because the transfers are late.*” (Facility3_provider72_interview)

Financial issues: Overall, the facilities’ management teams reported having more funds than before PBF. These were used, for example, to replenish medication stocks. However, many participants complained that payment delays caused financial issues over time, especially for small or vulnerable facilities. Under the PBF principle of managerial autonomy, many expenses previously covered by district management teams had been transferred to facilities, such as photocopying, telephone lines, meetings, training costs, mattresses, and carbonized receipts. PBF also generated specific expenses, such as food and drinks for auditors, copying of longer forms and new documentation, PBF meetings, and materials recommended by auditors. While payment delays were a challenge for some facilities, the financial gaps were covered once subsidies were transferred.

Tensions between managerial autonomy and top-down control: PBF increased facilities’ managerial autonomy by allowing providers to make expenditures up to 50 000 FCFA without the district managers’ approval and to recruit additional staff. At the same time, however, PBF was perceived as a directed, top-down approach. Some participants perceived PBF as a form of control. To prevent mismanagement, providers had to follow strict guidelines on how to spend revenues. Ninety percent of the facilities’ expenditure had to be on medications. Providers could only take premiums if the facility’s savings covered operating costs for 90 days. In theory, premiums for workers could not exceed 30% of revenues, but in reality, PBF auditors and district management teams constantly reduced this percentage in the index tool (e.g. to 12–17%) to increase bank reserves. They often modified the index tool content without consulting providers or obtaining consent. Moreover, coercive measures were used to force the adoption of PBF. Facilities that did not follow PBF guidelines or whose performance was not adequate were threatened with suspension. Also, head nurses were required to comply with PBF guidelines as part of their mandate. Thus, there was growing tension between the principle of managerial autonomy and control.

"*The monthly validation of the index tool is somewhat contrary to the principles of autonomy, but they were forced to go there because there were abuses.*" (Facility3_provider72_interview)

Activities delayed and reduced due to gradual withdrawal of other funding: PBF entailed a reduction in other sources of funding from the national level. Certain activities were now required to be covered by PBF subsidies and bonuses. The Programme d'appui au développement sanitaire (PADS - program to support health development) managed a common basket that combined funds from the government and financial partners to support district management teams. Facing financial difficulties, the PADS stopped allocating certain funds to PBF districts, reallocating them to non-PBF districts. Participants in multiple facilities perceived this as reduction of the state's commitment.

Participants believed the changes in funding modalities caused delays and reductions in the number and duration of activities, including meetings and training sessions. According to the PBF principle of managerial autonomy, district management teams and head nurses were expected to assess providers' needs and use subsidies to organize activities, but this did not happen. Many activities previously funded through the PADS were either not organized in a timely fashion or were shortened, possibly affecting their quality. This upset providers, who previously had received per diems when attending these activities.

A "budgetivorous" intervention: Participants in healthcare facilities and at the national level expressed concerns about high costs of PBF related to audits, meetings, registers, etc. Some described the intervention as "budgetivorous", arguing that it disproportionately consumed budgets. Many questioned its financial sustainability.

"*PBF is expensive! ... compared to non-PBF districts, budgets range from equal to 5 or 6 times higher. The results are not proportional. So, we may have to look at how PBF should be adapted to the State budget.*" (National_manager105_interview)

"*PBF eats up budgets. It's making us spend too much.*" (Facility1_provider19_interview)

Discussion

This theory-driven study makes a unique contribution to the literature by documenting a neglected topic, the unintended consequences of PBF in a low-income country. The vast amount of data analyzed will help give stakeholders a more comprehensive picture of PBF consequences in a real-life setting. Consistent with Rogers' diffusion of innovations theory, the results showed that PBF led to both desirable and undesirable unintended consequences, with the latter largely outweighing the former. This was partly due to the fact that some desirable consequences were considered to have been intended by program planners and were therefore outside the focus of this study. For example, we found some evidence that PBF was related to 1) feedback loops between supervisions and PBF audits, 2) some improvement in providers' knowledge, 3) increased social pressure for performance improvement and 4) improvement of staff's socioeconomic well-being. While these were not addressed in the implementation guides, PBF experts considered them to be intended according to the "spirit of the intervention" or its ideas[19]. This highlights the importance of going beyond implementation guides to decipher between intended and unintended consequences.

Moreover, the classification showed that almost all unintended consequences were primarily related to processes (i.e., intervention roll-out) rather than outcomes. This may be, to some extent, because the intervention model identified providers as the locus of behavioural changes[36], providers implemented few creative strategies that affected outputs, and communities were not well informed about or involved in PBF.

The results are consistent with the diffusion of innovations theory, which stipulates that while financial incentives may accelerate an innovation's adoption, the quality of adoption decisions may be low[17]. In this study, providers were incentivized to report increases in quality and quantity of care, but many services were not actually delivered as reported, limiting PBF's potential impact. Furthermore, some providers were fixated on performance measures and subsidies rather than on underlying objectives, again suggesting they did not always internalize the rationale linked to improving certain dimensions of services (e.g. patient confidentiality). These results suggest that, given providers' discretionary power in carrying out interventions, healthcare managers may have to find strategies to improve local actors' adherence to the underlying objectives of PBF to truly increase the quality and quantity of care[37].

Findings from this study raise important methodological considerations for the overall work of assessing PBF impact. While reported quantitative performance data suggested healthcare services had improved considerably, observations revealed that registers were often falsified to artificially enhance performance. The contrast between qualitative and quantitative data shown in this study highlights the risk of relying solely on one method to understand the effects of complex interventions such as PBF. The interpretation of quantitative performance data is more meaningful when implementation processes and local adaptations are considered. Like Cataldo and Kielmann[36], we believe PBF researchers should place more emphasis on spending time in the field, gaining trust, building rapport, conducting observation, and sustaining dialogue with participants to reap rich data that can further the understanding of how stakeholders respond to PBF and its impacts. Such approaches are crucial to send to the right policy signals to decision-makers.

An important question is whether PBF is responsible for the falsification of registers. Prior evidence indicates some falsification occurred in the absence of PBF[38]. During data collection, we did witness some falsification unrelated to PBF. Following the diffusion of innovations theory[17], we considered such behaviours to be part of past experiences and local practices that influence how local adopters re-invent innovations. Nevertheless, the rich data produced through our long-term involvement clearly suggested the existence of a link between PBF and the falsification of registers[39]. We were able to capture the link between PBF and falsification (as well as other unintended consequences) based on an in-depth understanding of meanings, contexts, and processes[40].

This study builds on our previous work on the unintended consequences of community verifications for PBF in Burkina Faso (reference blinded for review). Integrating both articles highlights that the anticipation of community verifications was not sufficient to dissuade providers from falsifying registers. Together, the articles also reveal weaknesses in the overall verification system. Providers were routinely falsifying data to increase performance scores, but community verifications were not able to clearly detect this falsification due to the numerous implementation challenges during the community verifications (e.g., difficulty retracing patients, falsification of community verifiers in charge of tracing patients). In 28 months, no sanctions were given for the falsification of registers. This is similar to what has been observed in Niger where impunity prevails for professional misconduct and the only "sanction" applied in practice is to move a provider to another site[41].

Many unintended consequences detected in our work resonate with studies conducted elsewhere. A study in Rwanda also reported a fixation on performance measures[10]. Participants argued that when an incentive is offered for a precise indicator, it becomes "dissociated" from its very meaning and loses its rationale. That study also noted paperwork overload. Participants explained that time limitations forced them to choose between essential activities and those required for rewards (e.g. paperwork). Consistent with our findings, performance indicators were often falsified to improve reported results[10]. High concordance between providers' declared numbers and PBF auditors' validated numbers, as studied by Kuunibe and colleagues[42], does not rule out falsification. Our results showed deliberate falsification often occurred upstream, directly in the medical registers, and could not easily be detected by PBF auditors, thereby raising questions regarding the effectiveness of audits.

Moreover, many undesirable consequences regarding the payment and distribution of subsidies are consistent with existing evidence. First, research suggests workers lose motivation when incentive agreements are not respected[43,44]. Such implementation lapses go against the intervention theory, which relies on financial incentives to motivate staff. Second, regarding the demotivation of community representatives who did not receive premiums, participants in a study in Tanzania warned against solely rewarding providers, as they often have to collaborate with community leaders[9]. Third, studies in Benin, Rwanda, and Burkina Faso showed that providers were concerned about "unfair" distribution of rewards[10,45,46]. As in our study, it was not always those producing the greatest results who obtained the highest compensation. Lastly, our study echoed findings in Benin, where providers suspected their hierarchic superiors of monopolizing premiums[46]. In Burkina Faso, even Ministry of Health senior executives requested and obtained PBF premiums.

This study has important implications for global health organizations and policy-makers in LMICs. In coming years, many LMIC governments will pursue PBF through new funding agreements with the Global Financing Facility and the World Bank. This is already underway in Burkina Faso[47]. Given their scope and breadth, we advise careful consideration of the undesirable consequences of PBF before pursuing or scaling up the intervention. When discussing preliminary results with high-level stakeholders in Burkina Faso, some revealed that they were already aware of many unintended consequences reported in this study. Deliberations were already underway to resolve some of them. For example, they planned to stop purchasing integrated household visits due to the falsification. They also planned to start paying CHW premiums to increase their satisfaction levels. However, the later was not materialized due to costs. This suggests that some unintended consequences may be addressed while others may have to be accepted as trade-offs if the intervention is to be pursued. Actions can be decided on a case by case basis.

The study has implications for future research. We hope the framework and methods will stimulate research on unintended consequences of PBF in other settings and of other complex health interventions in LMICs to produce more comprehensive evidence to improve population health. Future research could also examine the unintended consequences of PBF in other sectors (e.g., education) in LMICs.

Limitations

We recognize the potential limitations of the study. First, the six facilities were in only two districts, which limits the transferability of findings. Although prolonged observation limited the number of facilities we could include, it produced rich findings with high internal validity. Moreover, member checks with stakeholders at the national level confirmed many of the results. Second, we encountered a language barrier due to the large number of languages spoken in Burkina Faso. However, the researchers' background helped minimize this barrier. The first author's mother tongue is French, a language regularly spoken between providers. She also took courses to learn Dioula and conducted a 4-month immersion program in a Dioula-speaking area. The second author also spoke French and learned Dioula while living in Burkina Faso. We also had to use local interpreters to conduct 15 interviews. Third, the quantitative data were used for descriptive statistics only. We did not perform statistical tests, which limits the depth of these complementary analyses.

Conclusions

PBF is widely implemented in many LMICs to improve healthcare system performance. This multiple case study provided new insights into its unintended consequences and their contributing factors. Results showed PBF led to important unintended consequences in primary healthcare facilities, such as falsification of registers. Most unintended consequences were undesirable and could jeopardize the intervention. With this evidence, policy-makers may be able to develop strategies to avoid or minimize unintended consequences. Others may be accepted as trade-offs. More research-based evidence is needed on unintended consequences of complex interventions to help achieve universal health coverage.

Acknowledgements:

The authors are grateful to Dr. Manuela De Allegri for her valuable input to this manuscript. The authors also express their gratitude to the members of the PBF technical service, including Dr. Philippe Compaoré and Dr. Aloys Zongo, the district-level managers, and all participants for their generous contributions to this study. This study benefited from the data on the Ministry of Health's PBF portal, so we thank all actors involved in collecting and reporting this data (www.fbrburkina.org). The authors would also like to thank Donna Riley for her editing support on a previous version of the manuscript and Guillaume Edger for the graphic design.

This work was supported by a research award funded by the International Development Research Centre (Project n° 107759-99906075-009). AMTT received a training bursary from the Canadian Institutes of Health Research (CIHR). The research project is part of the "Community research studies and interventions for health equity in Burkina Faso" (ROH-115213), supported by CIHR. The authors also received funding from the University of Montreal Public Health Research Institute. Sponsors did not have a role in the study design; the collection, analysis, and interpretation of data; the writing of the report; and the decision to submit the article for publication.

Authors' contributions:

AMTT conceived the study protocol, collected and analyzed the data, and wrote the first draft of the manuscript. VR helped conceive the study protocol, interpret the results, and critically revised the manuscript. IAGG collected part of the data, contributed to the analysis, and reviewed the manuscript. All authors read and approved the final manuscript.

Conflict of interest:

VR was a co-researcher on the baseline study of the impact evaluation of PBF in Burkina Faso. However, he received no salary from the funder (World Bank) for this activity. The authors have no conflicts of interest regarding the publication of this paper.

Ethical issues:

The protocol was approved by the ethics committees in Burkina Faso (deliberation N_ 2015-12-07)) and at the University of Montreal Hospital Research Centre (CE 13.358).

References

1. World Health Organization. Health Financing for Universal Coverage. Moving from Passive to Strategic Purchasing. 2019; http://www.who.int/health_financing/topics/purchasing/passive-to-strategic-purchasing/en/. Accessed December 6, 2019.

2. Kutzin J, Witter S, Jowett M, Bayarsaikhan D. *Developing a National Health Financing Strategy: A Reference Guide.* Geneva: World Health Organization;2017.

3. Ash JS, Sittig DF, Poon EG, Guappone K, Campbell E, Dykstra RH. The extent and importance of unintended consequences related to computerized provider order entry. *J Am Med Inform Assoc.* 2007;14(4):415-423. doi:10.1197/jamia.M2373

4. Eijkenaar F, Emmert M, Scheppach M, Schöffski O. Effects of pay for performance in health care: a systematic review of systematic reviews. *Health Policy.* 2013;110:115-130. doi:10.1016/j.healthpol.2013.01.008

5. Cashin C, Chi Y-L, Smith P, Borowitz M, Thomson S. *Paying for Performance in Health Care. Implications for Health System Performance and Accountability*. Berkshire: Open University Press; 2014.

6. Gorter AC, Ir P, Meessen B. *Evidence Review. Results-Based Financing of Maternal and Newborn Health Care in Low- and Lower-Middle-Income Countries*: German Federal Ministry for Economic Cooperation & Development;2013.

7. Witter S, Fretheim A, Kessy FL, Lindahl AK. Paying for performance to improve the delivery of health interventions in low- and middle-income countries. *Cochrane Database Syst Rev.* 2012;2(CD007899). doi:10.1002/14651858.CD007899.pub2

8. Basinga P, Gertler PJ, Binagwaho A, Soucat ALB, Sturdy J, Vermeersch CMJ. Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: an impact evaluation. *Lancet.* 2011;377(9775):1421-1428.

9. Chimhutu V, Lindkvist I, Lange S. When incentives work too well: locally implemented pay for performance (P4P) and adverse sanctions towards home birth in Tanzania - a qualitative study. *BMC Health Serv Res.* 2014;14(23).

10. Kalk A, Paul FA, Grabosch E. Paying for performance in Rwanda: does it pay off? *Trop Med Int Health.* 2010;15(2):182-190. doi:10.1111/j.1365-3156.2009.02430.x

11. Sieleunou I, Turcotte-Tremblay A-M, De Allegri M, et al. How does performance-based financing affect the availability of essential medicines in Cameroon? A qualitative study. *Health Policy Plan.* 2019;34(Supplement_3):iii4-iii19. doi:10.1093/heapol/czz084

12. Antony M, Bertone MP, Barthes O. Exploring implementation practices in results-based financing: the case of the verification in Benin. *BMC Health Serv Res.* 2017;17(204). doi:10.1186/s12913-017-2148-9

13. Ssengooba F, McPake B, Palmer N. Why performance-based contracting failed in Uganda - an "open-box" evaluation of a complex health system intervention. *Soc Sci Med.* 2012;75(2):377-383. doi:10.1016/j.socscimed.2012.02.050

14. Turcotte-Tremblay A-M, Gali Gali IA, De Allegri M, Ridde V. The unintended consequences of community verifications for performance-based financing in Burkina Faso. *Soc Sci Med.* 2017;191:226-236. doi:10.1016/j.socscimed

15. Lohmann J, Muula AS, Houlfort N, De Allegri M. How does performance-based financing affect health workers' intrinsic motivation? A Self-Determination Theory-based mixed-methods study in Malawi. *Soc Sci Med.* 2018;208:1-8. doi:10.1016/j.socscimed.2018.04.053

16. Maini R, Lohmann J, Hotchkiss DR, Mounier-Jack S, Borghi J. What happens when donors pull out? Examining differences in motivation between health workers who recently had performance-based financing (PBF) withdrawn with workers who never received PBF in the Democratic Republic of Congo. *Int J Health Policy Manag.* 2019;8(11):646-661.

17. Rogers EM. *Diffusion of Innovations*. 5th ed. New York: Free Press; 2003.

18. Bloomrosen M, Starren J, Lorenzi NM, Ash JS, Patel VL, Shortliffe EH. Anticipating and addressing the unintended consequences of health IT and policy: a report from the AMIA 2009 Health Policy Meeting. *J Am Med Inform Assoc.* 2011;18(1):82-90. doi:10.1136/jamia.2010.007567

19. Béland D, Ridde V. Ideas and policy implementation: understanding the resistance against free health care in Africa. *Glob Health Gov.* 2016;X(3):9-23.

20. Jabeen S. Unintended outcomes evaluation approach: A plausible way to evaluate unintended outcomes of social development programmes. *Evaluation and Program Planning.* 2018;68:262-274. doi:10.1016/j.evalprogplan.2017.09.005

21. Koch D-J, Schulpen L. Introduction to the special issue ‘unintended effects of international cooperation’. *Eval Program Plann.* 2018;68:202-209. doi:10.1016/j.evalprogplan.2017.10.006

22. de Zwart F. Unintended but not unanticipated consequences. *Theory Soc.* 2015;44(3):283-297. doi:10.1007/s11186-015-9247-6

23. Burlyuk O, Noutcheva G. Unintended Consequences of EU External Action. *The International Spectator.* 2019;54(1):1-15.

24. Société d'études et de recherche en santé publique. *Rapport d'étape de progrès de la mise en oeuvre et du suivi du processus communautaire de sélection des indigents au Burkina Faso.* Ouagadougou: SERSAP;2014.

25. Ministère de la Santé. *Evaluation finale de la phase-test du financement basé sur les résultats dans les districts sanitaires de Boulsa, Léo et Titao.* Ouagadougou: Ministère de la Santé;2013.

26. Ministère de la Santé. *Guide de mise en oeuvre du financement basé sur les résultats dans le secteur de la santé.* Ouagadougou: Ministère de la Santé;2013.

27. Yin RK. *Case Study Research: Design and Methods*. 4th ed. Thousand Oaks, CA: SAGE Publications; 2009.

28. Patton MQ. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*. 4th ed. Thousand Oaks, CA: SAGE Publications; 2015.

29. Spicer E. *Human Problems in Technological Change. A Casebook*. New York: Russell Sage Foundation; 1952.

30. Warford MK. Testing a diffusion of innovations in education model (DIEM). *Innov J.* 2005;10(3):1-41.

31. Guba EG. Criteria for assessing the trustworthiness of naturalistic inquiries. *ECTJ.* 1981;29(2):75. doi:10.1007/BF02766777

32. Pluye P, Hong QN. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. *Annu Rev Public Health.* 2014;35:29-45. doi:10.1146/annurev-publhealth-032013-182440

33. Olivier de Sardan J-P, Tidjani Alou A. *Epistemology, Fieldwork, and Anthropology*. New York: Palgrave Macmillan; 2015.

34. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int J Qual Methods.* 2006;5(1):80-92.

35. ST-FBR. *Résultats Quantitatifs.* Ouagadougou: St-FBR;2016.

36. Cataldo F, Kielmann K. *Qualitative Research to Enhance the Evaluation of Results-Based Financing Programmes: the Promise and the Reality.* Washington, D.C.: World Bank Group;2016.

37. Lipsky M. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*. New York: Russell Sage Foundation; 1980.

38. Melberg A, Diallo AH, Storeng KT, Tylleskär T, Moland KM. Policy, paperwork and ‘postographs’: global indicators and maternity care documentation in rural Burkina Faso. *Soc Sci Med.* 2018;215:28-35. doi:10.1016/j.socscimed.2018.09.001

39. Maxwell JA. Using qualitative methods for causal explanation. *Field Methods.* 2004;16(3):243-264. doi:10.1177/1525822X04266831

40. Maxwell JA. The importance of qualitative research for causal explanation in education. *Qual Inq.* 2012;18(8):655-661. doi:10.1177/1077800412452856

41. Olivier de Sardan J-P, Bako MTA, Harouna A. *Les normes pratiques en vigueur dans les secteurs de l'éducation et la santé au Niger Une base pour des réformes ancrées dans les réalités ?*: Lasdel; 2018 2018. 127.

42. Kuunibe N, Lohmann J, Schleicher M, et al. Factors associated with misreporting in performance-based financing in Burkina Faso: implications for risk-based verification. *Int J Health Plann Manage.* 2019:1-21.

43. Deci EL, Ryan RM. *Intrinsic Motivation and Self-determination in Human Behavior*. New York: Plenum Press; 1985.

44. Kok MC, Dieleman M, Taegtmeyer M, et al. Which intervention design factors influence performance of community health workers in low- and middle-income countries? A systematic review. *Health Policy Plan.* 2014;30(9):1207-1227. doi:10.1093/heapol/czu126

45. Bodson O. *Analyse pré évaluative de la fidélité de l'implantation de la politique de financement basé sur les résultats au Burkina Faso.* Liège: Université de Liège;2014.

46. Paul E, Sossouhounto N, Eclou DS. Local stakeholders' perceptions about the introduction of performance-based financing in Benin: a case study in two health districts. *Int J Health Policy Manag.* 2014;3(4):207-214.

47. World Bank. *International Development Association Project Appraisal Document on a Proposed Grant in the Amount of SDR 55.7 Million (US$80 Million Equivalent) to the Republic of Burkina Faso for a Health Services Reinforcement Project.* Washington, DC: World Bank;2018. PAD2720.

Appendix 1 – Examples of unit prices for PBF indicators for quantity verifications

		Unit prices in CFA francs			
	Indicators	**Facility 1**	**Facility 2**	**Facility 3**	**Facility 4**
1	Number of new patients aged 5 years and older seen in curative consultation	140	120	140	130
2	Number of new patients aged less than 5 years seen in curative consultation	210	180	210	200
3	Number of days of patients place under observation	350	300	350	330
4	Number of counter-references obtained from hospitals accepting referred patients	1 400	1 200	1 400	1 300
5	Number of children completely vaccinated	430	360	320	390
6	Number of pregnant women who received tetanus toxoid at least twice during the month	350	300	350	330
7	Number of prenatal consultations conducted in the facility during the month	560	480	560	520
8	Number of postnatal consultations conducted in the facility during the month (D6-D8 and W6-W8)	700	600	700	650
9	Number of deliveries conducted with partograph during the month	2 100	1 800	2 100	1 950
10	Number of women (old and new) seen during the month in family planning consultation and using long-term contraceptive methods (DUI and implant)	700	600	700	650
11	Number of women (old and new) seen during the month in family planning consultation and using oral contraceptives or injectables	1 400	1 200	1 400	1 300

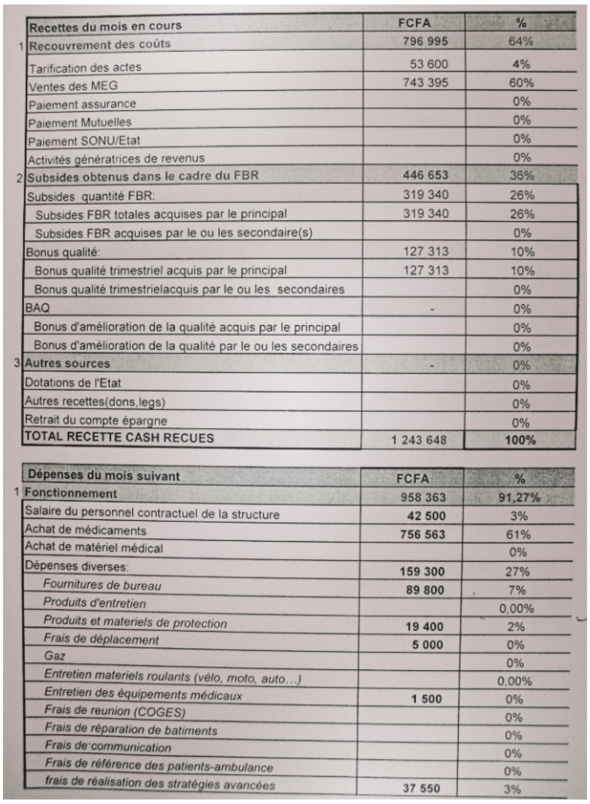
12	Number of newly enrolled patients 0-11 months seen in healthy child consultations	140	120	140	130
13	Number of children aged 12-23 months seen in healthy child consultations	350	300	350	330
14	Number of children aged 6-59 months treated for moderate acute malnutrition	420	360	420	390
15	Number of children aged 6-59 months treated for severe acute malnutrition	1 050	900	1 050	980
16	Number of integrated household visits conducted	4 200	3 600	4 200	3 900
17	Number of people who underwent voluntary screening for HIV (aside from those screened in the context of mother-to-child transmission)	700	600	700	650
18	Number of pregnant women and partners who underwent HIV screening in the context of mother-to-child-transmission	700	600	700	650
19	Number of HIV+ mothers treated with antiretrovirals	3 500	3 000	3 500	3250
20	Number of children born from HIV+ mothers followed	4 200	3 600	4 200	3 900
21	Number of people living with HIV/AIDS treated with antiretrovirals	1 400	1 200	1 400	1 300
22	Number of cases of smear-positive pulmonary tuberculosis (new cases or relapses)	8 400	7 200	8 400	7 800
23	Number of tuberculosis cases (all forms) treated and declared cured or treatment ended	11 900	10 200	11 900	11 050
Note: 1 USD = 581.18 CFA francs					

Appendix 2 – Dimensions of technical quality of care assessed every trimester

	Verified area or activity	**Maximum score possible (case 1)**
1	General indicators	70
2	Conditions of reception/waiting area for patients	50
3	Availability of medical technical equipment	110
4	Bookkeeping	60
5	Management of medications, consumables, and inputs	150
6	Respect of norms related to stocking and storing consumables and to conducting certain activities	140
7	Financial management	80
8	Prevention of infections (hygiene and sanitation)	120
9	Performance improvement plan	60
10	Household visits	80
11	Care for new patients aged 5 years and older seen in curative consultations	100
12	Care for sick children aged 2 months to less than 5 years	250
13	Patients under observation	100
14	Vaccination of children aged 0-11 months	50
15	Prenatal consultations	160
16	Postnatal consultations	90
17	Births	250
18	Family planning	80

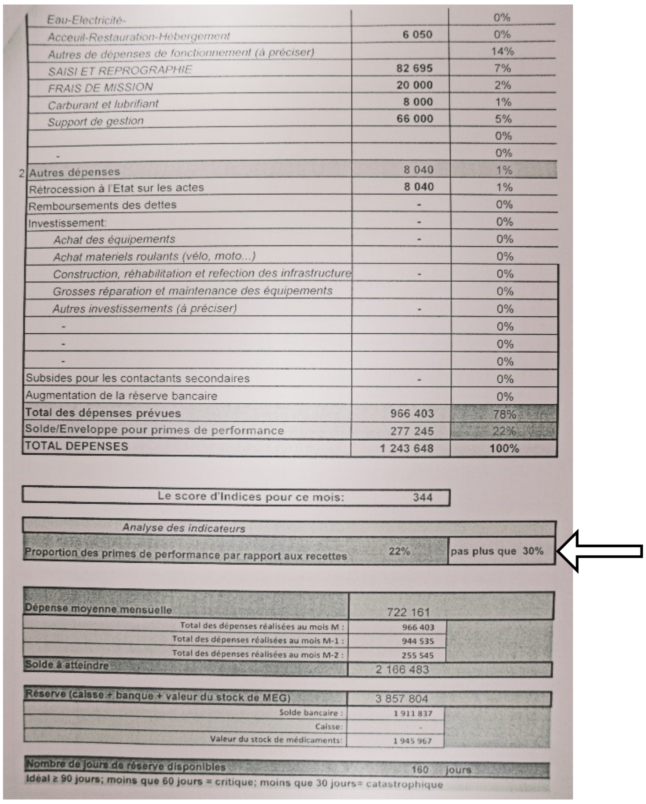
19	Consultations for healthy children aged 0-11 months	55
20	Consultations for healthy children aged 12-23 months	50
21	Care for children aged 6-59 months with moderate acute malnutrition	80
22	Care for children aged 6-59 months with severe acute malnutrition without complications	115
23	Screening for HIV infection	60
24	Care for HIV-positive pregnant women	25
25	Application of the PMTCT protocol for newborns of HIV positive mothers	25
26	Monitoring for people living with HIV under ARV	25
27	Screening for tuberculosis	50
28	Care for treated and cured cases of TB	100
	Total	**2 565**

Appendix 3 – Example of index tool completed



Recettes du mois en cours	FCFA	%
1 Recouvrement des coûts	796 995	64%
Tarification des actes	53 600	4%
Ventes des MEG	743 395	60%
Paiement assurance		0%
Paiement Mutuelles		0%
Paiement SONU/Etat		0%
Activités génératrices de revenus		0%
2 Subsides obtenus dans le cadre du FBR	446 653	36%
Subsides quantité FBR:	319 340	26%
Subsides FBR totales acquises par le principal	319 340	26%
Subsides FBR acquises par le ou les secondaire(s)		0%
Bonus qualité:	127 313	10%
Bonus qualité trimestriel acquis par le principal	127 313	10%
Bonus qualité trimestrielacquis par le ou les secondaires		0%
BAQ	-	0%
Bonus d'amélioration de la qualité acquis par le principal		0%
Bonus d'amélioration de la qualité par le ou les secondaires		0%
3 Autres sources	-	0%
Dotations de l'Etat		0%
Autres recettes(dons,legs)		0%
Retrait du compte épargne		0%
TOTAL RECETTE CASH RECUES	1 243 648	100%

Dépenses du mois suivant	FCFA	%
1 Fonctionnement	958 363	91,27%
Salaire du personnel contractuel de la structure	42 500	3%
Achat de médicaments	756 563	61%
Achat de matériel médical		0%
Dépenses diverses:	159 300	27%
Fournitures de bureau	89 800	7%
Produits d'entretien		0,00%
Produits et materiels de protection	19 400	2%
Frais de déplacement	5 000	0%
Gaz		0%
Entretien materiels roulants (vélo, moto, auto…)		0,00%
Entretien des équipements médicaux	1 500	0%
Frais de reunion (COGES)		0%
Frais de réparation de batiments		0%
Frais de communication		0%
Frais de référence des patients-ambulance		0%
frais de réalisation des stratégies avancées	37 550	3%



Eau-Electricité-		0%
Acceuil-Restauration-Hébergement	6 050	0%
Autres de dépenses de fonctionnement (à préciser)		14%
SAISI ET REPROGRAPHIE	82 695	7%
FRAIS DE MISSION	20 000	2%
Carburant et lubrifiant	8 000	1%
Support de gestion	66 000	5%
		0%
-		0%
2 Autres dépenses	8 040	1%
Rétrocession à l'Etat sur les actes	8 040	1%
Remboursements des dettes	-	0%
Investissement:	-	0%
Achat des équipements	-	0%
Achat materiels roulants (vélo, moto...)		0%
Construction, réhabilitation et refection des infrastructure	-	0%
Grosses réparation et maintenance des équipements		0%
Autres investissements (à préciser)	-	0%
-		0%
-		0%
-		0%
Subsides pour les contactants secondaires	-	0%
Augmentation de la réserve bancaire		0%
Total des dépenses prévues	966 403	78%
Solde/Enveloppe pour primes de performance	277 245	22%
TOTAL DEPENSES	1 243 648	100%

Le score d'Indices pour ce mois: 344

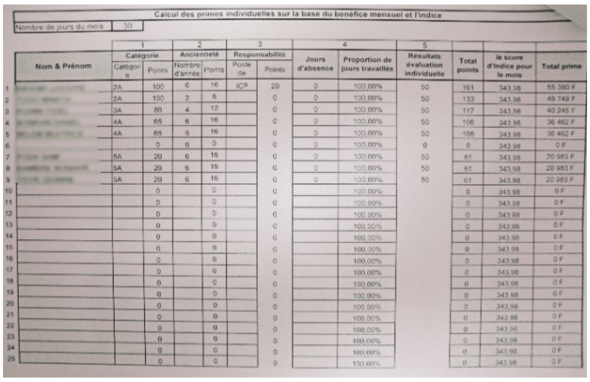
Analyse des indicateurs		
Proportion des primes de performance par rapport aux recettes	22%	pas plus que 30%

Dépense moyenne mensuelle	722 161
Total des dépenses réalisées au mois M :	966 403
Total des dépenses réalisées au mois M-1 :	944 535
Total des dépenses réalisées au mois M-2 :	255 545
Solde à atteindre	2 166 483

Réserve (caisse + banque + valeur du stock de MEG)	3 857 804
Solde bancaire :	1 911 837
Caisse:	-
Valeur du stock de médicaments:	1 945 967

Nombre de jours de réserve disponibles 160 jours

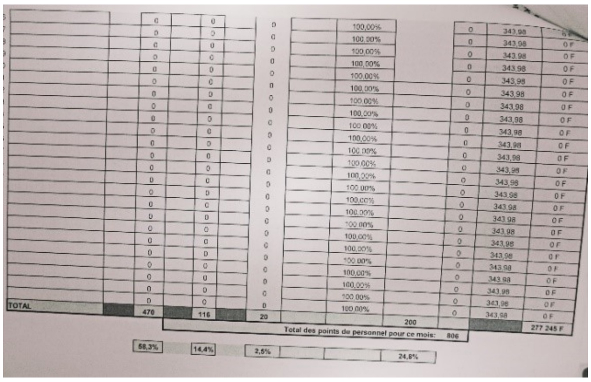
Idéal ≥ 90 jours; moins que 60 jours = critique; moins que 30 jours= catastrophique



Calcul des primes individuelles sur la base du benéfice mensuel et l'indice

Nombre de jours du mois 30

		1		2		3		4		5			
		Catégorie		Ancienneté		Responsabilité							
	Nom & Prénom	Catégorie	Points	Nombre d'année	Points	Poste de	Points	Jours d'absence	Proportion de jours travaillés	Résultats évaluation individuelle	Total points	le score d'Indice pour le mois	Total prime
1	[redacted]	2A	100	6	16	ICP	20	0	100,00%	50	161	343,98	55 380 F
2	[redacted]	2A	100	2	8		0	0	100,00%	50	133	343,98	45 749 F
3	[redacted]	3A	80	4	12		0	0	100,00%	50	117	343,98	40 245 F
4	[redacted]	4A	65	6	16		0	0	100,00%	50	106	343,98	36 462 F
5	[redacted]	4A	65	5	16		0	0	100,00%	50	106	343,98	36 462 F
6			0	0	0		0	0	100,00%	0	0	343,98	0 F
7	[redacted]	5A	20	6	16		0	0	100,00%	50	61	343,98	20 983 F
8	[redacted]	5A	20	6	16		0	0	100,00%	50	61	343,98	20 983 F
9	[redacted]	5A	20	6	16		0	0	100,00%	50	61	343,98	20 983 F
10			0		0		0		100,00%		0	343,98	0 F
11			0		0		0		100,00%		0	343,98	0 F
12			0		0		0		100,00%		0	343,98	0 F
13			0		0		0		100,00%		0	343,98	0 F
14			0		0		0		100,00%		0	343,98	0 F
15			0		0		0		100,00%		0	343,98	0 F
16			0		0		0		100,00%		0	343,98	0 F
17			0		0		0		100,00%		0	343,98	0 F
18			0		0		0		100,00%		0	343,98	0 F
19			0		0		0		100,00%		0	343,98	0 F
20			0		0		0		100,00%		0	343,98	0 F
21			0		0		0		100,00%		0	343,98	0 F
22			0		0		0		100,00%		0	343,98	0 F
23			0		0		0		100,00%		0	343,98	0 F
24			0		0		0		100,00%		0	343,98	0 F
25			0		0		0		100,00%		0	343,98	0 F



		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
		0		0		0		100,00%		0	343,98	0 F
TOTAL		470		116		20			200			277 245 F

Total des points du personnel pour ce mois: 806

58,3%	14,4%	2,5%			24,8%

Appendix 4 – Examples of interview questions

Characteristics and perception of actors

1. What is your current position?
2. How long have you been in this role?
3. What is your involvement in PBF?
4. Could you describe how PBF activities are carried out in your organization?
5. What do you think of PBF (e.g., advantages and disadvantages)?
    - nature of incentives
    - performance indicators
    - verification process
    - incentive distribution
6. How did you first hear about the PBF? What was your reaction when you first heard about it?
7. Who was involved in developing or adapting the intervention?
8. Does everyone support PBF? What reactions did you observe?

Nature of the social system

9. Is PBF adapted to the local context?
    - health system (e.g., organizational capacity, resources)
    - local needs
    - local beliefs and values
10. What are the similarities and differences between you and the promoters of PBF?
11. Do your previous experiences or relationships with donors or promoters of PBF influence your perception of the intervention?
12. What factors influence the implementation process of the intervention?
13. How is information about PBF communicated between the different actors involved?
14. Do you receive any comments or advice about your performance after quality and quantity verifications?

Nature and use of innovation

15. Before the implementation of PBF, what strategies were used to motivate health workers to improve care? Is the new system more useful than the old system?
16. Is PBF easy to understand and implement? What are the difficulties?
17. Did you have the opportunity to test the intervention and make suggestions for improvement?
18. Did the intervention solve some problems and meet your needs?
19. How well is the intervention adapted to what health workers are used to and expect?
20. Is the intervention flexible?
21. Have you changed the way that FBR activities are conducted?
22. How does the distribution of subsidies and bonuses work? What have you received so far?
23. Who benefits from the intervention? Who does not benefit?

Changes

24. What changes has the intervention caused?
    - organization of the health system
    - practices, behaviours or activities
    - governance or supervision
    - level of autonomy of health centres
    - working conditions (wages, number of hours worked) or staff absenteeism
    - quantity of healthcare (discuss possible increases, reductions or stagnation depending on the type of care)
    - quality of healthcare (discuss possible increases, reductions or stagnation depending on the type of care)
    - community outreach activities (e.g., home visits)
    - relationships between actors (e.g., conflict, collaboration)
    - power or social positions of actors
    - cultures, beliefs or values of actors
    - motivation
    - initiatives or strategies to increase the quality and quantity of care

- care for vulnerable groups
- the revenues and expenses of health centres
- availability of or access to medical equipment
- the availability of or access to medicines
- migration or rotation of health staff
- communication between the different actors of the health system
- the participation of community members in the health system (e.g., members of the management committee)
- population health
- laws
- technology
- unintended consequences (desirable or undesirable)

Temporality

25. How have the consequences of PBF changed over time?
26. How could PBF be improved in the future?
27. How do you think that PBF will have evolved in 10 years?

Appendix 5 - Justification of the classification of consequences as anticipated vs. unanticipated according to intervention guides

	Anticipated		Unanticipated	
	Direct (Process)	Indirect (Outcome)	Direct (Process)	Indirect (Outcome)
Desirable			Limits on medication sales without consultations ➢ Not addressed in guides	
Undesirable	Gaming ➢ “To strengthen the credibility of the auditing (quantitative and qualitative), a counter-audit is carried out at six-month intervals by an external entity recruited for this purpose. It consists in verifying, in a sample of health facilities, that the audit was properly conducted.” (Ref 1, p. 71) ➢ “The development of the spirit of initiative. Actors are incited to put in place innovative strategies to address the challenges they face in terms of barriers to access and to the use of services. (Ref 2. p. 22) Fixation on indicators and subsidies ➢ “Derelictions occur when service providers neglect services that are not supported by incentives” (Ref 1, p. 78) Falsification of medical registers and documents ➢ “The results of the survey on the veracity of services may, in the event that fraud is detected, result in sanctions for the offending health facilities” (Ref 1, p. 47)		Teaching trainees improper practices ➢ Not addressed in guides Overwhelming paperwork ➢ Not addressed in guides Pursuit of narrow performance indicators ➢ Not addressed in guides Manipulation of the index tools ➢ Not addressed in guides. An optimistic view of the index tool is presented. Tensions and conflicts related to index tools ➢ Not addressed in guides. An optimistic view of the index tool is presented. Staff’s dissatisfaction and demotivation due to payment delays	

➢ “Fraud refers to the observation, in the supporting documentation for an activity (curative, preventive, or promotional), of a different handwriting, a pen of a different tone or colour, excessive deletions, rampant use of "white-out", the creation of fictitious users, the reporting of acts or services that the user did not actually receive, etc.” (Ref 1, p.78) Complacency, collusion and complicity ➢ “To strengthen the credibility of the auditing (quantitative and qualitative), a counter-audit is carried out at six-month intervals by an external entity recruited for this purpose. It consists in verifying, in a sample of health facilities, that the audit was properly conducted.” (Ref 1, p. 71)		➢ Not addressed in guides. PBF is intended to improve motivation. Suboptimal planning due to payment delays ➢ Not addressed in guides Financial issues ➢ Not addressed in guides Frustrations for providers not eligible for quality points ➢ Not addressed in guides Tensions between managerial autonomy and trop-down control ➢ Not addressed in guides. PBF is intended to improve autonomy. Activities delayed and reduced due to gradual withdrawal of other funding ➢ Not addressed in guides A “budgetivorous” intervention ➢ Not addressed in guides	

References

1. Ministère de la santé (2013). Guide de mise en œuvre du financement basé sur les résultats dans le secteur de la santé. Burkina Faso.
2. Ministère de la santé (2010). Stratégie nationale de mise en œuvre du financement basé sur les résultats. Burkina Faso.

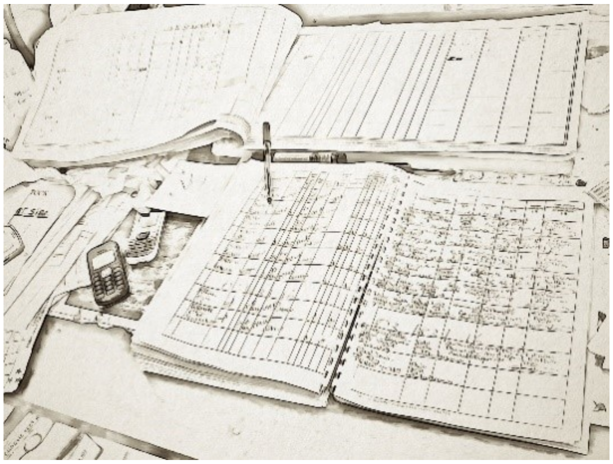
Appendix 6 – Additional examples of the falsification of medical registers and documentation

Identity and qualification of providers: PBF was aimed at improving quality of care by incentivizing the provision of services by certain types of qualified providers. Managers modified registers so that providers could indicate their names and qualification for each service provided. Services provided by certain types of providers (e.g. itinerant health workers) were not eligible for quality points, although they counted for quantity audits. This evaluation criterion clashed with the healthcare system context and local practices. Participants explained that nurses could not consult with all patients due to the quantity of work and the shortage of human resources. Thus, different types of providers were already in the habit of providing care beyond their official level of expertise to avoid leaving patients unattended.

In all but one facility, providers developed a range of strategies to falsify the identity and qualification of providers delivering services. Often, providers who did not qualify for quality points delivered the services as usual but left the signature and qualification columns blank in registers. Later, qualified providers (e.g. nurses for curative consultations or midwives for maternal care) signed their names and qualifications, despite not having been present for those consultations. In case 2, the itinerant health worker, who consulted alone when the head nurse was absent, wrote patient information on a sheet of paper or in an old register. Upon his return, the head nurse transcribed consultation information into the real register that was audited for PBF (see **Figure 9**). The head nurse corrected the information as needed and signed as the sole provider. This ensured that the handwriting was consistent for both the medical information and signature to avoid detection by PBF auditors. In case 3, the midwife and birth attendant systematically co-signed each service delivered in the maternity ward to ensure they met the evaluation criteria regarding provider qualification. These signatures were added subsequently even though they were not both present during consultations.

"The head nurse retranscribed all the consultations and signed as if he had provided the care. But he was in another town… He makes the corrections as he goes along." (case2_observation)

Figure 9. Illustration of the falsification of providers' identity and qualification



Description: Itinerant health workers used an old register (top) to record information on patient consultations. Later, the head nurse transcribed this medical information into the real register (bottom) and signed as the care provider to score quality points during PBF audits.

Dates: Consultations that did not respect the recommended appointment intervals received a score of zero in PBF audits. Consequently, providers across facilities sometimes falsified consultation dates. Participants explained that this rigid criterion was not adapted to the local context because illiterate patients often made mistakes on dates and the long distances made it difficult for them to return another day. Providers argued they should not be penalized for something over which they lacked control.

Prescriptions: Quality points were deducted if providers prescribed medications that did not comply with the diagnostics and treatments guide. Thus, the medication prescriptions reported in medical registers sometimes differed from the medication actually prescribed. This enabled providers to increase quality scores while prescribing what they wished.

"What we have found is that sometimes... in the register, [providers] prescribe what the Guide recommends, but in reality, they prescribe something else." (National manager_106, interview)

Other health data: The falsification of medical data to increase PBF subsidies affected health statistics beyond the intervention. First, providers falsified the number of consultations directly in the medical registers, which are also used to collect data for the National Health Information

System. Every month, providers used the falsified medical registers to fill out the facilities' monthly reports, which were transferred to the districts' Health Information and Epidemiological Surveillance Centres. District teams then entered the falsified statistics into their system, which is used to monitor population health and plan interventions.

Moreover, in case 3, providers who were not familiar with PBF evaluation criteria but who were under pressure to improve performance scores, falsified services that were not covered by PBF. For example, PBF only paid for newly enrolled healthy children seen in consultation but false consultations were also added for returning infants.

5.2 Article 3: The unintended consequences of community verifications for performance-based financing in Burkina Faso

Anne-Marie Turcotte-Tremblay, Idriss Ali Gali-Gali,
Manuela De Allegri and Valéry Ridde

Social Science & Medicine 2017, 191

Doi: 10.1016/j.socscimed.2017.09.007

Title: The unintended consequences of community verifications for performance-based financing in Burkina Faso

Anne-Marie Turcotte-Tremblay[a,b*], Idriss Ali Gali-Gali[c,d], Manuela De Allegri[e], Valéry Ridde[a,b]

a. University of Montreal Public Health Research Institute, Montreal, Canada

b. School of Public Health, University of Montreal, Montreal, Canada

c. Association Action Gouvernance Intégration Renforcement (AGIR), Ouagadougou, Burkina Faso

d. Association Zama Forum pour la Diffusion des Connaissances et des Expériences Novatrices en Afrique (Zama Forum / ADCE-Afrique), Bobo-Dioulasso, Burkina Faso

e. Institute of Public Health, Medical Faculty, Heidelberg University, Heidelberg, Germany

* Corresponding author: Anne-Marie Turcotte-Tremblay, University of Montreal Public Health Research Institute, 7101 Avenue du Parc, Room 3060, Montreal, QC H3N 1X9, Canada

Contributions of authors:

Anne-Marie Turcotte-Tremblay conceived the original idea, developed the protocol, collected and analyzed data, and wrote the first draft of the manuscript.

Manuela De Allegri reviewed the protocol, helped interpret results and critically revised the manuscript.

Idriss Ali Gali-Gali collected part of the data, contributed to the analysis, and revised the manuscript.

Valéry Ridde reviewed the protocol, helped interpret results, and critically revised the manuscript.

Abstract

Performance-based financing (PBF) is being widely implemented to improve healthcare services in Africa. An essential component of PBF involves conducting community verifications, wherein investigators from local associations attempt to trace samples of patients. Community surveys are administered to patients to verify whether healthcare workers reported fictitious services to increase their revenue. At the same time, client satisfaction surveys are administered to assess whether patients are satisfied with the services received. Although some global health actors are concerned that PBF can trigger unintended consequences, this topic remains neglected. The objective of this study was to document the unintended consequences of community verification. Guided by the diffusion of innovations theory, we conducted a multiple case study. The cases were the catchment areas of seven healthcare facilities in Burkina Faso. Data were collected between January 2016 and May 2016 using non-participant observation, 92 semi-structured interviews, and informal discussions. Participants included a wide range of stakeholders, such as community verifiers, investigators, patients, and healthcare providers. Data were coded using QDA Miner, and thematic analysis was conducted. Healthcare workers did not significantly disturb or try to influence community verifiers during patient selection for community verifications. Unintended consequences included stakeholders' dissatisfaction regarding compensation modalities, work overload for community verifiers, and falsification of verification data by investigators. Community verifications led to loss of patient confidentiality as well as fears and apprehensions, although some patients were pleased to share their views regarding healthcare services. Community verifications also triggered marital issues, resulting in conflicts with, or interference from, husbands. The numerous challenges associated with locating patients in their communities led stakeholders to question the validity and utility of the results. These unintended consequences could jeopardize the overall effectiveness of community verifications. Attention should be paid to these unintended consequences to inform effective implementation and refine future interventions.

Highlights

- Community verifications led to important unintended consequences (UC).
- PBF investigators were dissatisfied with compensation and falsified data.
- For patients, verification led to loss of confidentiality and some fear.
- The diffusion of innovations theory is useful to study UC.

Keywords

Burkina Faso, Performance-based financing, Verification, Implementation, Unintended consequences, Multiple case study

1. Introduction

Performance-based financing (PBF) is being widely implemented to improve healthcare services in low- and middle-income countries (LMICs). This approach represents a shift from input-based financing to output-based financing. In PBF, contracted healthcare facilities are paid according to the quantity and quality of services they provide, to motivate them to perform better. To promote accountability and transparency, the services delivered are verified by independent structures before payments are released. While verification is essential for any accountable system, it is a cornerstone of PBF interventions, as it helps ensure that services submitted for payment are actually provided and are of good quality (Fritsche et al., 2014). Thus, PBF aims to reinforce verification mechanisms already in place or set up new ones where they are missing.

Verification mechanisms tend to be structured similarly across PBF models, especially when supported by the same organization (e.g., World Bank). In Burkina Faso, for example, services provided by healthcare facilities contracted under PBF are verified at two levels: within facilities and within communities. For verifications within facilities, a medical verifier from a contractualization and verification agency (CVA) counts the quantity of services reported in registers. Then, an evaluation team evaluates the quality of services by inspecting the conditions of healthcare facilities and the content of registers. However, these two types of verifications are insufficient, because healthcare workers could falsify reports to increase their performance scores or could treat patients poorly even when technical quality scores are high (Gorter, et al., 2013, ST-FBR, 2016).

To address the shortcomings of verifications within facilities, two types 55 of verifications are conducted at the community level, which we refer to as community verifications. For these activities, a community verifier extracts identification and medical information from the consultation registers for a sample of patients who visited the facility in the previous trimester. That information is transmitted to investigators from a local association, who are charged with tracing the sample of patients to administer two surveys at the same time. First, *community surveys* are administered to assess the accuracy of the data provided by healthcare workers by comparing patients' declarations against the health facilities' data (Ministère de la Santé, 2016). This serves to deter healthcare workers from reporting false services as well as to detect fictitious patients or services reported, thereby increasing accountability and transparency, as well as the quality of routine information. Second, *client satisfaction surveys* are administered to determine patients' level of satisfaction with the services provided by the health facilities and to collect patients' suggestions for improving quality of care (Ministère de la Santé, 2016). The information collected through the satisfaction survey contributes to the calculation of the healthcare facility's overall quality score and thereby influences bonus payments that motivate providers. In Burkina Faso, the client satisfaction survey was also presented as a way to reinforce the voice of the community (ST-FBR, 2016). Similarly, some global health actors have argued it can empower communities, leading to a more equal and constructive relationship with providers (Renmans et al., 2017). Other global health actors, however, believe the verification process can create distrust and endanger the relationship between the community and providers (Renmans et al., 2017). Such divergence suggests that, to date, there is a lack of consensus regarding the theory of change and mechanisms at play.

Despite the growing interest around PBF in LMICs, little research has specifically focused on verifications in general or teased apart its multiple mechanisms (Falisse et al., 2012, Renaud and Semasaka, 2014, Renmans et al., 2016, Witter et al., 2013). To our knowledge, the community survey and the client satisfaction survey are neglected research topics, as little empirical data is available and certain useful data collection methods, such as non-participant observation, have not sufficiently been used. This is a significant gap in the literature, considering the importance of community verifications to ensure proper functioning of PBF in LMICs and to increase accountability of systems. One of the rare studies on this topic analyzed 79 community-based organizations (CBOs) contracted to verify health facilities' performance in Burundi (Falisse et al., 2012). The authors concluded that PBF does not necessarily give greater voice to communities and that more experiments are needed to develop efficient mechanisms of accountability in healthcare facilities. More recently, an action research in Benin showed that providers received limited feedback, despite the high costs and time invested in verifications (Antony et al., 2017).

Alongside these considerations, some global health actors are concerned that the different components of PBF can cause unintended consequences beyond the targeted objectives of the intervention (Fretheim et al., 2012). Unintended consequences are defined as changes that occur in a social system for which there is a lack of deliberate action following adoption of an innovation such as PBF (Ash et al., 2007a, Merton, 1936, Rogers, 2003). They can also be desirable or undesirable, as well as anticipated or unanticipated, depending on stakeholders' views. For example, disclosure of patient information during community verification could have consequences for patient confidentiality. To date, little research has examined the unintended consequences emerging from PBF or its verification mechanisms (Witter et al., 2013). This is an important gap in the literature because unintended consequences could have wide scope and breadth, equal to or surpassing intended consequences. Consequently, an evidence-based understanding of intended and unintended consequences could help stakeholders judge an intervention's overall value.

This paper is intended to fill two knowledge gaps simultaneously by using the innovative analytical lens of unintended consequences to study a neglected topic, community verifications of PBF. More specifically, we document the unintended consequences of a community verification process that coupled a community survey with a community client satisfaction survey in Burkina Faso.

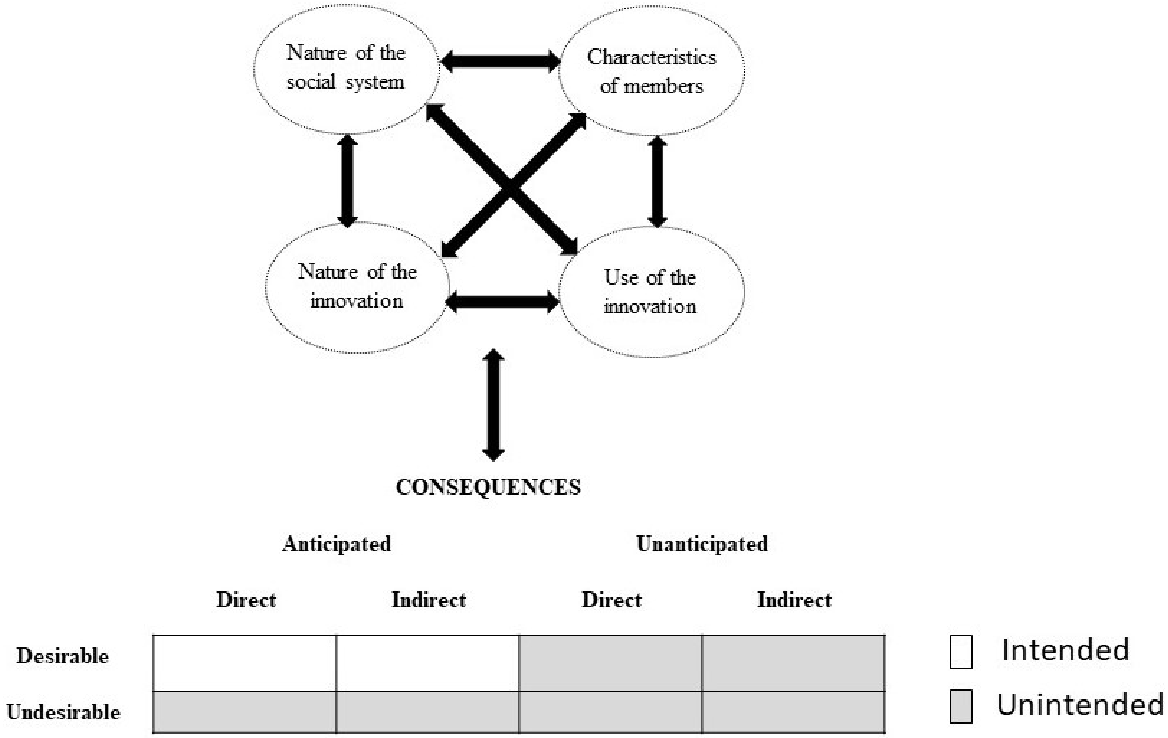
2. Methods

2.1 Theoretical model

We used Rogers' diffusion of innovations theory to study unintended consequences (Rogers, 2003). Innovations, such as community verifications and PBF in Burkina Faso, are ideas or practices that are perceived as new by members of a social system. Innovations are not fixed entities; rather, people shape them by giving them meaning. The theory posits four main dimensions that can influence the diffusion process of innovations, including the emergence of unintended consequences. These are: 1) the characteristics of the members of the social system (e.g. actors' perceptions and interests); 2) the nature of the social system (e.g. norms, culture, organizational capacity); 3) the nature of the innovations (e.g. compatibility, complexity, observability, relative advantage); and 4) the use of the innovations (e.g. reinvention). These dimensions interact to influence the emergence of consequences, although what these will be is uncertain. According to Rogers (2003), change agents are rarely able to predict the consequences of an innovation nor people's subjective perceptions of it. They often fail to consider cultural values, resulting in program failure or at least unforeseen consequences.

Rogers established three categories for classifying consequences of innovations: 1) desirable vs. undesirable, 2) anticipated vs. unanticipated, and 2) direct vs. indirect. In operationalizing these concepts, we considered consequences to be anticipated if they were addressed in the implementation guides. We integrated Ash's (Ash et al., 2007b) approach, by considering direct consequences to be related to processes and indirect consequences to outcomes. We also integrated Bloomrosen et al.'s work (2011), which refined Rogers' categorization of consequences to specify that intended consequences tend to be simultaneously desirable and anticipated, while unintended consequences tend to be undesirable and/or unanticipated. **Figure 10** illustrates our theoretical framework.

Figure 10. Theoretical Framework



Adapted from Rogers (2003) and Bloomrosen et al. (2011)

2.2 Study setting

The study was conducted in a rural district of Burkina Faso. According to the Human Development Index, Burkina Faso ranks 183rd of 188 countries (United Nations Development

Programme, 2015). There are 371 maternal deaths per 100,000 live births and 89 deaths of children under age five per 1000 births (World Bank, 2017).

In 2011, the government of Burkina Faso, supported by the World Bank, conducted a pre-pilot PBF test in three districts to improve maternal and child healthcare services. In 2014, this intervention was modified and expanded to an additional 12 districts. The intervention covers 4.5 million people and involves over 576 healthcare facilities. The intervention model calls for community verifications to be carried out every trimester (Ministère de la santé, 2013). Although PBF started in January 2014, the first and second community verifications were only conducted in June 2015 and May 2016. In both cases, the community verification coupled a community survey with a community client satisfaction survey to determine whether the patients reported in the medical registers actually existed, and if so, whether they had received the services declared and were satisfied with the services provided. Supplementary File 1 describes the actors involved in community verifications in Burkina Faso.

2.3 Research strategy

We conducted a multiple case study with several embedded levels of analyses (Yin, 2009). This research was nested within a larger longitudinal process evaluation of the PBF intervention (Ridde et al., 2014, Ridde et al., 2017).

2.4 Sampling of cases

The cases were seven healthcare facilities and their catchment areas. Six were *Centres de santé et de promotion sociale* (CSPS - centers for health and social promotion) and one was a *Centre médical avec antenne chirurgicale* (CMA – medical center with surgical satellite services, district hospital). For the overarching process evaluation, case selection followed a multistage screening procedure using mixed methods (Yin, 2009). We applied a series of criteria to select the healthcare facilities, including: 1) location (e.g. within a district that represents the normal healthcare system context, in a relatively safe zone); 2) facility type (i.e., CSPSs and hospitals); 3) performance level (e.g. high and low initial performance on key activity indicators); 4) intervention arms (i.e., intervention modalities using different financial incentives); and 5) perceptions of key local informants regarding the facilities' performances, the representativeness of cases, and opportunities for insight.

2.5. Sampling for interviews

Participants included a wide range of stakeholders, such as community verifiers, investigators in charge of tracing patients, community leaders, service users, healthcare providers, and representatives from the *Comité de gestion* (COGES – healthcare facility management committee). Participants were purposefully selected based on their ability to provide relevant information and their accessibility. Then, using the snowball approach, some key informants referred us to other potential participants who could shed light on the intervention. The diversification principle inherent in these approaches resulted in a sample of participants with a variety of intrinsic characteristics, such as different occupations, socio-economic status, and genders (Patton, 2015).

2.6 Data collection method

Data were collected during two sequential qualitative phases, with the first informing the methods used for the second. For the first phase, the first author conducted three months of fieldwork between January and April 2016. Data were collected on the first community verification conducted in June 2015 as well as on the PBF implementation. Field immersion provided a better understanding of context and helped establish trust with stakeholders. Semi-structured interviews, informal discussions, and non-participant observation were conducted in four facilities. Field notes on observations and informal discussions were systematically recorded in research diaries. Observation sites included healthcare facilities and social settings. The researcher also attended a national-level six-day annual PBF review meeting for 2015, where community verifications across the country were discussed.

For the second phase, the second author conducted 20 days of fieldwork in May 2016, specifically to deepen our assessment of the community verifications. While verifications were taking place, the researcher conducted semi-structured interviews, informal discussions, and non-participant observation in each of the seven cases. Field notes were recorded in research diaries. Observation sites included the seven healthcare facilities (during selection of patients), villages (during investigations), and other social settings (during meals and festivities). Due to limited resources, we gave more weight to the four facilities visited in the first phase (primary cases) than to the three cases added in the second phase (secondary cases). We conducted 92 semi-structured interviews: 76 during the first phase and 16 during the second phase, which was primarily devoted to non-participant observations. As well, 241 observation sessions were recorded in research diaries. Local community members not directly involved with the intervention served as interpreters during 15 interviews. Applying the principle of saturation, we stopped collecting data when interviews and observations no longer provided information that was sufficiently different to justify continuing. Research team members transcribed recordings of semi-structured interviews. **Table X** provides a breakdown of the data collected for each case and across cases.

Table X. Summary of data collected to study unintended consequences of community verifications

	Quantity
Non-participant observation	
Sessions reported in field notes	241
Interviews	
Facility level	
Healthcare providers	15
Other support staff (drug manager, janitor, security guard)	13
Volunteers & trainees	7
Community leaders (e.g. COGES & community health workers, counselor)	23
Service users	18
District level	
Administrative staff (e.g. manager, accountant)	3
CVA members	4
Members of local association conducting community verifications	7
National level	
Representative from the *Programme d'appui au développement en santé* (PADS - program to support health development)	1
Representative from the *Service technique - financement basé sur les résultats* (ST-FBR - results-based financing – technical service)	1
Total interviews	92

2.7 Instruments

We constructed semi-structured interview guides that drew upon previous questionnaires used for innovation diffusion research (Spicer, 1952, Warford, 2005) but were tailored to this study's needs and adapted to the different types of stakeholders (see Supplementary File 2). The guides assessed how factors—such as the social system, characteristics of the members, and the nature and use of the innovation—interacted to produce unintended consequences of community verifications over time.

2.8 Data analyses

The primary unit of analysis was each healthcare facility and its catchment area. We conducted thematic analysis on the data. Data were triangulated by comparing various information sources (Olivier de Sardan and Tidjani Alou, 2015). Using a hybrid deductive-inductive approach, we assigned data to predefined themes and derived new themes as we read through the data (Pluye and Hong, 2014). We used QDA Miner 4 to code and retrieve text segments. We integrated the results from both data collection phases and used a cross-case synthesis to draw general conclusions. Following a replication logic, we considered that results independently arising from more than one case are more powerful than those from a single case, and thus the former were given more importance in the results section (Yin, 2009).

2.9 Ethics

The protocol was approved by the research ethics committees in Burkina Faso (deliberation N° 2015-12-07) and at the University of Montreal Hospital Research Center (CE 13.358). Participants provided consent to participate, as required by the ethics committees.

3. Results

Results showed that community verifications led to important unintended consequences for implementers and service users. In **Table XI**, these unintended consequences are classified according to our conceptual model. The subsections below are labeled according to the unintended consequences that arose as the intervention process unfolded overtime. For each, we explain how the interactions between the nature and use of the intervention, the actors' characteristics, and the nature of the social system led to the emergence of these unintended consequences.

Table XI. Classification of unintended consequences of community verifications

	Anticipated		Unanticipated	
	Direct (Process)	Indirect (Outcome)	Direct (Process)	Indirect (Outcome)
Desirable	▨	▨	☐	☐
Undesirable	➢ Falsification of community verification records ➢ Staged supervisions ➢ Dissatisfaction and conflicts regarding payment modalities	➢ Loss of patient confidentiality	➢ Work overload created by patient sampling ➢ Inconclusive process for identifying fictitious services or providing feedback to healthcare workers	➢ Fear and apprehension among patients regarding community verification ➢ Fear of retaliation from healthcare workers ➢ Marital issues

▨ Intended
☐ Unintended

3.1 Work overload created by sampling of patients

The CVA's regular medical verifiers served as community verifiers to select patients from registers. This task required significant time and energy. For the sampling, community verifiers travelled by motorcycle to healthcare facilities up to 90 km away in rural areas, some of which had no paved roads. For the second community verification, two verifiers selected about 400 patients in almost 20 healthcare facilities in under 10 days. Upon arrival, they borrowed medical registers to select patients. The patient selection procedure became more complex between the first and second verifications. While the first used only random sampling, the second used a mixed approach that incorporated purposeful sampling to select particular types of patients (e.g. indigents, patients living within 10 km). Verifiers focused on 10 out of 23 indicators to select patients (e.g. assisted birth, prenatal consultation). Some indicators were omitted to protect confidentiality (e.g. family planning, HIV, tuberculosis) and for reasons of practicality (e.g. household visits).

Community verifiers manually completed a first form for each selected patient by transcribing their name, profession, sex, age, address, telephone number, symptoms, and treatments prescribed. Sample size depended on the number of patients who had visited the facility during the previous trimester but represented approximately 1% of consultations purchased through PBF during the trimester.

> *"[The community verifier] took more than four hours to compile the sample for this healthcare center."* (Field notes, case 5)

During evenings, community verifiers continued preparing the community verification process. They filled out a second form that was given to members of a local association, called investigators, whose role was to trace patients and assess the services received. To help protect patient confidentiality, this form contained no medical information. After the community verifications, verifiers entered the data and analyzed any discordance between both forms to determine whether patients reported the same information that healthcare workers reported.

The heavy workload involved in patient selection influenced the launch of the community verification by local associations. The morning the verification was supposed to start, the investigators were informed that the "forms weren't entirely ready" and that it would be postponed by almost a week. Ultimately, the local association received the samples in two batches, which disrupted the association's schedule and organization.

> "*We really were under pressure, physical and psychological, to get these samplings within the specified time frame*." (Community verifier_28, interview, across cases).

3.2 Little interference of healthcare workers during patient selection

Healthcare workers did not significantly disturb or try to influence community verifiers during patient selection. One head nurse did, however, openly express nervousness regarding the verification process, for example, by hovering around the community verifier during the selection of patients and stating that villagers might speak badly of him despite his hard work. At the district hospital, healthcare workers urged the community verifier to release the registers as soon as possible, as the lack of registers was slowing down their work.

3.3 Dissatisfaction and conflicts regarding funding and payment modalities

The different actors across the cases were dissatisfied with the funding and payment modalities.

3.3.1 Community verifiers

Community verifiers reported that the PADS, i.e., the organization in charge of managing the funds at the national level, did not transfer the money to support community verifications in 2016. Thus, the CVA had to pre-finance the activities, which caused delays, logistical complications, and motivational issues.

3.3.2 Local association

Representatives from the local association in charge of coordinating the investigators decried the lack of financial support for the association. To compensate, they withheld part of the sum paid to investigators for each patient found.

3.3.3 Investigators

Investigators from all cases complained about the cost–benefit ratio of tracing patients in their community. They invested time, energy, and money (including transportation and communication fees) for each patient they were tracking down. In return, they received up to \$4.19 USD (2 500 CFA francs) per patient found, but the local association deducted overhead fees (\$0.81–1.62 USD/patient; 500–1,000 CFA francs/patient). If investigators found that a patient had died or moved away permanently, they were paid. However, if they were unable to find a patient who, for example, was travelling temporarily or was simply absent during the verification, they were not paid. Also, investigators were not paid when patients reported different information than healthcare workers, because these forms were not validated. This was especially problematic with respect to elderly patients who did not remember for which disease they had sought care. This payment scheme was chosen in an attempt to prevent investigators from inventing verification data for patients. However, it caused tensions between the CVA and local association members because it was perceived as a lack of recognition for the work accomplished.

> "*If patients aren't found, their forms aren't validated. So, imagine if I had gone to a healthcare center today, spent my entire day in that area, and didn't find a single person—that would mean my day didn't count, even though it cost me money. That's my situation currently…it's really irritating!*" (Investigator_31, interview, cases 1, 2 & 5)

3.3.4 Community health workers (CHWs)

Some CHWs who helped investigators trace patients in the villages were disappointed by the lack of compensation. While many did not explicitly ask investigators for money, the non-compensation created uneasiness among the parties.

3.4 Fears and apprehensions regarding community verifications

Patients and family members had mixed reactions to community verification, regardless of the dominant ethnic group within the catchment area. Across cases, many patients showed signs of apprehension only in the beginning, while others stayed fearful of investigators throughout the process. Investigators reassured community members by stating that their visit was "cold"—a local expression implying that there was no problem. Some fearful community members even hid information that could have helped investigators find patients.

> "*When you said you were looking for [my name], that's when I got scared and asked myself lots of questions: 'Why are these people at our home, asking not for my husband, but me?' (…) I answered 'yes' in a low voice because I was afraid.*" (Patient_35, interview, case 5)

Once reassured, some patients expressed gratitude regarding the verification process, as it allowed them to share their views.

> "*It's true that it's a surprise, but it also allowed me to share my opinions.*" (Patient_35, interview, case 5)

Part of this apprehension was due to the novelty of the activity and lack of awareness in the communities. Community verifiers explained that, for the 2016 verification, they tried to collaborate more with local actors such as prefects/mayors and radio stations (mainly in the city) to better inform the population. However, one participant reported that they did not “feel the effect of that in the field.”

> *“Many people don’t have a radio at home (...) Because they didn’t get the information, that can have an influence, maybe make them reluctant (...). [Investigators] also don’t have badges to say they’re really authorized to be there.”* (Community verifier_28, interview, across cases)

As the population was generally uninformed about PBF, we examined whether, through community verifications, patients learned that healthcare workers received financial incentives to treat them and whether this influenced the patients’ levels of trust. Results showed that investigators did not directly discuss PBF with patients due to lack of time and fear of complications. Some investigators purposely avoided stating that the activity was a verification or survey, as they felt it would make it more difficult to convince patients to collaborate.

3.5 Fear of retaliation from healthcare workers

Some patients were worried that healthcare workers would find out what they reported and retaliate against them. Some participants at the local and national levels were concerned that patients’ fear influenced their responses.

> *“[the verification] is risky, because the healthcare workers, if they’re badly scored, might think that maybe those of us who talked with you are to blame. Maybe they’ll think we misinterpreted things or maybe, regarding the quality of the healthcare center, we were the ones who spoke badly about the place. Anyway, at the CSPS, they’ll have lots to say if they’re badly rated.” (Patient_35, interview, case 5)*

3.6 Loss of patient confidentiality

Investigators were trained to protect confidentiality when tracing patients and were expected to complete the survey with patients privately. In reality, however, the community verification led to significant breaches of patient confidentiality across cases. The majority of surveys were conducted in front of facility members, neighbors, and community health workers (see Supplementary File 3). The latter sometimes acted as interpreters. Investigators did not systematically ask to be alone with patients before conducting the survey. While many patients did not seem to mind the lack of confidentiality, others reported they were intimidated and would have preferred to be alone during the verification.

> *“The sound of our engines and the fact that we were clearly outsiders aroused the curiosity of neighbors, who approached to see what was happening. The questionnaire was administered to the woman in front of them, and she replied without hesitation.*” (Field notes, case 5)

Women who had consulted for prenatal care or deliveries were particularly embarrassed by the verification. In the local cultures, pregnancy is generally not discussed openly with strangers.

> "*During this time, her brothers-in-law, who were plowing their field just outside their concession, and her sisters-in-law all came near, out of curiosity. Her parents-in-law and her children or nephews were also sitting there listening attentively to her conversation with the investigator. But she seemed ashamed to answer in front of everyone, because she had visited the healthcare center in December to deliver a child. Throughout the interview, she kept her head down and replied in subdued tones.*" (Field notes, case 2)

3.7 Marital issues

Given that PBF targets maternal healthcare, women constituted a considerable portion of patients to be surveyed. All investigators hired for the verification in 2016 were men. We found that some husbands were concerned that strangers from the opposite sex (i.e., investigators) were contacting and meeting their wives without their authorization. In local cultures, wives are sometimes considered to be under the responsibility and authority of their husbands. As such, wives are not always free to be in contact with whomever they want, especially a man they do not know. These gender dynamics led to various consequences: 1) some wives and investigators had altercations with husbands who did not understand why a man was contacting their wives without their consent; 2) some husbands forbade their wives to answer surveys in their absence; and 3) some husbands actively participated in the verification, making suggestions and even influencing their wives’ responses.

> "*You know, last night, there were almost sparks flying here! (...) since I didn't answer, [my husband] said that it was surely my lover who called me.*" (Patient cited in field notes, case 7)

> "*Her husband had an influence on her by sometimes whispering to her, sometimes answering in her stead, especially making suggestions.*" (Field notes, case 2)

3.8. Falsification of community verification records

Multiple sources of evidence suggested that many actors conducting the community verification developed deliberate and organized strategies to falsify the surveys. Evidence suggested the verification data was falsified in the majority of cases, but the strategies used to do this differed according to the investigators in charge of conducting the verification.

Two investigators impersonated PBF officers and used false pretexts to access the registers in healthcare centers and retrieve medical information that would enable them to falsify the patient surveys. Healthcare workers reported that they had helped the investigators go through the registers to find information on patients. Simply using the names and ages of patients, investigators were able to find medical notes in the registers necessary to complete survey questions, such as reason for seeking care and services received.

> "*The investigator came to the CSPS and asked for the consultation registers to search for some missing information on patients selected for the community survey. We gave him the registers.*" (Healthcare worker cited in field notes, case 4)

One investigator who impersonated a PBF officer officially reported that he found all 24 patients (100%) in a single day with the help of CHWs. Yet the CHWs living in the area said they were never contacted. Such reported numbers are high, considering that the target recommended by intervention guidelines is about seven patients per day per investigator.

In another case, an investigator who had an unsuccessful day searching for patients filled out survey forms on his own, without any patients nearby. His recurrent falsification of surveys was confirmed when a patient whom he was supposed to have traced in another catchment area reported to us that she was never in contact with this investigator, despite the fact that the local association reported that 100% of selected patients had been traced in her village. The falsification of forms partly explains why, at the district level, 40% of verification forms reported information that was inconsistent with medical registers during the first verification. This percentage was higher than at the national level, where 28% of forms were inconsistent with medical registers, according to data presented at a national PBF meeting.

> "*The investigator sat on a chair under a mango tree (...) He opened his bag and took out the survey forms. Then, with his pen, he filled out two forms, one after the other.*" (Field notes, case 4).

Another indication that verification data was falsified was that the remarkable success rate (near 100%) reported in the local association's final report was inconsistent with data from nonparticipant observations and interviews (see Supplementary File 4). Observations showed that a large portion of patients were in fact never traced or were absent when investigators passed by. The investigators' statements during interviews, namely regarding the unrealistic working conditions to attain objectives, also clashed with the near-perfect performance reported in the local association's final report.

> "*Per day, we're supposed to contact at least ten people (...) Some days you go all around but can only find one person, and so you've lost both petrol and energy.*" (Investigator_46, interview, across cases)
>
> "*Ten days really aren't enough because there are people who aren't at home when we come by.*" (Investigator_39, interview, case 7)
>
> "*Lots of difficulties! Because, in the time allotted, I couldn't even finish what they asked of me (...) it's impossible to locate more than 80 people in nine days.*" (Investigator_31, interview, cases 1, 2 & 5)

Participants also described at great length how social system characteristics hindered the community verification process. There was a lack of compatibility between the community verification process and the social system. **Table XII** presents the characteristics of the social system and its members that made it difficult to trace patients. Although these obstacles are related to the implementation process, they help explain the context in which investigators falsified data to achieve high performance scores and get paid foreach patient found.

Table XII. Characteristics of the social system and its members that made it difficult to find patients and obtain valid data during community verifications

Characteristics	Examples of citations
CHALLENGES OF TRACING PATIENTS	
Lack of contact information	"*Of the 32 patients to be found, only three had cell phone numbers.*" (Field notes, case 4)
Names of parents missing for child consultations	"*You can't write the name of a 6-month or 12-month-old child without his parents' names, and then ask me to find this child. It will be nearly impossible when it's in a village. When a child is born today, everyone gives him his name (...) Sure, it's a survey, but the point isn't to make investigators suffer.*" (Investigator_31, interview, cases 1, 2 & 5)
Dysfunctional telephone networks in rural areas	*"The investigator tried to reach two patients successively, but their numbers didn't work.*" (*Field notes, case 5*)
Widespread use of nicknames (locally referred to as botanical names) instead of official names	"*If a mother and father fought on the day of the child's birth, they might decide to call the baby 'Big Mouth'."* (Birth attendant cited in field notes, case 1).
Lack of knowledge regarding one's own name or a family member's name	*"A son didn't know his mother's name. The head nurse laughed a little at the patient. The volunteer explained that they use nicknames in the village. They don't even try to know their names. Then, the community health worker arrived with a woman who didn't know her own name. She didn't give the same name as in her health booklet.*" (Field notes, case 1)
Spelling mistakes of names in registers	"*Healthcare workers often write names down phonetically.*" (Field notes, case 2) "*Do names like these exist in our community?!*" (CHW cited in field notes, case 5)
Frequent homonyms within villages	"*In the village we'll find more than 20 people with the same name.*" (Coordinator of local association_38, interview)
Names change over time	*"Someone born in the village, who has a botanical name, and who is then baptized—he goes to the healthcare center, gives his baptism name, and when*

	you go to the village, you'll search for him in vain. You won't find him." (Investigator_31, interview, cases 1, 2 & 5)
Mobility of patients	"*…the period isn't really good for this (…) they're actually farmers, so there I had lots of problems due to patients moving around in agricultural hamlets. That really messed things up.*" (Investigator #30, interview, cases 3 & 7) "*People are migrating.*" (Manager cited in field notes, across cases) "*Mine workers are difficult to locate.*" (Manager cited in field notes, across cases)
Rainy season	"*I started out with rain so I wasn't able to get the number of patients I wanted that day.*" (Investigator_32, interview, cases 4 & 7)
CHALLENGES OF OBTAINING VALID DATA	
Memory lapse	*"We had problems with the forms because some elderly people we found couldn't remember what disease they went there for, and others had visited the healthcare center more than four times with four different illnesses during the trimester, so which illness will you take?"* (Investigator_46, interview, across cases)
Misunderstanding of services provided	"*If we take postnatal consultations, for instance (…) She'll say, 'I was there, but it was for a delivery.' That's how she views it, because the postnatal consultation is seven days after delivery.*" (Verifier_28, interview, across cases)
Fear or apprehension of investigator or of retaliation from healthcare workers	"*Sometimes, patients say no just so the investigator will leave.*" (Manager cited in field notes, across cases)
Social desirability bias	*"The healthcare centers got excellent scores, but we know, based on experience, without studies, that people are complaining (…). If we ask them how they're doing, they'll say there's no problem."* (Manager cited in field notes, across cases)
Desire to protect secrecy	"*Sometimes women consult without their husbands knowing."* (Manager cited in field notes, across cases)

Local actors who intervene across cases (e.g. community verifiers and CVA coordinators, local association coordinator) agreed there were high risks of data falsification and collusion between stakeholders (e.g. healthcare workers and investigators). For example, one community verifier confirmed that he saw surveys for which the data appeared to have been falsified: "*a five-year-old had given birth*." Community verifiers and the CVA coordinator also expressed concerns regarding the falsification of verification records, especially given the difficulty of observing the community verification in action.

Different types of actors involved in community verifications had incentives to report high performance scores. On one hand, the local association obtained the contract through a competitive process, in a context where it had few other ongoing projects or funding opportunities. On the other, investigators were paid solely according to the number of patients found and therefore had a financial incentive to falsify reports, especially in a context where financial difficulties and corruption are common. Healthcare workers were also financially motivated to help investigators in order to increase their bonuses.

> "*Currently, all the investigators are unemployed.*" (Investigator_31, interview, cases 1, 2 & 5)
>
> "*We have the impression we're going to search for cheaters, but there are villages where everyone cheats.*" (Manager cited in field notes, across cases)

3.9. Staged supervisions

Community verifiers are responsible for conducting supervisions of the verification process to assess whether it follows the recommended procedures. However, observation showed some community verifiers were very close to the investigators from the local association. Some had developed friendships and called each other "relatives" (because they shared the same last name) and spent a lot of leisure time together. During one such social outing, a community verifier staged the next day's supervisions with two investigators in charge of finding patients in different catchment areas. He revealed exactly when and where the supervisions were going to occur.

3.10 Inconclusive process for identifying fictitious services or providing feedback to healthcare workers

One of the main purposes of community verification is to "ensure the veracity of the reported healthcare services by identifying fictitious users and fictitious services" (ST-FBR, 2016). According to the intervention model, community verification results can lead to sanctions for healthcare centers if fraud is detected (Ministère de la santé, 2013). When the data were collected, however, the community verification results had not been used to identify cases of fraud and to take appropriate sanctions. The community verification results had not been presented to healthcare workers to enable them to improve their practice. In fact, stakeholders at the district and national level found it quite difficult to interpret the data and conclude that healthcare workers had voluntarily falsified medical records. There was a lack of consensus among stakeholders regarding what decisions and actions to take regarding patients who were not found or services that were not confirmed by patients.

> "*If we're going to penalize [CSPSs], we want to make sure the process is objective.*" (Manager cited in field notes, across cases)
>
> "*We need to identify a reasonable threshold for defining cases as fraud.*" (Manager cited in field notes, across cases)

Participants argued that a missing patient or a person who denied receiving services did not necessarily imply that providers had committed fraud, given the numerous challenges encountered during the verification. The numbers of patients reported as missing also depended on the motivation and abilities of investigators in charge of tracing them. Due to these challenges, participants at the national PBF meeting questioned the methodology adopted for the verification as well as the validity and utility of results. Others questioned the value-for-money obtained. Managers reported that both community verifications jointly cost more than $316,839 USD (186,375,875 CFA francs) across all intervention districts. Some argued that budgets might be better invested in supporting district management teams.

4. Discussion

This paper presents one of the rare studies using Rogers' theory on the diffusion of innovations to study the unintended consequences of a public health innovation in a LMIC. Within and beyond PBF, it sheds light on what happens when new accountability measures and financial incentives are introduced into complex systems. As stipulated by the theory, we found that members' characteristics interacted with the social system and with the nature and use of the innovation, leading to the emergence of unintended consequences over time. The innovation was reinvented during the implementation partly due to the intervention's high level of complexity and its low levels of observability and compatibility with the local context.

All the unintended consequences were undesirable, as they did not promote proper functioning of PBF, or even of the broader health system, for that matter. This study of unintended consequences may have been biased towards undesirable consequences, because many desirable consequences were targeted by the program at the outset, resulting in their exclusion from this study. For example, we did not assess the extent to which the community verifications served as a powerful tool to dissuade healthcare providers from falsifying results or to motivate them to interact with patients more kindly. Moreover, many of these consequences were anticipated by program planners who, in the intervention guidelines, directly or indirectly addressed the risk that investigators would falsify verification records and the importance of preserving patient confidentiality (see Supplementary File 5) (Ministère de la Santé, 2015, Ministère de la Santé, 2016). To adapt our model to the evaluation of innovations in healthcare organizations, we found it useful to qualify direct consequences as those pertaining to process and indirect consequences as those pertaining to outcomes. This provided a meaningful distinction between consequences. From an analytical standpoint, we found that the findings of this study can be generalized to Rogers' theoretical propositions. More than one of the cases supported the theory, so we can claim replication (Yin, 2009).

This qualitative study is an original contribution to a field dominated by quantitative analyses done by health economists. To our knowledge, it is the first qualitative multiple case study to examine the unintended consequences of community verifications, a cornerstone of PBF. Overall, results were not really different between cases. The same set of factors shaped the implementation process and the unintended consequences of community verification. For patients, unintended changes included loss of confidentiality, fears, and apprehensions, as well as marital issues. For actors conducting the verification, unintended changes included work overload, dissatisfaction regarding compensation, and falsification of data. The results of the community verifications were difficult for local stakeholders to interpret due to the numerous challenges encountered during the verifications (e.g. difficult working conditions, population mobility). Some actors questioned the utility and validity of the verification results, which were not presented to healthcare workers or used to identify cases of fraud. Some local stakeholders were left wondering whether community verification offered good value for money, as the costs were perceived to be high. This highlights the importance of pursuing research on the efficiency of this mechanism (Turcotte-Tremblay et al., 2016).

Interestingly, we found that some unintended consequences were mainly the result of poor implementation processes, while others reflected potential weaknesses in the logic of the intervention model. Efforts to improve implementation, for example, could more easily address issues such as work overload, some dissatisfaction regarding payment modalities, and lack of feedback to healthcare providers. Other unintended consequences, however, such as breaches of patient confidentiality, fears of retaliation from healthcare workers, marital issues, and inability to identify fictitious services reported by healthcare providers, may continue to arise even when the intervention is implemented perfectly because the context may not allow for the application of the model in the first place. Program planners may find it more difficult to address these unintended consequences without adapting the intervention model to fit the context. This highlights the importance of examining how real-life contextual factors influence the implementation and effectiveness of intervention models, as few studies have focused on this (Belaid and Ridde, 2015, Shoveller et al., 2016).

Although Rogers' classification of desirable/undesirable consequences is dichotomous, it is important to consider that undesirable consequences are not all equivalent in their negative impacts. For example, some undesirable consequences, such as the falsification of verified data or staged supervisions, may hinder the functioning of the intervention without directly harming population health, at least in the short term. In contrast, other undesirable consequences, such as patients' loss of confidentiality or increased fears, may represent a greater threat to quality of care (given the way quality is measured within the PBF intervention) and may discourage healthcare seeking and hence negatively affect population health. Policy-makers involved in PBF should judge the potential harm of unintended consequences in relation to context and prioritize actions addressing those that are more likely to cause salient levels of damage or harm. This study highlights the importance of examining each component of the complex PBF intervention model in depth. Many components of PBF have yet to be closely scrutinized (Renmans et al., 2016). Past PBF studies, for example, have not sufficiently focused on: 1) healthcare workers' coaching by CVA agents; 2) performance improvement plans; 3) index tools; 4) counter-verifications; and 5) PBF at management levels. Unintended changes in any of these parts of the intervention can trigger significant changes in other parts and consequently should be given attention (Morin, 2006, Rogers, 2003). As described by the complexity approach, we cannot know the whole without knowing the parts (May et al., 2016, Morin, 2006).

The results are consistent with past studies of PBF in LMICs. Many studies have shown that workloads induced by different components of PBF are very burdensome (Antony et al., 2017, Kalk et al., 2010, Paul et al., 2014). Past research has also highlighted verification officers' conflicts of interest (Bertone and Meessen, 2013). For example, the fact that investigators are paid only for patients they trace triggers a clash between their self-interest (i.e., maximizing their personal income) and the public interest (i.e., verifying healthcare workers' performance). Studies of PBF in Rwanda and Burkina Faso found that actors sometimes filled out forms arbitrarily and retrospectively, particularly due to lack of time (Kalk et al., 2010). Similarly, investigators who falsified forms in this study highlighted the unrealistic time frames allotted in which to trace patients. Lastly, our results are consistent with studies that found that patients' comments regarding healthcare services were not presented to the medical staff, despite the amount of resources invested (Antony et al., 2017, Falisse et al., 2012). In Benin, for example, about 0.50 USD was spent on verifications for each 1 USD paid to providers (Antony et al., 2017).

Our study does bring forward some different findings than past literature. The study in Burundi found that community-based organizations contracted as verifiers had gotten in touch with the medical staff, something prohibited by their contract, in only two cases (Falisse et al., 2012). In contrast, our results suggested that community verification data were often falsified. This divergence may be due to the fact that our study included non-participant observations and informal discussions, which tend to reveal more authentic behaviors and beliefs over time (Olivier de Sardan and Tidjani Alou, 2015). Future studies on PBF and community verifications may find it beneficial to include data-gathering techniques inspired by anthropology.

The World Bank's PBF toolkit argues that, in well-designed PBF programs, fewer than 5% of service users cannot be traced back in the community (Fritsche et al., 2014). However, it is not clear where this data come from or what “well-designed” implies in complex settings where implementation is unlikely to be carried out exactly as intended. Results of the current study highlight the need to be skeptical of high performance scores for community verifications, as they may be falsified. Stakeholders should be wary of the potential false sense of security created by an ineffective verification mechanism in which everyone has an incentive to report positive results.

This study does have some limitations. All the PBF investigators observed during the verification were employed by the same local association. It is possible that including investigators from other local associations would have influenced the results. We do, however, believe the results may be transferable to a larger context within Burkina Faso for two main reasons. First, the results were replicated in different cases located in different villages, suggesting they were not due to a particular situation. Second, most of the unintended consequences, and their contributing factors, that emerged during the multiple case study were also reported during an annual PBF review meeting for 2015, which covered other health districts and over 70 local associations implementing community verifications across the country. Another limitation of the study is that it only captured unintended consequences that emerged in the short-term, given the timing of our data collection vis-à-vis the intervention. Research in countries that have more experience in conducting community verifications may be able to confirm the existence of these unintended consequences and provide insight into how they might evolve over time. For example, future research could further explore how verifications modify trust relationships among health workers, patients and managers over time.

5. Conclusion

Community verifications are mechanisms adopted to promote transparency and give greater voice to the population to improve healthcare services. This multiple case study examined the unintended consequences, and their contributing factors, of community verifications in the context of a PBF intervention. Results showed that community verifications led to a series of undesirable unintended consequences. These unintended consequences could jeopardize the overall effectiveness of community verifications and impede the success of PBF.

Acknowledgements

The authors express their gratitude to the members of the PBF technical service, including Dr Aloys Zongo, the district-level managers and the participants for their generous contribution to this study. The authors would also like to thank Donna Riley for her editing support.

Competing interests

MD and VR are researchers on the baseline and endline study of the impact evaluation of PBF in Burkina Faso, but they have not received any salary from the funder (World Bank) for this activity. The authors have no conflict of interests regarding the publication of this paper.

Funding

This work was supported by the Canadian Institutes of Health Research (CIHR), which funded the program (ROH-115213) “Community research studies and interventions for health equity in Burkina Faso.” AMTT received a training bursary from the Canadian Institutes of Health Research (CIHR) and a doctoral research award from the International Development Research Centre. VR holds a CIHR-funded Research Chair in Applied Public Health. Sponsors had no role in the study design; the collection, analysis and interpretation of data; the writing of the report; or the decision to submit the article for publication.

References

Antony, M., Bertone, M.P., Barthes, O., 2017. Exploring implementation practices in results-based financing: the case of the verification in Benin. BMC Health Serv. Res. 17 (204), 1-12. https://doi.org/https://doi.org/10.1186/s12913-017-2148-9.

Ash, J.S., Sittig, D.F., Poon, E.G., Guappone, K., Campbell, E., Dykstra, R.H., 2007a. The extent and importance of unintended consequences related to computerized provider order entry. J. Am. Med. Inf. Assoc. 14 (4), 415-423. https://doi.org/10.1197/jamia.M2373.

Ash, J.S., Sittig, D.F., Dykstra, R.H., Guappone, K., Carpenter, J.D., Seshadri, V., 2007b. Categorizing the unintended sociotechnical consequences of computerized provider order entry. Int. J. Med. Inf. 76 (Suppl. 1), S21-S27.

Belaid, L., Ridde, V., 2015. Contextual factors as a key to understanding the heterogeneity of effects of a maternal health policy in Burkina Faso? Health Policy Plan. 30 (3), 309-321. https://doi.org/10.1093/heapol/czu012.

Bertone, M.P., Meessen, B., 2013. Studying the link between institutions and health system performance: a framework and an illustration with the analysis of two performance-based financing schemes in Burundi. Health Policy Plan. 28 (8), 847-857. https://doi.org/10.1093/heapol/czs124.

Bloomrosen, M., Starren, J., Lorenzi, N.M., Ash, J.S., Patel, V.L., Shortliffe, E.H., 2011. Anticipating and addressing the unintended consequences of health IT and policy: a report from the AMIA 2009 Health Policy Meeting. J. Am. Med. Inf. Assoc. 18 (1), 82-90. https://doi.org/10.1136/jamia.2010.007567.

Falisse, J.-B., Meessen, B., Ndayishimiye, J., Bossuyt, M., 2012. Community participation and voice mechanisms under performance-based financing schemes in Burundi.

Trop. Med. Int. Health 674-682. https://doi.org/10.1111/j.1365-3156.2012.02973.x.

Fretheim, A., Witter, S., Lindahl, A.K., Olsen, I.T., 2012. Performance-based financing in low- and middle-income countries: still more questions than answers. Bull.World Health Organ. 90 (8), 559-559A. https://doi.org/10.2471/BLT.12.106468.

Fritsche, G.B., Soeters, R., Meessen, B., 2014. Performance-based Financing ToolKit. World Bank, Washington, DC. Available from: https://openknowledge.worldbank.org/handle/10986/17194.

Gorter, A.C., Ir, P., Meessen, B., 2013. Evidence review. Results-based financing of maternal and newborn health care in low- and lower-middle-income countries. Ger. Fed. Minist. Econ. Coop. Dev. Available from: https://www.oecd.org/dac/peer-reviews/Evidence-RBF-maternal-health.pdf.

Kalk, A., Paul, F.A., Grabosch, E., 2010. Paying for performance in Rwanda: does it pay off? Trop. Med. Int. Health 15 (2), 182-190. https://doi.org/10.1111/j.13653156.2009.02430.x.

May, C.R., Johnson, M., Finch, T., 2016. Implementation, context and complexity. Implement. Sci. 11 (141). https://doi.org/10.1186/s13012-016-0506-3.

Merton, R.K., 1936. The unanticipated consequences of purposive social action. Am. Sociol. Rev. 1 (6), 894-904.

Ministère de la santé, 2013. Guide de mise en oeuvre du financement basé sur les résultats dans le secteur de la santée. Available from: http://www.dev.bluesquare.org/openrbf/cside/contents/docs/Guide_de_mise_en_oeuvre_du_FBR_dans_le_secteur_de_la_sant%C3%A9_au_Burkina_Faso.pdf.

Ministère de la Santé, 2015. Guide de la contre vérification et contre évaluation de la performance des structures dans le cadre de la mise en oeuvre du financement basé sur les résultats. Burkina Faso.

Ministère de la Santé, 2016. Guide de la vérification communautaire dans le cadre de la mise en oeuvre du financement basé sur les résultats. Burkina Faso.

Morin, E., 2006. Restricted Complexity, General Complexity. Available from: http://cogprints.org/5217/1/Morin.pdf.

Olivier de Sardan, J.-P., Tidjani Alou, A., 2015. Epistemology, Fieldwork, and Anthropology. Palgrave Macmillan, New York, NY.

Patton, M.Q., 2015. Qualitative Research & Evaluation Methods: Integrating Theory and Practice, fourth ed. SAGE Publications, Thousand Oaks, CA.

Paul, E., Sossouhounto, N., Eclou, D.S., 2014. Local stakeholders' perceptions about the introduction of performance-based financing in Benin: a case study in two health districts. Int. J. Health Policy Manag. 3 (4), 207-2014. https://doi.org/10.15171/ijhpm.2014.93.

Pluye, P., Hong, Q.N., 2014. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. Annu. Rev. Public Health 35, 29-45. https://doi.org/10.1146/annurev-publhealth-032013-182440.

Renaud, A., Semasaka, J.-P., 2014. Verification of Performance in Results-based Financing (RBF). The Case of Community and Demand-side RBF in Rwanda. World Bank, Washington, DC cited 2016 Nov 7]. Available from: http://hdl.handle.net/10986/20791.

Renmans, D., Holvoet, N., Criel, B., Meessen, B., 2017. Performance-based financing: the same is different. Health Policy Plan. 32 (6), 860-868. https://doi.org/10.1093/heapol/czx030.

Renmans, D., Holvoet, N., Orach, C.G., 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 & Plan. 31, 1297-1309. https://doi.org/https://doi.org/10.1093/heapol/czw045.

Ridde, V., Turcotte-Tremblay, A.-M., Souares, A., Lohmann, J., Zombré, D., Koulidiati, J.L., et al., 2014. Protocol for the process evaluation of interventions combining performance-based financing with health equity in Burkina Faso. Implement. Sci. 9 (1), 149. https://doi.org/10.1186/s13012-014-0149-1.

Ridde, V., Yaogo, M., Zongo, S., Somé, P.-A., Turcotte-Tremblay, A.-M., 2017. Twelve months of implementation of health care performance-based financing in Burkina Faso: a qualitative multiple case study. Int. J. Health Plann Mgmt 1-15.

Rogers, E.M., 2003. Diffusion of Innovations, fifth ed. Free Press, New York.

Shoveller, J., Viehbeck, S., Ruggiero, E.D., Greyson, D., Thomson, K., Knight, R., 2016. A critical examination of representations of context within research on population health interventions. Crit. Public Health 26 (5), 487-500. https://doi.org/10.1080/09581596.2015.1117577.

Spicer, E.H. (Ed.), 1952. Human Problems in Technological Change. A Casebook. Russell Sage Foundation, New York.

ST-FBR, 2016. Résultats de l’enquête communautaire. Presented at the Atelier bilan 2015 de la mise en œuvre du FBR au Burkina, Ouagadougou, Burkina Faso.

Turcotte-Tremblay, A.-M., Spagnolo, J., De Allegri, M., Ridde, V., 2016. Does performance-based financing increase value for money in low- and middle-income countries? A systematic review. Health Econ. Rev. 6 (1), 30. https://doi.org/10.1186/s13561-016-0103-9.

United Nations Development Programme, 2015. Human Development Report 2015. Work for Human Development, New York. Available from: http://hdr.undp.org/en/2015-report.

Warford, M.K., 2005. Testing a diffusion of innovations in education model (DIEM). Innovat. J. Public Sect. Innovat. J. 10 (3), 1-41.

Witter, S., Fretheim, A., Kessy, F.L., Lindahl, A.K., 2013. Paying for performance to improve the delivery of health interventions in low- and middle-income countries. Cochrane Database Syst. Rev. 2 (CD007899), 1-82. https://doi.org/10.1002/14651858.CD007899.pub2.

World Bank, 2017. Health, Nutrition and Population Data and Statistics. HNP Sustainable Development Goals. Retrieved from. http://datatopics.worldbank.org/hnp/HNPSDG.

Yin, R.K., 2009. Case Study Research: Design and Methods, fourth ed. SAGE Publications, Thousand Oaks, CA.

Supplementary File 1. Description of actors involved in the community verifications

Actors	Brief description[1]
Programme d'appui au développement en santé [PADS[2] - program to support health development]	This national-level organization is in charge of managing and distributing World Bank funds to different PBF-related stakeholders.
Service technique FBR [ST-FBR - results-based financing –technical service]	This unit manages and coordinates PBF implementation at the national level. It develops guidelines, provides training to contractualization and verification agencies, conducts technical monitoring and analyzes data. This unit is part of the Ministry of Health.
Contractualization and verification agency (CVA)	This agency is responsible for establishing contracts with facilities within the district, verifying the quantity of services on a monthly basis, and coaching healthcare workers to reach PBF standards. For community verifications, they are in charge of recruiting the local association, training investigators, and entering the verification data.
Community verifiers	Community verifiers are responsible for selecting the samples of patients to be traced for community verification. They are hired by the CVA. In this case, the community verifiers were the same agents who conducted the verification for the quantity of healthcare services and provided the coaching in healthcare facilities.
Local association	This association is contracted by the CVA to recruit investigators from the community and conduct the verification. In this case, the local association was a relatively small organization that usually intervenes in the field of health promotion and social development. Some districts have more

	than one association for this work but, in this case, the local association covered the entire district.
Investigators	Investigators are hired by the local association to search for patients in the different communities and administer the surveys. They submit the verification data to the CVA.
Healthcare workers	Healthcare workers provide medical registers to community verifiers from which to sample the patients that need to be traced.
Community health workers (CHW)	CHWs are volunteers in the villages who assist healthcare workers during vaccination campaigns and other health-promotion activities. Investigators often seek their help to identify and trace patients for community verifications.
Patients selected to be traced and surveyed	These patients were selected by community verifiers to determine whether they had truly visited the healthcare facilities, whether they had received the services reported by the healthcare workers, and whether they were satisfied with services received.

1. The roles are mainly described in relation to the verifications.
2. PADS is the equivalent of the Health Sector Wide Approach (SWAp) in English.

Supplementary File 2. Examples of Interview Questions

General questions

1. What is your occupation?
2. How long have you had this role?
3. What is your role regarding the intervention?

Characteristics and perceptions of actors

4. How do the different actors (inc. community-based health workers) participate in the community verifications?
5. What do you think about the community verifications? What are the advantages or disadvantages?
6. Does everyone support this intervention? What responses have you observed?
7. Who benefits from the community verification? Who does not benefit?

Nature and use of the innovation

8. How are the community verifications are conducted (e.g. selection of participants, tracing patients, completing surveys)?
9. Is the community verification easy to understand and implement? Are there difficulties, namely to trace patients?
10. Which strategies were used to find patients?
11. What happens when patients are not found? How is it interpreted?
12. How are ressources managed and used for community verifications?
13. How does money circulate between actors involved in the community verification?
14. Did local actors receive feedback regarding the healthcare center's performance after the community verifications?
15. Which local association conducts the verification and how does that influence its implementation?
16. What are the work conditions of the contractualisation and verification agents, the community verifiers and investigators for the community verfication?
17. Has the community verification process been modified overtime?

Nature of the social system

18. Which factors facilitated or hindered the implementation process of community verifications?
19. What are the relations between the different types of actors before, during and after the community verification?
20. How is the information about the intervention communicated between the different actors involved?
21. Is the community verification suited to the local healthcare system?
22. Is the community verification suited to local needs?
23. Is the community verification suited to local beliefs or values?

Consequences

24. Did the community verification cause changes? Which ones?
25. Did this intervention lead to desirable or undesirable consequences that were not intended in the beginning?

26. How did the different types of actors react to the community verification?
    - Healthcare workers
    - Patients
    - Community members (e.g., health workers, village chiefs, leaders)
    - Contractualisation and verification agents, community verifiers
    - Investigators
27. Did the community verification change relations between actors (e.g., collaboration, conflitcs)?
28. Did the community verification influence behaviours, beliefs or practices of different actors?
29. Did the community verification influence the level of confidentiality of patients?
30. Did the community verification stigmatise certain patients?
31. How does the community verification influence indigents or vulnerable groups?
32. How did the community verification influence the work conditions and activities of healthcare providers?
33. Did investigators talk about PBF with the populations during the community verification? If so, how did these people react?
34. Did the community verification process affect community outreach or awareness?
35. Is the power or the social position of some individuals changed by the community verification?
36. Does the intervention change the participation of the community in the healthcare system?
37. Did the consequences of community verification change over time?

Supplementary File 3. Illustrations of the community verification[11]

Description: A community verification conducted in front of extended family members and neighbors. During the entire interview, the woman kept her head down and answered in a subdued voice. Her husband influenced the verification by whispering answers in her ear and answering in her stead.

Description: Community verification conducted in front of family members and neighbors.

---

[11] Images were modified to preserve anonymity.

Description. A community verification conducted by an investigator with a patient. A member of the COGES served as interpreter during the interview while the husband, a co-wife, and a neighbor listened attentively.

Supplementary File 4. Number of patients to trace and patients found

	1st community verification			2nd community verification		
Anonymized Facilities	Number of patients to trace	Number of patients found	Percentage found	Number of patients to trace	Number of patients found	Percentage found
Facilities included in study						
Case 1	15	15	100	24	24	100
Case 2	9	9	100	13	13	100
Case 3	18	18	100	24	24	100
Case 4	11	11	100	10	10	100
Case 5	35	34	97	32	31	99
Case 6	11	11	100	17	17	100
Case 7	19	17	89	21	21	100
Other facilities in the district						
A	26	26	100	22	22	100
B	12	12	100	19	19	100
C	14	14	100	19	19	100
D	25	24	96	35	35	100
E	10	10	100	11	11	100
F	13	13	100	18	18	100
G	52	52	100	64	64	100
H	1	1	100	3	3	100
I	13	12	92	13	13	100
J	16	16	100	21	20	95
K	7	7	100	10	10	100
L	22	20	91	26	26	100
M	17	17	100	24	24	100
N	29	30	103	37	37	100
O	14	14	100	18	18	100
TOTAL	**390**	**384**	**98**	**481**	**479**	**100**

Supplementary File 5. Justification for the classification of consequences as anticipated or unanticipated based on implementation guides

	Anticipated		Unanticipated	
	Direct	**Indirect**	**Direct**	**Indirect**
Desirable				
Undesirable	- Falsification of community verification records: “*The contract between the CV agency and the local organisation can be terminated in case of (…) cheating in the completion of questionnaires*” (1). - Staged supervisions : “*counter-verification is necessary to prevent (…) acquaintances, false results, frauds and cheating*” (2). - Dissatisfaction and conflicts regarding payment modalities: “*it can be necessary to use ressource people such as CHWs to help investigators in their work to collect data from service users. In that case, it is necessary that local associations provide a compensation for the work that they will have conducted*" (1).	- Loss of patient confidentiality: “*although it may be necessary to get assistance from CHWs (or other ressource people) to locate service users, their participation during interviews is strongly discouraged*” (1).	- Work overload created by patient sampling: this potential change is not addressed in guides. - Inconclusive process for identifying fictitious patients or providing feedback to healthcare workers : The intervention aimed to provide to identify fictitious patients and to provide feedback but the probability that this would not occur was not addressed in guides.	- Fear and apprehension among patients regarding community verification : this potential change is not addressed in guides. - Fear of retaliation from healthcare workers:: this potential change is not addressed in guides. - Marital issues : this potential change is not addressed in guides.

■ Intended

□ Unintended

References

1. Ministère de la Santé. Guide de la vérification communautaire dans le cadre de la mise en oeuvre du financement basé sur les résultats. Burkina Faso; 2016.

2. Ministère de la Santé. Guide de la contre vérification et contre évaluation de la performance des structures dans le cadre de la mise en oeuvre du financement basé sur les résultats. Burkina Faso; 2015.

5.3 Article 4: The unintended consequences of combining equity measures with performance-based financing in Burkina Faso

Anne-Marie Turcotte-Tremblay, Manuela De Allegri,
Idriss Ali Gali-Gali and Valéry Ridde

International Journal for Equity in Health 2018, 17:109

DOI 10.1186/s12939-018-0780-6

Title: The unintended consequences of combining equity measures with performance-based financing in Burkina Faso

Anne-Marie Turcotte-Tremblay[1,2*], Manuela De Allegri[3], Idriss Ali Gali-Gali[4,5], Valéry Ridde[1,6]

[1] University of Montreal Public Health Research Institute, Montreal, Canada

[2] School of Public Health, University of Montreal, Montreal, Canada

[3] Institute of Public Health, Medical Faculty, Heidelberg University, Heidelberg, Germany.

[4] Association Action Gouvernance Intégration Renforcement (AGIR), Ouagadougou, Burkina Faso.

[5] Association Zama Forum pour la Diffusion des Connaissances et des Expériences Novatrices en Afrique (Zama Forum / ADCE-Afrique), Bobo-Dioulasso, Burkina Faso.

[6] IRD (French Institute for Research on Sustainable Development), CEPED (IRD-Université Paris Descartes), Universités Paris Sorbonne Cités, ERL INSERM SAGESUD, Paris, France.

* Corresponding author: Anne-Marie Turcotte-Tremblay, 7101 Avenue du Parc, Room 3060, Montreal, QC H3N 1X9, Canada.

Contributions of authors:

Anne-Marie Turcotte-Tremblay conceived the study protocol, collected and analyzed the data, and wrote the first draft of the manuscript.

Idriss Ali Gali-Gali collected part of the data, contributed to the analysis, and revised the manuscript.

Manuela De Allegri and Valéry Ridde helped conceive the study protocol, interpret the results, and critically revise the manuscript.

Abstract

Background: User fees and poor quality of care contribute to low use of healthcare services in Burkina Faso. The government implemented an innovative intervention that combines equity measures with performance-based financing (PBF). These health equity measures included a community-based selection of indigents to receive user fee exemptions and paying healthcare centres higher purchase prices for services provided to indigents. Research suggests complex interventions can trigger changes not targeted by program planners. To date, however, there is a knowledge gap regarding the unintended consequences that can emerge from combining PBF with health equity measures. Our objective is to document unintended consequences of the equity measures in this complex intervention.

Methods: We developed a conceptual framework using the diffusion of innovations theory. For the design, we conducted a multiple case study. The cases were four healthcare facilities in one district. We collected data through 93 semi-structured interviews, informal discussions, observation, as well as intervention documents. We conducted thematic analysis using a hybrid deductive-inductive approach. We also used secondary data to describe the monthly evolution of services provided to indigent and non-indigent patients before and after indigent cards were distributed. Time series graphs were used to validate some results.

Results: Local actors, including members of indigent selection committees and healthcare workers, re-invented elements of the PBF equity measures over which they had control to increase their relative advantage or to adapt to implementation challenges and context. Some individuals who did not meet the local conceptualization of indigents were selected to the detriment of others who did. Healthcare providers believed that distributing free medications led to financial difficulties and drug shortages, especially given the low purchase prices and long payment delays. Healthcare workers adopted measures to limit free services delivered to indigents, which led to conflicts between indigents and providers. Ultimately, selected indigents received uncertain and unequal coverage.

Conclusions: The severity of unintended consequences undermined the effectiveness and equity of the intervention. If the intervention is prolonged and expanded, decision-makers and implementers will have to address these unintended consequences to reduce inequities in accessing care.

Keywords: Performance-based financing, User fee exemption, Universal health coverage, Indigents, Unintended consequences, Burkina Faso, Multiple-case study

Background

Achieving health equity remains a challenge in many low- and middle-income countries (LMICs). User fees significantly limit access to services, especially for the poor, while quality of care is often considered to be insufficient. In the pursuit of universal health coverage (UHC), governments are adopting a range of interventions to provide access to high-quality health services without exposing patients to financial hardship [1, 2]. Some approaches are primarily directed at service providers (supply side) to improve the quality of healthcare services, while others focus on beneficiaries (demand side) to reduce financial obstacles that limit access to care. Interventions that combine measures to improve equity in service use, quality of care, and financial protection may be promising, as they provide a more comprehensive response to health needs [1, 2].

In this vein, performance-based financing (PBF) is increasingly being adopted to improve the quantity and quality of healthcare services. However, few attempts have been made to combine PBF with equity measures that target vulnerable groups, in spite of emerging evidence suggesting PBF is not inherently pro-poor [3, 4]. In Cameroon, for example, a PBF program with specific measures to target the poorest found under-coverage was a concern. Indigents who attended the facility constituted only a tiny proportion of the population (maximum 0.7%) [5]. According to Renmans and colleagues [6], consensus exists on the fact that "*PBF is not adapted to tackle social determinants or health inequities*." More broadly, it is possible that any purchasing mechanism, by being primarily focused on the supply side, has difficulty producing equity changes. Global health actors are consequently calling for strategic purchasing reforms such as PBF to be reoriented by linking them with additional measures that can promote equity and achieve universal health coverage by 2030 [7].

Innovating in this field, the government of Burkina Faso received financial and technical support from the World Bank to test PBF with different equity measures specifically targeting indigents [8]. Health equity measures included: a) a community-based selection of indigents, b) user fee exemption measures for indigents at point of service, and c) higher purchase prices to healthcare centres for some services delivered to indigents than for those provided to non-indigents. To select indigents, a local consultancy firm was contracted to adapt and reproduce the process described by Ridde, whereby village committees proposed lists of indigents that were then validated by the health centres' management committees [9]. This method was chosen by the Ministry of Health based on evidence of its effectiveness [9, 10]. Committees of community representatives relied on their knowledge of the population and living conditions to select indigents based on locally accepted definitions: individuals who are extremely disadvantaged socially and economically, unable to look after themselves, and devoid of internal or external resources [9]. The definitions of indigence could be heterogeneous across communities because they were intended to be adapted to local realities. According to intervention reports, 15–20% of the population in the selected healthcare centres' catchment areas were supposed to receive indigent cards to access free healthcare services and medication [11, 12].

For the PBF component of this intervention, healthcare centres were paid a unit purchase price for each targeted service delivered (e.g. curative consultation for adults). Healthcare centres that met quality-related performance targets following verifications were also eligible to receive bonus payments. Quality scores of over 50% were used to inflate PBF payments. PBF payments were used to fund expenditures, increase bank reserves, and pay bonuses to employees of the healthcare centres [13].

The intervention described above is complex, given the number of interacting components, the number of groups and organizational levels targeted, and the number outcomes [14, 15]. Many global health actors are concerned that implementing such a complex intervention could produce unintended consequences that are outside the targeted objectives of the intervention [16, 17, 18]. These unintended consequences are defined as changes for which there is no purposeful action or causation and that occur in a social system as a result of adopting, adapting, or rejecting an innovation such as PBF [19]. These changes can be desirable or undesirable, depending on the stakeholders' perspectives. They can affect various actors, such as service users, providers, donors, community members, and government representatives.

To our knowledge, the intervention implemented in Burkina Faso presents a unique opportunity to develop scientific knowledge because no study has been conducted to date on the unintended consequences of combining PBF with equity measures for indigents in

Africa. Although program planners believe these approaches may have a synergistic potential, the combination may not work out as planned. Interaction between the different rationales, goals, and operating procedures may produce unintended consequences. Thus, our objective is to document the unintended consequences of equity measures integrated into the complex PBF intervention in Burkina Faso.

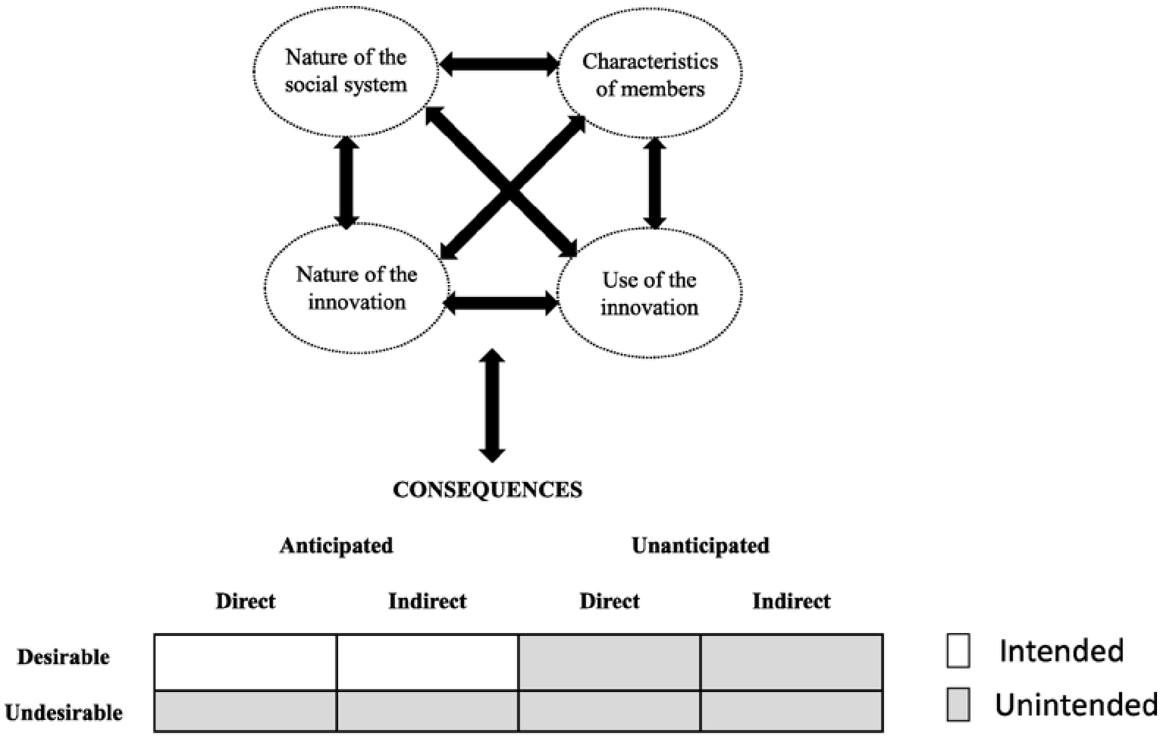
Methods

Theoretical framework

This study was based on Rogers' diffusion of innovations theory [19]; our aim was to focus on the intervention's adoption and adaptation from a broad perspective, in order to capture unintended consequences. While the theory provides an original approach to the study of PBF in a low-income setting, it has also been used in the past to analyze the consequences of health innovations [20, 21, 22]. According to the theory, combining PBF with health equity measures constitutes an innovation because both practices are perceived as new by adopters. The theory stipulates that diffusion of innovations usually widens the socioeconomic gap. However, when special efforts are made by a diffusion agency, it is possible to narrow or at least not to widen it.

To understand an innovation's diffusion process and consequences, we can examine four main dimensions: 1) the characteristics of the members of the social system (e.g. their knowledge and beliefs about the intervention, attitude towards change); 2) the nature of the social system (e.g. norms, culture, characteristics of the organization); 3) the nature of the innovation (e.g. relative advantage, compatibility, triability, complexity); and 4) the use of the innovation (e.g. its re-invention) [19]. These dimensions can interact to influence the emergence of various types of consequences. Rogers classified consequences as: 1) desirable or undesirable, 2) direct or indirect, and 3) anticipated or unanticipated. To operationalize these concepts, we considered desirable consequences to be those that are functional (positive) for the social system and undesirable consequences to be those that are dysfunctional (negative). A consequence could potentially be both desirable and undesirable, depending on the point of reference [21]. We considered consequences as anticipated if they were explicitly or implicitly addressed in the implementation guides. In accordance with Ash et al.'s [21] approach, we considered direct consequences to be related to processes and indirect consequences, to outcomes. Like Bloomrosen et al. [20], we considered that intended consequences tend to be those that are simultaneously desirable and anticipated. In contrast, unintended consequences tend to be those that are undesirable and/or unanticipated. Our rationale for these assumptions is that program planners generally intend to make changes they consider desirable and that they can anticipate. We also assume program planners do not purposefully target changes they consider undesirable or have not anticipated. We have shown the applicability of this typology elsewhere [23]. **Figure 11** illustrates our theoretical framework [23].

Figure 11. Theoretical framework



Adapted from Rogers (2003) and Bloomrosen et al. (2011)

Study setting

Burkina Faso is a low-income country where health needs are a major concern. The maternal mortality ratio is 371/100,000 live births [24]. The under-5 mortality rate is 89/100,000 live births [25]. In the country's National Health Development Plan for 2011–2020 [26], priority issues include: 1) poor performance of the health system, especially in terms of governance and service delivery; 2) lack of human resources; 3) inadequate quality and supply of health products such as medication and vaccines; 4) insufficient coverage and maintenance of infrastructure, equipment, and logistics; 5) poor health information systems management; and 6) inadequate funding for health and poor management of resources.

To address these issues, the government of Burkina Faso conducted a pre-pilot PBF test in 2011 in three districts [27]. In 2014, this intervention was modified to include the health equity measures described in the introduction. It was also expanded to an additional 12 districts. To conduct an impact evaluation, funded by the World Bank, four intervention modalities were implemented across 15 districts [8]: 1) PBF1: healthcare centres were paid fixed unit prices for activity indicators achieved; 2) PBF2: PBF1 coupled with a community-based selection of indigents to be exempted from user fees at point of service; services provided to indigents were purchased at a higher unit price than those provided to non-indigents to compensate healthcare centres for financial loss due to unpaid user fees; 3) PBF3: PBF2 with higher unit prices for services provided to indigents, to motivate healthcare workers to treat indigents and to better compensate healthcare centres for financial loss (see Additional file 1); and 4) PBF4: PBF1 linked with a community-based health insurance program and a community-based selection process for indigents. In this article, for reasons of feasibility, we focus on the PBF1 and PBF3 intervention modalities (see details in the section *Sampling of cases*).

The present study took place in a district of Burkina Faso where achieving equitable use of healthcare services remains a challenge. The district population was estimated at 135,740 in 2016, with more than 50% living in poverty [28]. Of the 19 primary healthcare centres in this district, five were allocated to PBF1, seven to PBF2, and seven to PBF3. Although PBF started in January 2014, cards to identify the selected indigents were only available for distribution in November 2015. Healthcare workers, however, were encouraged to begin applying user fee exemptions for indigents before then. Implementation guides describe the planned intervention model and the different actors supposed to be involved in the selection process [13, 29].

Research strategy

This research was nested within a larger longitudinal process evaluation of the intervention [8]. For the design, we conducted a contrasted multiple case study with several embedded levels of analyses [30]. The cases were four primary healthcare centres, called *Centres de santé et de promotion sociale* (CSPS – centres for health and social promotion).

Sampling of cases (facilities)

Case selection was done shortly after the intervention launch and followed a multistage screening procedure [30, 31]. First, we identified a district that represented the normal healthcare system context and was located in a relatively safe area for researchers. Within this district, we assessed the CSPSs' levels of performance on key activity indicators for maternal and child health. We ranked the CSPSs into quintiles to select centres with contrasting levels of performance. We then asked key informants (i.e., members of the district management teams) in each district to help us select facilities that were representative of their performance category and that offered opportunities for significant insight [30, 32, 33]. This dialogue with local informants helped us avoid selecting cases that were outliers or unrepresentative. For this analysis specifically, we decided to focus on facilities in the first and third intervention arms only. We selected the first intervention arm (PBF1) because it represents a common PBF model that is being widely implemented in low-income countries, thus increasing the pertinence of the results. We selected the third intervention arm (PBF3) because it is an innovative PBF model with health equity measures. The final set of cases consisted of two high- and low-performing PBF3 facilities and two high- and low-performing PBF1 facilities. The data collected in the PBF3 facilities were primarily used to understand the implementation and various changes related to the equity measures integrated within the PBF intervention, while the data collected in the PBF1 facilities were primarily used for triangulation purposes and to better understand the overall context, while avoiding over-attributing relevance to the equity measures. We did not include PBF2 facilities, as the targeting intervention was comparable and only unit prices differed. We also excluded PBF4 facilities because the intervention model combining insurance with PBF is radically different and rarely used in other countries, thereby limiting the utility of results. **Table XIII** describes each facility included.

Table XIII. Description of four cases included to study equity measures with PBF

Descriptors	Facility 1	Facility 2	Facility 3	Facility 4
Intervention arm	PBF3	PBF3	PBF1	PBF1
Initial performance	Low	High	Low	High
Type of facility	CSPS, public, not-for-profit	CSPS, public, not-for-profit	CSPS, public, not-for-profit	CSPS, public, not-for-profit
Healthcare workers	1 head nurse 2 itinerant health workers (IHW) 1 auxiliary midwife 1 IHW volunteer	1 head nurse 1 IHW 1 auxiliary midwife	1 head nurse 1 nurse 2 IHWs 1 midwife 1 auxiliary midwife 3 trainees (temporary)	1 head nurse 1 IHW 1 auxiliary midwife 4 trainees (temporary)
Support staff	1 drug depot manager 1 guard 1 janitor	1 drug depot manager 1 guard 1 janitor	1 drug depot manager 1 guard 1 janitor	1 drug depot manager 1 guard 2 janitors
Number of villages in catchment area	5	8	22	6

Population in catchment area	~ 8,000	~ 3,600	~ 11,000	~ 3,700
Easy access to paved road	No	No	Yes	Yes
Ethnic majority	Dagara	Lobi	Lobi	Birifor, Djan
Number of indigents selected (coverage rate)	829 (10,4%)	566 (15,7%)	0 (0%)	0 (0%)
Economic activities	Agriculture Livestock farming Production of local alcohol	Agriculture Livestock farming Production of local alcohol	Agriculture Livestock farming Production of local alcohol	Agriculture Livestock farming Production of local alcohol
Distinctive features	Gardening during dry period	High migration rate	High migration rate	

*Itinerant health workers are employees in charge of promoting health, hygiene, and vaccination, notably through household visits and community gatherings. In practice, they also deliver healthcare services due to the shortage of healthcare workers.

Data collection method

We collected qualitative data during two sequential phases, with the first informing the methods used for the second. For the first phase, the first author conducted 3 months of fieldwork between January and April 2016. The researcher’s immersion in the milieu provided a better understanding of the context and helped create a relationship of trust with stakeholders. We visited each healthcare facility for a two-week period to conduct semi-structured interviews, informal discussions, and non-participant observation. Participants included a wide range of stakeholders, such as indigents, non-indigents, members of indigent selection committees, representatives from the *Comité de gestion* (COGES – healthcare facility management committee), community-based health workers (CHWs), healthcare workers, and patients. Participants were purposefully selected based on their ability to provide relevant information and their accessibility. Then, following the snowball approach, some key informants referred us to other potential participants who could shed light on the intervention. Using these approaches, we followed the diversification principle to select participants with a variety of intrinsic characteristics, such as different indigent statuses, occupations, and genders [33]. For the interviews, we constructed guides that drew on previous questionnaires used for research on the diffusion of innovations [34, 35]. We systematically recorded field notes on observations and informal discussions in research diaries. Observation sites included healthcare facilities, villages, and other social settings (both public and private). The first author also participated in a six-day annual PBF review meeting at the national level to triangulate data regarding unintended consequences, better understand the different contexts, and assess the potential transferability of results to other facilities in intervention districts.

For the second phase, the third author conducted 20 days of fieldwork in May 2016 to deepen our assessment of the relations between community verifications and equity measures for indigents. He conducted semi-structured interviews, informal discussions, and non-participant observation in each of the four facilities. To provide complementary data, he conducted an additional interview in December 2016 with a key stakeholder involved in indigent selection. The same procedure was used to select participants as described above. He recorded field notes in research diaries.

In total, we conducted 93 semi-structured interviews and recorded 241 observation sessions in research diaries. Applying the principle of saturation, we stopped collecting data when interviews and observations no longer provided information that was sufficiently different to justify continuing. Research team members produced verbatim transcriptions of interview recordings. **Table XIV** provides a breakdown of the qualitative data collected and analyzed. It should also be noted that the last author has in-depth understanding of the context, having participated in workshops to define the intervention process for the equity measures and taken part in follow-up meetings on this topic.

Table XIV. Summary of data collected to study equity measures with PBF

	Quantity
Non-participant observation	
Sessions reported in field notes	241
Interviews	
At facility level	
Healthcare providers	15
Other support staff (drug depot manager, janitor, security guard)	13
Volunteers & trainees	7
Community leaders (e.g., COGES, selection committees & community health workers)	23
Service users (e.g. patients, indigents)	18
At district level	
Administrative staff (e.g. manager, accountant, data collection agent/photographer)	4
Members of contractualization and verification agency	4
Members of local association conducting community verifications	7
At national level	
Representative from the *Programme d'appui au développement en santé* (PADS – program to support health development)	1
Representative from the *Service technique – financement basé sur les résultats* (ST-FBR – results-based financing – technical service)	1
Total semi-structured interviews	93

We also used secondary data on healthcare services delivery that are publicly available on the Ministry of Health's PBF portal (www.fbrburkina.org). These longitudinal data are collected monthly in each healthcare centre for PBF verifications. Healthcare workers report the quantity of healthcare services delivered to indigent and non-indigent patients, based on the medical registers. Then PBF officers verify the reported data by manually recounting the quantity of services. They enter the data into an electronic platform. We used the data collected between October 2015 and September 2016, that is, before and after fee-exemption cards were distributed to indigents starting in November 2015. The main sample for the quantitative component consisted of the two facilities with equity measures (PBF3) included in the qualitative phase. To assess the transferability of the findings across the study district, however, we examined all seven facilities within the district that were assigned to the same intervention arm as the two selected for inclusion in the qualitative component (PBF3) and for which data were available. To assess the transferability of findings more widely, we also examined all 196 facilities in the intervention districts that belonged to intervention arms with similar measures for indigents (PBF2 and PBF3) and for which data were available.

Data analyses

The primary unit of analysis was the healthcare facilities and their catchment areas. We combined deductive and inductive thematic analysis [36, 37]. We began by developing a template of themes based on our theoretical framework. Then we carefully read the transcripts and field notes to assign the raw data to the predefined themes. At the same time, we derived new themes that were not included in the initial template but that emerged from the data and were judged relevant to our research topic. In some cases, we narrowed down and provided more focus to the initially defined themes to enhance their applicability to the data. We used QDA Miner 4 to code and retrieve text segments.

We also used descriptive statistics to examine how the quantity of services provided to indigents evolved over time, compared to those provided to non-indigents. We used Excel to create graphs and conducted a visual analysis to highlight patterns that emerged over time [38]. This complementarity information was used to triangulate some of the findings.

To classify the various unintended consequences, we followed a procedure previously developed and applied [23]. During the data analysis, we classified the different types of consequences based on the definitions of anticipated/unanticipated, desirable/undesirable, and direct/indirect presented above. To determine whether a consequence was anticipated or unanticipated by program planners, we reviewed intervention documents (e.g. guides, midterm reports) to better understand the design of the intervention model and its implementation. The document review enabled us to compare the program planners' intended processes and outcomes to what actually emerged in real life. The titles of the documents reviewed are available in the references [12, 13, 29, 39]. In addition, we classified consequences as desirable or undesirable depending on whether we considered these changes to be functional (positive) or dysfunctional (negative) for the social system. Lastly, we classified consequences as direct or indirect depending on whether we considered these changes to be related to processes or outcomes.

We used a cross-case synthesis to draw general conclusions [30]. Following a replication logic, we considered that results arising independently from more than one facility are more powerful than those coming from a single facility, and thus gave the former more importance in the results section [30].

Results

The results showed that community-based selection of indigents for user fee exemptions within a PBF program led to unintended consequences. **Table XV** summarizes the results.

Table XV. Classification of unintended consequences of combining equity measures with PBF

	Anticipated		Unanticipated	
	Direct (Process)	Indirect (Outcome)	Direct (Process)	Indirect (Outcome)
Desirable	■	■		➢ Increased awareness regarding health equity within the community ➢ Sense of empowerment among indigents to access healthcare
Undesirable	➢ Selection of individuals who did not meet the local conceptualization of indigents to the detriment of others who did ➢ Ossification	➢ Dissatisfaction regarding the selection of indigents	➢ Financial difficulties and drug shortages ➢ Withholding of indigent cards ➢ Capping the cost of medications ➢ Failure to follow the percentage cap of indigents covered monthly ➢ Triage of indigents during consultation ➢ Fixation on quality indicators	➢ Conflicts between indigents and healthcare system actors ➢ Uncertain and unequal coverage for indigents

■ Intended
☐ Unintended

Additional file 2 specifies how the anticipated consequences were explicitly addressed in the intervention guides.

The subsections below describe in detail how interactions between the nature and use of the intervention's different components (i.e., indigent selection, user fee exemption measures, and pro-poor purchase prices), the actors' characteristics, and the nature of the social system led to the emergence of a wide array of unintended consequences.

Selection of individuals who did not meet the local conceptualization of indigents to the detriment of others who did

Approximately 10 and 15% of the population were selected as indigents within the catchment areas of the two facilities in PBF3. Study participants and stakeholders who attended the annual PBF review meeting strongly affirmed that a portion of people selected as indigents did not meet the local conceptualization of indigents, that is, they were not individuals with no means to support themselves and not receiving assistance, such as widows, elders without children, handicapped persons, or orphans. Based on their knowledge of the communities and living conditions, some participants claimed that many individuals who were selected and obtained cards were not indigents. During interviews, some of these selected 'indigents' openly recognized that they did not truly qualify as such. These individuals received a card despite having income-generating activities, social support, ability to work, relatively high social status, belongings, etc. Examples of indigent card holders encountered during this study included the mother of the president of the COGES, a security guard and a janitor of a healthcare centre, a shop owner, a village chief, and a member of the village development committee. The latter benefited from the indigent card to the detriment of other unselected individuals who were considered worst-off.

> *"They didn't identify those who should have been…. Some people were selected, and others said [about them], 'no, that person is working and has means!'" (Community leader_50, interview, facility 2)*
>
> *"There are some indigents who do not have a card because it is not the real indigents who were selected." (Healthcare worker_16, interview, facility 1).*

Numerous factors contributed to the selection process drift. First, study participants revealed that part of the selection was based on personal gain, affinity, social relations, and social status. For example, numerous CHWs and village councillors who sat on indigent selection committees obtained indigent cards for themselves or their immediate family members (see Additional file 3).

> *"The treasurer [of the CSPS] is an indigent. Is that normal? ...she works and has support. The old community health worker also has his indigent card. They wanted to pull a scam and play politics." (Healthcare worker cited in field notes, facility 2).*
>
> *"... this is a situation in which acquaintances and relations were used to distribute the indigent cards." (Healthcare worker_23, interview, facility 1)*

Some CHWs in PBF3 facilities justified obtaining indigent cards for themselves or their family by arguing they had been doing volunteer work for the community without sufficient compensation. At the same time, some villagers and healthcare providers accused selection committee members of using the selection process to gain political influence for local elections. Others believed the high *relative advantage* of possessing a card played a role in the selection of individuals with questionable indigent status:

> *"Because they're saying everything will be free, everyone wants to be on the list." (Photographer for indigent cards_39, interview, across facilities)*

Another important factor contributing to the selection of individuals not locally perceived as indigents was the confusion and misunderstanding regarding the number of indigents to select. After the selection of indigents had been completed, supervisors asked the committees to increase the numbers of indigents to reach a targeted number per village. As such, in both PBF3 facilities, a second selection was conducted, and people who did not meet the local conceptualization of indigents, including CHWs with revenues who sat on selection committees, were added to the lists.

> *"We were identifying indigents and not reaching the [targeted] number. We were tired, and we just had to get it done.... [So] each one of us doing the selection decided to register himself...." (CHW_15, interview, facility 1)*
>
> *"...they told us to stop because there were problems with the numbers in the register. We had to add, then we had to remove. At the same time, they told us to*

> *stop, so there were problems between the supervisors...." (Photographer for indigent cards_39, interview, across facilities)*

Moreover, the 'photographers'—workers assigned to take indigents' photos for the identification cards and record their GPS location—arrived unannounced in the villages to conduct their work. Not having been informed, some indigents had left the village with their families—for example, to cultivate, or to attend funerals—so the photographers were not able to take their photos. So, to reach the targeted number of indigents, the 'photographers' and CHWs in the first healthcare centre quickly replaced some of the absent indigents with other villagers available that day.

> *"One day, we were all surprised to see the team with the photographer arrive in the village to take the pictures of the selected indigents. Because no one knew they were coming, some of the people selected as indigents were absent... I didn't want to leave a void, so I simply replaced the people who were absent with others. When these people came back, they complained. I told them that I replaced them because they were not there and that it is not my fault because [the photographer] came without informing us in advance." (CHW_27, interview, facility 1)*

> *"The day of the selection, we went to his place and didn't see him. So, we said [in the village], we need at least 200 people. So they had to just take whoever they found because the decision-makers were pressuring us." (Healthcare worker_16, interview, facility 1)*

One 'photographer' reported that the remuneration modality, which was based on performance, also contributed to selecting individuals not on the initial list of indigents. The data collection agents were reportedly paid about 320 CFA francs (0.57 USD) for each indigent identified.

> *"The clever ones, you'll notice, started taking [photos of] all of the children who were at home to facilitate their work... It's a strategy they made up." (Photographer_66, interview, across facilities)*

Some selection committee members argued that the conditions under which the selection was conducted affected the quality of their work, especially due to the *complexity* of the task. Some participants noted, for example, that individuals doing the selection were not sufficiently trained, that the communication system was deficient, that not enough time was provided for the selection, and that they received no financial compensation for their hard work. Participants also revealed deficiencies within the committees involved in the selection process. For example, one CHW stated that he conducted the selection of indigents alone in his village. Meanwhile, in another centre, two members of the selection committee at the facility level revealed that they had not seen the final list of indigents, and one was unaware that indigent cards had been distributed in the catchment area during this study, claiming that *"the bureau didn't do its job."* Moreover, a midterm report [12] confirmed the committees that were initially supposed to be in charge of validating the lists of indigents (referred to as the local validation groups) were not implemented: *"...this structure was never created in the villages, given its relevance to realities on the ground. The main observation was that the community leaders held multiple responsibilities. Thus, the people who were part of the indigent selection committees were mostly the same people who were in the local validation groups"* (p. 15). Although these obstacles relate to the implementation process, they help explain the context in which gaming occurred for the selection of indigents.

Our observations and interviews suggested that selecting indigents based on personal affinity and personal gain was consistent with the broader social system and local stakeholders' past experience. Study participants reported that relationships and informal networks are important for survival and prosperity, especially in a context of widespread poverty. They spoke often about the high rate of corruption within and outside the healthcare sector. As one participant described, malfeasance is not uncommon in new projects implemented by international organizations with limited funding and timeframes.

> *"I see projects that come to the village, and the chief is asked to bring forward the indigents. Everyone gathers up their own family, even if they're able to cover their own care." (Patient_10, interview, facility 1)*
>
> *"The country is corrupt! Here, everything depends on relationships." (Student midwife cited in field notes, across cases)*

Ossification

According to a midterm report [12], consideration was given to setting up a system to update indigent lists after the initial selection: *"This approach makes it possible to regularly update the list of indigent persons selected"* (p. 7). However, no update mechanism had been implemented at the time of this study. Thus, indigents who were absent when the photographer came to their village or people who fell into poverty after the selection were unable to obtain an indigent card. After the photos were taken, selection committees were unable to modify indigent lists. Many study participants did not know how long indigent cards were valid, and some believed changes would not be possible for the next 3 years. Indigent cards with identification errors could not be corrected, as they were manufactured in Vietnam. The selection process had a low level of *adaptability*, that is, stakeholders did not formally have the opportunity to make modifications according to their needs and constraints over time. Thus, the intervention led to a certain level of ossification, that is, organizational paralysis brought about by a rigid system and the presence of a centralized decision-making structure, as illustrated by the following citations:

> *"Because they [decision-makers] say we can only review this in three years, we'll go along with it to see what happens over the next three years and how they'll select the indigents next time.... We'll bear with it and keep advocating to see whether they can shorten that three-year period." (Healthcare worker_17, interview, facility 1)*
>
> *"If PBF [officials] don't come back, how can we get that card for him? It's a problem." (COGES_60, interview, facility 2)*
>
> *"We don't know how we'll get through this." (Head nurse cited in field notes, facility 1)*

Dissatisfaction regarding selection of indigents

In both facilities, study participants reported that the selection process led to frustrations, conflicts between actors, and a sense of injustice. Indigents omitted from the selection or absent when the photographer came demanded that the situation be rectified. Some individuals demanded to be selected as indigents due to the *relative advantage* of having free healthcare services, the perceived inequity of the selection process, and the lack of understanding regarding the definition of 'indigents'. *"Why hasn't anyone from my household been selected? Not a single person!? How is it that some benefit and others don't?"* asked one member of the committee in charge of coordinating the selection at the facility level during a heated COGES meeting (facility 2). To appease these types of frustrations, selection committee members sometimes made false promises to the population, made apologies and distanced themselves from the selection process, arguing that it was the 'community' that chose the indigents.

> *"…if I'd known, I wouldn't even have gotten involved in this work. It caused us a lot of problems. In fact, every morning people would come to my house to ask whether a new list had opened up so they could register. This bothered me a lot. Also, it made me uncomfortable when some people scowled and got angry." (CHW_14, interview, facility 1)*
>
> *"People are envious. Some people want to really force their way into getting a spot, but it's not for them." (COGES_59, interview, facility 2)*

Despite these complaints, community members generally remained in favour of user fee exemptions for indigents.

> *"In any case, the villagers said it's a good project for the whole village." (CHW_27, interview, facility 1)*
>
> *"The people actually appreciated the idea of covering indigents. They even said that, if it were to really happen…then everyone will start to believe in 'the white man's paper'." (Volunteer IHW_11, interview, facility 1)*

Increased awareness regarding health equity with the community

The intervention triggered discussions and reflections within the community on health equity and the issue of indigence. For example, community members not selected as indigents engaged in discussions with healthcare workers and selection committee members to better understand the selection process and the reasons for their exclusion. This provided opportunities to explain the concept of indigence and the importance of providing access to services to the most vulnerable individuals.

> *"In the community, some have welcomed it. Then there are others who say, no, if that's how it is, then everyone is an indigent, even though they're not indigents. So we explain often… it's just to help the poorest…. Some understand, but others don't." (Healthcare staff_17, interview, facility 1)*

> *"Some said, the entire village is made up of indigents, so we should select everyone. We said, no, it's not like that. We explained to those people that there are selection criteria. We have to select the old widows who have no support, people with no support. Those are the people we chose." (CHW_27, interview, facility 1)*

Withholding of indigent cards

A major concern for study participants in the second facility was that some indigent cards were missing and never distributed to their owners. Healthcare workers and CHWs put the blame for these missing cards on the 'photographers' and technical difficulties with the equipment used to identify and photograph indigents (i.e., digital tablets). However, observation revealed that a head nurse—who did not approve of the selection of certain indigents and was concerned this process would negatively influence the medications stock—had surreptitiously removed some indigent cards before their distribution in the community. A district supervisor reported that this strategy had been used in other healthcare centres and recommended this approach to healthcare workers in the first facility to lower the number of indigents and limit the healthcare centres' financial difficulties (as discussed in the next subsection).

> *"Some head nurses filtered the cards, and when people ask for them, say they haven't arrived. They say that every time. You just had to do the same thing." (Supervisor cited in field notes, facility 1)*

Financial difficulties and drug shortages

Healthcare centres in PBF3 received higher unit purchase prices for some targeted services provided to indigents (see Additional file 1). For example, in the first facility, a consultation for an indigent adult was purchased at 1020 F CFA (1.72 USD), and for a non-indigent adult, 140 F CFA (0.24 USD). In exchange for these subsidies, healthcare centres were required to provide free services and free medications to indigents. If the cost of the medication prescribed was higher than the lump sum provided through the unit purchase price, the COGES had to absorb the difference using their other sources of revenues (user fees and sales of medication to non-indigent patients). If the cost was lower, the COGES retained the profit.

> *"A district supervisor said, 'It's not just the white man's money. The COGES also has to contribute to the indigents' medications.' In response, the healthcare workers shook their heads in disapproval." (Field notes, facility 1)*

Numerous participants, including healthcare workers and COGES members, argued that delays in PBF payments caused financial difficulties for healthcare centres and led to drug shortages. At the time of the study, these delays were more than 6 months for quantity-related payments and more than 1 year for quality-related payments. Participants complained that, without the revenues from medications provided to indigents, it was difficult to replenish the centres' drug depots. Some participants feared this would lower the quality of care for patients, who would have to obtain their medications elsewhere.

> *"We have to wait for PBF to come pay for the products the indigents used before placing another order. I find it a bit difficult." (Healthcare worker_51, interview, facility 2)*

There was also consensus among healthcare workers and COGES members in the first healthcare centre that the unit purchase prices for services to indigent patients were insufficient to cover the cost of their medications and that the healthcare centres did not have enough *slack in resources* to ensure proper functioning of the user fee exemption for indigents. Participants believed the insufficiency of compensation was causing financial difficulties and could lead to drug shortages in the healthcare centre.

> *"If we stubbornly continue to treat people with prescriptions costing up to 3,000 francs and the system only pays 800 francs, who loses in that case? It's the COGES that will suffer, and over time, we risk not even having products here at the depot…. Ultimately the healthcare facility could be at risk for closure. People will prefer to*

> *consult where they can find the products." (Healthcare worker_11, interview, facility 1)*

According to the intervention guide [13], the purchase prices were intended to "*encourage healthcare workers*" to provide services to the poor. In practice, however, the financial incentive was perceived as insufficient to trigger proactive strategies on their part. For many healthcare workers in the first facility (PBF3), the *relative advantage* of providing user fee exemptions to indigents was mitigated by the fact that the healthcare centre lost money when the value of the medication provided for free was higher than the unit purchase price. Consequently, no additional efforts or innovative strategies were deployed to provide more services to indigents specifically, as explained by this healthcare worker:

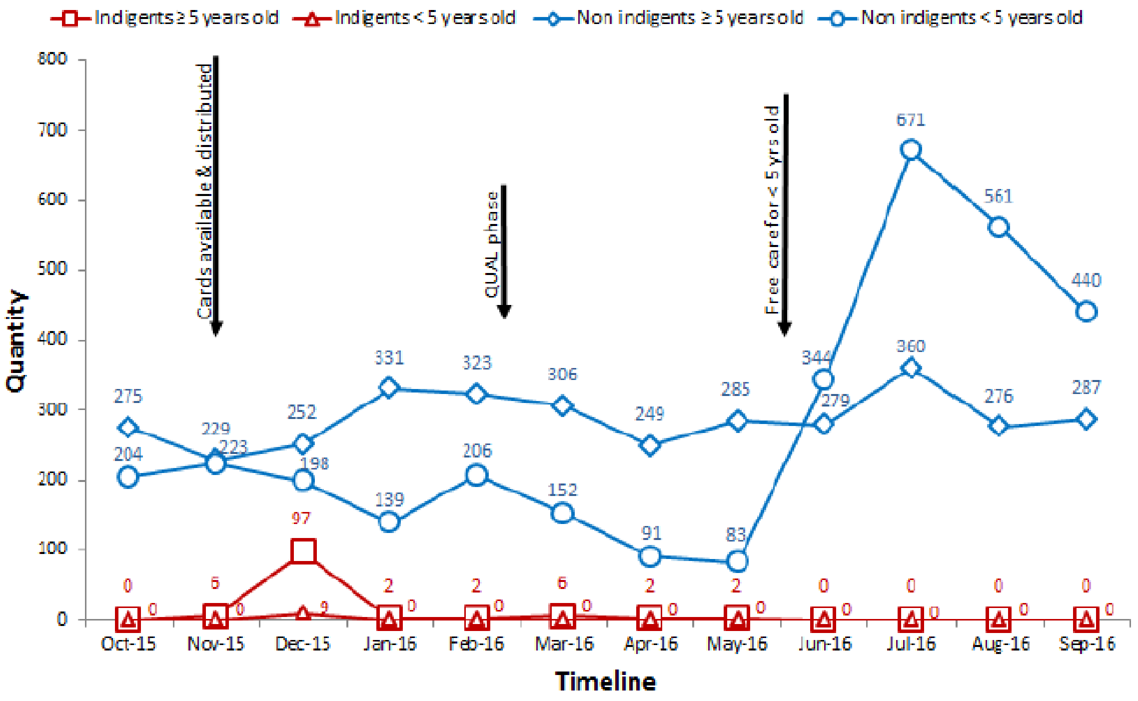
> *"We didn't think of doing that. When an indigent person comes in, we treat him, and that's all…. We know that with this [intervention], sometimes we make money, and sometimes we lose." (Healthcare worker_17, interview, facility 1)*

Healthcare workers from facilities without equity measures for indigents (PBF1) also expressed lack of support for intervention models that provide user fee exemptions for indigents, for fear that those caused financial difficulties.

Multiple strategies adopted to limit services to indigents

Qualitative data showed that, shortly after the distribution of indigent cards, healthcare workers in the first facility (PBF3) adopted a series of strategic measures to limit the services and medications provided for free to individuals with indigent cards (as described in the subsections below). Secondary data on the quantity of services provided to indigent patients before and after indigent cards were distributed were consistent with these findings. **Figure 12** and **Figure 13** show that, in both facilities with indigent targeting, the number of new consultations for patients classified as indigents increased considerably after indigent cards became available in November 2015. However, the following months saw rapid declines in the number of new consultations for patients classified as indigents. Since these declines are unlikely to have been due to sudden changes in morbidity prevalence or to the rapid cure of all indigents, these data support the findings that healthcare workers limited free services delivered to indigents. This is relatively consistent with the evolution of care in other healthcare centres belonging to the same intervention arm (PBF3) within the study district (**Figure 14**), supporting the transferability of findings.

Figure 12. Facility 1 - Total number of new consultations for patients classified as indigents or non-indigents seen in curative care



Note: As shown by the qualitative data, the curves representing non-indigents are likely to include individuals who should have received user fee exemptions but were requested to pay, either because they did not possess an indigent card or because healthcare workers refused to recognize their indigent status. Similarly, the curves representing indigents may include individuals who received indigent cards even though they did not truly meet the local conceptualization of indigents

Figure 13. Facility 2 - Total number of new consultations for patients classified as indigents or non-indigents seen in curative care

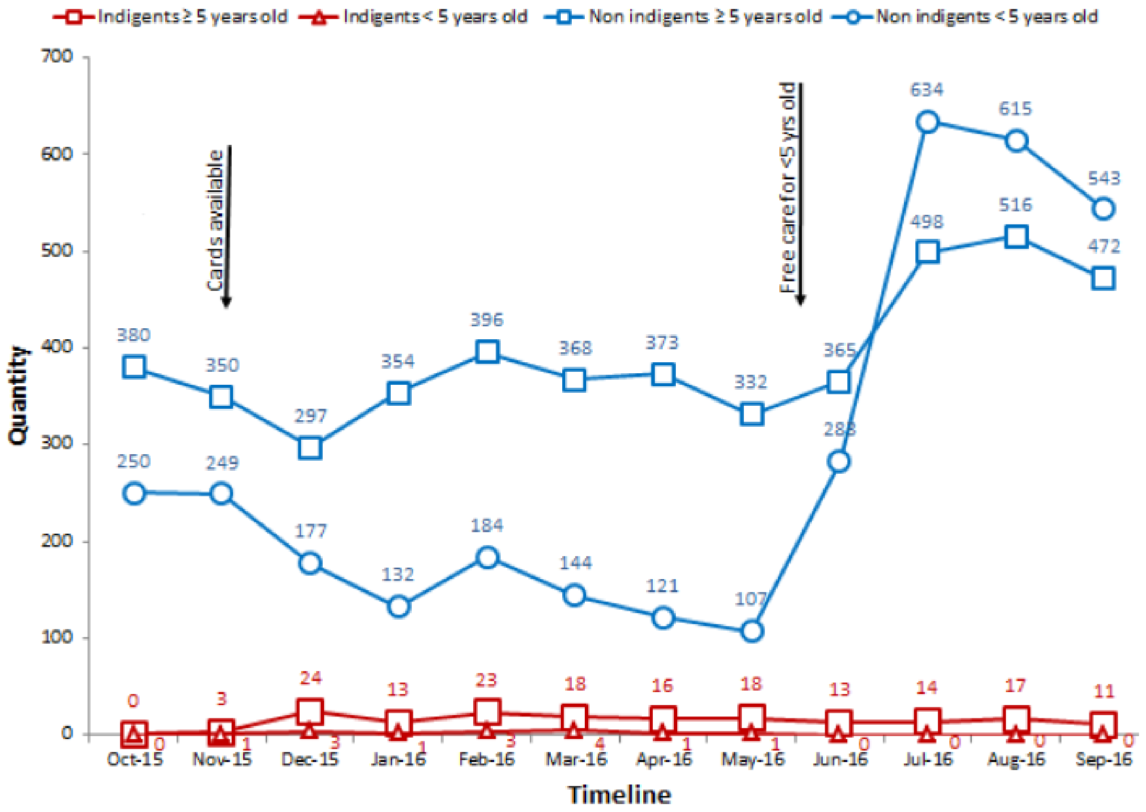
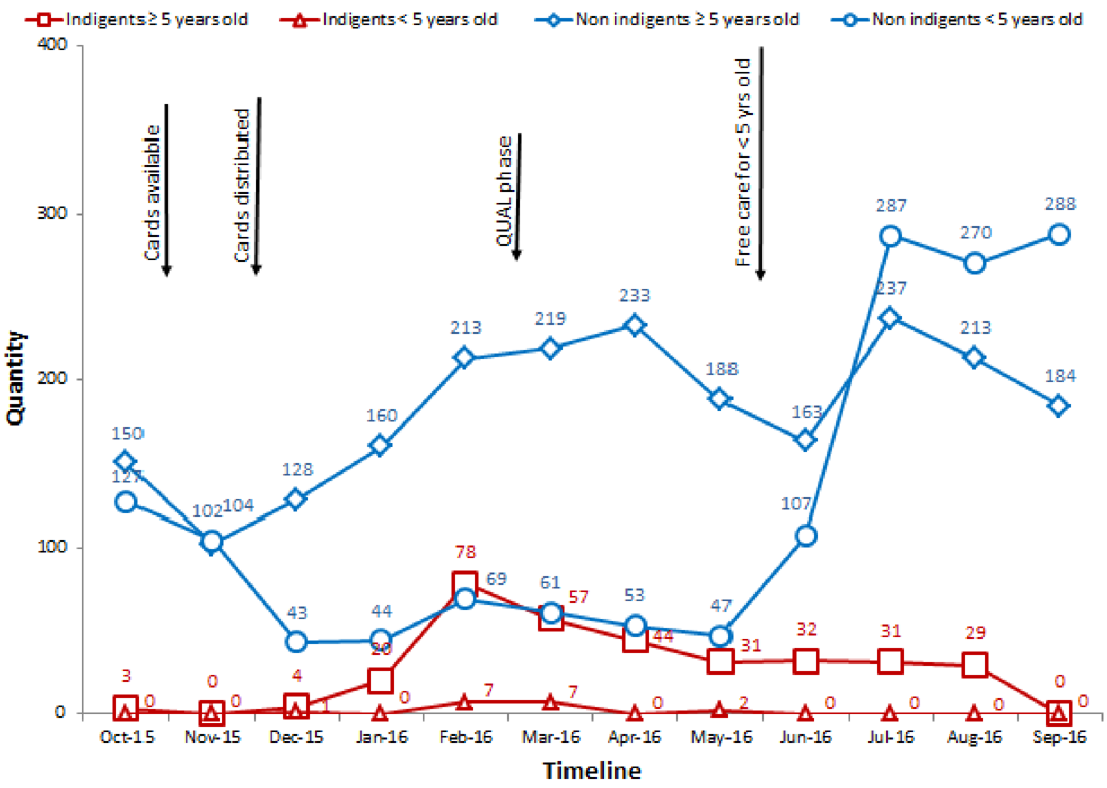


Figure 14. Average number of new consultations for patients classified as indigents or non-indigents seen in curative care in PBF3 facilities (n = 7) within the study district



Note: Healthcare centres (*n* = 7) for which the intervention arm was not specified in the database were excluded. PBF ceased to fund services delivered to indigent children under five once the free healthcare policy began in June 2016 because it covered their user fees and medications

Moreover, the number of new consultations for patients classified as indigents did not follow seasonal patterns. Rainfalls generally influence the spread of diseases such as malaria, causing an increase in use of care between June and August. As expected from seasonal patterns, the number of new consultations for adult patients not classified as indigents peaked in July 2016. However, the number of new consultations for adult patients classified as indigents during this period actually followed the opposite pattern and gradually decreased or remained null. There is no reason to believe indigent adults would not be similarly affected by seasonal patterns. These secondary data also support the finding suggesting healthcare workers limited the delivery of free services to indigents.

Within the study district (**Figure 14**) and across intervention districts (Additional file 4), we also observed that the average number of new consultations for adult patients classified as indigents did not follow seasonal patterns.

The decline in new consultations for patients classified as indigents was more gradual for facility 2 than facility 1. This observation is consistent with the findings. In comparing the two facilities, we found that healthcare workers in facility 2 displayed a weaker understanding of PBF payment modalities for services and medications provided to indigents. They erroneously believed the PBF intervention would reimburse the entire cost of medications provided to indigents in addition to paying a unit purchase price for each consultation. Thus, at the time of the qualitative data collection, we found that, even though some indigent cards had been retained and there were delays in distributing cards (as shown in **Figure 13**), healthcare workers in facility 2 delivered healthcare services to indigents. They expressed less disapproval of the indigent component of the intervention compared to workers in facility 1, whose disapproval was relatively high. Over time, however, the patterns in the number of new consultations for patients classified as indigents gradually moved in the same direction in both facilities.

It should be noted, however, that the peak in new consultations for non-indigent children coincided with the implementation of a new national policy for free healthcare to children under five in June 2016. At that time, the PBF intervention stopped purchasing services delivered to indigent children because children's medications became covered by the national policy.

Capping the value of medication prescribed

To limit financial difficulties and protect the drug depot, healthcare workers and COGES members in facility 1 limited the prescribing of medications for indigents. They tried to keep the costs of those prescriptions under the PBF's lump sum purchase prices. Interviews and examination of the indigent registry confirmed that prescriptions for indigents tended to cost around 1000 F CFA. Some supervisors and healthcare workers were concerned this practice was not rational and could have negative consequences on treatment effectiveness, healthcare system efficiency, and patients' health.

> *"If someone [indigent] comes in with simple malaria, we've been told we're not supposed to exceed 850 F CFA for a prescription. ACTs for adults are 300 francs, the consultation is 200 francs, two paracetamol tablets are at least 150 francs. That's 650 francs. At 850 francs, they say the healthcare centre makes a profit, but you've used gloves to examine the patient! Who covers the price of the gloves? What does the healthcare centre gain? Nothing! And, for example, if someone comes in with malaria plus pneumonia, whether you like it or not, the prescription costs more because you have to give an antibiotic, Amoxine, at least three tablets, plus Carbotoux [cough syrup], which goes for around 650 francs. Already that doubles or even triples their 850 francs. And if you don't do that, the sick person will come back!... So they have to either increase the coverage or suspend their indigent business.... Now that they've imposed this on us, we're obliged to do what they want." (Healthcare worker_16, interview, facility 1)*

Failure to follow the percentage cap on indigents covered monthly

According to the intervention guidelines [13], free consultations for indigents should not constitute more than 10% of the total quantity of consultations to *"avoid the moral hazard"* (p. 53). Both the qualitative and quantitative data (see **Figure 12** and **Figure 13**) showed this policy was not being systematically applied in either of the healthcare centres with user free exemptions. In facility 1, for example, 19% of curative consultations in December 2015 were provided to indigents. This proportion fell to under 1% in January 2016, when measures were taken to limit free consultations to indigents. Although healthcare workers in both centres knew about the percentage cap, there was misunderstanding regarding the correct percentage of patients that could be treated for free as indigents each month. Some participants also disagreed with applying a percentage cap.

> *"At one point they [supervisors] had given us a monthly target rate. We exceeded it, and the indigents kept coming. We tried telling them we had to stop [for the month] and start again later, but they [indigents] didn't accept that! They said I didn't want to give the products for free." (Drug depot manager_22, interview, facility 1)*

> *"…if you reach the 40th person, will you tell the others not to come?! Ah, no!" (Healthcare worker_16, interview, facility 1)*

Triage on indigents during consultations

Fearing financial difficulties due to the user fee exemption, the COGES of facility 1 requested that healthcare workers triage patients during consultations, then provide free services only to those they believed were 'true' indigents and require 'false' indigents to pay. The healthcare workers' triage was based on their knowledge and perceptions of patients' current socio-economic situation. COGES members and healthcare workers were confident they could accurately identify genuine indigents.

> *"…we told them we would stop the system and verify for ourselves who the true indigents are. Currently, when an elderly person comes in and we see he doesn't even have enough to pay for products, we qualify him as indigent. A blind person is an indigent, as is someone who lost their children and is alone without support. We take these people as indigents, and we make sure the prescription doesn't exceed 800 francs." (Healthcare worker_11, interview, facility 1)*
>
> *"When healthcare workers take the cards from indigents, they ask them certain questions…. like, does he have anyone who can give him a hand and help him with his expenses? Questions like that." (COGES_14, interview, facility 1)*

In contrast, a participant from facility 2 explained that they did not conduct any triage during the consultation because that would cause too much conflict with the local population, who traditionally are known to be warriors: *"the healthcare worker wouldn't be able to work here anymore!"* This helps explain the more gradual decrease in the quantity of curative care to indigents in **Figure 13**

Uncertain and unequal coverage for indigents

Both observation and interviews suggested that the selected indigents in facility 1 did not know in advance whether their healthcare would be free of charge. Upon consultation, some indigents had to decide whether to pay for the services they needed or leave without treatment. A number of factors influenced indigents' access to free services, such as the healthcare workers' triage of 'true' and 'false' indigents, the monthly percentage cap on indigents, the cap on the value of medications prescribed, sudden interruptions of the user fee exemption due to financial difficulties, indigents' reactions to these measures, etc.

> *"The first time, it was free, the second time it was free again, but the third time they told me to pay…. Ah, really, it discouraged me... If I don't have money, I won't come back [to the CSPS]. Now I know it's not free." (Indigent patient_20, interview, facility 1)*

> *"There was an indigent man one time who went to the healthcare centre, and even though he was an indigent with a card, he paid a certain sum of money." (CHW_26, interview, facility 1)*

Fixation on PBF quality indicators

To promote orderliness, PBF evaluators deducted performance points if information in the registers was erased or crossed out. In both facilities with user fee exemptions for indigents, the staffs' fixation on such performance indicators occasionally prevented indigents from receiving free care. For example, indigents who were accidentally listed in the wrong register were required to pay for services, as mistakes could not be erased or scratched out without risking losing PBF points.

> *"It often happens that people have nicknames. If someone gives a name that isn't on the indigent card, we'll tell him he has to come back another day, because PBF doesn't like it when we cross things out or erase things." (Drug depot manager_22, interview, facility 1)*

> *"When I arrived, I didn't present the [indigent] card right away and they recorded my information in the register. After I showed them the card, they said I still had to pay for the medication, and I paid." (Indigent_18, interview, facility 1)*

Conflicts between indigents and health system actors regarding user fee exemptions

In facility 1, indigents expressed great dissatisfaction and lack of trust regarding healthcare workers and selection committee members because of the strategies used to limit their access to free healthcare and medications; indigents accused them of cheating and scamming. This experience was discouraging for some indigents.

> *"They don't understand why they were promised free healthcare services through these indigent cards and then later told they had to pay for these services. So they said it's the healthcare workers who are playing politics on them…. many people came here to complain, saying that I had told them that with the card they would have full and free healthcare services and that the products prescribed to them would also be free, and yet that's not the case at all."(COGES_14, interview, facility 1)*

Sense of empowerment for indigents to access healthcare

Many participants in both facilities argued that the user fee exemptions initially increased the selected indigents' sense of empowerment to access healthcare services. It facilitated their decisions and actions to seek healthcare services more quickly. This finding was consistent with healthcare workers' reports and the quantitative data indicating that the user fee exemption policy triggered a steep rise in attendance at health centres, at least until the services were curtailed.

> *"It's better because the [decision to seek] healthcare is in the hands of the indigent person. Under the previous system, the indigent was objectified. The person's relatives decided everything." (COGES_59, interview, facility 2)*

> *"Their morale improved and they became brave…. All the old sick people who had been hiding came out." (COGES_21, interview, facility 1)*

However, this initial sense of empowerment did not always translate into greater access to free healthcare services over time, due to the curtailing of services described above.

Discussion

As postulated by the diffusion of innovations theory, we found that the nature and use of the intervention interacted with the social system and the characteristics of the different actors to trigger unintended consequences. One of the main findings was that different types of actors deliberately re-invented elements of the intervention over which they had control to strategically increase its relative advantage and cope with implementation challenges, thereby triggering unintended consequences. More specifically, many selection committee members partly re-invented the selection process to benefit personally from access to free healthcare services. In contrast, for many healthcare workers and COGES members, the relative advantage of providing free healthcare services and medications to indigents was insufficient due to the perceived low unit purchase prices for services to indigents (as conceived in the initial intervention model), the late payments (the implementation of the innovation), and the healthcare centre's financial constraints (the nature of the local context). Healthcare workers deliberately modified the intervention model in different ways to make it more compatible with local resources and their own needs by retaining indigent cards, capping the value of medications provided, triaging patients into 'true' and 'false' indigent categories, etc. Although such re-invention was perceived as desirable by some local actors, it can also threaten the theoretical basis and equity implications of the intervention depending on the nature of the modifications to essential components [40].

Application of the theory

Rogers suggests that one way to better understand the consequences of innovations is to classify them in a taxonomy [19]. Program evaluators and researchers tend to focus on certain types of consequences (e.g. desirable and anticipated) while neglecting others (e.g. undesirable and unanticipated) [41]. Thus, conceiving an inclusive typology *ex ante* compels stakeholders to consider the possibility that interventions can produce consequences that are not intended. In this study, we found Roger's classification useful for conceptualizing different types of consequences, broadening our focus beyond intended consequences during data collection, and organizing the presentation of results.

One challenge we encountered, however, was in determining whether consequences were anticipated or unanticipated, since this could vary depending on the perspectives of the different types of stakeholders (e.g. researchers, policy makers, healthcare workers). As described in the theoretical framework, we classified consequences as anticipated if they were addressed in the intervention's implementation guides (see Additional file 2). However, guides were sometimes unclear and imprecise regarding anticipated consequences outside the targeted objectives. Change agents hired to develop intervention guidelines do not always have a comprehensive understanding of the scientific knowledge and do not always openly disclose undesirable consequence that could undermine intervention models. Thus, while the concept of 'anticipation' was useful to guide our focus during data collection, its application was more problematic for a clear-cut classification of consequences.

Targeting and user fee exemption policies

Our findings are consistent with past research suggesting that user fee exemption policies can lead to unintended consequences. With regard to indigent selection, for example, a study in Madagascar reported that village workers' own interests influenced the selection and that individuals who were not selected complained [42]. Multiple studies have also found that user fee exemption policies can lead to reimbursement delays, revenue losses for health centres, or the unavailability of drugs [16, 43, 44, 45, 46]. In line with our findings, studies in Mali, Senegal, and Madagascar found that healthcare workers adopted various strategies to reduce the scope of free care for targeted groups due to implementation dysfunctions, sometimes leading to complaints from the targeted population [42, 44, 46]. One study on targeting the poorest in a PBF program in Cameroon also found negative reactions among community members, such as jealousy [5]. The finding that user fee exemptions increase indigents' sense of empowerment has also previously been documented [47].

Unlike in other studies, however, healthcare workers in Burkina Faso did not explicitly report feeling exploited or overworked with regard to providing services to indigents, although they did report an increase in the use of services [16]. This difference may be due to the limited number of indigents covered by the user fee exemption policy, the healthcare centres' available capacity, the staff's strategies to limit free services, and the timing of the data collection, as the use of services varies over the year.

Combining PBF with equity measures

Innovation clusters, such as combining PBF with user fee exemption measures for indigents, may be useful to respond to the growing concern that PBF inadequately addresses inequity in access to care [7]. We found, however, that packaging PBF with user fee exemption measures for indigents can lead to unintended consequences. First, the healthcare workers' fixation on PBF quality indicators sometimes hindered access to free healthcare services for indigents. Second, the low purchase prices for services were perceived as insufficient to motivate providers to target indigents and to cover the costs of the medication. A similar result was found in Cameroon, where some healthcare workers complained that the costs of treating indigents within a PBF program often surpassed the amount received [5].

However, we do not have quantitative data to confirm or challenge the participants' claims that services to indigents were in fact costing more than the unit purchase price. Caution is advised when interpreting this result, as healthcare workers may have ulterior motives. Past studies have reported that user fee exemption policies may result in a loss of income for healthcare workers [16]. Thus, in the present study, it is unclear whether such practices might have influenced the healthcare workers' negative discourse regarding the purchase prices for services to indigents. Those prices were supposed to be higher than simple cost recovery, to motivate healthcare workers to actively seek out the very poor, but as these services were not costed before the intervention was implemented, prices were set arbitrarily. The healthcare workers' views and beliefs regarding the intervention were important because they translated into concrete actions that affected the quality and adequacy of services delivered to indigents. This highlights the pivotal role of "street-level bureaucrats", who have a wide scope of discretion when implementing policies [48, 49]. As was found in this study, capping the value of medications given to indigents can lead to ineffective treatments and exacerbation of diseases, although a proper audit would be useful to evaluate the threat to quality of care. Program planners need to calculate adequately the real costs of treating indigents with medications to ensure they are not putting healthcare workers in a conflict of interests, where they have to choose between providing needed medications to indigents and protecting their own or the healthcare centre's financial interests [42]. The importance of adequately calculating fixed purchase prices for user fee exemption policies in Burkina Faso has been highlighted in the past [50].

Policy implications

The results of this study have policy implications, as governments in LMICs and funders search for strategies to promote the human right to health and to achieve universal health coverage [51]. Global health authorities have affirmed that it would be an unacceptable trade-off to *"first include in the universal coverage scheme only those with the ability to pay and not include informal workers and the poor, even if such an approach would be easier"* [52]. Thus, even if PBF funding ceases, global health actors must consider intervention models that can reach those most in need. In Burkina Faso, for example, a law on universal health insurance (n° 060–2015/CNT) stipulates that the state is the debtor for the indigents' subscriptions (article 48) [53]. In this context, this study's results may help decision-makers appreciate the implementation challenges and unintended consequences that can emerge from a community-based selection of indigents. The results are also pertinent with respect to the implementation of the country's national health financing strategy for universal health coverage (2016–2030), which established strategic purchasing as one of its pillars [54].

Combining user fee exemption policies with PBF is likely to continue to be met with criticism and resistance from local healthcare workers if program planners do not resolve implementation challenges such as lack of starting funds, long reimbursement delays, or insufficient incentives. Past experiences suggest that, for a financing policy to be implemented successfully, budgets must be realistic and lost revenues need to be replaced in a timely manner to ensure a smooth flow of resources [55]. As McPake et al. [56] argue, quick action without sufficient preparation could lead to a deterioration in service quality. The practical issues of UHC implementation need more attention and research [1].

Limitations of the study

Despite our rigorous design, this study does present potential limitations. First, some participants may have tried to portray the intervention positively, either to attract more international aid or due to fears of loss of confidentiality. However, this would have led to an underestimation of undesirable consequences rather than an overestimation. The high number of participants and the researcher's immersion in the milieu reduced the risk of such potential biases. Second, while long observation periods within a few healthcare centres increased the credibility of results, they may have limited the degree to which findings may be generalized to other contexts or settings. It is possible that the implementation process differed between districts and intervention modalities, challenging transferability of results. However, we triangulated the results from our study facilities with multiple sources of qualitative and quantitative data based on broader samples (e.g. routine data, discussions during a one-week national PBF meeting, intervention documents). Moreover, we worked with local management teams to carefully select healthcare centres that were considered representative of the normal context. A third limitation is that we did not conduct observations during the training of local actors and the indigent selection process. This may have introduced potential biases in data collection (e.g. memory bias) and may have resulted in our capturing only a portion of the unintended consequences. Finally, we found that dealing with language differences was a challenge. In interviews, some participants spoke in their second language (French), while others spoke in their native language and relied on an interpreter. While this may have limited some participants' ability to express themselves, we do not believe it affected the validity of results, given the large number of participants and the triangulation of data.

Directions for future research

This study suggests numerous paths for future research. It would be interesting to use quantitative methods to conduct complementary statistical analyses. This could enable us, for example, to: 1) compare the value of medications prescribed to indigent vs. non-indigent patients; 2) assess the number of false inclusions and exclusions on indigent rolls; and 3) assess the cost-effectiveness of pro-poor targeting in comparison to other health equity measures. It would also be interesting to examine how the leadership and management of the intervention at the national and district levels influenced the implementation challenges that emerged. Studies have found that management and leadership practices, including personal initiatives of district leaders, effective supervision, and commitment of the district health management team and local government officials, are critical for successful implementation of exemption policies and UHC reforms [1, 57, 58].

Conclusion

In the pursuit of universal health coverage, international organizations and governments of LMICs are increasingly considering strategies to combine PBF with health equity measures. Using the diffusion of innovations theory, we found that implementing PBF combined with user fee exemptions for indigents led to considerable unintended consequences in Burkina Faso. These unintended consequences can significantly undermine the overall effectiveness and equity of the intervention. To promote successful implementation, program planners need to ensure that local actors, such as healthcare workers, truly adhere to user fee exemption policies. This requires calculating the real costs of treating indigents with medications so that purchase prices paid to healthcare centres are adequate. Moreover, when combining PBF with equity measures, program planners should ensure that healthcare workers' fixation on performance indicators does not undermine free services to indigents. Future research and evaluation of promising health interventions should focus well beyond intended consequences to consider unintended changes that may be less discernible but equally important.

Abbreviations

CHW: Community-based health workers

COGES: *Comité de gestion* (healthcare facility management committee)

CSPS: *Centre de santé et de promotion sociale* (centre for health and social promotion)

F CFA: Franc from the *Communauté Financière d'Afrique*

GPS: Global positioning system

LMICs: Low- and middle-income countries

PBF: Performance-based financing

UHC: Universal health coverage

USD: United States dollar

Declarations

Acknowledgements

The authors express their gratitude to the members of the PBF technical service, the district-level managers, and all participants for their generous contributions to this study. This study would not have been possible without the availability of data on the Ministry of Health's PBF portal, so we thank all actors involved in collecting and reporting this data (www.fbrburkina.org). The authors would also like to thank Stephanie Topp, Quan Nha

Hong and Pierre Pluye for their constructive comments on a previous version of the manuscript, as well as Donna Riley for her editing support.

Funding

This work was supported by a research award funded by the International Development Research Centre. AMTT received a training bursary from the Canadian Institutes of Health Research (CIHR). The research project is part of the "Community research studies and interventions for health equity in Burkina Faso" (ROH-115213), supported by CIHR. The authors also received funding from the University of Montreal Public Health Research Institute. Sponsors did not have a role in the study design; the collection, analysis, and interpretation of data; the writing of the report; and the decision to publish the article.

Availability of data and materials

The data generated and analyzed during the current study are not publicly available to protect confidentiality of participants. They may be available from the corresponding author on reasonable request if the ethical guidelines of both committees involved can be respected.

Authors' contributions

AMTT conceived the study protocol, collected and analyzed the data, and wrote the first draft of the manuscript. IAGG collected part of the data, contributed to the analysis, and critically revised the manuscript. MD and VR helped conceive the study protocol, interpret the results, and critically revise the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The protocol was approved by the ethics committees in Burkina Faso and at the University of Montreal Hospital Research Centre (CRCHUM). Participants provided consent to participate, as required by the ethics committees. The names of the study district and facilities are kept confidential due to the sensitivity of the research topic.

Consent for publication

We obtained consent to use images.

Competing interests

MD is a co-researcher on the baseline and endline studies of the impact evaluation of PBF in Burkina Faso. VR is a co-researcher on the baseline study of the impact evaluation of PBF in Burkina Faso. However, they have received no salary from the funder (World Bank) for this activity. The authors have no conflict of interests regarding the publication of this paper.

References

1. Nicholson D, Yates R, Warburton W, Fontana G. Delivering universal health coverage. A Guide for Policymakers. World Innovation Summit for Health; 2015.
2. Kutzin J, Yip W, Cashin C. Alternative financing strategies for universal health coverage. In: Scheffler RM, editor. World Scientific Handbook of Global Health Economics and Public Policy; 2016. p. 267–309.
3. Priedeman Skiles M, Curtis SL, Basinga P, Angeles G. An equity analysis of performance-based financing in Rwanda: are services reaching the poorest women? Health Policy Plan. 2013;28:825–37.
4. Lannes L, Meessen B, Soucat A, Basinga P. Can performance-based financing help reaching the poor with maternal and child health services? The experience of rural Rwanda. Int J Health Plann Manag. 2015;31:309–48.
5. Flink IJ, Ziebe R, Vagaï D, van de Looij F, van ‘T Riet H, Houweling TA. Targeting the poorest in a performance-based financing programme in northern Cameroon. Health Policy Plan. 2016;31:767–76.
6. Renmans D, Holvoet N, Criel B, Meessen B. Performance-based financing: the same is different. Health Policy Plan. 2017;32:860–8.
7. Ridde V, Gautier L, Turcotte-Tremblay A-M, Sieleunou I, Paul E. Performance-based financing in Africa: time to test measures for equity. Int J Health Serv. 2018:1-12.
8. Ridde V, Turcotte-Tremblay A-M, Souares A, Lohmann J, Zombré D, Koulidiati JL, et al. Protocol for the process evaluation of interventions combining

performance-based financing with health equity in Burkina Faso. Implement Sci. 2014;9:149.

9. Ridde V, Haddad S, Nikiema B, Ouedraogo M, Kafando Y, Bicaba A. Low coverage but few inclusion errors in Burkina Faso: a community-based targeting approach to exempt the indigent from user fees. BMC Public Health. 2010;10:631.
10. Atchessi N, Ridde V, Zunzunegui MV. Is the process for selecting indigents to receive free care in Burkina Faso equitable? BMC Public Health. 2014;14:1158.
11. Société d'études et de recherche en santé publique. Processus communautaire de sélection des indigents dans huit districts sanitaires au Burkina Faso. 2014.
12. Société d'études et de recherche en santé publique. Rapport d'étape de progrès de la mise en oeuvre et du suivi du processus communautaire de sélection des indigents au Burkina Faso. Ouagadougou, Burkina Faso; 2014.
13. Ministère de la Santé. Guide de mise en oeuvre du financement basé sur les résultats dans le secteur de la santé. Ouagadougou, Burkina Faso; 2013.
14. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ. 2008;337:a1655.
15. Petticrew M. When are complex interventions "complex"? When are simple interventions "simple". Eur J Pub Health. 2011;21:397–8.
16. Ridde V, Robert E, Meessen B. A literature review of the disruptive effects of user fee exemption policies on health systems. BMC Public Health. 2012;12:1–7.
17. Cashin C, Chi Y-L, Smith P, Borowitz M, Thomson S. Paying for performance in health care. Implications for health system performance and accountability. Berkshire. UK: Open University Press; 2014.
18. Dkhimi F, Soors W, Criel B. Foreword. BMC Health Serv Res. 2015;15(Suppl 3):I1.
19. Rogers EM. Diffusion of innovations. 5th ed. New York: Free Press; 2003.

20. Bloomrosen M, Starren J, Lorenzi NM, Ash JS, Patel VL, Shortliffe EH. Anticipating and addressing the unintended consequences of health IT and policy: a report from the AMIA 2009 health policy meeting. J Am Med Inform Assoc. 2011;18:82–90.

21. Ash JS, Sittig DF, Dykstra RH, Guappone K, Carpenter JD, Seshadri V. Categorizing the unintended sociotechnical consequences of computerized provider order entry. Int J Med Inf. 2007;76(Suppl 1):S21–7.

22. Ash JS, Sittig DF, Poon EG, Guappone K, Campbell E, Dykstra RH. The extent and importance of unintended consequences related to computerized provider order entry. J Am Med Inform Assoc. 2007;14:415–23.

23. Turcotte-Tremblay A-M, Gali Gali IA, De Allegri M, Ridde V. The unintended consequences of community verifications for performance-based financing in Burkina Faso. Soc Sci Med. 2017;191:226–36.

24. World Health Organization. Trends in maternal mortality: 1990 to 2015. Estimates by WHO, UNICEF, UNFPA, The World Bank and the United Nations Population Division Geneva, Switzerland; 2015. Accessed 6 June 2018.

25. World Health Organization. Atlas of African Health Statistics 2016. Health Situation Analysis of the African Region. Brazzaville: World Health Organization; 2016.

26. Ministère de la Santé. Plan national de développement sanitaire 2011–2020. Ouagadougou, Burkina Faso; 2011. Accessed 6 June 2018.

27. Steenland M, Robyn PJ, Compaoré P, Kaboré M, Tapsoba B, Zongo A, et al. Performance-based financing to increase utilization of maternal health services: evidence from Burkina Faso. SSM - Popul Health. 2017;3:179–84.

28. Ministère de la Santé. Plan d’action 2016 du district sanitaire. Burkina Faso: Ministère de la Santé; 2016.

29. Société d’études et de recherche en santé publique. Stratégie de mise en place des CSI/GVL et de sélection des indigents. Ouagadougou, Burkina Faso; no date.

30. Yin RK. Case study research: design and methods. 4th ed. Thousand Oaks: SAGE Publications; 2009.

31. Zombré D, Turcotte-Tremblay A-M, Sangaré A, Zongo S, Somé P-A, Coulibaly A, et al. Méthode mixte et participative de sélection d'études de cas multiples pour l'évaluation d'intervention de santé au Burkina Faso et au Mali: de l'éthique à l'émique. Poster presented at the 84th Congress of the Association francophone pour le savoir; 2016.

32. Stake RE. Multiple case study analysis. New York: The Guilford Press; 2005.

33. Patton MQ. Qualitative research & evaluation methods: integrating theory and practice. 4th ed. Thousand Oaks: SAGE Publications; 2015.

34. Spicer EH, editor. Human problems in technological change. A casebook. New York: Russell Sage Foundation; 1952.

35. Warford MK. Testing a diffusion of innovations in education model (DIEM). Innov J Public Sect Innov J. 2005;10:1–41.

36. Pluye P, Hong QN. Combining the power of stories and the power of numbers: mixed methods research and mixed studies reviews. Annu Rev Public Health. 2014;35:29–45.

37. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods. 2006;5:1–11.

38. Shadish WR, Cook TD, Campbell DT. Experimental and Quasi-Experimental Designs. Belmont: Wadsworth Cengage Learning; 2002.

39. Ministère de la Santé. Note d'information et d'orientation sur la prise en charge des indigents dans le cadre de la mise en oeuvre du processus de financement basé sur les résultats dans le secteur de la santé Ouagadougou, Burkina Faso; 2014.

40. Pérez D, Van der Stuyft P, Zabala MC, Castro M, Lefèvre P. A modified theoretical framework to assess implementation fidelity of adaptive public health interventions. Implement Sci IS. 2016;11:91.

41. Sveiby K-E, Gripenberg P, Segercrantz B, Eriksson A, Aminoff A. Unintended and undesirable consequences of innovation. Vienna; 2009. p. 16.

42. Honda A. L’analyse d’un fonds d’équité à Madagascar. In: Ridde V, Jacob J-P, editors. Les indigents et les politiques de santé en Afrique. Expériences et enjeux conceptuels. Louvain-la-Neuve: Academia - L’Harmattan; 2013.

43. Hatt LE, Makinen M, Madhavan S, Conlon CM. Effects of user fee exemptions on the provision and use of maternal health services: a review of literature. J Health Popul Nutr. 2013;31(Suppl 2):67–80.

44. Seck A. Les femmes âgées et l’accès aux soins de santé au Sénégal. In: Ridde V, Jacob J-P, editors. Les indigents et les politiques de santé en Afrique. Expériences et enjeux conceptuels. Louvain-la-Neuve: Academia - L’Harmattan; 2013.

45. Diarra A, Ousseini A. The coping strategies of front-line health workers in the context of user fee exemptions in Niger. BMC Health Serv Res. 2015;15:S1.

46. Touré L. User fee exemption policies in Mali: sustainability jeopardized by the malfunctioning of the health system. BMC Health Serv Res. 2015;15(Suppl 3):S8.

47. Samb O. La gratuité des soins et ses effets sociaux : entre renforcement des capabilités et du pouvoir d’agir (empowerment) au Burkina Faso. Thesis. Université de Montréal; 2014. https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/11925/Samb_Oumar_2014_These.pdf?sequence=4&isAllowed=y.

48. Lipsky M. Street-level bureaucracy: dilemmas of the individual in public service. New York: Russell Sage Foundation; 1980.

49. Ridde V, Leppert G, Hien H, Robyn PJ, De Allegri M. Street-level workers’ inadequate knowledge and application of exemption policies in Burkina Faso jeopardize the achievement of universal health coverage: evidence from a cross-sectional survey. Int J Equity Health. 2018;17(1):5.

50. Ridde V, Richard F, Bicaba A, Queuille L, Conombo G. The national subsidy for deliveries and emergency obstetric care in Burkina Faso. Health Policy Plan. 2011;26(Suppl 2):ii30–40.

51. General Assembly of the United Nations. Universal Declaration of Human Rights 1948. www.un.org/en/documents/udhr/index.shtml.

52. WHO. Making fair choices on the path to universal health coverage. Final report of the WHO Consultative Group on Equity and Universal Health Coverage. Geneva: WHO Press; 2014.

53. Conseil national de la transition. Loi N° 060–2015/CNT portant régime d'assurance maladie universelle au Burkina Faso. JO N°07 DU 18 FEVRIER 2016.
2016. https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/101522/122349/F1485776767/BFA-101522.pdf. Accessed 6 June 2018.

54. Ministère de la Santé. Stratégie nationale de financement de la santé pour la couverture sanitaire universelle 2016–2030. Ouagadougou: Ministère de la santé; 2017.

55. Ganaba R, Witter S, Storeng K, De Brouwere V, Ilboudo P, Cresswell J, et al. Burkina Faso introduces free maternal and child healthcare. Health Financing in Africa. 2016; http://www.healthfinancingafrica.org/home/march-20th-2016

56. McPake B, Brikci N, Cometto G, Schmidt A, Araujo E. Removing user fees: learning from international experience to support the process. Health Policy Plan. 2011;26(suppl 2):ii104–17.

57. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol. 2008;41:327–50.

58. Maluka SO. Why are pro-poor exemption policies in Tanzania better implemented in some districts than in others? Int J Equity Health. 2013;12:80.

Additional file 1. Examples of unit costs for healthcare services provided to indigents and non-indigents paid through the intervention (PBF3)

		Unit price (F CFA)	
	Indicators	**Non-indigents***	**Indigents**
1	Number of new consultants of 5 years and older seen in curative care	140	1020
2	Number of new consultants of less than 5 years seen in curative care	210	1270
3	Number of days of patients placed under observation	350	1350
4	Number of counter-referrals received	1400	1400
5	Number of children completely vaccinated	420	420
6	Number of pregnant women who received the tetanus toxoid vaccine 2 or more times during the month	350	350
7	Number of prenatal consultations conducted in the health facility during the month	560	560
8	Number of postnatal consultations conducted in the health facility during the month (Day6–Day8 and Week6–Week8)	700	700
9	Number of births conducted with partographs during the month	2100	3230
10	Number of women (old and new) seen during the month for family planning and using oral or injectable contraceptives	700	1330
11	Number of women (old and new) seen during the month for family planning and using long-term contraceptive planning methods (intrauterine device & implant)	1400	5150
12	Number of new registrants aged 0-11 months seen in consultation for healthy infants	140	140
13	Number of children aged 12–23 months seen in consultation for healthy infants	350	350
14	Number of children aged 6–59 months being treated for moderate acute malnutrition (MAM)	420	420
15	Number of children aged 6–59 months being treated for severe acute malnutrition (SAM) without complication	1050	1050
16	Number of integrated household visits conducted	4200	4200
17	Number of people who underwent voluntary HIV testing (outside of people tested in the context of the prevention of mother-to-child transmission of HIV (PMTCT)	700	700
18	Number of pregnant women and partners who benefited from HIV testing in the context of the PMTCT	700	700
19	Number of HIV+ mothers who benefited from antiretroviral therapy	3500	3500

20	Number of children born from HIV+ mothers followed	4200	4200
21	Number of people living with HIV under ARV followed	1400	1400
22	Number of cases of smear-positive pulmonary tuberculosis (new cases and relapses) tested	8400	8400
23	Number of tuberculosis cases (all kinds) treated and declared cured or completed treatment	11900	11900

***Non-indigent patients also pay user fees in addition to these unit costs paid through PBF. A curative consultation costs 200 F CFA while being in observation costs a flat fee of 500 F CFA.**

Additional file 2. Explanations for the classification of anticipated vs. unanticipated consequences

	Anticipated consequences		Unanticipated consequences	
	Direct (Process)	Indirect (Outcome)	Direct (Process)	Indirect (Outcome)
Desirable				Increased awareness regarding health equity within the community ➢ Not addressed in guides Sense of empowerment among indigents to access healthcare ➢ Not addressed in guides
Undesirable	Selection of individuals who did not meet the local conceptualization of indigents to the detriment of others who did *“External validation was organized.”* (1) ➢ *“Strategies need to be developed to minimize targeting errors as much as possible.”* (2) Ossification ➢ *“This approach makes it possible to regularly update the list of selected indigents”* (2)	Dissatisfaction regarding the selection of indigents ➢ *“The majority of people supported it, but not institutional actors such as health workers, or from other sectors, for whom financial management processes are routinely seen as opportunities to be exploited.”* (2)	Financial difficulties and drug shortages ➢ Not addressed in guides Retention of indigent cards ➢ Not addressed in guides Capping medication costs	Conflicts between indigents and healthcare system actors regarding user fee exemptions ➢ Not addressed in guides Uncertain and unequal coverage for indigents ➢ Not addressed in guides

➢ *“The notion of indigence… raised some concerns regarding certain cases known in the society, particularly cases of temporary disability.”* (2) ➢ *“There was some question about whether any process had been put in place to update the lists, to be able to take into account people who became indigent after the selection.”* (2)		➢ Not addressed in guides Failure to respect the monthly percentage cap of indigents covered ➢ Not addressed in guides Triage of indigents during consultation ➢ Not addressed in guides Fixation on quality indicators ➢ Not addressed in guides	Risk of stigmatization of indigents ➢ Discussed during training sessions but not addressed in guides

■ Intended

☐ Unintended

References

SERSAP. Stratégie de mise en place des CSI/GVL et de sélection des indigents. Ouagadougou.

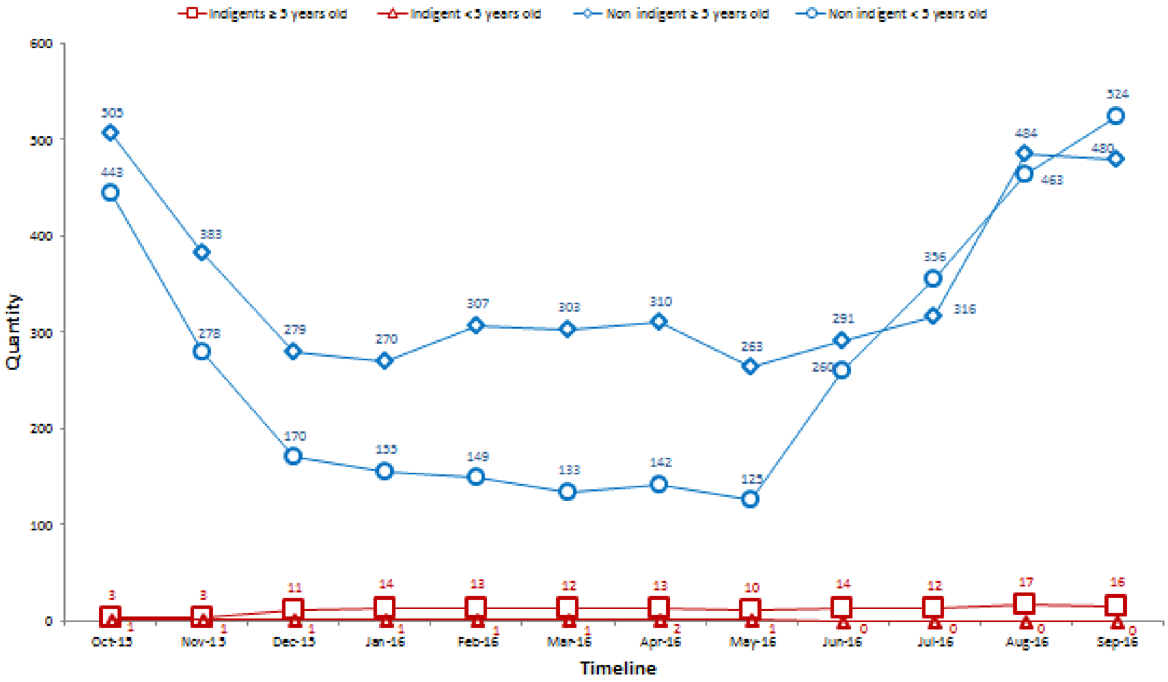
Société d’études et de recherche en santé publique. Rapport d’étape de progrès de la mise en oeuvre et du suivi du processus communautaire de sélection des indigents au Burkina Faso. 2014.

Ministère de la Santé. Note d’information et d’orientation sur la prise en charge des indigents dans le cadre de la mise en oeuvre du processus de financement basé sur les résultats dans le secteur de la santé. 2014.

Additional file 3. A member of the selection committee with indigent cards for his children

Description: A community health worker, member of the indigent selection committee, showing the indigent cards he obtained for himself and his immediate family members.

Additional file 4. Average number of new consultations for patients classified as indigents or non-indigents seen in curative care for healthcare centres in PBF2 and PBF3 across eight districts



Note: Only data from healthcare centres belonging to the intervention arms with user fee exemptions for indigents (PBF2 & PBF3) were used for this graph to facilitate comparison. Districts with intervention arms combining PBF with mutual health organizations (PBF4) or who did not have user fee exemptions for indigents were excluded (PBF1).

Chapter 6 – Discussion

The main purpose of this chapter is to discuss the results of this thesis and propose future avenues for research and practice. First, I will describe how this thesis contributes to the advancement of empirical, theoretical, and methodological knowledge (section 6.1). Second, I will discuss a few cross-cutting themes that emerged from the findings and link them to the broader literature (section 6.2). Third, I will reflect on the study's framework and methods (section 6.3). Fourth, the strengths and limitations of this thesis will be laid out (section 6.4). Following this, I will review the knowledge translation strategies used to disseminate results (section 6.5). This will include a reflection on the challenge of disseminating results locally while protecting the participants' confidentiality (Article 5). I will end this chapter by discussing implications for policy-makers and identifying new research avenues (6.6).

6.1 Empirical contributions and synthesis of results

Until now, the unintended consequences of PBF in LMICs have remained a neglected topic. Scientific evidence has primarily focused on the interventions' intended outcomes, targeted at the outset. This thesis contributes to the development of scientific knowledge by uncovering how a PBF intervention, combined with equity measures, triggered considerable unintended consequences in a low-income setting. Each of the three scientific articles presented in the Results section demonstrated how the nature and implementation of different components of the intervention interacted with the context and the characteristics of local actors to produce a vast array of (mostly undesirable) unintended consequences. Below, I synthesize the results of each article, making links with a World Bank-funded qualitative evaluation. The latter was conducted following an impact evaluation to explain why the PBF intervention fell short of expectations in Burkina Faso (De Allegri, Lohmann, et al., 2019). This add-on study was conducted in 2018, one-and-a half years after the impact evaluation endline, by collecting data through in-depth interviews and focus group discussions across eight districts.

Synthesis of Article 2: Article 2 met the first objective of this thesis by analyzing PBF's unintended consequences in six primary healthcare facilities. A desirable unintended consequence was that some healthcare staff refused to sell medication to individuals who did not consult in order to increase the quantity of patients seen during consultations. This practice was perceived as desirable for its potential to reduce self-medication. This outcome was not reported in the qualitative program evaluation, probably because the latter focused on understanding why results fell short of expected impact.

Regarding undesirable unintended consequences, strong evidence showed that staff members adopted numerous gaming strategies to obtain PBF points. Before PBF evaluations, staff cleaned the facilities more thoroughly, temporarily improved storage conditions for medication, filled out medical registers for former patients, and put on their uniforms. Moreover, PBF verifiers and managers pursued some performance indicators that were locally perceived as irrelevant, unrealistic, or too costly simply to increase performance scores. Providers were sometimes fixated on performance indicators and subsidies rather than on underlying objectives. For example, providers installed curtains to meet PBF's confidentiality criteria, but did not hesitate to breach patient confidentiality during consultations. Staff members also appeared to be fixated more on paperwork than on care provision. These findings are consistent with the program's qualitative evaluation, which underscored that *"many health workers' understanding of PBF had not been as intended, focusing on the monetary aspects of the intervention (…) with the performance improvement aspect of relatively low salience to many*" (De Allegri, Lohmann, et al., 2019, p. 41). As incomplete medical registers automatically received a quality score of zero, providers in the selected facilities spent considerable time and energy filling out registers retrospectively, often inventing information to fill in blanks, instead of serving patients. Medical registers were routinely falsified to artificially increase the quantity and quality of care. The identity and qualifications of providers were systematically falsified to increase PBF scores. Supervisors and trainees were also involved in data falsification to improve performance scores. While the qualitative program evaluation did not elaborate on the issue of falsification, the report does state that managers "*reported registering fake patients, registering children in utero as tested for HIV, recording VAD which had not or only very insufficiently happened*" (De Allegri, Lohmann, et al., 2019, p. 48). In the present study, factors explaining these practices included the pressure to perform, competition between facilities, implementation challenges, strict PBF criteria unadapted to local realities, lack of adherence to official norms, and desires for premiums.

The results show that dissatisfaction emerged over time, in relation to the distribution of premiums, the chronic lateness of PBF funds, and the non-attribution of quality points for some services delivered by itinerary health workers and auxiliary midwives, considered “unqualified”. Moreover, tensions arose between the concepts of managerial autonomy and top-down control. The present study suggests that investment quotas were implemented earlier than suggested by the qualitative program evaluation (De Allegri, Lohmann, et al., 2019). In addition, local actors reported that PBF triggered reductions in other sources of funding from the national level, which negatively influenced activities in the intervention districts. Due to financial difficulties, some funds from the national level were redirected to other non-intervention districts. While this issue was not addressed in the qualitative program evaluation, it may have undermined the interventions’ impact and certainly played on providers’ satisfaction levels.

Local actors also expressed concerns regarding the high costs of PBF and questioned its financial sustainability. According to the qualitative program evaluation, such concerns for sustainability generated substantial reluctance towards to the intervention (De Allegri, Lohmann, et al., 2019)**.**

Synthesis of Article 3: Achieving the second objective of this thesis, Article 3 was the first to decipher the unintended consequences specific to community verifications and client satisfaction assessments. These findings represent an original empirical contribution, as even the World-Bank funded program evaluation did not address these two core components of the PBF intervention model. The results showed that, unlike what was initially planned, the regular PBF verifiers assumed the added role of community verifiers, causing them to experience work overload as well as physical and psychological pressure. Moreover, the intervention’s payment modalities led to dissatisfaction and conflicts for various actors involved. For example, the local association in charge of coordinating community verifications decried the lack of financial support for their organization and withheld part of the sum that was supposed to be paid to investigators who traced patients. Moreover, investigators complained about the cost-benefit ratio of tracing patients. As per the intervention design, investigators were not paid for patients that they were not able to find, despite resources invested to search. CHWs who helped investigators trace patients were also disappointed by the lack of compensation. The intervention's unintended consequences also affected patients. Across cases, patients showed signs of apprehension due to the novelty of the verification and the lack of information, although some were pleased to share their views. Some patients also expressed fears of retaliation from providers. Contrary to instructions, patient confidentiality was consistently breached, with surveys being conducted publicly. Women who had consulted for prenatal care or deliveries were particularly embarrassed by the verification, under the pressure of cultural norms. Due to the social context, community verification also triggered several marital issues. Some husbands were concerned that male investigators were contacting and meeting their wives without their authorization. This caused altercations and influenced the women's participation in the verifications. Moreover, many investigators developed deliberate and organized strategies to falsify the surveys. For example, some investigators impersonated PBF officers and used false pretexts to access the registers in healthcare facilities and retrieve medical information that would enable them to falsify the patient surveys. Healthcare providers reported helping the investigators go through the registers to find information on patients. Some supervisions were orchestrated. Factors that helped explain the falsification of verification data include the payment mechanism based on performance, the unrealistic working conditions to attain objectives, and the numerous challenges associated with locating patients in their communities. Finally, participants questioned the validity and utility of the verification results as well as the value-for-money obtained.

Synthesis of Article 4: Aligned with the third objective of this thesis, Article 4 was one of the first scientific publications to examine the unintended consequences of health equity measures within a PBF program. The results showed that a portion of people selected by community-based committees did not meet the local conceptualization of indigents. Managers, providers, and community members decried the fact that some indigent cardholders had income-generating activities, social support, ability to work, relatively high social status, and belongings. The drift in the selection process was explained by factors related to the nature of the intervention (e.g., high relative advantage), the context (e.g., past experiences with development projects, poverty), the use/implementation of the intervention (e.g., deficiencies in organization and coordination) and the characteristics of actors (e.g., volunteers looking for compensation and political influence). These results are consistent with the qualitative program evaluation, which highlighted that healthcare workers "*were largely unhappy about the outcome of the targeting process, reporting that many individuals were chosen for political reasons rather than for necessity*" (De Allegri, Lohmann, et al., 2019, p. 55). However, whereas the program evaluation reports no major conflict in communities, our results show that the perceived inequity in the selection process did lead to some frustrations, especially between the members of the selection committees and individuals in the community. People regularly demanded to be added to the indigent lists. On the bright side, the intervention provided opportunities to explain the concept of indigence as well as the need to provide services to the most vulnerable. The lack of mechanisms to update the lists or correct selection errors caused a certain level of ossification locally. Within facilities, healthcare staff and members of the management committees agreed that delays in PBF payments and the insufficient unit fees for services delivered to indigents were causing financial difficulties and eventually drug shortages. They claimed that, without the revenues from medications provided to indigents, it was difficult to replenish the facilities' drug depots. Providers adopted a series of strategies to limit the services and medications provided for free to indigents, including surreptitiously removing some indigent cards before their distribution, capping the value of medication prescribed to indigents, and conducting a triage of indigents during consultations. Respondents of the qualitative program evaluation reported similar strategies, such as retracting indigent cards when providers perceived the person as a "faux indigent" (De Allegri, Lohmann, et al., 2019). We found that indigents expressed great dissatisfaction and lack of trust regarding healthcare staff and selection committee members because of the strategies used to limit their access to free care. Moreover, staff members, fixated on PBF quality indicators, asked indigents who were accidentally entered in the wrong registers to pay for their services in order to avoid losing points for erasing information (cleanliness). While the intervention initially increased the selected indigents' sense of empowerment to access healthcare service, it ultimately led to uncertain and unequal coverage for indigents. These findings help explain why the combination of PBF with user fee exemptions did not significantly affect the utilization of healthcare services among the ultra-poor (Beaugé & De Allegri, submitted).

6.2 Theoretical contributions and reflections

6.2.1 Developing a framework based on the diffusion of innovations theory

This thesis contributed to the development of a framework to study the unintended consequences of complex health interventions. For this, we relied on the diffusion of innovations theory as a starting point, but had to move beyond it. While the diffusion of innovations theory proposes an interesting typology of consequences, it presents some limitations to studying unintended consequences. First, the theory examines the entire diffusion process of an innovation, starting from its emergence and going all the way to its consequences. This broad approach makes it tricky for researchers to determine how to circumscribe their focus to study unintended consequences. Second, the theory does not sufficiently elaborate on consequences, especially unintended ones. For example, it does not offer operational definitions of the consequences it proposes in the typology, making their application rather unclear. A literature review showed that only 0.2% of research based on the diffusion of innovations theory focused on consequences (Rogers, 2003). Furthermore, Sveiby et al. (2009) found that only 0.1% of articles were devoted to *unintended* consequences of innovations. Rogers recognized this limitation by stating that "*diffusion studies display a kind of sameness as they pursue a small number of research issues with somewhat stereotyped approaches*" (2003, p. 40).

By stepping outside the typical mold of diffusion research, we helped widen its perspectives. We proposed an analytical framework that can be used to examine an innovation's unintended consequences and their contributing factors. To do so, we synthesized and pieced together literature from diffusion research and unintended consequences. We also contributed to the operationalization of some key concepts to facilitate their application in research (e.g., anticipated vs. unanticipated consequences, direct vs. indirect). Moreover, we proposed a new graphic design to convey the key concepts of this framework to study unintended consequences (see **Figure 1**). The parsimony of the design enabled us to communicate the framework to researchers and knowledge users in effective ways. While some researchers may appreciate its openness and flexibility, others may feel that it does not provide enough guidance or sufficiently inform about precise causal relations between variables. We did, however, demonstrate the feasibility and added value of using this framework by applying it to different components of an innovations (e.g., equity measures, community verification, PBF). The three articles presented in our results section served as proofs of concept for this framework. This application is an important contribution in a field where researchers rarely use explicit theories and frameworks to underpin their analysis (Walt et al., 2008). According to Walt et al. (2008), the advancement of health policy analysis requires that researchers use such frameworks and theories more extensively.

6.2.2 Using the theory to study a global health intervention

The diffusion of innovation theory was initially developed in the U.S. to understand the reluctance of farmers to use innovative technologies (Rogers, 2003). Believed to be valid cross-culturally, the theory was applied in different regions of the world (Rogers, 2003). About 24% of publications in diffusion research concern Latin America, Asia, or Africa (Valente & Rogers, 1995). For illustrative purposes, Barker (2004) summarizes three development projects that applied the theory's tenets in LMICs. In Mali, a study evaluated communication channels to transmit reproductive health information to youth and improve education activities. In Haiti, voodoo practitioners (i.e., trusted opinion leaders) were trained to conduct HIV/AIDS education in remote areas. In Nepal, nutrition demonstrations in households promoted the consumption of vitamin A-rich foods and persuaded neighbors to emulate this behaviour.

More recently, however, the transfer of this theory and other evaluative approaches has raised concerns about their applicability in LMICs (Banque africaine de développement, 2019; Rogers, 2003). Does the transfer of this theory promote the domination of Western worldviews? Is the theory able to capture the worldviews, values, and local practices of people in other settings? Are the concepts consistent with dynamics in Africa? Below, I use my research experience to provide some valuable insight regarding the theory's applicability to study a global health intervention in Burkina Faso, a francophone, low-income country in West Africa.

The applicability of the theory in Burkina Faso

I found the narrative of the diffusion of innovation theory to be well aligned with the field of global health. First, the theory thoroughly describes the roles of actors often involved in global health (e.g., the World Bank acting as change agencies, providers acting as adopters). Second, the diffusion process explained in detail by the theory is consistent with the travelling models often described in the field of global health (Olivier de Sardan et al., 2017). Interventions are developed by international experts and introduced across numerous countries using various strategies to improve health systems in LMICs. Local adopters sometimes reinvent these interventions to fit with their context and needs. Third, I found the diffusion of innovations theory to be suitable in LMICs because it specifically addresses issues of inequality. The theory describes the processes by which innovations widen socio-economic gaps within social systems. It explains that specific strategies can be introduced to narrow these gaps. The theory's explicit discussion of socio-economic inequalities improves its pertinence in a country like Burkina Faso, which has notable social inequalities. It also strengthens its compatibility with the field of global health, which places a high priority on achieving equity for all people worldwide (Koplan et al., 2009). This contrasts with other theories used in public health, such as Talcott Parsons' structural-functionalism, which has been criticized for overlooking or downplaying the issue of inequality (Little, 2012).

I also found the diffusion of innovations theory to be suitable in LMICs because it is aligned with the divergence school of thought (Amankwah-Amoah, 2014). According to this approach, research must pay attention to the different norms, culture, rules, and unwritten rules that shape human behaviour and organizational culture. In this vein, the diffusion of innovations theory stipulates that the consequences of an innovation are highly contingent on local factors, such as the perceptions of actors, the "indigenous knowledge systems", values, norms, past experiences, etc. This contrasts with the convergence school of thought, which contends that theories are applicable universally and yield similar results in any given condition, irrespective of environmental differences and organizational circumstances (Amankwah-Amoah, 2014; Mom, 2018). Such theories have little regard for local concepts or approaches. According to Mom (2018), management theories have failed, particularly in the context of Africa, because they are applied indiscriminately without being circumstance/environment contingent.

For this doctoral study, I contributed to the development of a framework, inspired by the diffusion of innovations theory, with very broad categories of factors that influence consequences: characteristics of actors, the social system, and the nature and use of the innovation. These broad categories enabled our team to capture factors that are particular, yet highly important, in the context of Burkina Faso. For example, the article on community verifications highlighted how kinship may influence verification procedures within PBF. Results showed that an agent from the contractualization and verification agency staged a supervision with local investigators who had the same family name and whom he called “brothers”. Similarly, our results showed how gender inequalities can influence community verifications in Burkina Faso. Some husbands were displeased that male strangers contacted and met with their wives in their absence. Other local specificities emerged as important in our analyses, such as chieftaincy, widespread poverty, small corruption in facilities, informal networks, the lack of compatibility between the evaluation criteria and the context, the history of mutual protection between providers, lack of sanctions, poor infrastructure, staffing, institutional layers, and past experiences with externally-funded interventions. Our open and flexible framework enabled us to incorporate these elements.

In various disciplines, Western-based theories have at times been found to be inadequate to understand African societies because they pay too much attention to internal processes within the person, independent of social and contextual determinants (Amankwah-Amoah, 2014). For some African societies, for example, it is important to consider the concept of “Ubuntu”, which refers to the existence of the person through his or her relations with other people (Frehiwot, 2019). However, I found that the framework based on the diffusion of innovations theory enabled us to integrate both an individualistic perspective (traditionally associated with Western worldviews) and a communalism perspective (traditionally associated with African worldviews) (Amankwah-Amoah, 2014; Mkhize, 2006). Concretely, this was done by adopting a well-balanced analytical lens that considered factors related to both the characteristics of individuals and the social system. According to the theory, diffusion is fundamentally a social process that involves interpersonal networks (Rogers, 2003). Thus, the classic individualism-versus-communalism dichotomy did not emerge as an issue when applying this framework in Burkina Faso.

Another key reason why the diffusion of innovations theory was useful in the context of this global health study was that it encouraged the adoption of methods that can capture local specificities (Rogers, 2003). Methods commonly used in anthropology, such as prolonged immersion, observation, and informal discussions, enabled us to gain an in-depth understanding of local beliefs and practices. They also provided us, external researchers, the flexibility needed to collect data on themes that emerged as important in the field, increasing their pertinence in the context of Burkina Faso.

The challenges of applying the framework in Burkina Faso

I found that the ability of the framework to reveal local particularities in Burkina Faso was contingent on the cultural knowledge and competence of the two qualitative researchers (AMTT and IAG) involved in the data collection. By serving as data collection tools, we made the bridge between the conceptual framework and the empirical world. Thus, our ability to detect particularities and analyze the local social systems was crucial. Pointing to this issue, a Burkinabè proverb says, "*A stranger looks but does not see*". For example, outsiders do not always understand the symbolic meaning behind objects or actions. To overcome this challenge, we actively tried to be attentive to the social systems (e.g., culture, norms, meanings) and used cultural interpreters to validate our interpretation of data. It is possible, however, that some important elements remained undetected and therefore could not emerge in our analyses, despite our open and flexible framework. It should be noted, however, that the use of outside researchers was also a strength in some cases. For example, as an outsider, I could question and bring attention to what local actors already take for granted. Moreover, because of their naivety, outside researchers are sometimes be able to trespass certain forbidden areas or behave in unconventional ways (e.g., regarding gender relations) to collect pertinent data that can feed the framework. For example, as a female Canadian researcher, I was able to reside, eat and drink with both men and women participants which insiders may not have been able to do as easily.

Another challenge related to the application of the framework was determining the weight to give to cultural explanations in our analyses. Rogers highlights the importance of culture to understand the diffusion of innovations within a given context. However, we encountered a risk of becoming "culturalist" by placing too much emphasis on culture (Olivier de Sardan, 2010). This approach can be based on stereotypes and be too deterministic. Ultimately, we tried to reach a balance by referring to culture, when needed, without falling into the trap of focusing too much on that. A collaboration with local researchers helped minimize the risk of putting too much emphasis on culture.

When applying the framework in Burkina Faso, I also encountered a few challenges related to language. First, most of the literature on the diffusion of innovations theory is available in English. This language is rarely spoken in Burkina Faso, making it more difficult to communicate the underlying concepts of the study to local research assistants. Out team provided translated summaries to local collaborators to overcome this issue and guide the data collection.

Another challenge in using the diffusion of innovations theory is its use of jargon that is uncommon, not only in the fields of public health or global health, but also in Burkina Faso. For example, simple terms such as "innovation" and "use" had to be translated to "intervention" and "implementation" in the field of public health. Moreover, terms such as "relative advantage", "triability", and "observability" are not commonly used in public health. This caused a dilemma because, on one hand, I wanted to stay true to the diffusion of innovations theory and ensure that its followers recognized the basic concepts of the theory. On the other hand, I wanted to ensure that a broader audience in the field of global health could understand the study. Ultimately, I tried to reach a balance in the articles.

This challenge was amplified by the fact that the study was conducted in Burkina Faso, which also uses different terminology. For example, a collaborator criticized the use of expressions such as "the intervention's implementation" with participants. This type of language was associated with "white academic" people and not well understood locally. Even the idea of using theories was misunderstood by local participants. After discussing the use of theories with a provider, participants in another village asked "*how can you be just a theoretical person? What does that mean?*" In some instances, we had to adapt our language to communicate with illiterate participants living in rural zones by adopting their language and using terminology such as “gifts” to refer to the intervention. Thus, adapting our language to interact with different types of stakeholders proved to be necessary in conducting the study.

6.3 Methodological contributions and reflections

This thesis contributes to advancement of methodological knowledge by illustrating how to design and conduct a study on the unintended consequences of health innovations. In the past, many theoreticians and researchers have argued that unintended consequences are difficult to study (Koch & Schulpen, 2018; Rogers, 2003). The difficulty of predicting and measuring unintended consequences has hampered this area of research. Over the course of this doctoral study, numerous researchers and practitioners privately reported witnessing important unintended consequences resulting from interventions in their respective lines of work but felt ill-equipped to assess them. These interventions included social accountability for health programming, HIV self-testing kits, and demand-side results-based financing. Many inquired about how the methods used in this study could be applied in their area of work. One researcher[12] from a European institution contacted me by email stating, “*We are big fans of your unintended effects paper (especially the figure) published in SS&M. We are even using it for a study related to HIV self testing kits in Uganda. The tests are being used in all kinds of unexpected ways!*” Our subsequent collaboration led to the submission of a manuscript in a peer-reviewed journal. This successful example highlights how this doctoral research can inspire others to engage in this area of research.

6.3.1 Insights on how to study unintended consequences

Our experience conducting research on the unintended consequences of a complex health intervention has brought about some learnings that can benefit other researchers in this field. We developed a series of considerations that can be useful to design and conduct studies on the unintended consequences of innovations (see **Box 2**). These build on the work of Jabeen (2018), who recently proposed an evaluation process to study unintended

---

[12] This researcher provided written consent to disclose the content of this e-mail.

consequences. Still limited in its practical guidance, Jabeen's three-step process includes: a) outlining program intentions, b) forecasting likely unintended effects, and c) mapping the anticipated and understanding the unanticipated unintended outcomes. Below, I discuss some of Jabeen's propositions while providing more detailed and comprehensive insights that can guide the study unintended consequences. Of course, these considerations should be validated and refined through future research and input from other researchers in this area.

Box 2. How to study the unintended consequences of complex health interventions

- Set an explicit objective or research question targeting unintended consequences
- Choose and define your terminology
- Clarify your point of view
- Adopt a theory or conceptual framework
- Clarify the intervention's logic model
- Forecast potential unintended consequences
- Focus on desirable, undesirable, and even neutral unintended consequences
- Use flexible methods
- Conduct observation
- Cast a wide net during the data collection
- Present findings to stakeholders to validate the classification of unintended consequences

Set an explicit objective or research question targeting unintended consequences: Few studies and evaluations explicitly aim to study unintended consequences from the outset (Bonell et al., 2015; Hageboeck et al., 2013; Jabeen, 2016; Merton, 1936; Mittelmark, 2014; Morell, 2005). We found that it is useful to elaborate an explicit objective and research question in the protocol to study unintended consequences. This increases the likelihood that sufficient resources, including time and money, will be dedicated to the examination of unintended consequences. Having an explicit objective to be met may also encourage the use of appropriate methods to collect pertinent data on this theme and ensure that the data are analyzed with this objective in mind. Nevertheless, we found that it is still possible to examine unintended consequences by conducting *post hoc* analyses. In some cases, data revealing unintended consequences have emerged from the field and stakeholders have been compelled to address the issue even though it was not in their initial objectives. Researchers in this situation may have to re-analyze all their data to meet this new objective, which ultimately may turn out to be more time-consuming and costly. For such *post hoc* analyses, it is also possible that not all unintended consequences are reflected in the data, given that the researchers were not actively collecting data on the topic. As a popular proverb says, what we see depends mainly on what we look for (Lubbock, 2011).

Choose and define your terminology: In the scientific literature, a wide array of terms related to unintended consequences are used interchangeably without clear definitions (de Zwart, 2015). For example, studies refer to unanticipated, unexpected, unintended, or spillover effects, without any clear conceptual distinction. To adequately conceptualize the research object, it is important to select and define what the research team will consider to be unintended consequences. This will influence the researchers' focus during the data collection and analyses.

Clarify your point of view: In operationalizing the concepts selected, researchers should clarify their point of view. Intended/unintended by whom? Anticipated/unanticipated by whom? Desirable/undesirable for whom? Foreseeable/unforeseeable by whom? For example, the level of anticipation of consequences can vary depending on stakeholders' position or even imagination. Similarly, the level of desirability can vary when we compare the perceptions of decision-makers with those of street-level workers, who have different goals and needs. Operationalizing these concepts requires difficult choices for research teams, but examples on how to do it are available in the literature (Turcotte-Tremblay et al., 2017; Turcotte-Tremblay, De Allegri, et al., 2018).

Adopt a theory or conceptual framework: Social science theories can be useful to study unintended consequences. According to Jabeen, these theories can help: 1) predict unintended outcomes and the likely mechanisms generating them; 2) explain findings about unintended effects and the mechanisms producing them; and 3) attribute identified unintended outcomes to the program. However, our experience has shown that selecting a theory to study unintended consequences can be difficult. In his book, Morell (2010, p. 57) argues that "*it does not matter which ones are chosen*" because, in any case, the program, the stakeholders, and the goals are the same. Thus, there will be a great deal of similarity across all the evaluations. Contrary to this view, I argue that the selection of a theory can be important because it influences the analytical lens that researchers use to collect and analyze data. Each theory employs a unique vocabulary to articulate specific concepts that are interrelated.

I found several promising theories to study unintended consequences. The *diffusion of innovations theory* has already been shown to be useful in the present thesis (Turcotte-Tremblay et al., 2017; Turcotte-Tremblay, De Allegri, et al., 2018). The *strategic actor theory* also provides interesting constructs to understand unintended effects (Crozier & Friedberg, 1977). That theory suggests that, to analyze organisations, researchers should distinguish between formal and informal structure. The former corresponds to the part of the structure that is official and codified. The latter corresponds to practices, interactions, and relationships that are not officially planned and that may be forbidden. Because of the actors' margins of liberty, actions in informal structures can lead to effects that were not originally intended within formal structures. Rather than considering unexpected behaviours as exceptions, the authors of the strategic actor theory suggest that it would be more fruitful to use these as the point of departure for an analysis (Crozier & Friedberg, 1977). The *complexity approach* also proposes concepts that can be useful to understand the mechanisms by which interventions produce unintended consequences, including feedback loops, the relation between the parts and the whole (systemic principle), the level of fit between an intervention and its context, and the dialogical principle, whereby notions that appear to be contradictory can be part of a unique whole (Morell, 2018; Morin & Le Moigne, 1999). Moreover, a number of authors have proposed different typologies to classify unintended consequences (see section 2.1.3). While they tend to lack the richness of social science theories to understand relations between concepts, these typologies may still be useful to orient researchers' or evaluators' focus during the data collection and analyses. Adopting an inclusive typology *ex ante* compels researchers and evaluators to consider the possibility that interventions can produce consequences that are not intended, instead of merely focusing on intended ones. Alternatively, it is also possible to study unintended consequences by using a more inductive approach without relying on specific theories or framework during the data collection.

Clarify the intervention's logic model: In order to correctly identify an intervention's unintended consequences, research teams must first understand the intended processes and outcomes. This knowledge will help the researchers orient their focus during the data collection and analysis. This is akin to Jabeen's (2018) first step, which aims to outline program intention by: 1) reviewing program documents and social science theory; 2) interviewing stakeholders; and 3) developing or reviewing the program theory/outlining the intended outcomes. While we consider this step to be highly useful to orient the data collection and enable consequences to be labelled as unintended, it should be noted that researchers using other approaches, such as goal-free evaluation, remain deliberately unaware of program intentions to avoid a narrow view (Jabeen, 2016).

Forecast potential unintended consequences: Like Jabeen (2018), we found that conducting a literature review was useful to identify a preliminary list of potential unintended consequences that could provide a starting point. However, we did not let this list limit our focus, to be able to capture unintended consequences that were context specific or unexplored in the literature. Jabeen (2018) also suggests that researchers can predict unintended outcomes by interviewing experts and involving stakeholders in forecasting outcomes. Before beginning the data collection, I did ask local stakeholders whether they had ideas regarding the potential unintended consequences of the intervention, but found that some had little to say about this issue, probably due to the sensitive nature of the topic and the lack of reflexivity among actors on this issue.

Focus on desirable, undesirable, and even neutral unintended consequences: Past studies have found that unintended consequences tend to be mostly undesirable or negative. However, researchers should try to overcome this tendency by collecting data on desirable, undesirable, and even neutral unintended consequences. Explicitly widening the focus of the study to encompass both desirable and undesirable unintended consequences may help researchers be accepted in the milieu. Local stakeholders may feel less threatened if the researcher can explain that she or he is not only there to report negative findings. Moreover, presenting both desirable and undesirable unintended consequences will provide a more complete picture. It may also make it easier for stakeholders with vested interests to accept the findings. Although this study did not find neutral consequences, future research should remain open to this possibility.

Use flexible methods: Researchers cannot foresee all unintended consequences when developing their protocol. Thus, they should remain flexible and open during the data collection to capture data on all unintended consequences that emerge. Exploratory qualitative methods can be powerful tools to study unintended consequences. In our experience, we found that interview guides and observation grids had to be adapted and refined as unintended consequences became perceptible to the research team over time.

Conduct observation: Conducting observation *in situ* over a long period of time was crucial to develop trusting relationships and be able to study hidden behaviours in a more natural context. As Jabeen (2018) explains, the narratives people tell and the reality of an event do not always align.

Cast a wide net during the data collection: During interviews and informal conversations, I found that general questions, such as "*did the intervention lead to any unintended consequences?*" or "*did the intervention cause changes that surprised you?*" generally did not yield interesting results. In most cases, respondents simply answered "no". Street-level workers did not fully comprehend the intervention's logic model or intended outcomes, especially due to the complexity of the design. Thus, to capture pertinent data, we cast a wide net by asking a lot of questions about different aspects of the intervention. We also tried to be creative and think outside the box to detect potential spillover effects that were not expected, even by the research team. We pursued leads that were triangulated through various sources or types of data. Then, we deciphered the intended and unintended consequences during the data analysis. The main challenge with this approach was that it required us to collect a colossal amount of data on both intended and unintended consequences.

Present findings to stakeholders to validate the classification of unintended consequences: Even if researchers diligently tried to understand an intervention's logic model, some of the consequences they identified as unintended may actually have been considered to be intended by program planners. Thus, we found that it was useful to present the findings on unintended consequences to stakeholders in order to verify their classification. This approach, however, differs from that of Jabeen (2018), who argued that any program outcomes not identified when the program intentions were first outlined should be considered unintended. I disagree with this position for two reasons. First, some consequences that emerged during this doctoral study were clearly identified as intended when discussing results with stakeholders, even though they had not been initially identified as such during our preliminary work to outline the intervention's complex logic model. Maintaining that these consequences were unintended despite this feedback could have discredited the whole study. Second, some interventions are dynamic and have evolving logic models. Thus, researchers should leave open the possibility of adapting to this as the intervention unfolds, if they see fit.

6.3.2 Article 5: Oscillating between passive and active roles during non-participant observation in global health research

This doctoral study relied heavily on observation sessions to collect empirical data. In the field, however, I found that defining my role as an observer was quite challenging. I was sometimes required to take a more proactive role than anticipated. For example, in one facility, staff members strongly insisted that I help them falsify medical records for a PBF verification. The providers were running late and urgently needed to fill out the documents before the PBF verifier arrived. This placed me in a delicate position, because falsifying medical information is a serious act that clashed with my observer status. However, assisting them more actively could help me assess the cognitive processes they were going through (e.g., imagining versus remembering patients). Partaking in this activity helped me confirm that providers were truly inventing medical information for the PBF verification and not simply writing medical information retrospectively. This was useful to assess the causal relation between the PBF intervention and unintended consequences. This reflection on the role of observers was developed in the article below and published as part of a special issue organized by the REALISME Chair [*REcherches AppLiquées Interventionnelles en Santé Mondiale et Equité*].

Article 5: Oscillating between passive and active roles during non-participant observation in global health

Anne-Marie Turcotte-Tremblay

Cahiers REALISME 2018, Hors-série

Available online in open access:

http://www.equitesante.org/reflexivite-sante-mondiale/

Title: Oscillating between passive and active roles during non-participant observation in global health

Anne-Marie Turcotte-Tremblay[a,b*]

[a] University of Montreal Public Health Research Institute, Montreal, QC, Canada

[b] School of Public Health, University of Montreal, Montreal, QC, Canada

* Corresponding author: Anne-Marie Turcotte-Tremblay

Contributions of author:

Anne-Marie Turcotte-Tremblay conceived the original idea, developed the reflection and and wrote the manuscript.

Introduction

The researcher’s immersion in a milieu can be useful in global health to understand local contexts, actors, and intervention processes. Through observation, the researcher can be in direct contact with the research object in a real-life setting. This method may be particularly pertinent for global health researchers, who generally come from outside the study site and need to understand local practices and norms. As a female Canadian researcher with a background in psychology and public health, I found this data collection method highly useful for my doctoral research on the unintended consequences of a health intervention in Burkina Faso (Turcotte-Tremblay, Gali-Gali, De Allegri, & Ridde, 2017).

Upon adopting observation as a method, I had to take strategic decisions about how to behave *in situ* because my presence in the field could interact with the object of study. Different models of observation with varying levels of integration have been documented (Olivier de Sardan, 2015; Poupart, Groulx, Deslauriers, Laperrière, & Mayer, 1997). For my doctoral study, I initially decided to conduct “non-participant” observation, which involved observing participants without becoming actively involved with them. My rationale, developed with my supervisors, was that because I am not trained to assist healthcare workers in their professional activities, I could not actively participate. My intention was simply to observe daily activities, like a bird on a branch, capturing every moment with as little interference as possible to allow natural behaviours to emerge in a natural setting. I believed my role was to stay passive and in the background to reduce potential biases caused by my presence.

In the field, however, I found I was sometimes invited or tempted to abandon my position as a passive observer and adopt a more active role in the social system. It was not always possible or even desirable, for the study and also for me, as an ethical individual, to remain a mere observer on the sidelines. I was occasionally called to intervene actively, that is, to take action within the environment where observation was being conducted. In this context, a series of questions arose : To what extent can a non-participant observer actively intervene in the local context? When are these interventions more or less useful? How could they affect the data collection? There may be no universal answer to these questions, as different disciplines (e.g. anthropology, medicine) may conceive observation differently. Nevertheless, engaging in a reflexive process may be useful to understand and improve our practices as researchers. As such, my objective in this paper is two fold: 1) to describe different types of interventions I conducted during my observation sessions, and 2) to present a reflexive analysis of these interventions that led to personal lessons learned. While social scientists such as Chauvin and Jounin (2012) have previously raised some of the issues discussed in this text, I hope these concrete examples from doctoral research in global health will help others better prepare and position themselves regarding some of the questions they may encounter during non-participant observation.

Context of the reflexive analysis

This reflexive analysis is based on my experience conducting an exploratory study for my doctoral research. The objective of the study was to understand the unintended consequences of combining an intervention called performance-based financing (PBF) with health equity measures in Burkina Faso (Ridde et al., 2014). In this intervention, healthcare workers receive financial incentives to improve healthcare services. They also receive compensation to provide free healthcare services to indigents. Our research team developed a conceptual framework based on Rogers' theory on the diffusion of innovations. Applying this framework, we examined how the social system (including local norms and culture), the characteristics of its members, and the nature and use of the innovation (i.e., the PBF intervention) interacted to influence the emergence of unintended consequences over time. The conceptual framework was deliberately broad, in order to capture all relevant data. For the data collection, I conducted three months of non-participant observation in healthcare centres, held semi-structured interviews, and participated in informal discussions with local actors. Observation sites included healthcare centres and their catchment areas. To integrate the field, I first obtained approval from ethics committees and district management teams. Then I called the head nurses of the selected healthcare centres to explain that I would be visiting in the upcoming days to conduct a study on the implementation and consequences of PBF. All the head nurses welcomed me cordially. In each case, I lodged in one of the facility's buildings, often with one of the healthcare workers. This enabled me to conduct observation around the clock for a wide spectrum of daily activities. I ate meals and spent all my leisure time with participants such as healthcare workers, villagers, and community leaders. Each activity represented an opportunity to collect relevant data on the nature of the social system, the characteristics of local actors, and the implementation of the intervention. As a consequence, however, it was difficult to make clear-cut distinctions between the object of study and my participation in the community. I recorded field notes systematically in a research diary. During this field work, I identified four main types of situations in which I was voluntarily or involuntarily incited to intervene. The section below presents each type, along with a general statement based on my reflexive analysis.

The non-participant observer's four types of interventions

1. *Intervening to promote health*

While conducting non-participant observation, I encountered vulnerable individuals in dire need. For example, a hut collapsed onto a baby whom I had previously met at the healthcare centre. With an eyeball bulging from its socket, this baby required healthcare, but his mother could not afford to bring him to the hospital. During observation sessions, I also witnessed healthcare workers mistreating girls in the facility. Moreover, I encountered elderly indigents who urgently required financial support. « When I'm hungry, I simply cry », said one such person during a heartbreaking conversation. Observing suffering and vulnerability triggered internal debates on whether I should intervene out of a moral solidarity for humankind. From an ethical perspective, how can one passively observe such suffering without providing immediate assistance?

Some of these situations were directly related to the intervention I was studying. For example, healthcare workers at one facility misunderstood the PBF intervention. They mistakenly believed the government would reimburse 100% of the cost of the medication they were distributing free of charge to poor patients. This misunderstanding could negatively impact the healthcare centre's financial situation and medication stock, ultimately hindering access to care. Realizing this conundrum, I had to decide whether to intervene by informing the healthcare workers that the medication would not be entirely reimbursed or instead continue to observe the situation silently as it unfolded. Intervening implied influencing the implementation process of the intervention I was studying, thereby changing its course permanently, while not intervening implied possibly allowing the healthcare centre to sink into a financial hole. In a similar situation, I met a blind indigent woman who did not know she had been selected to obtain free healthcare services. Again, I had to decide whether I should intervene to inform her.

Of course, each of these situations should be handled on a case-by-case basis. In general, however, I concluded that a non-participant observer such as myself should avoid intervening to promote health during fieldwork. Researchers should focus on the primary objectives of their presence in the field and carry out activities approved by ethics committees. Intervening in the situations described above could lead to research biases if participants come to see the researcher as a donor. It could also create financial expectations regarding researchers, thereby undermining research and its possible benefits for local populations. More importantly, researchers can involuntarily create more harm than good by intervening locally. For example, an observer could negatively influence the coping strategies, empowerment, or resilience of local actors. To make their decision, researchers can do risk-benefit assessments. If they determine an intervention is necessary and unlikely to cause perverse consequences or negatively influence the study, they can try to seek external help without intervening directly. In such difficult situations, researchers may find it useful to remind themselves of the long-term benefits of research that require sacrifices in the short term. This calls for emotional disengagement from immediate situations to prioritize a long-term development perspective.

2. *Intervening to integrate the social system and become accepted by local actors*

While conducting non-participant observation, I was sometimes invited to partake in interactions that would enable me to better integrate the social system and be accepted by local actors. These included helping to prepare meals and conversing with local actors. In some of these situations, however, local actors implicitly or explicitly invited me to adopt behaviours that conflicted with my principles or values. Resisting could negatively influence the relationship I was trying to build, while consenting would involve behaving in ways that were contrary to my principles or values. In the middle of medical consultations and birth deliveries, for example, healthcare workers sometimes interrupted their professional activities to engage in excessive laughter, drink tea, yell, and discuss their personal plans for the evening. My initial attempts to distance myself from such unprofessional behaviour in the presence of patients clashed with the social dynamic. Healthcare workers were outgoing and friendly towards me, while I responded quietly and reservedly so as not to encourage them. A similar situation occurred with a PBF verification officer who continually interrupted the verification process with healthcare workers to chitchat with me about the possibilities of acquiring a Canadian wife, etc. Responding would have interrupted the natural flow of the verification process I was observing, while ignoring him might have been perceived as being hostile.

In general, I realized a non-participant observer such as myself must sometimes adapt to the local setting to facilitate integration, even if this requires setting aside one's own values and principles. To be accepted, it is sometimes necessary to "play the game" with local actors, as long as no additional harm is caused by the observer's participation. The researcher should not assume responsibility for other people's behaviours, as this is outside of his or her control. Refusing to conform to local practices (even unprofessional ones) may hinder the relationship with actors and create tensions that could, among other things, cause potential biases in the data production. Local actors may not feel comfortable opening up and displaying natural behaviours if the non-participant observer is not able to show that he or she is "one of them". Observers may have to make concessions regarding their "non-participation" to find a middle ground where everyone is comfortable.

3. *Intervening to understand the research object*

One challenge I encountered as a non-participant observer was to determine how best to intervene during informal conversations to obtain information. At first, during informal conversations between healthcare workers, I interjected questions for clarification. I soon realized, however, that this influenced the natural progression of the conversations and sometimes inhibited conversations, as participants appeared wary of my proactive questions. I learned it was more useful to simply listen to conversations without interfering too much. Patience often paid off, as clarifications and explanations sometimes arose naturally in due time.

On some occasions, I was invited to participate actively in healthcare centre activities and social events to gain a better understanding of the local context. During a consultation for healthy children, for example, a woman placed her malnourished child in my care, inadvertently breaking my status as a non-participant observer. I waited one hour in the company of other mothers, with this girl on my lap; then I placed the harness on her and measured her before consulting with the midwife. This experience enabled me to feel and relate to mothers' experience when bringing their children to healthcare centres. I faced some of their daily challenges, namely regarding excessive heat, lack of water, and wait times. On another occasion, healthcare workers invited me to conduct outreach activities in nearby villages. They asked me to help them measure the babies' heights as they weighed them. For a brief moment, I was transformed into a healthcare worker facing ordinary challenges, such as the lack of safe roads to rural villages, long work hours without food, the heat, the dust, etc. By experiencing this firsthand, I could better account for those challenges in my analyses.

Overall, I found that intervening occasionally with moderation could be useful to understand the actors' points of view within the social system. Although I positioned myself formally as a non-participant observer, I still had to decipher when it was acceptable and useful to intervene to better understand the research object. Moreover, I realized it was sometimes necessary to set limits regarding healthcare workers' demands, such as those that involved performing medical acts (e.g. inserting a syringe, interpreting malaria tests). Even though it may be helpful to take part in local activities, non-participant observers should continuously assess the potential risks and benefits of their actions, not only for the research, but also for local participants, especially when the observers have no medical training.

4. *Intervening to serve an instrumental purpose to local actors*

In some cases, local actors tried to persuade me to intervene to achieve their own ends. For example, one healthcare worker who was concerned that the low use of healthcare services was negatively influencing his performance asked me to talk to the population to convince them to seek care. "If you tell them to come, it will stay in their heads," he pleaded. I politely declined this request, as it was not my mandate as a researcher and could influence the healthcare centre's performance score in the upcoming months, which our research team was investigating. In another situation, a healthcare worker strongly insisted that I ask a PBF officer to "excuse" his team for not having complied with the verification process. He also insisted that I tell this PBF officer that I had seen the worker and his colleagues “work as a team” to prepare the verification documents. I explained that I was only present as an observer, but after much insistence from the healthcare worker I eventually just nodded quietly. Not complying with this request was tricky, as we had developed a friendship. In another case, a midwife tried to obtain my explicit support to refuse a patient who arrived late at night. Despite her insistence, I refused to explicitly condone this behaviour and remained unresponsive to her request. The midwife felt some pressure and provided treatment to the woman in pain. Thus, my reactions as an observer potentially influenced the healthcare centres’ activities, despite my intentions. In general, nonparticipant observers should avoid being instrumentalized to further local actors’ purposes, as this may influence the research process and possibly the results of the study. However, this must be done delicately, as refusing to be instrumentalized can influence relations with local actors.

Conclusion

This reflexive analysis highlights how difficult and sometimes impossible it is to observe passively from a distance without intervening when conducting non-participant observation during extended field immersion. Despite their best intentions, non-participant observers can find themselves oscillating between passive and active roles. They may intervene to : 1) promote health, 2) integrate the social system and become accepted by local actors, 3) understand the research object, or 4) serve an instrumental purpose to local actors. Finding the appropriate distance for different situations requires some adjustment. It is important for observers to analyze their interventions *in situ* and to be conscious of them, as they are an integral part of how data are collected and can influence the object of study to varying degrees. Each of the different types of interventions has the potential to trigger a chain of events leading to other changes (Reardon, 1994). Ethical implications should be considered, as they can directly and indirectly affect the health of community members. Students in global health research should be encouraged and trained to adopt and reflect on different types of observations, as this can add value to our field of practice. Having support groups and safe spaces in which to engage in reflexive discussions before, during, and after data collection could help researchers produce sound findings.

Acknowledgements

The author is grateful to the Chair REALISME for its ongoing support. Special thanks go to Emilie Robert, Esther Mc Sween-Cadieux, and Kadidiatou Kadio for their comments on an earlier version of this paper. The author would also like to thank Donna Riley for her editing support. AMTT received a training bursary from the Canadian Institutes of Health Research (CIHR) and a doctoral research award from the International Development Research Centre.

References

Chauvin, S., & Jounin, N. (2012). L'observation directe. In S. Paugam (Ed.), *L'enquête sociologique* (pp. 143–165). Paris: Presses Universitaires de France.

Olivier de Sardan, J.-P. (2015). *Epistemology, fieldwork, and anthropology*. New York, NY: Palgrave Macmillan.

Poupart, J., Groulx, L. H., Deslauriers, J. P., Laperrière, A., & Mayer, R. (1997). *La recherche qualitative: enjeux épistémologiques et méthodologiques*. Montreal: Gaëtan Morin Éditeur.

Reardon, J. (1994). Treehouse of horror V. *The Simpson*s. 20th Century Fox. Retrieved from www.youtube.com/watch?v=oUbzxPpaxjw

Ridde, V., Turcotte-Tremblay, A.-M., Souares, A., Lohmann, J., Zombré, D., Koulidiati, J. L., … Allegri, M. D. (2014). Protocol for the process evaluation of interventions combining performance-based financing with health equity in Burkina Faso. *Implementation Science, 9*(1), 149. https://doi.org/10.1186/s13012-014-0149-1

Turcotte-Tremblay, A.-M., Gali-Gali, I. A., De Allegri, M., & Ridde, V. (2017). The unintended consequences of community verifications for performance-based financing in Burkina Faso. *Social Science & Medicine, 191*, 226–236.

6.4 Cross-cutting themes of the study

The three articles in the Results section share some transversal themes. In the subsection below, I address the following: 1) the gap between the intervention's official norms and practical norms; 2) the complementarity of articles for understanding the weaknesses of verification mechanisms; 3) the predominance of undesirable unintended consequences; and 4) the notable absence of some unintended consequences.

6.4.1 Persistent gaps

Official norms vs. practical norms: Together, the articles presented in the Results section highlight a persistent gap between official and practical norms, that is, between what people are supposed to do and what they actually do. According to Olivier de Sardan (2015), *official norms* are the formalized and explicit "rules of the game". They can refer to an intervention's manuals, contracts, procedures, and organizational structures. In contrast, *practical norms* refer to behaviour that is considered improper, deceitful, or illegal from the perspective of official norms, but which nonetheless follows an identifiable pattern. Within the healthcare system, professional practices that transgress official norms are widespread, affecting the quality and quantity of care. According to Olivier de Sardan (2015), interventions based on the New Public Management (NPM) approach, such as PBF, are introduced in an attempt to reduce divergences between formal and practical norms as much as possible. Various management monitoring methods are used to ensure that the practices of healthcare workers comply with professional norms (Olivier de Sardan, 2015). In this implementation process*,* new layers of official norms are imported and superimposed on the official norms already in place (and already far from being respected), thereby changing the rules of the game (Olivier de Sardan et al., 2018).

In Burkina Faso, PBF introduced new official norms, such as new verification mechanisms, payment modalities, registers, and documents to fill out. However, the three articles in the Results section of this thesis show that important divergences emerged between the formal rules introduced by PBF and the real practices of actors, including providers, community verifiers, and community health workers. In the face of new official norms, these local stakeholders adjusted their old practical norms and created new ones. These divergences were not merely individual and marginal. They were recurrent and widespread. In the first article, for example, healthcare providers developed new gaming strategies and ways of falsifying performance data to meet the PBF evaluation criteria. Similarly, the second article showed that providers helped investigators from local associations use medical registers to falsify verification data. Meanwhile, the third article showed that providers developed strategies to limit the free healthcare services delivered to indigents. Interestingly, Olivier de Sardan (2015) argues that program planners should start from the local practical norms when developing new interventions: "*Rather than introduce new official norms that will gain no more respect than the previous ones, should the pragmatic objective of a reform strategy not be to 'start' with and from the practical norms, to make some of them shift a bit, to change some of them sometimes, or to 'play' certain practical norms against others?*" (Olivier de Sardan, 2015, p. 32). In the context of PBF, for example, this may require developing evaluation criteria that are more closely aligned with the stakeholders' realities and behaviour rather than with ideal standards, some of which are hardly achievable due to structural constraints (e.g., pre-existing infrastructure). Along this line of thought, the case of Mali might offer some interesting avenues. A PBF "*à la Malienne*" was built by adapting the PBF principles to the local context, without simply "copying and pasting" the PBF programs of other countries or introducing new structures (Seppey et al., 2017; SNV, 2013).

Implementation gaps: Overall, this doctoral study highlights how several *implementation gaps* contributed to the emergence of unintended consequences (Derthick, 1974; Pressman & Wildavsky, 1984). These gaps refer to the distance between what is planned and what is implemented. All three articles presented in the Results section revealed that the content of the intervention was formulated, but then not implemented accordingly, such that potentially essential mechanisms were left out. Some of these implementation gaps appear to have threatened the theoretical basis of the intervention. For example, our analysis of the health equity measures revealed three main implementation gaps that contributed to unintended consequences. First, committees in charge of validating the lists of indigents (called the local validation groups) were planned but not effective. Second, instructions were modified during the selection of indigents to recruit a specific number of indigents. Together, these implementation gaps contributed to the selection of individuals who did not meet the local conceptualization of indigents, ultimately leading to dissatisfaction within communities. A third implementation gap was related to the initial intention to create financial incentives that would motivate providers to treat indigents. In reality, however, the unit fees offered for services delivered to indigents were not sufficient to drive their motivation and the payments were very late. Thus, while the intervention theory for the health equity measures might have been sound, its implementation was not carried out properly. Similarly, the article on PBF in primary healthcare facilities revealed that delays in payments of subsidies and the cancellation of the quality improvement bonuses caused suboptimal planning. Facilities did not have the funds available to carry out their performance improvement plans. Going forward, local stakeholders should reflect on the reasons why these processes were not implemented as planned and determine whether it is possible to take actions to close the implementation gaps within their context.

6.4.2 Complementarity of articles for understanding the weaknesses of verification mechanisms

The articles presented in the Results section are highly complementary. Together, they reveal the intervention's weaknesses in detecting and regulating non-compliant practices such as the falsification of data. More specifically, the first manuscript revealed that providers in the selected facilities regularly falsified medical registers in their day-to-day practice to increase performance scores. PBF verifiers, who were not present during the consultations, explicitly reported the difficulty of detecting whether services were falsified because the data were modified in the original registers using standard information. In 28 months, no sanctions were given for the falsification of registers, which is consistent with common practices observed in other countries and settings (Olivier de Sardan et al., 2018). In response to this, however, some PBF stakeholders quickly ensured that the issue of falsification would be addressed through the community verifications and client satisfaction surveys. In the local discourse, community verifications were sometimes presented as an effective solution that would ensure the proper functioning of the PBF intervention and punishment of guilty parties. In reality, however, the second article raised serious questions regarding the effectiveness of community verifications and client satisfaction surveys. We found that the investigators in charge of tracing patients falsified some data and managers did not know how to interpret or use the results of the community verification. Thus, the results of the two articles should be considered together to comprehensively grasp the complexity of the issue of falsification. It is crucial to address the systematic falsification of registers because it undermines the intervention's objectives to improve transparency and accountability within the healthcare sector. This refers to what pharmacologists call "paradoxical effects", that is, interventions increasing adverse outcomes they seek to prevent (Bonell et al., 2015).

6.4.3 Predominance of undesirable unintended consequences

The unintended consequences that emerged across the three articles presented in the Results section were predominantly undesirable. This finding is consistent with other research and expert opinions on unintended consequences (Jabeen, 2018; Morell, 2018). Koch and Schulpen (2018) question whether the preponderance of negative unintended consequences in the academic literature is a bias or represents the reality that, indeed, more negative than positive side effects occur.

In this study, we deliberately attempted to capture desirable unintended consequences in order to avoid a negative bias. However, as previous authors have explained, researchers are unlikely to encounter many positive unintended effects because "*reforms trying to sell a new programme are likely to have listed and exhausted all the positive results possible*" (Weiss, 1998 cited in Jabeen 2018). I believe this was the case for PBF in some LMICs. The complex intervention model, initially developed and promoted by collaborators of the World Bank and later implemented in Burkina Faso, targets an impressive array of processes and outcomes, many of which are conveyed in the PBF toolkit and PBF trainings. Since these desirable consequences are intended by program planners, they were excluded from our analyses. In high-income countries, a literature review found that the main positive spillover effects were the strengthening of health sector governance and the improvement of staff knowledge regarding recent norms (Cashin et al., 2014). However, both of these outcomes were excluded from our analyses because they are considered to be intended changes according to the intervention model implemented in Burkina Faso and some other LMICs. A similar approach was adopted in a recent PBF study in Malawi (Petross et al., 2020). The reported unintended consequences were negative because "*positive consequences were viewed as both expected of and intended by the program*" (Petross et al., 2020, p. 13).

A second plausible explanation for the predominantly undesirable unintended consequences may lie in the nature of the intervention, given the context. The intervention design provides strong financial motivations in a low-income setting where stakeholders perceive high levels of public sector corruption and are accustomed to developing coping strategies to thrive (Transparency International, 2018). Burkina Faso obtained a score of 40 on the Corruption Perception Index, which uses a scale where zero is highly corrupt and 100 is very clean (Transparency International, 2020). In 2017, the national anti-corruption network in Burkina Faso found that 70% of people interviewed believed that corruption was frequent (REN-LAC, 2017). About 39% of respondents reported having been victims of or witnessing corruption. Corruption scandals publicly denounced by the press remained without sanctions, highlighting the lack of political will to combat the phenomenon (REN-LAC, 2017). In a context of impunity for economic crimes, social conflicts, and growing terrorist acts, people are trying to ensure their survival and save money in case a crisis erupts. Thus, it is not surprising that part of the PBF intervention was reinvented to meet the financial needs of various stakeholders and increase its relative advantage.

A third possible reason for the sum of undesirable unintended consequences found in our PBF study may be related to the types of data collection methods employed. The methods that researchers use shape the evidence that they generate. By conducting observation *in situ* over a long period of time, this study revealed information that participants do not generally speak about with outsiders. A study on the consequences of PBF that did not use direct observation over a prolonged period of time is less likely to have identified some of the undesirable consequences that emerged in this study. For example, we found that observation was much more fruitful than recorded interviews to study conflicts or the falsification of medical registers.

Placing a magnifying glass on the undesirable consequences of PBF does have one major advantage. It can help balance out the pro-PBF bias in the literature, in a context where researchers are often accused of publishing only positive results. According to Ireland (2011), the absence of evaluation on the possible negative consequences of PBF reflects a favourable bias for PBF in the literature. This bias likely arises because: 1) studies showing successful implementation are more likely to be published; and 2) published authors are often involved in the implementation of PBF (Ireland et al., 2011). Similarly, Barnes, Brown and Harman (2015) argue that the World Bank continues to engage in the reproduction of a positive bias around PBF. Where criticism exists in the literature, it is around how to make PBF work better, not a critique of the appropriateness of the model itself (Barnes et al., 2015). These authors argue that positive research on PBF is contentious because it has been used by global health institutions to justify further investment in this approach to reform health systems in Africa, while obfuscating studies that highlight its limitations. According to Rogers (2003), one way of shedding pro-innovation biases, such as the one surrounding PBF, is to conduct different kinds of studies from those in the past. "*For balance, we need a number of diffusion researches with an "anti-innovation bias" in order to correct past tendencies*" (Rogers, 2003, p. 112). In this sense, the current study has the advantage of swinging the pendulum the opposite way to produce a more comprehensive understanding of the processes and impact of PBF.

6.4.4 Absence of some unintended consequences expected from the literature

When we examine the results across the articles, there is a notable absence or lack of evidence regarding some unintended consequences that might have been expected based on the literature. First, our study showed no evidence that providers cherry-picked patients who were the healthiest (adverse selection), as found in other programs where providers were allowed to select patients (Chen et al., 2011). Second, we did not find strong evidence showing that providers, in their day-to-day practice, focused more on healthcare services that paid more, as happened in Tanzania (Binyaruka et al., 2015). In fact, providers displayed little knowledge of the specific fees paid for each type of service. The absence of these two unintended consequences is likely due to the way in which the PBF scheme was elaborated. Broad categories of healthcare services were incentivized (e.g., consultations for adults, consultations for children under age five), making it unlikely that providers would be motivated to exclude or focus on certain types of patients based on their characteristics (e.g. age, comorbidities, severe conditions). Also, the study did not detect spillover effects within the healthcare facilities under PBF. This is likely due to the comprehensive evaluation grids that scrupulously examined and incentivized multiple dimensions of quantity and quality of care. However, there is a possibility that the intervention triggered spillover effects in facilities that were not under PBF. For example, the quantity and quality of care in facilities not under PBF may have increased due to improved supervision at the district or national level as well as the reorientation of funds to non-PBF districts. Such potential spillover effects, however, were not included in this study because they would be attributable to the design of the randomized control trial rather than to the PBF intervention model per se. Lastly, we did not find that providers used coercive strategies to attract patients to healthcare facilities, as was found in Tanzania (Chimhutu et al., 2014). While some providers included in this study told women that they would be sanctioned if they missed appointments or delivered at home, we were not able to attribute this behaviour to PBF specifically. Such discourses existed before the implementation of PBF.

6.5 Strengths and limitations

As with any study, it is important to acknowledge the strengths and limitations that may affect the results and conclusions. Some study limitations were already presented in the scientific articles above, including: 1) a potential social desirability bias to portray the intervention in a positive way; 2) a potential memory bias due to the lack of data collection during staff training; 3) the limited number of healthcare facilities included in the study; 4) a focus on short-term term consequences; and 5) the lack of statistical tests conducted on the quantitative data. In this section, I deepen this discussion by reviewing Guba’s (1991) four criteria for judging a qualitative study’s trustworthiness: credibility, transferability, dependability, and confirmability. These criteria are still highly cited and considered to be the gold standard by many authors (Anney, 2014; Bourgeois, 2016; Whittemore et al., 2001). Below, I explain how I attempted to satisfy each of these criteria in the present study, despite the potential limitations.

Credibility

The credibility criterion is often compared to the concept of internal validity in quantitative research (Bourgeois, 2016). More specifically, it refers to the confidence that can be placed in the truth of the research findings. It determines whether the research findings represent plausible information drawn from the original data and constitute an adequate interpretation of the participants’ original views (Korstjens & Moser, 2018). It helps answer the following question: “*Are we facing an authentic picture of what was observed*?” (Bourgeois, 2016).

In the current study, we used several strategies recommended to increase the credibility of findings (Egon G. Guba, 1981). First, we adopted a case study design, which is recognized for its high credibility or internal validity (Contandriopoulous et al., 2005; Yin, 2009). The explanatory power of case studies comes from the depth of the analyses, rather than the high number of cases (Contandriopoulous et al., 2005).

In each healthcare facility, we spent a prolonged period of time observing people around the clock to understand the context, local practices, intervention, etc. (Egon G. Guba, 1981). As Creswell and Miller (2000) explain, this prolonged engagement in the field enabled us to build trust with participants, find gatekeepers to access sites and people, and establish rapport to make participants feel more comfortable disclosing information. It helped us overcome potential distortions produced by our presence in the field. In fact, participants displayed more authentic behaviours over time, especially as the number of facilities visited grew.

Moreover, our prolonged engagement in the field helped us control for potential history biases, which could undermine the validity of findings. We were able to consider other potential interventions or events that could influence the emergence of unintended consequences. For example, in some of the facilities included in this study, a foreign donor subsidized the implementation of an intervention aimed at increasing treatments for child malnutrition. Community-based health workers were paid to screen underweight children and refer them to local healthcare facilities for treatment. The objectives of this intervention overlapped with PBF, which also incentivized treatments for malnutrition. In other facilities, an intervention subsidized services for facility planning, which could have interfered with our interpretation of data. Thus, our prolonged presence in the field enabled us to make clear distinctions between the processes and effects of these various interventions.

Another strategy that we adopted to increase the credibility of findings is called peer debriefing (Creswell & Miller, 2000). This requires that someone familiar with the phenomenon being explored reviews the data and the research process. I regularly conducted peer debriefing with a local researcher, part of the overarching study, who collected data on the PBF implementation process in some of the primary healthcare facilities included. Conversations on the research sites, telephone calls, emails, and Skype meetings helped us validate our interpretation of some data and verify the meaning of local expressions, etc. The feedback obtained during the peer debriefing sessions also enabled me to better integrate into the local context. The local researcher, for example, helped me adapt my academic vocabulary to the local context.

A second type of peer review was conducted during the publication process in scholarly journals (Willis, 2007). Anonymous scholars, some of whom seemingly rely on different paradigms and possess a certain expertise on the topic, reviewed the manuscripts presented in the Results section and made suggestions to improve the study.

During our field work, we also collected what Guba (1981) refers to as “slice-of-life” data that can later be used to test findings and interpretations. More specifically, we collected photographs and audio recordings to help demonstrate that findings were grounded in the raw data. Some photographs and citations were presented in the scientific articles to enable readers to access and interpret the raw data.

To further increase the credibility of results, we triangulated data (Creswell & Miller, 2000). We searched for convergence among multiple sources of information (e.g., different types of participants) and methods (e.g., formal interviews, informal discussions, observations, intervention documentation) to ensure the strength of the evidence.

Lastly, the credibility of findings was tested by doing “member checks,” one of the most important strategies to test the credibility of findings (Egon G. Guba, 1981). During a face-to-face meeting and numerous email exchanges, members of the ST-FBR at the national level stated that the findings confirmed what they have been observing in the field or what they already believed was going on. Thus, I am confident that the findings are credible.

Transferability

The transferability criterion is akin to external validity in quantitative studies (Bourgeois, 2016). It assesses whether our findings have applicability in other contexts. It aims to answer the following question: “*can the results obtained ‘there’ be expected ‘here’?*” (Woolcock, 2013). Our study primarily focused on public primary healthcare facilities in rural districts in Burkina Faso. Thus, knowledge users may find that the transferability of results to other settings (e.g., urban) with different types of facilities (e.g., private facilities, hospitals) or different PBF interventions models is limited.

Nevertheless, we used several strategies to increase the transferability potential of findings. First, we sought the input of local district management teams to help us select primary healthcare facilities that were representative of other local healthcare facilities (Zombré et al., 2016).

According to the principle of similarity, this selection process should increase the likelihood that results are relevant in similar empirical settings (Contandriopoulous et al., 2005).

Second, our design included multiple facilities, which enabled us to examine whether the unintended consequences were replicated in diverse contexts (e.g., high vs. low performing facilities). We also collected data in two districts and participated in a one-week national PBF review meeting to examine whether the unintended consequences were consistent across different districts. According to the principle of robustness, the elevated number of cases and districts should decrease the risk that the results are unique to a specific context and increase the generalization potential (Contandriopoulous et al., 2005). However, some prudence is warranted because the qualitative program evaluation did note some heterogeneity between districts, which may limit the transferability of results (De Allegri, Lohmann, et al., 2019). According to this evaluation, the heterogeneity was explained by the dynamism and motivation of leaders, the presence of staff trained in PBF, the facility's baseline conditions, and how much time the contractualization and verification agencies had to coach healthcare staff (De Allegri, Lohmann, et al., 2019). We did take these factors into consideration in our analyses and gave more weight to the findings that were replicated in many settings to increase the transferability of results. For example, we found that the falsification of medical registers occurred across facilities, regardless of the level of leadership of the head nurses, the presence of staff trained in PBF, and the facilities' baseline conditions.

Third, we provided thick and rich descriptions of the context, the cases selected, the intervention model, and the participants (Creswell & Miller, 2000; Egon G. Guba, 1981). We also followed the principle of explanation by describing, as much as possible, the processes by which the PBF intervention led to unintended consequences (Contandriopoulous et al., 2005). Such information should help knowledge users evaluate the relevance and similarity between the study's context and their own environment (Egon G. Guba, 1981).

Fourth, we used quantitative data from other healthcare facilities when possible to show that some results may potentially be applicable to a wider sample (e.g., see article on indigents).

It is possible, however, that the strategy adopted to select the facilities influenced the results. The facilities were selected using a multi-stage screening procedure (see description in the Methods section). Using a mixed and participatory approach, we selected facilities with contrasting levels of performance before the intervention was implemented. In retrospect, however, the study might have revealed more unintended consequences if we had selected facilities with contrasting levels of PBF performance scores. Staff members in facilities with very high or low performance scores might have developed more or different coping strategies. However, this hypothesis suggests that the strategy that we used to select facilities led to an underestimation of unintended consequences.

Dependability

The dependability criterion is the counterpart to the concept of reliability in quantitative research (Bourgeois, 2016). It assesses whether the findings are consistent and could be repeated over time. It aims to answer the following questions: “If different instruments had been used, would the results of the study be different?” or “Would the findings be repeated consistently with similar participants in the same context?” (Anney, 2014).

To increase the dependability of results, we followed Guba’s (1981) recommendation to combine multiple methods. We were able to compare the data collected through observation, semi-structured interviews, informal discussions, etc. The strengths of one method compensated for the weakness of another. We considered results to be more dependable when similar results were found using different methods.

However, as Poupart and colleagues (1997) describe, the search for concordance between different data collection methods does not exclude tensions or contradictions between data. For example, we found that data collected during observation sessions and informal discussions sometimes contradicted the more politically correct discourses expressed during recorded interviews. These discrepancies could be explained by the pressure to present the intervention in a positive way and the fear of loss of confidentiality due to recordings.

Another factor influencing the dependability of our study concerns the instrumental shifts stemming from developing insights on the part of the investigator-as-instrument (Egon G. Guba, 1981). Over the course of our field work, we became more familiar with the local context, the PBF intervention and local practices. During this process, the unintended consequences, as well as their contributing factors, began to emerge in the data, becoming more “visible” to the eye of the researchers. Our growing understanding enabled us to refine the data collection instruments to better capture unintended consequences (Contandriopoulous et al., 2005). For example, we were able to ask more pertinent follow up questions during interviews and report more pertinent data in observation notes. It is important to note, however, that such changes in the instruments throughout this multiple case study may have caused a bias related to the measurement of effects (Contandriopoulous et al., 2005). I believe it may have led to underestimation of unintended consequences in the first healthcare facilities visited (especially the two pilot cases). The absence of an unintended consequence in one of the first cases does not imply that it was not present: I may simply not have been able to detect it as a new observer.

The timing of the data collection is likely to have influenced the dependability of results. In Burkina Faso, the utilization of healthcare services varies considerably depending on the season, thereby influencing the healthcare providers' workload. Thus, the collection of data during the peak of the rainy season may have minimized or amplified unintended consequences as well as triggered new ones. For example, staff members may have developed new coping strategies. Similarly, our qualitative data collections took place before the introduction of the national free healthcare policy for maternal and child services. This policy influenced some of the unit fees paid for targeted services. According to the qualitative program evaluations, the introduction of this policy in 2016 overshadowed and interacted with PBF both positively and negatively (De Allegri, Lohmann, et al., 2019). Thus, it is possible that the interaction between the PBF intervention and the free healthcare policy would have influenced the unintended consequences observed.

Confirmability

The confirmability criterion is equivalent to the concept of objectivity in quantitative research (Bourgeois, 2016; Egon G. Guba, 1981). It raises concerns for neutrality by posing the following question: "To which degree are the findings a function of participants and conditions of the inquiry and not of the biases, motivations, interests, perspectives of the inquirer?" (Egon G. Guba, 1981). A researcher bias can occur if the person performing the study influences the results in order to portray a certain outcome.

In this study, for example, the mere formulation of the research question was based on the researchers' underlying assumption that complex interventions such as PBF could trigger unintended consequences. The research team's expectations could have influenced the results (Contandriopoulous et al., 2005). However, we used a number of techniques to minimize the risk of a researcher bias.

First, we practised reflexivity by keeping journals in which personal reflections, events that happened in the field, and potential biases were recorded on a daily basis (Anney, 2014). The content of the reflexive journal was coded using qualitative data analysis software to keep track of potential biases and consider them during the analysis.

Second, we collected data from a variety of perspectives and using a variety of methods so that any preconceived notions we had could be tested as strenuously as possible. Our prolonged engagement in each facility provided us ample time to test our biases and consider alternative possibilities. To preserve the neutrality of the study, we made conscious efforts to remain open-minded and value-free during the entire data collection.

Third, we presented evidence to substantiate. For example, we used citations from interviews, excepts from field notes, and photographs to show that the results were supported. Unfortunately, many citations and excerpts had to be deleted from the original manuscripts due to the journals' word limits, but they are available upon request.

Fourth, our research team was composed of researchers from different backgrounds and disciplines to balance out potential predispositions (Egon G. Guba, 1981). These researchers had an in-depth understanding of the context and the intervention implemented. They led or participated in larger studies examining the process and impact of the intervention. Throughout the study, we were able to discuss and compare our interpretations of the data and conclusions (Bourgeois, 2016).

Fifth, face-to-face meetings and email exchanges with local stakeholders confirmed that interpretations of the findings were derived from the data rather than simply figments of the researchers' imagination.

6.6 Knowledge translation

6.6.1 Strategies to disseminate results

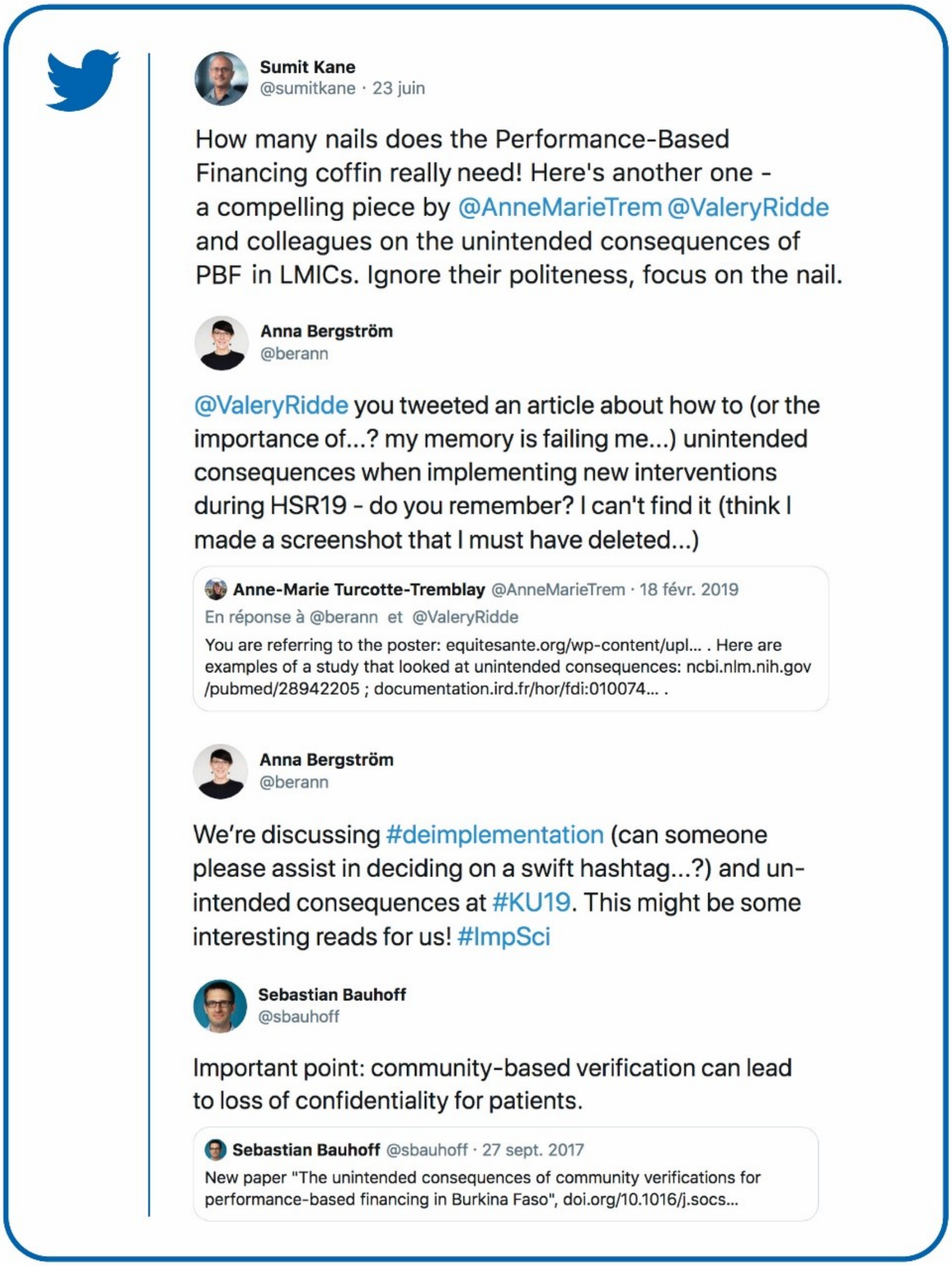
We adopted complementary strategies to disseminate the research findings. As recommended by the Canadian Institutes of Health Research (CIHR), various modes of communication were used to exchange with different types of audiences, including the scientific community and local and international knowledge users who can influence health policies (Canadian Institutes of Health Research, 2016). Each strategy is presented below.

Publications in open access: The content of this thesis was disseminated through five scientific articles published in peer-reviewed journals. The journals selected (e.g., Social Science & Medicine, International Journal for Equity in Health) targeted international readerships composed of researchers, policy-makers, and health practitioners. All articles were published in open access to reach a broad audience and increase research uptake, especially in LMICs, where access to subscription-based journals may be limited. We used dissemination bursaries, research grants, and personal funds to cover the publication charges for open access, which amounted to more than CAD 10,266. At the time of writing this thesis, the articles published had been cited more than 80 times, which highlights their high visibility and pertinence. According to available metrics, the three articles published through the publisher BioMed Central had been accessed more than 6,788 times. A World Bank Senior Health Specialist working on PBF in West Africa reported having read some of the articles, showing that high-profile stakeholders have been able to access them.

Conference presentations: We presented the study in provincial, national, and international conferences that bring together the full range of actors involved in health systems research. In total, we conducted seven oral and seven poster presentations to share our theoretical approach, methods, empirical findings, and ethical challenges encountered.

Internet and social media: Social media and the Internet are emerging as important tools to exchange knowledge (Choo et al., 2015; Wetsman, 2019). Thus, we did not hesitate to immediately share our research outputs (e.g., PowerPoint presentations and posters) online to reach a broader audience while counteracting the slow publication process. We used popular social media platforms such as Twitter and Facebook to promote their diffusion to health experts, researchers, and the general public. Through Twitter, for example, our team was able to share articles with high profile stakeholders from the World Health Organization as well as other researchers interested in our research topic or angle of analyses (see **Figure 15** below). We also shared key results with specific networks of knowledge users (e.g., SHAPES: social science approaches for research and engagement in health policy & systems, the Pan American Health Organization)

Figure 15. Examples of tweets on research outputs[13]



Sumit Kane
@sumitkane · 23 juin

How many nails does the Performance-Based Financing coffin really need! Here's another one - a compelling piece by @AnneMarieTrem @ValeryRidde and colleagues on the unintended consequences of PBF in LMICs. Ignore their politeness, focus on the nail.

Anna Bergström
@berann

@ValeryRidde you tweeted an article about how to (or the importance of...? my memory is failing me...) unintended consequences when implementing new interventions during HSR19 - do you remember? I can't find it (think I made a screenshot that I must have deleted...)

Anne-Marie Turcotte-Tremblay @AnneMarieTrem · 18 févr. 2019
En réponse à @berann et @ValeryRidde
You are referring to the poster: equitesante.org/wp-content/upl... . Here are examples of a study that looked at unintended consequences: ncbi.nlm.nih.gov/pubmed/28942205 ; documentation.ird.fr/hor/fdi:010074... .

Anna Bergström
@berann

We're discussing #deimplementation (can someone please assist in deciding on a swift hashtag...?) and un-intended consequences at #KU19. This might be some interesting reads for us! #ImpSci

Sebastian Bauhoff
@sbauhoff

Important point: community-based verification can lead to loss of confidentiality for patients.

Sebastian Bauhoff @sbauhoff · 27 sept. 2017
New paper "The unintended consequences of community verifications for performance-based financing in Burkina Faso", doi.org/10.1016/j.socs...

---

[13] Written consent was obtained from the authors to share the Tweets.

Training the next generation: I was invited to present the study to global health graduate students in four masters' level classes as well as two multidisciplinary seminars for PhD and postdoctorate trainees in universities across the country. These served to raise the awareness of the next generation of public health actors regarding the unintended consequences of complex health interventions.

Interactions with decision-makers: Our team collaborated with local decision-makers at various stages of the study to produce research findings that are more likely to be relevant to and used by knowledge users. We contacted representatives from the ST-FBR and district management teams on different occasions to either refine research questions, select the cases (i.e., health facilities), discuss potential outcomes of interest, interpret findings, or disseminate key results. For example, before leaving Burkina Faso, I organized a meeting with a representative from the Ministry of Health to present and discuss preliminary findings. I also presented the findings at the Ministry of Foreign Affairs in France. The seminar occurred at an important time when France's regional advisors in global health were questioning themselves on their position regarding PBF. My contribution helped provide evidence-based information on the unintended consequences of PBF, which nicely complemented an earlier presentation from the World Bank's Senior Director of Health, Nutrition and Population.

6.6.2 Article 6: An ethical challenge during the local dissemination of results

One of the challenges encountered during the dissemination of results was protecting the confidentiality of participants. The article below presents a reflection on this topic. It was published as part of a special issue organized by the Global Health Research Capacity Strengthening Program (GHR-CAPS).

6.4.2 Article 6: A reflection on the challenge of protecting confidentiality of participants while disseminating research results locally

A reflection on the challenge of protecting confidentiality of participants while disseminating research results locally

Anne-Marie Turcotte-Tremblay and Esther Mc Sween-Cadieux

BMC Medical Ethics 2018, 19, S1

DOI https://doi.org/10.1186/s12910-018-0279-0

Title: A reflection on the challenge of protecting confidentiality of participants while disseminating research results locally

Anne-Marie Turcotte-Tremblay[a,b*]

Esther Mc Sween-Cadieux[c]

[a] University of Montreal Public Health Research Institute, Montreal, QC, Canada

[b] School of Public Health, University of Montreal, Montreal, QC, Canada

[c] Department of Psychology, University of Montreal, Montreal, Qc, Canada

* Corresponding author: Anne-Marie Turcotte-Tremblay

Contributions of authors:

<ins>Anne-Marie Turcotte-Tremblay</ins> conceived the original idea, developed the reflection and wrote the first draft of the manuscript.

<ins>Esther Mc Sween-Cadieux</ins> critically reviewed the manuscript to deepen the analysis.

ABSTRACT

Background: Researchers studying health systems in low-income countries face a myriad of ethical challenges throughout the entire research process. In this article, we discuss one of the greatest ethical challenges that we encountered during our fieldwork in West Africa: the difficulty of protecting the confidentiality of participants (or groups of participants) while locally disseminating results of health systems research to stakeholders.

Methods: This reflection is based on experiences of authors involved in conducting evaluative research of interventions aimed at improving health systems in West Africa. Our observation and collaboration with the research projects' stakeholders informed our analysis. Examples from two research projects illustrate the issues raised.

Results: We found that in some cases there is a risk that local stakeholders may be able to identify research participants, or at least groups of participants, during the dissemination of results, even if they are anonymized. Four factors can interact and influence this challenge: 1) hierarchical structure, 2) small milieu, 3) immersion in a few sites, and 4) vested interests of decision-makers. For example, local stakeholders can sometimes find out when and where the data were collected. Moreover, health systems, especially rural healthcare centres, in West African countries can be small settings, so people often know each other. Some types of participants have unique characteristics or positions in the health system that may make them more easily identifiable by local stakeholders familiar with the environment. We identified a number of potential strategies that can help researchers minimize this difficulty and improve ethical research practices. These strategies pertain to the development of the study design, the process of obtaining informed consent, the dissemination of results, and the researchers' reflexivity.

Conclusion : Researchers must develop and adopt strategies that enable them to respect their promise of confidentiality while effectively disseminating sometimes sensitive results. Reflections surrounding ethical issues in global health research should be deepened to better address how to manage competing ethical responsibilities while promoting valuable research uptake.

BACKGROUND

Researchers studying health systems in low-income countries (LICs) face a myriad of ethical challenges throughout the entire research process. Distinctive features of this field of research can colour ethical issues, such as the balancing of risks and benefits for individuals, groups, and communities [1, 2]. In this article, we discuss one of the greatest ethical challenges we encountered during our fieldwork in West Africa: the difficulty of protecting the confidentiality of participants, or groups of participants, while locally disseminating results of health systems research. On one hand, researchers have to preserve the confidentiality of participants, or groups of participants, especially when the research focuses on sensitive issues such as negative or perverse effects of interventions. Breaches of confidentiality could harm participants, hinder the trust relationship between participants and researchers, and even hurt the reputation of a group or community. On the other hand, global health researchers are increasingly encouraged to disseminate research findings to local stakeholders such as decision-makers, managers, practitioners, and community members. Our experiences have sparked in us a growing concern that local stakeholders may be able to identify research participants, or at least groups of participants, during the dissemination of results. For example, local stakeholders can sometimes find out information on when and where the data were collected. In West African countries, health systems—especially rural healthcare centres—can be rather small settings, so people often know each other. Some types of participants have unique characteristics or positions in the health system that render them more easily identifiable by local stakeholders familiar with the environment. Together, such factors make it difficult to protect participants' confidentiality during the dissemination of results locally. The objectives of this reflection are to 1) show how researchers conducting health systems research in LICs can experience difficulty in protecting the confidentiality of participants locally and 2) present some potential strategies to minimize this difficulty.

METHODS

Past experiences

Our experiences conducting evaluative research on interventions aimed at improving health systems in West Africa have contributed to this reflection. One of our research projects was an analysis of the unintended consequences of a performance-based financing intervention in Burkina Faso. In this intervention, healthcare workers were paid for the quantity and quality of healthcare services they provided, which resulted in certain fraudulent practices being committed to increase financial gains [3]. We conducted prolonged field work in healthcare centres to collect data through observation, interviews, and discussions with participants (e.g. healthcare workers, community leaders, verifiers, healthcare users). Before leaving the host country, we wanted to disseminate the results of the study widely to local stakeholders (e.g. district management teams, representatives of the Ministry of Health and participants) while respecting ethical norms such as confidentiality. However, it became obvious that reporting research findings to local stakeholders could result in breach of confidentiality.

Our reflection was also inspired by a research project that evaluated the implementation and effects of an intervention aimed at promoting the use of research findings to influence health practices and policies in Burkina Faso. This unique intervention was implemented at the local level and only involved a few stakeholders (e.g. local and international researchers, local consultants, representatives of non-governmental organizations, community associations, and district-level decision-makers. As several difficulties arose during the intervention's implementation process, that research shed light on what led to its failure. Thus, that study also posed some challenges in terms of identifying the stakeholders in charge of the intervention when the results were disseminated, as will be further explained below.

In hindsight, we find that our reflexive process was iterative and corresponded to the phases of Schön's reflective practitioner model, as described by Tremblay and colleagues [4]. First, we went through the assessment phase, in which we formulated an initial understanding of a new and problematic situation, that is, the difficulty of protecting the confidentiality of participants while disseminating research results locally. In one study, for example, this issue became apparent and was explicitly discussed during the development of study protocols. Then, in the action phase, we tested this understanding and its implications in the field, for example, during data collection and dissemination of results. Lastly, in the reassessment phase, we revisited the terms of the problem, looked at it critically, and proposed solutions. The present article synthesizes this last stage.

RESULTS AND DISCUSSION

An ethical responsibility to protect confidentiality

One of our major and ongoing concerns was the need to protect the confidentiality of participants[14]. The ethical duty of confidentiality refers to researchers' obligation to safeguard entrusted information [5, 6]. Breaches of confidentiality can have negative repercussions on participants if other stakeholders are able to identify them [7]. For example, during the dissemination of research results, if a supervisor were to discover that a specific type of health worker participating in a study engaged in prohibited behaviour, such as falsifying consultation registers, the supervisor could be tempted to take actions against them. These participants could forfeit future job opportunities, lose professional credibility, or become socially ostracized within their environment. Breaches of confidentiality could also harm the reputation of a specific community by increasing stigmatization towards them (e.g. prejudice, marginalization). As researchers, we experienced difficulties in fully grasping or anticipating the nature and amplitude of such potential repercussions, due to cultural differences and our limited understanding of the complex social structure. For example, during the risk–benefit assessment of disseminating results locally, it was difficult to determine which specific results could lead to a breach of confidentiality, to what extent such a potential breach could harm participants, and how local stakeholders could react. Thus, we feel serious consideration should be given to the protection of confidentiality upstream.

Ensuring confidentiality is also essential to build trust relationships with participants [5]. Without the assurance of confidentiality, they might refuse to share data or hide data that are important to answer research questions, especially when the study focuses on sensitive issues (e.g. hidden behaviours, controversial views, perverse effects). Lack of trust from participants could increase the risk of biases in research (e.g. social desirability).

---

[14] There are some exceptions regarding the participants' willingness to be identified [7]. For example, in some forms of participatory action research, participants want their voices to be widely heard and insist that their identity be revealed.

In the context of global health research, respecting confidentiality is crucial because some participants already feel apprehensive towards researchers, who tend to be outsiders in relation to the local context. Indeed, global health researchers, whether they reside in West Africa or elsewhere, often come from different backgrounds (e.g. nationality, culture, socio-economic status)[15] from the participants and have different mandates from each other. In our experiences, research participants were sometimes skeptical of our true affiliation and mission in their organization. Our presence was sometimes perceived as surveillance from international development agencies or program funders to evaluate an intervention. Communities who have negative experiences with an outside researcher may be less welcoming towards future researchers. In this respect, protecting confidentiality is essential to promote people's openness towards global health researchers in participating communities and to facilitate future research.

Furthermore, power asymmetries can exist between researchers and communities as well as between communities and actors in the health system. Some participants may feel pressured to take part in a study or may unintentionally reveal sensitive information to researchers. Atchessi et al. [8] found that hierarchical authority can interfere with free and informed consent in global health research. Thus, protecting confidentiality is paramount to avoid causing undue harm to vulnerable populations who do not necessarily have sufficient means to protect their own interests. Although participatory action research may be a way to help rebalance power inequalities in research, it can lead to a range of particular ethical issues [9].

An ethical responsibility to disseminate results

Dissemination of results is a "planned process that involves consideration of target audiences and the settings in which research findings are to be received and, where appropriate, communicating and interacting with wider policy and health service audiences in ways that will facilitate research uptake in decision-making processes and practice" [10]. The dissemination of research findings is considered to be a researcher's ethical obligation. Ethics committees and funding agencies increasingly require that researchers conduct knowledge translation activities,

---

[15] The differences in characteristics between global health researchers and participants are partly related to the fact that most global health centres are located in high-income countries, that the field transcends national boundaries, and that it must reach across socio-economic boundaries to address health equity and disparities [16].

including dissemination of results [2, 11, 12]. The Declaration of Helsinki stipulates that "researchers have a duty to make publicly available the results of their research on human subjects" (World Medical Association, 2013, item 36). For Canadian researchers, the Tri-Council Policy Statement [5] states that researchers should provide copies of publications and research reports to organizations that are best suited to disseminate the results within participating communities. According to that statement, this is especially important in settings where the results are not easily accessible, such as LICs. However, the definition of community is somewhat labile, so it was sometimes ambiguous for us whether we had the responsibility to disseminate results at the village, district, region, or even country level. Still, we felt our duty was to report our research results in appreciation of the community and the research participants' involvement, and we thought our results could improve local practices and policies in health.

Difficulties in protecting confidentiality while collaborating with local stakeholders

We found it difficult to protect the confidentiality of participants while collaborating with local stakeholders, particularly for the dissemination of results. These local stakeholders included representatives from the Ministry of Health, intervention implementers, district management teams, local leaders, participants, etc. We identified four factors that may interact and influence this challenge: 1) hierarchical structure; 2) small milieu; 3) researcher immersion in one or just a few sites; and 4) vested interests of stakeholders. These four factors emerged from our experience, but the list is not intended be exhaustive. While these factors are not ethical problems per se, they may raise ethical concerns for researchers trying to protect confidentiality while collaborating with local stakeholders. We use examples from our research experiences in West Africa to illustrate the issue.

Hierarchical structure: Global health researchers cannot conduct research within the health system without informing various local authorities of their presence. Both local and international researchers must conduct courtesy visits to the Ministry of Health and/or the health district offices to inform decision-makers that research will be conducted within their jurisdiction. For example, to collect data within healthcare centres, authorization has to be obtained from the person responsible for each healthcare centre in question and from authorities at the district level. These courtesy visits are important because they significantly facilitate access to the research sites. They enable superiors (e.g. chief medical officer in a district) to inform their subordinates (e.g. chief head nurses in healthcare centres) that a study will be conducted and that they can collaborate, if they wish to do so. These visits are also an important step in building collaboration between the researchers and decision-makers to increase the relevance and use of results by stimulating their interest and allowing them to provide input into research questions, methods, etc. However, during these visits, local stakeholders inevitably ask researchers what healthcare centres they plan to visit. Thus, when the results are disseminated locally, these stakeholders may be more easily able to identify research sites and participants. Even if participants' identities are masked during results dissemination, local stakeholders may be able to infer the likely source.

Small milieu: Many West African countries (e.g. Benin, Burkina Faso) are relatively small in terms of population size and distances. Consequently, actors within a health system tend to know each other. To promote primary healthcare, small centres have been established across many West African countries [13]. Due to limited resources, many of these centres have only about seven workers, with distinct roles. Moreover, many roles or positions, especially at the higher levels of the health system, are distinctive or singular (e.g. director, supervisor, program planner). For example, the performance-based financing intervention we studied in West Africa only had two supervisors in the district who travelled to each healthcare centre to count the quantity of healthcare services delivered. Thus, results relevant to that aspect of the intervention could be traced back to them more easily than could more generic results relevant to a wide array of actors. In such a small milieu, staff from small healthcare centres or with distinctive roles are more easily identifiable when results are disseminated locally. Because these types of participants have access to particular information and details, it can be difficult to simply aggregate their data with the rest of the data to protect their confidentiality. Thus, we were often concerned that, even if researchers attempted to protect the participants' confidentiality, local stakeholders who knew the context, its actors, and past events might have been able to infer the likely source. As Richards et al. [7] explain, even after protocols for anonymization are applied, "quotations, speech mannerisms and context may provide enough information for participants to be identified" by local stakeholders, and it is not always easy to predict which data will lead to identification.

Researcher immersion in one or a few sites: Studying health systems in global health can require long immersions in the field to understand the behaviours, discourses, and obstacles that emerge in real life. Prolonged fieldwork can be useful, especially for researchers coming from a different background, to develop a more profound understanding of the local context, cultures, norms, etc. It is also useful for building relationships and trust between researchers and local actors, which lead to more authentic behaviours and discourses, thereby reducing potential biases. However, as long-term immersion in an organization (e.g. a healthcare centre) is time-consuming, it significantly limits the number of research sites a global health researcher can target to collect data. With fewer research sites, it can be easier for local stakeholders to make links between specific data collection sites or participants once the study results are disseminated.

Moreover, prolonged immersion in just a few healthcare centres means that local stakeholders, such as members of the district medical team, supervisors, and healthcare workers from other organizations, can spontaneously witness a researcher collecting data during observation sessions or interviews. Thus, they can know from whom the researcher collected data for the study. Again, it may be easier for these local stakeholders to make links between the results and participants during the dissemination of results, especially since they know the context well.

Vested interests of stakeholders: Local stakeholders sometimes have vested interests in promoting interventions such as performance-based financing. The salaries of those who are employed to manage or implement an intervention sometimes depend on its success. Local stakeholders may be wary of or intrigued by independent researchers who evaluate interventions, because a study's results could influence decisions to pursue or renew an intervention's funding. Studies on sensitive topics (e.g. unintended consequences of an intervention or the causes of its failure) may be more threatening for local stakeholders, as they may reveal hidden information. If their interests are at stake, local stakeholders may be more likely to keep track of a researcher and the data collection process as it unfolds. For example, they can seek information in their network on where data are being collected. During our own research, we found it difficult to hide a researcher's tracks during the data collection process. Ultimately stakeholder attention renders it more difficult for the researcher to hide the identities of participants during the dissemination of results.

Potential strategies

Beyond the usual method of anonymizing data, we found there is not enough discussion on how researchers can better protect the confidentiality of participants or groups of participants in different contexts while disseminating results to local stakeholders [14]. Based on our experience, we have identified a number of potential strategies to help researchers in this endeavour:

Strategies related to this study design:

- Adopt a design that allows the researcher to have a sufficient number of sites to protect confidentiality, while still being able to spend enough time in each site to develop a profound understanding of the context. The design can include primary cases with longer periods of fieldwork and secondary sites with shorter stays in other regions to “muddy the waters” and avoid stigmatizing specific groups of individuals.
- For qualitative studies, nest the study within a larger study to triangulate results with other locations and cover the researcher’s data collection sites.
- For quantitative studies, include a sufficient sample size to aggregate data for subgroups of participants.

Strategies related to informed consent

- Discuss risks of breaching confidentiality with the participants before obtaining informed consent.
- Engage in a dialogue with participants to determine whether they are comfortable with the results—anonymized—being shared with local stakeholders. However, this may bias the results of the study if participants subsequently choose to share mainly positive results and to conceal negative results.

Strategies related to the dissemination of results

- Discuss the issue of confidentiality with stakeholders before the dissemination of results and have them sign a confidentiality agreement.

- Target specific stakeholders with whom to share results locally. After due ethical consideration, researchers may decide it is in the best interests of the local community and future research to avoid presenting results to specific actors (e.g. immediate supervisors of participants). Instead, they may decide to present results to higher-level stakeholders for whom it may be more difficult to make direct links between the results and the participants and who would have the ability to use the results constructively. This strategy requires an in-depth risk–benefit assessment, as it may go against best practices in ethics and knowledge translation.
- Present results in a more general manner. When in doubt, researchers can omit information that may potentially be indirectly identifying. For multiple case studies in different healthcare organizations, researchers can choose to report only cross-case analyses [15].
- Reflect on how to describe the context in a way that protects collective confidentiality but is detailed enough to consider external validity. Researchers can attempt to hide the study's location, such as the country, district, or healthcare centre. However, preserving the anonymity of the case 1) prevents people from recollecting previous information about the case when interpreting it and 2) makes the entire case more difficult to review [15]. Contextual data are often an essential component of the analysis and interpretation [7].
- Wait for some time before disseminating results to help protect confidentiality, relying on the high mobility of stakeholders within the health system. In some cases, ongoing dissemination of results throughout the research process (e.g. in developmental evaluation) may have to be avoided to make it easier to hide the researchers' trail, although this may go against best practices in knowledge translation.
- Collaborate with a local knowledge broker to disseminate results to stakeholders. This third party may be more apt to disseminate sensitive results in a socially acceptable manner while maintaining participants' confidentiality. The knowledge broker could also organize dialogues and act as a mediator to address these sensitive issues.

Strategies related to the researchers' reflexivity

- Discuss ethical concerns and potential strategies with other colleagues who understand the context to find a solution that is adapted to the case and meets local needs.
- To orient action, continuously evaluate the risks and benefits not only for individual participants but also for social groups. The researchers' understanding of risks and benefits may change over the course of the study.
- Use practical judgment and reflexivity to develop strategies that are adapted to the study's context. Codes of practice cannot replace practical judgment and reflexivity, especially in the context of qualitative research on health services [7]. Several models on reflexivity can be useful for researchers to engage in this process [4].

CONCLUSION

Researchers conducting evaluative research of health systems in LICs sometimes find themselves in an ambiguous situation wherein their research results could improve health interventions, but the dissemination of these results could have negative consequences for the participants. For example, in this reflexive analysis, we have shown how having to obtain permission from high-level authorities to collect sensitive data in a few small facilities that local stakeholders knew well rendered it more difficult to preserve the confidentiality of participants during results dissemination. While this ethical issue may not be exclusive to health systems research in global health, we repeatedly found that it can be a challenge in this field. This issue may be relevant for any researcher working within a context of high inter-knowledge where people know each other and where it is difficult to separate the research setting from the results dissemination setting [14]. Given the above, what positions and actions should researchers take when faced with the conflicting imperatives of a) disseminating results to improve policies and practices and b) protecting individuals and groups at the local level? Researchers must develop and adopt strategies that enable them to respect their promise of confidentiality while effectively disseminating results that can sometimes be sensitive. Future research should examine the strategies that global health researchers from different methodological traditions are adopting in the field to reconcile both obligations. Moreover, research should attempt to better understand how 'confidentiality' is understood and operationalized as a concept in the sociopolitical and cultural contexts of these Western African countries. Going forward, reflections surrounding ethical issues in global health research should be deepened to better address how to manage competing ethical responsibilities while promoting valuable research uptake.

REFERENCES

1. Hyder AA, Pratt B, Ali J, Kass N, Sewankambo N. The ethics of health systems research in low- and middle-income countries: a call to action. Glob Public Health. 2014;9:1008–22.
2. Pratt B, Ali J. How can we strengthen "ethics" in health systems research? Health Systems Global. 2016. http://go.shr.lc/2fIPe02. Accessed 1 Dec 2017.
3. Turcotte-Tremblay A-M, Gali-Gali Idriss Ali, De Allegrie M, Ridde V. The unintended consequences of community verifications for performance-based financing in Burkina Faso. Social Science & Medicine. 2018;191: 226-36.
4. Tremblay M-C, Richard L, Brousselle A, Beaudet N. Learning reflexively from a health promotion professional development program in Canada. Health Promot Int. 2014;29:538–48.
5. Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada. Tri-council policy statement: ethical conduct for research involving humans. 2014. http://www.pre.ethics.gc.ca/pdf/eng/tcps2-2014/TCPS_2_FINAL_Web.pdf. Accessed 1 Dec 2017.
6. World Medical Association. Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310:2191–4.
7. Richards HM, Schwartz LJ. Ethics of qualitative research: are there special issues for health services research? Fam Pract. 2002;19:135–9.
8. Atchessi N, Niamba L, Zombré D. Le principe d'autonomie individuelle et le respect de la hiérarchie : La situation particulière des agents de santé participant à une recherche au Burkina Faso. BioéthiqueOnline 2013;2. http://www.bioethiqueonline.ca/2/22. Accessed 7 Oct 2017.
9. Hunt MR, Gogognon P, Ridde V. Ethical considerations related to participation and partnership: an investigation of stakeholders' perceptions of an action-research project on user fee removal for the poorest in Burkina Faso. BMC Med Ethics. 2014;15:13.

10. Wilson PM, Petticrew M, Calnan MW, Nazareth I. Disseminating research findings: what should researchers do? A systematic scoping review of conceptual frameworks. Implement Sci. 2010;5:91.

11. Canadian Institutes of Health Research. Knowledge translation at CIHR. 2016. http://www.cihr-irsc.gc.ca/e/29418.html#1. Accessed 1 Dec 2017.

12. Turcotte-Tremblay A-M, Fregonese F, Alam N. Ethical considerations of dissemination and restitution of findings in global health research. BioéthiqueOnline. 2013;1–4.

13. World Health Organization, UNICEF. Primary health care : report of the International Conference on Primary Health Care, Alma-Ata. USSR; 1978. http://www.who.int/publications/almaata_declaration_en.pdf. Accessed 1 Dec 2017.

14. Béliard A, Eideliman J-S. Au-delà de la déontologie. Anonymat et confidentialité dans le travail ethnographique - document. In: Les politiques de l'enquête. La découverte. Alban Bensa & Didier Fassin; 2008. p. 123–41. https://hal-descartes.archives-ouvertes.fr/hal-01241977/document. Accessed 10 Oct 2017.

15. Yin RK. Case study research: design and methods. 4th ed. Thousand Oaks: SAGE Publications; 2009.

16. Beaglehole R, Bonita R. What is global health? Glob Health Action. 2010;3 https://doi.org/10.3402/gha.v3i0.5142.

ACKNOWLEDGEMENTS

The authors are grateful to the organizing committee and participants of the GHR-CAPS winter school for their comments on an earlier version of this manuscript. Special thanks go to Slim Haddad, Béatrice Godard, Daniel Weinstock, Thomas Druetz, Lara Gautier, and Emmanuel Sambieni. The authors also thank all of the research team members and other stakeholders with whom they have collaborated.

FUNDING

Publication of this manuscript is sponsored the Global Health Research Capacity Strengthening Program (GHR-CAPS) funded by the Canadian Institutes of Health Research (CIHR) (Strategic Training Initiative in Health Research (Grant # 96123)). The authors also thank the Quebec Population Health Research Network (QPHRN) for its contribution to the financing of this publication. AMTT received a doctoral training bursary from the Canadian Institutes of Health Research (CIHR) and a research award from the International Development Research Centre (IDRC). She was also funded by the Strategic Training Program in Global Health Research, a partnership of the Canadian Institutes of Health Research and the Québec Population Health Research Network. EMC received a doctoral training bursary from the Fonds de Recherche du Québec – Santé.

AVAILABILITY OF DATA AND MATERIALS

This section is not applicable because the manuscript does not present study data.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This section is not applicable because the manuscript does not report results of studies.

CONSENT FOR PUBLICATION

This section is not applicable because the manuscript does not contain data from any individual person.

COMPETING INTERESTS

The authors declare that they have no competing interests.

6.7 Practical Implications

6.7.1 Avenues for policy

The findings of this doctoral study can play an important role in shaping high-level decision-making. Each article resulted in key recommendations, summarized in **Box 3**.

In LMICs, PBF is often presented as a way of injecting much-needed money into under-funded healthcare systems to increase staff income, equipment, etc. (Witter et al., 2019). While any health financing intervention is likely to entail some unintended consequences, the severity of the unintended consequences found in this study raises serious concerns about the PBF intervention under study. Decision-makers should consider these results to determine whether to pursue PBF as the best investment to strengthen healthcare systems in their context or rather to proceed with de-implementation (van Bodegom-Vos et al., 2017). Such considerations should incorporate the overall body of scientific literature, including other process and impact evaluations, which have revealed important implementation challenges and mixed effects in Burkina Faso (De Allegri, Lohmann, et al., 2018, 2019; Ridde et al., 2017b; Zizien et al., 2019). I call for careful consideration of the current findings before adopting, pursuing, or scaling up the intervention. Precaution is justified, given that reasonable scientific evidence has demonstrated that the intervention can lead to significant undesirable consequences on health systems.

The results of this study can also be used to refine the intervention model and its implementation. Project planners do have some control over the contributing factors that influenced the unintended consequences identified, such as the selection and pricing of performance indicators and the design of management tools. Thus, the results can be used to devise and test new strategies to minimize undesirable consequences and maximize desirable ones.

The results also demonstrate the utility of placing unintended consequences on our radar before, during, and after the implementation of interventions. As suggested by Mittlemark (2014), attempts to make changes in complex systems should be preceded by consideration and anticipation of possible effects that are not intended by the intervention. For example, Bonell et al. (2015) developed a new process by which evaluators develop “dark logic models” to guide the evaluation of potential harms and underlying mechanisms. These can inform empirical evaluations. Furthermore, our results highlight the added value of actively and continuously monitoring the unintended consequences of PBF schemes and other types of health interventions. This can help evaluators gain a more comprehensive understanding of the overall effects of interventions and may help program planners respond in a timely manner. The fates of lives, aid funds, and careers depend on how the unintended consequences are addressed.

Box 3. Summary of key recommendations for policy

- Consider the breadth and scope of unintended consequences before pursuing or scaling up PBF interventions
- Refine the PBF intervention to transfer subsidies and bonuses on time, ensure that their distribution is perceived as equitable among actors involved, improve the staff members' internalization of quality standards, adapt performance indicators to the local context, and seek independent PBF auditors.
- Be careful of incentives and performance pressures that can encourage the deliberate and systematic falsification of medical registers.
- Be wary of the potential false sense of security created by verification mechanisms in which everyone has an incentive to report positive results.
- Adequately calculate the real costs of treating indigents with medications to ensure that the unit fees paid do not place providers in a situation of conflict of interest where they must choose between providing needed medications to indigents and protecting their own financial interests.
- Adequately calculate the real costs of treating indigents with medications to ensure that unit fees replace the facilities' lost revenues.
- Avoid quick action without enough preparation for important activities such as the selection of indigents.
- Increasingly monitor desirable and undesirable unintended consequences of complex interventions to gain a more comprehensive understanding of their overall impact on health systems.
- Develop and test strategies to avoid or minimize some of the undesirable unintended consequences and to promote desirable ones in a timely manner.
- Address stakeholders' concerns on whether the intervention provides good value for money and is sustainable.
- Determine whether the overall costs and benefits of the intervention justify the wide array of undesirable consequences.

6.7.2 Avenues for research

Given the presence of important unintended consequences, this doctoral research invites further inquiry. First, there is a need to replicate this research protocol in other settings. Different PBF models are simultaneously being implemented in high-, middle-, and low-income countries. The ways in which these interventions are implemented vary depending on the stakeholders and contexts. Thus, it would be useful to determine whether similar or different unintended consequences emerged in other settings in order to learn from a broad range of experiences.

Future research could also expand and deepen the body of knowledge on the unintended consequences of PBF. Many core and ancillary components of the intervention model remain understudied (e.g., counter-verifications, coaching). It would also be beneficial to assess the unintended consequences of PBF at different levels of healthcare systems. Thus far, few studies have examined the implementation process and consequences of PBF at the district and national levels. For example, it is possible that international exchanges surrounding PBF, as reported by Gautier (2019), could lead to the adoption of higher accountability and integrity standards at the national level. Such effects have been found in other studies on foreign aid (Dijkstra, 2018). Moreover, the evaluation criteria used to assess the performance of organizations at the district and national levels are completely different, which could result in other types of unintended consequences.

The evolution of unintended consequences in the long term could also be assessed. In Burkina Faso, PBF was often described as “dynamic” and subject to changes as needed. Thus, it would be useful to examine the different unintended consequences that may emerge as the intervention and context evolve over time. For example, researchers could further examine what happens to a healthcare system after a PBF project’s external funding is suspended or complete. This is consistent with Rogers’ (2003) recommendation to include the notion of time in diffusion research.

Future research on PBF could also explore practical strategies to respond to unintended consequences. How can we react quickly and positively to ameliorate untoward effects when these are observed (Mittelmark, 2014)? This enterprise could begin with an examination of the mechanisms that program planners have already considered or used to minimize undesirable unintended consequences and promote desirable ones.

Beyond PBF, the scientific community should continue research on the nature and scope of unintended consequences of different types of health interventions. The greater the chances that undesirable unintended consequences could emerge, the more we should promote research that can reduce the uncertainty surrounding these risks. The results of these studies should be disclosed in transparent manners.

It is also important to continue to test the applicability of this study's conceptual framework in various fields of intervention, that is, beyond PBF. Further research could help us refine the framework and improve analyses of unintended consequences of public health interventions. Researchers interested in contributing to the development and application of the conceptual framework used in this thesis are invited to contact our research team.

Chapter 7 – Conclusion

Global health organizations are widely promoting the implementation of PBF in LMICs, despite concerns regarding the numerous potential unintended consequences that could emerge. Until now, researchers and evaluators have consistently overlooked this issue, focusing primarily on the interventions' targeted objectives. This thesis widened the analytical lens by examining changes that were not initially planned by policy-makers but that are essential to determine the overall value of the intervention. We developed a framework based on the diffusion of innovations theory that enabled us to document the unintended consequences of the intervention. To achieve the objectives for this thesis, we: 1) analyzed the unintended consequences of the PBF program in primary healthcare facilities in Burkina Faso; 2) deciphered the unintended consequences of community verifications coupled with client satisfaction assessments; and 3) examined the unintended consequences of the intervention's health equity measures.

Empirical results showed that the nature and use of the intervention interacted with the social system and the characteristics of its actors to cause unintended consequences. Mostly undesirable, these unintended consequences affected various stakeholders such as providers, community members, healthcare managers, and patients. More specifically, we found that providers were fixated on performance measures rather than on underlying objectives, falsified medical registers, and taught trainees improper practices. As a desirable unintended consequence, we found that some facilities limited the sale of non-prescribed medication to encourage patients to consult. Community verifications, aimed at tracing patients to verify the authenticity of reported services and assess patient satisfaction, also led to unintended consequences, such as the falsification of verification data, the loss of patient confidentiality, and fears among patients, although some were pleased to share their views. Lastly, health equity measures also triggered changes that were not intended by program planners. For example, providers limited the free services delivered to indigents, which led to conflicts between parties.

The results of this thesis can inform discussions regarding the development of the intervention. Strategies to minimize undesirable unintended consequences and promote desirable ones could be devised and tested. Lastly, this thesis underscores the need to conduct more research on the unintended consequences of complex health interventions. We hope that it will inspire and guide future researchers to broaden their analytical angle to capture both intended and unintended consequences of complex health interventions aimed at increasing access to high quality healthcare for all.

Bibliography

Abomo Kele, P. (2018). La diffusion du paiement à la performance dans la réforme des systèmes de santé du Nord au Sud: Promotion de l'efficacité ou dynamique de dépolitisation de l'action publique en santé ? [Doctoral dissertation]. Université Paris 1 Panthéon-Sorbonne.

Allen-Scott, L. K., Hatfield, J. M., & McIntyre, L. (2014). A scoping review of unintended harm associated with public health interventions: Towards a typology and an understanding of underlying factors. International Journal of Public Health, 59(1), 3–14. https://doi.org/10.1007/s00038-013-0526-6

Amankwah-Amoah, J. (2014). Coming of age, seeking legitimacy: The historical trajectory of African management research. Critical Perspectives on International Business, 63625. https://mpra.ub.uni-muenchen.de/63625/

Aninanya, G. A., Howard, N., Williams, J. E., Apam, B., Prytherch, H., Loukanova, S., Kamara, E. K., & Otupiri, E. (2016). Can performance-based incentives improve motivation of nurses and midwives in primary facilities in northern Ghana? A quasi-experimental study. Global Health Action, 9(32404). https://doi.org/10.3402/gha.v9.32404

Anney, V. N. (2014). Ensuring the quality of the findings of qualitative research: Looking at trustworthiness criteria. Journal of Emerging Trends in Educational Research and Policy Studies, 5(2), 272–281.

Antony, M., Bertone, M. P., & Barthes, O. (2017). Exploring implementation practices in results-based financing: The case of the verification in Benin. BMC Health Services Research, 17(204). https://doi.org/10.1186/s12913-017-2148-9

APHA. (2020). What is public health? https://www.apha.org/what-is-public-health

Ash, J. S., Sittig, D. F., Dykstra, R. H., Guappone, K., Carpenter, J. D., & Seshadri, V. (2007). Categorizing the unintended sociotechnical consequences of computerized provider order entry. International Journal of Medical Informatics, 76(Suppl. 1)(Bardach), S21–S27. https://doi.org/10.1016/j.ijmedinf.2006.05.017

Ash, J. S., Sittig, D. F., Poon, E. G., Guappone, K., Campbell, E., & Dykstra, R. H. (2007). The extent and importance of unintended consequences related to computerized provider order entry. Journal of the American Medical Informatics Association, 14(4), 415–423. https://doi.org/10.1197/jamia.M2373

Bamberger, M., Tarsilla, M., & Hesse-Biber, S. (2016). Why so many "rigorous" evaluations fail to identify unintended consequences of development programs: How mixed methods can contribute. Evaluation and Program Planning, 55, 155–162. https://doi.org/10.1016/j.evalprogplan.2016.01.001

Banque africaine de développement. (2019). Des évaluations adaptées à l'Afrique. Volume 1: Approches théoriques. EVALUation Matters, 3rd trimester. https://idev.afdb.org/fr/document/des-%C3%A9valuations-adapt%C3%A9es-%C3%A0-lafrique-volume-1-approches-th%C3%A9oriques

Bardach, N. S., & Cabana, M. D. (2009). The unintended consequences of quality improvement. Current Opinion in Pediatrics, 21(6), 777–782. https://doi.org/10.1097/MOP.0b013e3283329937

Barker, K. (2004). Diffusion of innovations: A world tour. Journal of Health Communication, 9(S1), 131–137. https://doi.org/10.1080/10810730490271584

Barnes, A., Brown, G. W., & Harman, S. (2015). Global politics of health reform in Africa: Performance, participation, and policy. Palgrave Pivot.

Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L. B., Sturdy, J., & Vermeersch, C. M. J. (2011). Effect on maternal and child health services in Rwanda of payment to primary health-care providers for performance: An impact evaluation. The Lancet, 377(9775), 1421–1428. https://doi.org/10.1016/S0140-6736(11)60177-3

Beaugé, Y., & De Allegri, M. (submitted). Effects of user fee exemptions with built-in provider financial incentives on ultra-poor—A panel data analysis in Burkina Faso.

Beaugé, Y., Koulidiati, J.-L., Ridde, V., Robyn, P. J., & De Allegri, M. (2018). How much does community-based targeting of the ultra-poor in the health sector cost? Novel evidence from Burkina Faso. Health Economics Review, 8(19). https://doi.org/10.1186/s13561-018-0205-7

Benjamin-Chung, J., Abedin, J., Berger, D., Clark, A., Jimenez, V., Konagaya, E., Tran, D., Arnold, B. F., Hubbard, A. E., Luby, S. P., Miguel, E., & Colford, J. M. (2017). Spillover effects on health outcomes in low- and middle-income countries: A systematic review. International Journal of Epidemiology, 46(4), 1251–1276. https://doi.org/10.1093/ije/dyx039

Bertone, M. P., Falisse, J.-B., Russo, G., & Witter, S. (2018). Context matters (but how and why?) A hypothesis-led literature review of performance based financing in fragile and conflict-affected health systems. PLOS ONE, 13(4), e0195301. https://doi.org/10.1371/journal.pone.0195301

Bertone, M. P., & Meessen, B. (2013). Studying the link between institutions and health system performance: A framework and an illustration with the analysis of two performance-based financing schemes in Burundi. Health Policy and Planning, 28(8), 847–857. https://doi.org/10.1093/heapol/czs124

Biesma, R. G., Brugha, R., Harmer, A., Walsh, A., Spicer, N., & Walt, G. (2009). The effects of global health initiatives on country health systems: A review of the evidence from HIV/AIDS control. Health Policy and Planning, 24(4), 239–252. https://doi.org/10.1093/heapol/czp025

Binyaruka, P., Patouillard, E., Powell-Jackson, T., Greco, G., Maestad, O., & Borghi, J. (2015). Effect of paying for performance on utilisation, quality, and user costs of health services in Tanzania: A controlled before and after study. PLOS ONE, 10(8), e0135013. https://doi.org/10.1371/journal.pone.0135013

Binyaruka, P., Robberstad, B., Torsvik, G., & Borghi, J. (2018). Does payment for performance increase performance inequalities across health providers? A case study of Tanzania. Health Policy and Planning, 33(9), 1026–1036. https://doi.org/10.1093/heapol/czy084

Bloomrosen, M., Starren, J., Lorenzi, N. M., Ash, J. S., Patel, V. L., & Shortliffe, E. H. (2011). Anticipating and addressing the unintended consequences of health IT and policy: A report from the AMIA 2009 Health Policy Meeting. Journal of the American Medical Informatics Association, 18(1), 82–90. https://doi.org/10.1136/jamia.2010.007567

Bodson, O. (2014). Analyse pré évaluative de la fidélité de l’implantation de la politique de financement basé sur les résultats au Burkina Faso (p. 78) [Master’s thesis]. Université de Liège. http://hdl.handle.net/2268/235261

Bodson, O., Barro, A., Turcotte-Tremblay, A.-M., Zanté, N., Somé, P.-A., & Ridde, V. (2018). A study on the implementation fidelity of the performance-based financing policy in Burkina Faso after 12 months. Archives of Public Health, 76(4). https://doi.org/10.1186/s13690-017-0250-4

Bonell, C., Jamal, F., Melendez-Torres, G. J., & Cummins, S. (2015). 'Dark logic': Theorising the Harmful Consequences of Public Health Interventions. Journal of Epidemiology and Community Health, 69(1), 95–98. https://doi.org/10.1136/jech-2014-204671

Bonfrer, I., Soeters, R., Van de Poel, E., Basenya, O., Longin, G., van de Looij, F., & van Doorslaer, E. van. (2014). Introduction of performance-based financing in Burundi was associated with improvements in care and quality. Health Affairs, 33(12), 2179–2187. https://doi.org/10.1377/hlthaff.2014.0081

Bonfrer, I., Van de Poel, E., & Van Doorslaer, E. (2014). The effects of performance incentives on the utilization and quality of maternal and child care in Burundi. Social Science & Medicine, 123, 96–104. https://doi.org/10.1016/j.socscimed.2014.11.004

Borgès Da Silva, R., Fiset-Laniel, J., Hazra, A., & Strumpf, E. C. (2015). Analyse d'une politique de santé visant l'amélioration de l'accès à un médecin de famille. Les guichets d'accès pour la clientèle orpheline au Québec (p. 50). CIRANO.

Borghi, J., Little, R., Binyaruka, P., Patouillard, E., & Kuwawenaruwa, A. (2015). In Tanzania, the many costs of pay-for-performance leave open to debate whether the strategy is cost-effective. Health Affairs, 34(3), 406–414. https://doi.org/10.1377/hlthaff.2014.0608

Bourgeois, L. (2016). Assurer la rigueur scientifique de la recherche-action. In I. Carignan, M.-C. Beaudry, F. Larose, & L. Lafontaine, La recherche-action et la recherche-développement au service de la littératie (Les Éditions de l'Université de Sherbrooke, pp. 6–20).

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to improve health. BMJ, 321, 694–696. https://doi.org/10.1136/bmj.321.7262.694

Canadian Institutes of Health Research. (2016). Knowledge translation at CIHR. http://www.cihr-irsc.gc.ca/e/29418.html#1

Canadian Institutes of Health Research (CIHR), Natural Sciences and Engineering Research Council of Canada (NSERC), & Social Sciences and Humanities Research Council of Canada (SSHRC). (2014). Tri-council policy statement: Ethical conduct for research involving humans. http://www.frqnt.gouv.qc.ca/documents/10191/186009/TCPS2.+pdf/6a8ab915-431b-428d-aa86-b22ca5c78053

Canadian Medical Association. (2015). Appropriateness in Health Care [CMA Policy]. https://policybase.cma.ca/documents/policypdf/PD15-05.pdf

Canavan, A., & Swai, G. (2008). Payment for performance (P4P) evaluation: 2008 Tanzania country report for Cordaid (p. 74). Royal Tropical Institute, KIT Development Policy & Practice. http://www.bibalex.org/Search4Dev/files/305432/135295.pdf

Cashin, C., Chi, Y.-L., Smith, P., Borowitz, M., & Thomson, S. (Eds.). (2014). Paying for performance in health care: Implications for health system performance and accountability. Open University Press.

Chalkley, M., Mirelman, A., Siciliani, L., & Suhrcke, M. (2016). Paying for performance for health care in low- and middle-income countries: An economic perspective (Working

Papers 140cherp). Centre for Health Economics, University of York. https://ideas.repec.org/p/chy/respap/140cherp.html

Chang, R.-E., Lin, S.-P., & Aron, D. C. (2012). A pay-for-performance program in Taiwan improved care for some diabetes patients, but doctors may have excluded sicker ones. Health Affairs, 31(1), 93–102. https://doi.org/10.1377/hlthaff.2010.0402

Chen, T.-T., Chung, K.-P., Lin, I.-C., & Lai, M.-S. (2011). The unintended consequence of diabetes mellitus pay-for-performance (P4P) program in Taiwan: Are patients with more comorbidities or more severe conditions likely to be excluded from the P4P program? Health Services Research, 46(1 Pt 1), 47–60. https://doi.org/10.1111/j.1475-6773.2010.01182.x

Cherkaoui, M. (2004). Le réel et le rationnel. Rationalité et conséquences inattendues chez Max Weber. European Journal of Social Sciences, XLII(129), 73–86. https://doi.org/10.4000/ress.372

Chimhutu, V., Lindkvist, I., & Lange, S. (2014). When incentives work too well: Locally implemented pay for performance (P4P) and adverse sanctions towards home birth in Tanzania—A qualitative study. BMC Health Services Research, 14(23). https://doi.org/10.1186/1472-6963-14-23

Chimhutu, V., Songstad, N. G., Tjomsland, M., Mrisho, M., & Moland, K. M. (2016). The inescapable question of fairness in Pay-for-performance bonus distribution: A qualitative study of health workers' experiences in Tanzania. Globalization and Health, 12(77). https://doi.org/10.1186/s12992-016-0213-5

Choo, E. K., Ranney, M. L., Chan, T. M., Trueger, N. S., Walsh, A. E., Tegtmeyer, K., McNamara, S. O., Choi, R. Y., & Carroll, C. L. (2015). Twitter as a tool for communication and knowledge exchange in academic medicine: A guide for skeptics and novices. Medical Teacher, 37(5), 411–416. https://doi.org/10.3109/0142159X.2014.993371

Cole, M. S., Boydell, V., Hardee, K., & Bellows, B. (2019). The extent to which performance-based financing programs' operations manuals reflect rights-based principles: Implications for family planning services. Global Health: Science and Practice, 7(2), 329–339. https://doi.org/10.9745/GHSP-D-19-00007

Contandriopoulous, A.-P., Champagne, F., Potvin, L., Denis, J.-L., & Boyle, P. (2005). Savoir préparer une recherche. La définir, la structurer, la financer. Gaëtan Morin Éditeur.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2013). Developing and evaluating complex interventions: The new Medical Research Council guidance. BMJ, 337, a1655. https://doi.org/10.1136/bmj.a1655

Creswell, J. W., & Miller, D. L. (2000). Determining Validity in Qualitative Inquiry. Theory Into Practice, 39(3), 124–130. https://doi.org/10.1207/s15430421tip3903_2

Creswell, J. W., & Plano Clark, V. L. (2007). Designing and conducting mixed methods research (Sage Publications).

Crozier, M., & Friedberg, M. W. (1977). L'acteur et le système. Les contraintes de l'action collective. Éditions du Seuil.

De Allegri, M., Bertone, M. P., McMahon, S., Mounpe Chare, I., & Robyn, P. J. (2018). Unraveling PBF effects beyond impact evaluation: Results from a qualitative study in Cameroon. BMJ Global Health, 3(2), e000693. https://doi.org/10.1136/bmjgh-2017-000693

De Allegri, M., Lohmann, J., Koulidiati, J. L., & Somé, P.-A. (2019). Qualitative research to explain and unpack quantitative findings from the performance-based financing impact evaluation in Burkina Faso. World Bank. https://www.rbfhealth.org/resource/qualitative-research-explain-and-unpack-quantitative-findings-performance-based-financing

De Allegri, M., Lohmann, J., & Schleicher, M. (2018). Results-based financing for health impact evaluation in Burkina Faso. Results report (p. 103). World Bank Group. https://www.rbfhealth.org/impact-evaluation/burkina-faso-impact-evaluation

De Allegri, M., Makwero, C., & Torbica, A. (2019). At what cost is performance-based financing implemented? Novel evidence from Malawi. Health Policy and Planning, 34(4), 282–288. https://doi.org/10.1093/heapol/czz030

de Walque, D., Robyn, P. J., Saidou, H., Sorgho, G., & Steenland, M. (2017). Looking into the performance-based financing black box. Evidence from an impact evaluation in the health sector in Cameroon (Policy Research Working Paper No. 8162; p. 82). World Bank, Health Nutrition and Population Global Practice Group. http://documents.worldbank.org/curated/en/834601502391015068/pdf/WPS8162.pdf

de Zwart, F. (2015). Unintended but not unanticipated consequences. Theory and Society, 44(3), 283–297. https://doi.org/10.1007/s11186-015-9247-6

Derthick, M. (1974). Implementation: How great expectations in Washington are dashed in Oakland; Or, why it’s amazing that federal programs work at all, this Being a saga of the economic development administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes. By Jeffrey L. Pressman and Aaron Wildavsky. (Berkeley: University of California Press, 1973. Pp/ xviii, 182.

$7.50). American Political Science Review, 68(3), 1336–1337. https://doi.org/10.2307/1959201

Desclaux, A., & Boye, S. (2014). Mondialisation des effets indésirables et construction locale de la communitation médecin-patient autour des antiviraux au Sénégal. Journal Des Anthropologues, 138–139, 89–111. https://doi.org/10.4000/jda.4423

Dijkstra, G. (2018). Aid and good governance: Examining aggregate unintended effects of aid. Evaluation and Program Planning, 68, 225–232. https://doi.org/10.1016/j.evalprogplan.2017.09.004

Donabedian, A. (1988). The quality of care. How can it be assessed? JAMA, 260(12), 1743–1748. https://doi.org/doi: 10.1001/jama.260.12.1743

Eberhard, D. M., Simons, G. S., & Fennig, C. D. (2019). Burkina Faso. Ethnologue: Languages of the World. https://www.ethnologue.com/country/BF

Eijkenaar, F., Emmert, M., Scheppach, M., & Schöffski, O. (2013). Effects of pay for performance in health care: A systematic review of systematic reviews. Health Policy, 110(2–3), 115–130. https://doi.org/10.1016/j.healthpol.2013.01.008

Eldridge, C., & Palmer, N. (2009). Performance-based payment: Some reflections on the discourse, evidence and unanswered questions. Health Policy and Planning, 24(3), 160–166. https://doi.org/10.1093/heapol/czp002

Ellis, R. P. (1998). Creaming, skimping and dumping: Provider competition on the intensive and extensive margins1This is a substantially rewritten version of a paper entitled `Creaming, Skimping, and Dumping: Provider Competition for Patients'.1. Journal of Health Economics, 17(5), 537–555. https://doi.org/10.1016/S0167-6296(97)00042-8

Engineer, C. Y., Dale, E., Agarwal, A., Agarwal, A., Alonge, O., Edward, A., Gupta, S., Schuh, H. B., Burnham, G., & Peters, D. H. (2016). Effectiveness of a pay-for-performance intervention to improve maternal and child health services in Afghanistan: A cluster-randomized trial. International Journal of Epidemiology, 45(2), 451–459. https://doi.org/10.1093/ije/dyv362

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International Journal of Qualitative Methods, 5(1), 80–92. https://doi.org/10.1177/160940690600500107

Fillol, A., Lohmann, J., Turcotte-Tremblay, A.-M., Somé, P.-A., & Ridde, V. (2019). The importance of leadership and organizational capacity in shaping health workers' motivational reactions to performance-based financing: A multiple case study in Burkina Faso. International Journal of Health Policy and Management, 8(5), 272–279.

Flink, I. J. E., Ziebe, R., Vagaï, D., van de Looij, F., van 'T Riet, H., & Houweling, T. A. J. (2016). Targeting the poorest in a performance-based financing programme in northern Cameroon. Health Policy and Planning, 31(6), 767–776. https://doi.org/10.1093/heapol/czv130

Frehiwot, M. (2019). L'évaluation adaptée à l'Afrique: Décoloniser l'évaluation en Afrique. EVALUation Matters, 3rd trimester, 26–35.

Fretheim, A., Witter, S., Lindahl, A. K., & Olsen, I. T. (2012). Performance-based financing in low- and middle-income countries: Still more questions than answers. Bulletin of the World Health Organization, 90(8), 559-559A. https://doi.org/10.2471/BLT.12.106468

Fritsche, G. B., Soeters, R., & Meessen, B. (2014). Performance-based financing toolkit. World Bank. https://openknowledge.worldbank.org/handle/10986/17194

Gaudet, S., & Robert, D. (2018). L'aventure de la recherche qualitative: Du questionnement à la rédaction scientifique. University of Ottawa Press. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1885644

Gautier, L. (2019). Des idées à la prise de décision: L'économie politique de la diffusion du financement basé sur la performance aux niveaux global, continental, et national [Unpublished doctoral dissertation]. Université de Montréal & Sorbonne Paris Cité.

Gilson, L., Hanson, K., Sheikh, K., Agyepong, I. A., Ssengooba, F., & Bennett, S. (2011). Building the Field of Health Policy and Systems Research: Social Science Matters. PLOS Medicine, 8(8). https://doi.org/10.1371/journal.pmed.1001079

Gorter, A. C., Ir, P., & Meessen, B. (2013). Evidence review. Results-based financing of maternal and newborn health care in low- and lower-middle-income countries. German Federal Ministry for Economic Cooperation & Development (BMZ). https://www.oecd.org/dac/peer-reviews/Evidence-RBF-maternal-health.pdf

Greenhalgh, T., & Papoutsi, C. (2018). Studying complexity in health services research: Desperately seeking an overdue paradigm shift. BMC Medicine, 16(95). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6009054/

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations. The Milbank Quarterly, 82(4), 581–629.

Guba, E G, & Lincoln, Y. S. (2005). Paradigmatic Controversies, Contradictions, and Emerging Confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), The Sage handbook of qualitative research (Sage Publications Ltd, pp. 191–215).

Guba, Egon G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. ECTJ, 29(75). https://doi.org/10.1007/BF02766777

Hageboeck, M., Frumkin, M., & Monschein, S. (2013). Meta-evaluation of quality and coverage of USAID evaluations 2009 – 2012. http://www.ghpro.dexisonline.com/sites/default/files/Meta-Evaluation%20of%20Quality%20and%20Coverage%20of%20USAID%20Evaluations%202009-2012.pdf

Heider, C. (2017). Rethinking Evaluation—What is Wrong with Development Effectiveness? http://ieg.worldbankgroup.org/blog/rethinking-evaluation-development-effectiveness

Hogan, D. R., Stevens, G. A., Hosseinpoor, A. R., & Boerma, T. (2018). Monitoring universal health coverage within the Sustainable Development Goals: Development and baseline data for an index of essential health services. The Lancet Global Health, 6(2), e152–e168. https://doi.org/10.1016/S2214-109X(17)30472-2

Huillery, E., & Seban, E. (2014). Performance-Based Financing,Motivation and FinalOutput in theHealth Sector: Experimental Evidence from the Democratic Republic of Congo (p. 57). Blavatnik School of Government.

Hurley, J. (2010). Health Economics (First Edition). McGraw-Hill Ryerson.

Hysong, S. J., SoRelle, R., Broussard Smitham, K., & Petersen, L. A. (2017). Reports of unintended consequences of financial incentives to improve management of hypertension. PloS One, 12(9), e0184856. https://doi.org/10.1371/journal.pone.0184856

Independent Evaluation Group. (2017). ML-Strengthening Reprod Health (FY12) (P124054) (Implementation Completion Report (ICR) Review No. ICRR0020893; p. 19). World Bank.

Ireland, M., Paul, E., & Dujardin, B. (2011). Can performance-based financing be used to reform health systems in developing countries? Bull World Health Organ, 695–698.

Jabeen, S. (2016). Do we really care about unintended outcomes? An analysis of evaluation theory and practice | Elsevier Enhanced Reader. Evaluation and Program Planning, 55, 144–154. https://doi.org/10.1016/j.evalprogplan.2015.12.010

Jabeen, S. (2018). Unintended outcomes evaluation approach: A plausible way to evaluate unintended outcomes of social development programmes. Evaluation and Program Planning, 68, 262–274. https://doi.org/10.1016/j.evalprogplan.2017.09.005

Kalk, A. (2011). The costs of performance-based financing. Bull World Health Organ, 319.

Kalk, A., Paul, F. A., & Grabosch, E. (2010). Paying for performance in Rwanda: Does it pay off? Tropical Medicine & International Health, 15(2), 182–90. https://doi.org/10.1111/j.1365-3156.2009.02430.x

Koch, D.-J., & Schulpen, L. (2018). Introduction to the special issue ‘unintended effects of international cooperation.’ Evaluation and Program Planning, 68, 202–209. https://doi.org/10.1016/j.evalprogplan.2017.10.006

Koplan, J. P., Bond, T. C., Merson, M. H., Reddy, K. S., Rodriguez, M. H., Sewankambo, N. K., & Wasserheit, J. N. (2009). Towards a common definition of global health. The Lancet, 373(9679), 1993–1995. https://doi.org/10.1016/S0140-6736(09)60332-9

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. The European Journal of General Practice, 24(1), 120–124. https://doi.org/10.1080/13814788.2017.1375092

Kuunibe, N., Lohmann, J., Hillebrecht, M., Nguyen, H., Tougri, G., & De Allegri, M. (2020). What Happens When Performance Based Financing Meets Free Healthcare? Evidence From An Interrupted Time-Series Analysis With Independent Controls. Health Policy and Planning, 1–12. https://doi.org/10.1093/heapol/czaa062

Kuunibe, N., Lohmann, J., Schleicher, M., Koulidiati, J.-L., Robyn, P. J., Zigani, Z., Sanon, A., & Allegri, M. D. (2019). Factors associated with misreporting in performance-based financing in Burkina Faso: Implications for risk-based verification. International Journal of Health Planning and Management. https://doi.org/10.1002/hpm.2786

Lannes, L., Meessen, B., Soucat, A., & Basinga, P. (2015). Can Performance-Based Financing Help Reaching the Poor with Maternal and Child Health Services? The Experience of Rural Rwanda. The International Journal of Health Planning and Management. https://doi.org/10.1002/hpm.2297

Lemière, C., de Walque, D., Ayivi-Guedehoussou, N., & Juquois, M. (2015). Rapport d’analyse de l’enquête de base. Evaluation d’impact du financement basé sur les résultats.

Lester, H. E., Hannon, K. L., & Campbell, S. M. (2011). Identifying unintended consequences of quality indicators: A qualitative study. BMJ Quality & Safety, 20(12), 1057–1061. https://doi.org/10.1136/bmjqs.2010.048371

Little, W. (2012). Chapter 1. An Introduction to Sociology. Introduction to Sociology – 2nd Canadian Edition.

https://opentextbc.ca/introductiontosociology2ndedition/chapter/chapter-1-an-introduction-to-sociology/

Lohmann, J., Houlfort, N., & De Allegri, M. (2016). Crowding out or no crowding out? A Self-Determination Theory approach to health worker motivation in performance-based financing. Social Science & Medicine (1982), 169, 1–8. https://doi.org/10.1016/j.socscimed.2016.09.006

Lohmann, J., Muula, A. S., Houlfort, N., & De Allegri, M. (2018). How does performance-based financing affect health workers' intrinsic motivation? A Self-Determination Theory-based mixed-methods study in Malawi. Social Science & Medicine, 208, 1–8. https://doi.org/10.1016/j.socscimed.2018.04.053

Lorenc, T., & Oliver, K. (2013). Adverse effects of public health interventions: A conceptual framework. Journal of Epidemiology and Community Health, jech-2013-203118. https://doi.org/10.1136/jech-2013-203118

Lubbock, J. (2011). The beauties of nature and the wonders of the world we live in. Kirk Press.

Lutwama, G. W., Roos, J. H., & Dolamo, B. L. (2013). Assessing the implementation of performance management of health care workers in Uganda. BMC Health Services Research, 13, 355. https://doi.org/10.1186/1472-6963-13-355

Maini, R., Lohmann, J., Hotchkiss, D. R., Mounier-Jack, S., & Borghi, J. (2019). What Happens When Donors Pull Out? Examining Differences in Motivation Between Health Workers Who Recently Had Performance-Based Financing (PBF) Withdrawn With Workers Who Never Received PBF in the Democratic Republic of Congo.

International Journal of Health Policy and Management, 0. http://www.ijhpm.com/article_3647.html

Ma-Nitu, S. M., Tembey, L., Bigirimana, E., Dossouvi, C. Y., Basenya, O., Mago, E., Salongo, P. M., Zongo, A., & Verinumbe, F. (2018). Towards constructive rethinking of PBF: Perspectives of implementers in sub-Saharan Africa. BMJ Global Health, 3(5), e001036. https://doi.org/10.1136/bmjgh-2018-001036

McQueen, D. V. (2014). Evidence and harm: Time for reflection. International Journal of Public Health, 59(1), 1–2. https://doi.org/10.1007/s00038-013-0534-6

Medical Research Council. (2000). Framework for the Development and Evaluation of RCTs for Complex Interventions to Improve Health (p. 19).

Medical Research Council. (2008). Developing and evaluating complex interventions (p. 39). MRC. https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/

Merton, R. K. (1936). The unanticipated consequences of purposive social action. American Sociological Review, 1(6), 894–904.

Ministère de la Santé. (2011). Plan national de développement sanitaire 2011-2020 (p. 50).

Ministère de la Santé. (2013a). Guide de mise en oeuvre du financement basé sur les résultats dans le secteur de la santé.

Ministère de la Santé. (2014). Note d’information et d’orientation sur la prise en charge des indigents dans le cadre de la mise en oeuvre du processus de financement basé sur les résultats dans le secteur de la santé.

Ministère de la Santé. (2017). Profile sanitaire complet du Burkina Faso. Module 2. Système de santé du Burkina Faso.

Ministère de la Santé, A. (2013b). Évaluation finale de la phase test du financement basé sur les résultats dans les districts sanitaires de Boulsa, Leo et Titao. Rapport final. SERSAP.

Ministère de la Santé et des Services sociaux. (2015). Programme nationale de santé publique 2015-2025. https://publications.msss.gouv.qc.ca/msss/fichiers/2015/15-216-01W.pdf

Mittelmark, M. B. (2014). Unintended effects in settings-based health promotion. Scandinavian Journal of Public Health, 42(15 Suppl), 17–24. https://doi.org/10.1177/1403494814545108

Mkhize, N. (2006). Psychology: An African perspective. In D. Hook (Ed.), Introduction to Critical Psychology (pp. 24–50). University of Capetown (UCT) Press.

Mom, A.-R. (2018). Why do management theories fail? Reasons and solutions: A perspective of African context. International Journal of Economics & Management Sciences, 7(3). https://doi.org/https:/doi.org/10.4172/2162-6359.1000520

Morell, J. A. (2005). Why Are there Unintended Consequences of Program Action and What Are the Implications for Doing Evaluation? American Journal of Evaluation, 26(4), 444–463.

Morell, J. A. (2010). Evaluation in the Face of Uncertainty. Anticipating Surprise and Responding to the Inevitable. (The Guildford Press).

Morell, J. A. (2018). Systematic iteration between model and methodology: A proposed approach to evaluating unintended consequences. Evaluation and Program Planning, 68, 243–252. https://doi.org/10.1016/j.evalprogplan.2017.09.006

Morin, E., & Le Moigne, J.-L. (1999). L’intelligence de la complexité (L’Harmattan).

Musgrove, P. (2011). Financial and Other Rewards for Good Performance or Results: A Guided Tour of Concepts and Terms and a Short Glossary. World Bank. https://www.rbfhealth.org/sites/rbf/files/RBFglossarylongrevised_0.pdf

Njagi, P., Arsenijevic, J., & Groot, W. (2018). Understanding variations in catastrophic health expenditure, its underlying determinants and impoverishment in Sub-Saharan African countries: A scoping review. Systematic Reviews, 7. https://doi.org/10.1186/s13643-018-0799-1

Norton, R. (2008). Unintended consequences. In The Concise Encyclopedia of Economics (pp. 505–506). Liberty Fund.

OECD. (2002). Glossary of Key Terms in Evaluation and Results Based Management (OECD PUBLICATIONS). http://www.oecd.org/dac/evaluation/2754804.pdf

Oliver, K., Lorenc, T., & Tinkler, J. (2019). Evaluating unintended consequences: New insights into solving practical, ethical and political challenges of evaluation. Evaluation, 1356389019850847. https://doi.org/10.1177/1356389019850847

Olivier de Sardan, J.-P. (2010). Le culturalisme traditionaliste africaniste: Analyse d’une idéologie scientifique. Cahiers d’études Africaines, 50, 419–453.

Olivier de Sardan, J.-P. (2015). Practical norms: Informal regulations within public bureaucracies (in Africa and beyond). LASDEL.

Olivier de Sardan, J.-P., Bako, M. T. A., & Harouna, A. (2018). Les normes pratiques en vigueur dans les secteurs de l’éducation et la santé au Niger Une base pour des réformes ancrées dans les réalités ? (Études et Travaux Du LASDEL No. 127). LASDEL.

https://www.pseau.org/outils/ouvrages/lasdel_les_normes_pratiques_en_vigueur_dans_les_secteurs_de_l_education_et_la_sante_au_niger_2019.pdf

Olivier de Sardan, J.-P., Diarra, A., & Moha, M. (2017). Travelling models and the challenge of pragmatic contexts and practical norms: The case of maternal health. Health Research Policy and Systems, 15(Suppl 1), 60. https://doi.org/10.1186/s12961-017-0213-9

Olivier de Sardan, J.-P., & Tidjani Alou, A. (2015). Epistemology, Fieldwork, and Anthropology. Palgrave Macmillan.

Ooms, G., Latif, L. A., Waris, A., Brolan, C. E., Hammonds, R., Friedman, E. A., Mulumba, M., & Forman, L. (2014). Is Universal Health Coverage the Practical Expression of the Right to Health Care? BMC International Health and Human Rights, 14(3).

Oxman, A. D., & Fretheim, A. (2008). An overview of research on the effects of resultsbased financing (No. 16–2008; p. 81). Norwegian Knowledge Centre for the Health Services.

Oxman, A. D., & Fretheim, A. (2009). Can paying for results help to achieve the Millennium Development Goals? Overview of the effectiveness of results-based financing. Journal of Evidence-Based Medicine, 2(2), 70–83. https://doi.org/10.1111/j.1756-5391.2009.01020.x

Patton, M. Q. (2011). Developmental Evaluation Applying Complexity Concepts to Enhance Innovation and Use. The Guilford Press.

Patton, M. Q. (2015). Qualitative Research & Evaluation Methods: Integrating Theory and Practice (4th edition). SAGE Publications.

Paul, E., Albert, L., Bisala, B. N., Bodson, O., Bonnet, E., Bossyns, P., Colombo, S., Brouwere, V. D., Dumont, A., Eclou, D. S., Gyselinck, K., Hane, F., Marchal, B., Meloni, R., Noirhomme, M., Noterman, J.-P., Ooms, G., Samb, O. M., Ssengooba, F., … Ridde, V. (2018). Performance-based financing in low-income and middle-income countries: Isn’t it time for a rethink? BMJ Global Health, 3(1), e000664. https://doi.org/10.1136/bmjgh-2017-000664

Paul, E., Sossouhounto, N., & Eclou, D. S. (2014). Local stakeholders’ perceptions about the introduction of performance-based financing in Benin: A case study in two health districts. International Journal Health Policy Managing, 3(4), 207–214. https://doi.org/10.15171/ijhpm.2014.93

Pelto, P. J. (1973). The Snowmobile Revolution: Technology and Social Change in the Arctic. Waveland Press.

Petross, C., McMahon, S. A., Lohmann, J., Chase, R. P., Muula, A., & De Allegri, M. (2020). Intended and unintended effects: Community perspectives on a performance-based financing programme in Malawi. BMJ Global Health, 5(4), e001894.

Pluye, P., & Hong, Q. N. (2014). Combining the power of stories and the power of numbers: Mixed methods research and mixed studies reviews. Annual Review of Public Health, 35, 29–45. https://doi.org/10.1146/annurev-publhealth-032013-182440

Pommier, J., & Grimaud, O. (2007). Les fonctions essentielles de santé publique: Histoire, définition et applications possibles. Santé Publique, 19(hs), 9–14. https://doi.org/10.3917/spub.070.0009

Poupart, J., Groulx, L. H., Deslauriers, J. P., Laperrière, A., & Mayer, R. (1997). La recherche qualitative: Enjeux épistémologiques et méthodologiques. Gaëtan Morin Éditeur.

Pressman, J. L., & Wildavsky, A. (1984). Implementation: How great expectations in Washington are dashed in Oakland; Or, why it’s amazing that federal programs work at all, this being a saga of the economic development administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hope (3rd ed.). University of California Press.

Priedeman Skiles, M., Curtis, S. L., Basinga, P., & Angeles, G. (2013). An equity analysis of performance-based financing in Rwanda: Are services reaching the poorest women? Health Policy and Planning, 28(8), 825–837. https://doi.org/10.1093/heapol/czs122

REN-LAC. (2017). État de la corruption au Burkina Faso. https://secureservercdn.net/184.168.47.225/118.c3a.myftpupload.com/download/rapports/Rapport-REN-LAC-2017.pdf

Renmans, D., Holvoet, N., & Criel, B. (2017). Combining Theory-Driven Evaluation and Causal Loop Diagramming for Opening the “Black Box” of an Intervention in the Health Sector: A Case of Performance-Based Financing in Western Uganda. International Journal of Environmental Research and Public Health, 14(9). https://doi.org/10.3390/ijerph14091007

Ridde, V., Gautier, L., Turcotte-Tremblay, A.-M., Sieleunou, I., & Paul, E. (2018). Performance-based Financing in Africa: Time to Test Measures for Equity. International Journal of Health Services.

Ridde, V., Robert, E., & Meessen, B. (2012). A Literature Review of the Disruptive Effects of User Fee Exemption Policies on Health Systems. BMC Public Health, 12(289), 1–7. https://doi.org/10.1186/1471-2458-12-289

Ridde, V., Turcotte-Tremblay, A.-M., Souares, A., Lohmann, J., Zombré, D., Koulidiati, J. L., Yaogo, M., Hien, H., Hunt, M., Zongo, S., & Allegri, M. D. (2014). Protocol for the process evaluation of interventions combining performance-based financing with health equity in Burkina Faso. Implementation Science, 9(1), 149. https://doi.org/10.1186/s13012-014-0149-1

Ridde, V., Yaogo, M., Zongo, S., Somé, P.-A., & Turcotte-Tremblay, A.-M. (2017a). Twelve months of implementation of health care performance-based financing in Burkina Faso: A qualitative multiple case study. Int J Health Plann Mgmt, 1–15.

Ridde, V., Yaogo, M., Zongo, S., Somé, P.-A., & Turcotte-Tremblay, A.-M. (2017b). Twelve months of implementation of health care performance-based financing in Burkina Faso: A qualitative multiple case study. Int J Health Plann Mgmt, 1–15.

Rogers, E. M. (2003). Diffusion of Innovations (5th ed). Free Press.

Rouleau, L. (2007). Théories des organisations. Approches classiques, contemporaines et de l'avant garde. Presses de l'Université du Québec.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. Contemporary Educational Psychology, 25(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Sadaphal, S., & Bongiovanni, A. (2016). Impact Evaluation Results-based Financing in the Democratic Republic of Congo (AID-660-M-13-00001/GS-10F-0309P; p. 63). International Business & Technical Consultants.

Scott, S. D., Rotter, T., Flynn, R., Brooks, H. M., Plesuk, T., Bannar-Martin, K. H., Chambers, T., & Hartling, L. (2019). Systematic review of the use of process evaluations in knowledge translation research. Systematic Reviews, 8(1), 266. https://doi.org/10.1186/s13643-019-1161-y

Seppey, M., Ridde, V., Touré, L., & Coulibaly, A. (2017). Donor-funded project's sustainability assessment: A qualitative case study of a results-based financing pilot in

Koulikoro region, Mali. Globalization and Health, 13(86). https://doi.org/10.1186/s12992-017-0307-8

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and Quasi-Experimental Designs. Wadsworth Cengage Learning.

Sharp, L. (1952). Steel Axes for Stone Age Australians. In Edward Holland Spicer (Ed.), Human Problems in Technological Change. RUSSELL Sage Foundation.

Sieleunou, I., Turcotte-Tremblay, A.-M., De Allegri, M., Fotso, J.-C. T., Yumo, H. A., Tamga, D. M., & Ridde, V. (2019). How does performance-based financing affect the availability of essential medicines in Cameroon? A qualitative study. Health Policy & Planning, 1(34), iii4–iii19.

Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. International Journal of Public Administration, 18(2–3), 277–310. https://doi.org/10.1080/01900699508525011

SNV. (2013). Manuel des procédures pour la mise en oeuvre de la phase pilote du financement basé sur les résultats dans la région de Koulikoro (Mali).

Société d’études et de recherche en santé publique. (2014). Rapport d’étape de progrès de la mise en oeuvre et du suivi du processus communautaire de sélection des indigents au Burkina Faso.

Souares, A., Sauerborn, R., Haidara, O. D., & Robyn, P. J. (2013). Impact evaluation for health performance-based financing in Burkina Faso. Concept Note (p. 58).

Spicer, Edward H. (Ed.). (1952). Human Problems in Technological Change. A Casebook. Russell Sage Foundation.

Ssengooba, F., McPake, B., & Palmer, N. (2012). Why performance-based contracting failed in Uganda-an “open-box” evaluation of a complex health system intervention. Social Science & Medicine, 75(2), 377–383. https://doi.org/10.1016/j.socscimed.2012.02.050

Stake, R. E. (2005). Multiple Case Study Analysis. The Guilford Press.

Sveiby, K.-E., Gripenberg, P., Segercrantz, B., Eriksson, A., & Aminoff, A. (2009, June 21). Unintended and undesirable consequences of innovation. XX ISPIM Conference, Vienna, Austria. https://pdfs.semanticscholar.org/79d0/23c50f70082c5b472e3fc1183f08bb7a3ddf.pdf

The World Bank. (2012). Concept note for the impact evaluation of a PBF pilot in Cameroon. The World Bank. https://microdata.worldbank.org/index.php/catalog/2047/related-materials

The World Bank. (2018). Financing agreement (Health Services Reinforcement Project) between Burkina Faso and International Development Association. http://documents.worldbank.org/curated/en/310951536348591377/pdf/ITK425962-201808071524.pdf

The World Bank. (2019). Projects & Operations. BF-Reproductive Health Project (FY12). http://projects.worldbank.org/P119917/reproductive-health-project-burkina-faso?lang=en&tab=financial

Transparency International. (2018). Corruption Perceptions Index 2018. Corruption Perceptions Index 2018. https://www.transparency.org/cpi2018

Transparency International. (2020). Corruption perception index 2019. https://www.transparency.org/cpi2019

Turcotte-Tremblay, A.-M., De Allegri, M., Gali Gali, I. A., & Ridde, V. (2018). The unintended consequences of combining equity measures with performance-based financing in Burkina Faso. International Journal for Equity in Health, 17(109). https://doi.org/10.1186/s12939-018-0780-6

Turcotte-Tremblay, A.-M., Gali Gali, I. A., De Allegri, M., & Ridde, V. (2017). The unintended consequences of community verifications for performance-based financing in Burkina Faso. Social Science & Medicine, 191, 226–236. https://doi.org/10.1016/j.socscimed

Turcotte-Tremblay, A.-M., Gautier, L., Bodson, O., Sambieni, N. E., & Ridde, V. (2018). Le rôle des acteurs de la santé mondiale dans l'expansion du financement basé sur la performance dans les pays à faible et à moyen revenu. Journal de Gestion et d'Économie Médicales, 36(5), 261–279. https://doi.org/10.3917/jgem.185.0261

Turcotte-Tremblay, A.-M., Spagnolo, J., De Allegri, M., & Ridde, V. (2016). Does performance-based financing increase value for money in low- and middle- income countries? A systematic review. Health Economics Review, 6(1), 30. https://doi.org/10.1186/s13561-016-0103-9

United Nations. (1948). Universal Declaration of Human Rights. https://www.un.org/en/universal-declaration-human-rights/

United Nations. (2008). The right to health (Fact Sheet No. 31). https://www.ohchr.org/Documents/Publications/Factsheet31.pdf

United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development. https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

The Right to the Highest Attainable Standard of Health, no. General Comment No. 14. Article 12 of the International Covenant on Economic, Social, and Cultural Rights (2000). https://www.refworld.org/pdfid/4538838d0.pdf

United Nations Development Programme. (2018). Human Development Indices and Indicators. 2018 Statistical Update (p. 113).

Université de Montréal. (2018). Guide d'information sur le consentement libre, éclairé et continu. https://recherche.umontreal.ca/fileadmin/recherche/documents/BCRR/communs_hum/Guide_FCLE.pdf

Valente, T. W., & Rogers, E. M. (1995). The origins and development of the diffusion of innovations paradigm as an example of scientific growth. Science Communication, 16(3), 242–273. https://doi.org/10.1177/1075547095016003002

van Bodegom-Vos, L., Davidoff, F., & Marang-van de Mheen, P. J. (2017). Implementation and de-implementation: Two sides of the same coin? BMJ Quality & Safety, 26(6), 495–501. https://doi.org/10.1136/bmjqs-2016-005473

van de Ruit, C. (2019). Unintended Consequences of Community Health Worker Programs in South Africa. Qualitative Health Research, 1049732319857059. https://doi.org/10.1177/1049732319857059

Wagstaff, A., Flores, G., Hsu, J., Smitz, M.-F., Chepynoga, K., Buisman, L. R., Wilgenburg, K. van, & Eozenou, P. (2018). Progress on catastrophic health spending in 133 countries: A retrospective observational study. The Lancet Global Health, 6(2), e169–e179. https://doi.org/10.1016/S2214-109X(17)30429-1

Walt, G., Shiffman, J., Schneider, H., Murray, S. F., Brugha, R., & Gilson, L. (2008). ‘Doing’ health policy analysis: Methodological and conceptual reflections and challenges. Health Policy and Planning, 23(5), 308–317. https://doi.org/10.1093/heapol/czn024

Warford, M. K. (2005). Testing a diffusion of innovations in education model (DIEM). The Innovation Journal: The Public Sector Innovation Journal, 10(3), 1–41.

Wetsman, N. (2019). How Twitter is changing medical research. Nature Medicine, 26, 11–13. https://doi.org/10.1038/s41591-019-0697-7

Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. Qualitative Health Research, 11(4), 522–537. https://doi.org/10.1177/104973201129119299

WHO. (2018a). Atlas of African Health Statistics 2018: Universal health coverage and the Sustainable Development Goals in the WHO African Region (WHO Regional Office for Africa).

WHO. (2018b). World Health Statistics 2018: Monitoring health for the SDGs. https://www.who.int/gho/publications/world_health_statistics/2018/en/

WHO. (2020a). Countries. Burkina Faso. https://www.who.int/countries/bfa/en/

WHO. (2020b). Performance-based financing as an instrument to introduce strategic purchasing to move towards universal health coverage. Health Financing for Universal Coverage. https://www.who.int/health_financing/topics/performance-based-financing/universal-health-coverage/en/

Willis, J. W. (2007). Foundations of qualitative research. Interpretive and critical approaches. SAGE Publications. https://us.sagepub.com/en-us/nam/foundations-of-qualitative-research/book228788

Witter, S., Bertone, M. P., Namakula, J., Chandiwana, P., Chirwa, Y., Ssennyonjo, A., & 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, 4(3). https://doi.org/10.1186/s41256-019-0094-2

Witter, S., Fretheim, A., Kessy, F. L., & Lindahl, A. K. (2012). Paying for performance to improve the delivery of health interventions in low- and middle-income countries. The Cochrane Database of Systematic Reviews, 2(CD007899). https://doi.org/10.1002/14651858.CD007899.pub2

Woolcock, M. (2013). Using case studies to explore the external validity of 'complex' development interventions. Evaluation, 19(3), 229–248. https://doi.org/10.1177/1356389013495210

World Bank. (2014). RBFHealth. Projects. https://www.rbfhealth.org/project/our-projects

World Health Organization. (2005). Fifty-eight World Health Assembly.

World Health Organization. (2010). World Health Report. Health Systems Financing. The Path to Universal Coverage.

World Population Review. (2019). Burkina Faso Population 2019. http://worldpopulationreview.com/countries/burkina-faso-population/

Yin, R. K. (2009). Case Study Research: Design and Methods (4th edition). SAGE Publications.

Zizien, Z. R., Korachais, C., Compaoré, P., Ridde, V., & De Brouwere, V. (2019). Contribution of the results-based financing strategy to improving maternal and child health indicators in Burkina Faso. The International Journal of Health Planning and Management, 34(1), 111–129. https://doi.org/10.1002/hpm.2589

Zombré, D., Turcotte-Tremblay, A.-M., Sangaré, A., Zongo, S., Somé, P.-A., Coulibaly, A., Touré, L., Yaogo, M., & Ridde, V. (2016). Méthode mixte et participative de sélection d'études de cas multiples pour l'évaluation d'intervention de santé au Burkina Faso et au Mali: De l'éthique à l'émique [Poster]. https://docs.google.com/viewerng/viewer?url=http://www.equitesante.org/wp-content/uploads/2016/05/Poster.David_.ACFAS_.pdf&hl=fr

Appendices

Appendix A. Article 7 : Does Performance-Based Financing Increase Value for Money in Low- and Middle-Income Countries? A Systematic Review

Anne-Marie Turcotte-Tremblay, Jessica Spagnolo,
Manuela De Allegri and Valéry Ridde

Health Economics Review 2016, 6:30

Doi: 10.1186/s13561-016-0103-9

Title: The unintended consequences of community verifications for performance-based financing in Burkina Faso

Anne-Marie Turcotte-Tremblay[a,b*], Jessica Spagnolo[a,b,c], Manuela De Allegri[c], Valéry Ridde[a,b]

a. University of Montreal Public Health Research Institute, Montreal, Canada

b. School of Public Health, University of Montreal, Montreal, Canada

c. Douglas Mental Health University Institute, 6875 LaSalle Blvd., Montreal, QC, Canada H4H 1R3. 4Institute of Public Health, Medical Faculty, Heidelberg

d. 4Institute of Public Health, Medical Faculty, Heidelberg University, Im Neuenheimer Feld 324, 69120 Heidelberg, Germany

* Corresponding author: Anne-Marie Turcotte-Tremblay, University of Montreal Public Health Research Institute, 7101 Avenue du Parc, Room 3060, Montreal, QC H3N 1X9, Canada

Contributions of authors:

Anne-Marie Turcotte-Tremblay conceived the review, conducted the search, synthesized the literature and wrote the first draft of the manuscript.

Jessica Spagnolo helped conceive the review, conducted the search, synthesized the literature and wrote the first draft of the manuscript.

Manuela De Allegri helped analyse and interpret the literature and critically revised the manuscript to improve its content.

Valéry Ridde helped analyse and interpret the literature and critically revised the manuscript to improve its content.

ABSTRACT

Governments of low- and middle-income countries (LMICs) are widely implementing performance-based financing (PBF) to improve healthcare services. However, it is unclear whether PBF provides good value for money compared to status quo or other interventions aimed at strengthening the healthcare system in LMICs. The objective of this systematic review is to identify and synthesize the existing literature that examines whether PBF represents an efficient manner of investing resources. We considered PBF to be efficient when improved care quality or quantity was achieved with equal or lower costs, or alternatively, when the same quality of care is achieved using less financial resources. A manual search of the reference lists of two recent systematic reviews on economic evaluations of PBF was conducted to identify articles that met our inclusion and exclusion criteria. Subsequently, a search strategy was developed with the help of a librarian. Four databases were searched: PubMed, EconLit, Google Scholar and Google. Experts on economic evaluations were consulted for validation of the selected studies. A total of seven articles from five LMICs were selected for this review. We found the overall strength of the evidence to be weak. None of the articles were full economic evaluations; they did not make clear connections between the costs and effects of PBF. Only one study reported using a randomized controlled trial, but issues with the randomization procedure were reported. Important alternative interventions to strengthen the capacities of the healthcare system have not been considered. Few studies examined the costs and consequences of PBF in the long term. Important costs and consequences were omitted from the evaluation. Few LMICs are represented in the literature, despite wide implementation. Lastly, most articles had at least one author employed by an organization involved in the implementation of PBF, thereby resulting in potential conflicts of interest. Stronger empirical evidence on whether PBF represents good value for money in LMICs is needed.

KEYWORDS: Performance-based financing (PBF); economic evaluation; efficiency; low- and middle- income countries (LMICs); systematic review

INTRODUCTION

Governments and international organizations are investing resources to reduce preventable deaths and diseases across low-and middle- income countries (LMICs). Still, the World Health Organization (WHO) [1] reports that between 20–40% of resources spent on health are being wasted. Inefficiency is caused by inappropriate use of medicine and equipment, medical errors, suboptimal quality of care, costly staff mix, unmotivated healthcare workers, and corruption [1]. Faced with these issues, program planners must make difficult decisions about the best ways to invest limited resources to improve healthcare services and population health.

In recent years, many governments, donors, consultancy firms and non-governmental organisations (NGOs) have started transforming the funding mechanisms of healthcare systems in LMICs, namely by implementing performance-based financing (PBF) to link payments to results. In this model, healthcare facilities are paid based on the extent to which providers meet pre-defined quantity- and quality-related performance targets, following an independent verification [2]. Examples of quantity-related performance indicators include the number of consultations for children under the age of five or the number births per month. Examples of quality-related performance indicators include the healthcare center's cleanliness or completeness of patient registries. Healthcare centers sometime have to reach a minimal quality score (e.g., at least 50%) in order to be eligible for bonuses. Quality scores are also used as inflator or deflator of bonus payments.

The implementation of PBF is rapidly expanding. For example, the World Bank reports that the number of African countries using PBF increased from four to 21 between 2006 and 2013 [3]. Despite the rapid implementation of PBF, it is unclear whether given the same amount of resources, PBF can buy more healthcare services or health than the status quo or other interventions aiming to strengthen the healthcare system in LMICs. Existing systematic reviews on economic evaluations of PBF mainly draw their conclusions from studies conducted in high-income countries (HICs) [4], [5]. The results of these systematic reviews therefore cannot be generalized to LMICs, seeing that contexts and resources differ significantly. Distinctive characteristics of LMICs may influence the relations between the costs of PBF and the outcomes observed in HICs. For instance, the initial fixed costs associated with building data infrastructure or monitoring systems may require different investments. According to Fritsche et al. [3], PBF programs tend to require about five percent of additional financing in Organization for Economic Cooperation and Development (OECD) countries compared to 30-40 percent of additional financing in LMICs. Moreover, factors unrelated to the motivation of health workers or outside of their control may affect healthcare services to a greater degree in LMICs compared to HICs. On the provider side, these factors may be related to the lack of continuous training, drug supplies, tools and the availability of other resources. On the service-user side, these factors can be related to the difficulty of paying direct and indirect user-fees [6], [7]. Thus, it is important to evaluate whether PBF represents good value for money specifically within the context of LMICs.

The objective of this systematic review is to identify and synthesize the existing literature that examines whether PBF represents an efficient manner of investing resources. In line with Emmert et al.'s approach [4], pay-for-performance (P4P) is considered efficient when improved care quality or quantity is achieved with equal or lower costs, or alternatively, when the same quality or quantity of care is achieved using less financial resources.

REVIEW

Methods

Protocol and registration

We conducted a systematic review to identify and synthesize literature on economic evaluations of PBF in LMICs. This review is in line with the PRISMA statement [8]. The initial protocol was not registered.

Eligibility criteria

Inclusion criteria: In this systematic review, we included: 1) studies conducted in LMICs, as define by the World Bank [9]; 2) studies using experimental or observational designs to assess the costs (or inputs) and consequences (or outputs); and 3) studies in which a comparison between alternatives was made (including the status quo). We included studies that were primarily impact evaluations only if they also presented results on the costs of PBF. Following Drummond et al.’s [10] categorization scheme, we differentiates studies depending on whether costs, consequences, or both were considered. This approach results in a classification that distinguishes: ‘‘Type I’’ studies as full economic evaluations that make a clear connection between the costs and consequences of two or more alternatives (e.g., cost-effectiveness analyses, cost-utility analyses or cost-benefit analyses); ‘‘Type II’’ as partial economic evaluations that describe the costs and consequences of initiatives without making a clear connection between the two; ‘‘Type III’’ as studies that compare the costs of the initiatives without providing an effectiveness analysis regarding the health services or health outcomes; ‘‘Type IV’’ as studies that provide information on the costs of a PBF initiative without any description of changes in healthcare services or health outcomes [4]. To avoid overlooking important literature, we included articles belonging to these four types of economic evaluation studies.

Exclusion criteria: In this systematic review, we excluded: 1) studies conducted in HICs, as defined by the World Bank [9]; 2) publications that did not provide empirical evidence, such as editorials and interviews; 3) non-comparative evaluations because full economic evaluations require the comparison of two alternatives; 4) studies that only described a PBF program or solely evaluate their effectiveness; and 5) studies that focused only on demand-side financial incentives, such as financial compensations or bonuses for people who seek healthcare.

Table 1 Inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
Language	English, French	Other languages
Publication type	All documents presenting empirical data (e.g., peer-reviewed articles)	Protocols, editorials, guidelines and interviews
Study type	Experimental or observational studies including a quantitative assessment of 1) costs and effects, or 2) costs alone	Qualitative studies or studies that only examine effects
Economic evaluation type	Comparative evaluations: full economic evaluations and partial economic evaluations	Non-comparative evaluations
Targeted entity	Healthcare providers	Solely patients
Country	LMICs	HICs

Information sources

Searching in previous systematic reviews: We began our search by manually screening the reference lists of two recent systematic reviews to find economic evaluations of PBF focusing specifically on LMICs. A well-cited review, conducted by Emmert et al. [4], covered economic evaluations of PBF published between January 2000 and April 2010. The authors did not impose location-related restrictions. Meacock et al. [5] repeated the same search in September 2012 to ensure that no recent articles were omitted. We also screened the reference lists of additional relevant reviews that came to our attention during our search [11]–[14]. By reviewing past systematic reviews, we were able to identify pertinent studies published between January 2000 and September 2012.

Searching in databases: As Rethlefsen et al. [15] recommend, we collaborated with a professional librarian from the University of Montreal. We adapted Emmert et al.'s [4] search strategy to find more recent literature on economic evaluations of PBF in LMICs. Our search differed from Emmert et al. [4]'s in that we: 1) added Mesh terms and descriptors to expand the search; 2) modified the list of search terms by using more truncated terms (e.g., "cost*" includes "cost-effectiveness"); 3) deleted currency-related terms (e.g., dollars, yen) to better target pertinent results, given the rapid expansion of PBF worldwide; and 4) updated the inclusion and exclusion criteria (see below).

We conducted electronic searches in two databases: PubMed and Econlit. Search limits included studies written in English and French, published between January 2012 and June 2014. These dates allowed us to have an overlap with the timeframe covered by previous systematic reviews to avoid missing any pertinent articles [5]. The complete search history is available in Appendix 1.

In addition to the two databases listed above, we used Google and Google Scholar to identify other potentially relevant documents such as books, unpublished studies, study protocols, conference articles, and new PBF initiatives. We consulted the websites of governmental and scientific institutes concerned with PBF (e.g., the World Bank's website on RBF, www.rbfhealth.org; the Global Fund, http://www.theglobalfund.org). In addition, we contacted health economics experts to request information on additional ongoing or recently completed studies. We provided them with a list of the articles selected for this review and invited them to identify any missing article.

Study selection

One investigator judged titles and abstracts of potentially relevant studies according to inclusion and exclusion criteria (Table 1). When the investigator could not reach a final decision based on the abstract solely, she proceeded to review the full text. If a decision was still unattainable a second investigator reviewed the article before reaching a consensual decision. Two investigators read and appraised the articles selected.

Data items and extraction

Two members of the research team performed data extraction. The data extraction forms were custom-designed. The following information was extracted to summarize the articles: first author, publication year, country where study was conducted, characteristics of the PBF program, study objective (implicit or explicit), sample size, data gathering techniques, primary data analysis approach and main results of the study in relation to our focus.

Summary measures and data synthesis

The studies selected used a variety of principal summary measures (e.g., technical efficiency scores, Malmquist productivity index, difference in costs). Where possible, we present the effects of the interventions as the difference between the intervention and control groups at baseline and follow up percentages or scores (see Table 2). We could not perform a meta-analysis due to heterogeneity of studies and presentation of results.

Appraising methodological and reporting quality of included studies

We appraised the results of the studies by examining the relation established between the costs and consequences; the alternative interventions that were considered; the costs and consequences that were included or omitted; the study limitations; and potential conflicts of interests (see Table 3). To help us synthesize our assessment of the overall strength of the evidence, we developed a concise list of questions, adapted from Drummond et al. [10].

1. Was a clear relation between costs and consequences demonstrated empirically?
2. Which types of designs were used to assess the effectiveness of PBF?
3. Were different types of interventions considered as alternatives?
4. Were the costs (or inputs) and consequences measured longitudinally to examine change over time?
5. Were all important costs (or inputs) and consequences considered?
6. Were the studies conducted in different countries and contexts?
7. Did the authors report potential conflicts of interest?

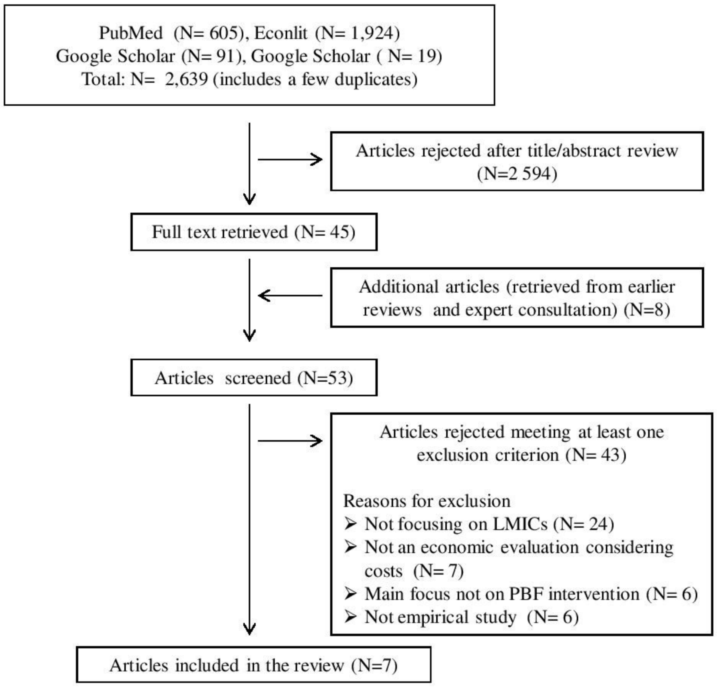
Results

Study selection

In total, we identified 2, 639 potentially relevant articles throughout PubMed, Econlit, Google Scholar and Google. After eliminating duplicates and reviewing the remaining abstracts, 45 studies were retained for more detailed analysis. Screening reference lists from earlier reviews and expert consultations yielded 8 additional articles. Thus, 53 full texts were assessed. Of these, seven studies met our inclusion criteria (Fig. 1) and were included in the review.

Appendix 3 presents a list of articles that were screened, but then excluded. The most common reason for exclusion was that the articles did not focus on LMICs.

Figure 1: Search flow and results



Study characteristics and appraisal

We present a summary of each study's characteristics in Table 2. The table highlights the diversity of intervention designs, study methods and outcomes. We also provide a summary of our appraisal for each study in Table 3.

Table 2. Characteristics of included studies

Author (year) Country	PBF program	Objectives	Sample	Data gathering	Data analysis	Main results
Bowser **(2014)** **Belize**	National Health Insurance (NHI) using performance contracts. Implemented in 2001. Expanded in 2006.	To assess trends in financial sustainability, efficiency payments, bonuses and health system and health outcomes.	*Contracted facility areas*: 3 private, 5 public. *Non-contracted facility areas*: providers in three districts financed by the MOH.	Data obtained from databases at the facility, district and national levels.	Descriptive trend analysis.	Per capita spending on health services provided by the NHI program decreased from approximately BZ$177 (i.e., US$ 89) to BZ$ 136 (i.e., US$ 68) between 2006 and 2009.
					Difference-in-difference approach (technical efficiency indicators).	NIH-contracted facility areas had greater improvements in facility births, nurse density, reducing maternal mortality, diabetes deaths, and morbidity compared to non-contracted areas. However, NIH-contracted facility areas had worst outcomes for physician density and death per hypertension between 2006 and 2010.
Gok **(2014)** **Turkey**	Pay-for-performance (P4P) program implemented in public and private hospitals implemented since 2004.	To analyze the effects of the P4P system on the hospitals' efficiencies.	251 hospitals of which 25 are private and 226 are public.	Data obtained from the Annual Statistical Health Report (2001–2008) and the Statistical Institute.	Data envelopment analysis (technical efficiency scores).	In public hospitals, the average efficiencies increased from 0.68 in 2005 to 0.73 in 2008, after the P4P system was adopted. In private hospitals, the average efficiencies decreased from 0.75 in 2005 to 0.61 in 2008.
					Productivity trends (Malmquist Productivity Index).	In public hospitals, the efficiency trend increased from 0.981 in the pre-P4P period to 1.018 after the implementation of the PFP system. In private hospitals, the efficiency trend decreased from

						1.016 in the pre-PFP period to 0.967 after the implementation of the P4P system.
Zeng **(2013)** **Haiti**	PBF program initiated in 1999 and scaled-up in 2005. Funded by USAID.	To evaluate the costs of implementation as well as the impact of PBF and/or international support (training & monitoring) on primary healthcare services.	15 health centers with PBF and 202 without PBF.	Routine data on the quantity of services provided & 12 interviews with NGO and health facility management staff.	Difference-in-differences approach (growth of incentivized vs non-incentivized services).	Incentive payments added 6% to base costs of PBF while international support added 39%. Incentives alone were associated with a 39% increase in health services. Support alone was associated with a 35% increase in health services. Support combined with incentives was associated with an 87% increase compared with health facilities that did not receive either. Non-incentivized services did not perform significantly lower than incentivized services.
Basinga **(2011)** **Rwanda**	P4P scheme providing incentives to providers for improvements in utilisation and quality of care. National program gradually	Assess the effect of P4P on the use and quality of child and maternal care services.	80 health facilities were assigned to a P4P program and 86 health facilities were assigned to be control facilities. 2 158	Facilities and households were surveyed at baseline and after 23 months.	Descriptive statistics from annual reports at the national level.	The administrative costs associated with P4P were estimated to be US$ 0.3 per person in total, representing 0.8% of total health expenditures per person and 1.2% of public and donor expenditures combined.

	implemented since 2005, after pilot schemes by NGOs.		households were also included.		Difference-in-difference model (multivariate regression).	The intervention group had a 23% increase in institutional deliveries, a 56% increase in preventive care visits by children aged 23 months or younger, and a 132% increase in preventative care visits by children between 24 and 59 months, compared to the control group. However, there were no improvement in the number of women receiving any prenatal care, the number of women completing four or more prenatal visits, and the number of children receiving full immunisation schedules.
Rusa **(2009)** **Rwanda**	PBF (reimbursement mechanism with 'indicator purchasing' linked to formative supervision). Implemented in 2005. Expanded in 2006. Funded by the Belgian Cooperation.	To evaluate the effect of PBF on the performance of healthcare centers.	74 health centers that implemented PBF in 2005 and 85 health centers that implemented PBF in 2006.	Data on services were collected on a monthly basis by the district supervisors.	Time-series with a two-staged implementation but only descriptive statistics.	The part of the subsidies spent on the functioning of the health facility, grew from approximately 8% in 2005, to 23% in 2006 and to 38% in 2007. Overall, the budget allocated to the implementation of a PBF program amounted to US$ 0.25/cap/year, of which US$ 0.20/cap/year for subsidies and an estimated US$ 0.05 /cap/year for administration, supervision and training. Results showed a positive effect for activities that were less organized (i.e., monitoring services and institutional deliveries). No effects were found on curative consultations, family planning, antenatal consultations and

						vaccinations. Compliance rates with norms rose in both groups.
Sabri (2007) Afghanistan	3 NGO contracting programs with capitation payments to providers for each individual enrolled. Implemented since 2001. Funded by World Bank, USAID or European Commission.	To analyze the financing and costs of contracting healthcare services.	No description provided.	Statistics from government and NGO reports.	Descriptive statistics.	The reference cost used to negotiate the delivery of a basic package of health services with contracted NGOs was estimated to be US$ 4.5 for 2002. The cost varied among the different donors. The annual per-capita cost was US$ 3.8 for the World Bank, US$ 4.2 for USAID and US$ 5.1 for the European Commission. The population coverage for basic health services increased from 9% in 2002 to 82% in 2006. However, the quality of services provided appeared to be poor (ex.: long waiting times, absence of laboratory services, shortage of drugs, and disrespect for patients). Facilities run under the ministry's strengthening mechanism and NGO contracts under the World Bank and the USAID performed better than

						contracts held by the European Commission due to cumbersome administrative procedures. Authors discuss the preliminary results of an Afghanistan household survey suggesting that under five child and infant mortality rates improved.
Soeters (2006) Rwanda	P4P program introduced in 2002 by Cordaid.	To present Rwanda's P4P experience.	240 and 320 households in province with P4P.	Household surveys in 2003 and 2005.	Difference-in-difference approach (no clear description of analyses).	Out-of-pocket health expenditure decreased by 62%, from US$ 9.05 to US$ 3.45. The percentage of respondents who experienced a catastrophic user fee payments decreased from 2.5% in 2003 to 0.7% in 2005. The proportion of women delivering in a health facility increased from 25% to 60%. In the discussion, authors indicate that the administrative costs of the fundholder were about 25% of the total contracting costs, according to Cordaid data.

Table 3. Appraisal of included studies

Author (year)	Was a clear relation between costs & consequences established?	Which alternative intervention was considered?	The costs (or inputs) and consequences were measured over which timeframe?	Were important costs (or inputs) and consequences omitted?	What were the limitations?	Are there potential conflicts of interest?
Bowser (2014)	No	Status quo (traditional salaries and line-item budgets).	2006 to 2010	No clear description of the included and omitted costs.	1) absence of pre-intervention data; 2) possibility that other factors influenced the costs per capita; 3) difficulty of teasing apart the effects due to the incentives from those related to other components of the reform.	- None declared - 1 author affiliated with the organisation involved in the implementation
Gok (2014)	No	Status quo (before vs after P4P).	2001-2008	Yes, for example, the costs of implementing the program were not included.	1) the absence of randomization; 2) the lack of a control group; and 3) the use of aggregate input and output variables.	- None declared
Zeng (2013)	No	International support (including procurement procedures, minor renovations, advice on community mobilization, communication, public relations & promotion of family planning).	2008-2010	No clear description of the included and omitted costs for the "international support".	1) absence of randomization; 2) the absence of pre-intervention data ; 3) the lack of control for the quality of the data in the 202 health centers without PBF.	- Declaration that one co-author was employed by an organisation involved in PBF.
Basinga (2011)	No	Input-based budgets in the control group were increased by the average P4P payments made to the intervention group	June, 2006 to Avril 2008 (~ 23 months)	Lack of detailed information on the costs of PBF. Health outcomes were not included.	1) the absence of pre-intervention data; 2) problems identified with allocation to treatment and control-groups (see Witter et al., 2013).	- None declared - Some authors affiliated with organisations involved in the

						funding and implementation.
Rusa (2009)	No	Status quo for performance data (3 months of pre-intervention data). No alternative intervention was used to compare costs.	2005 to 2007 for costs. October 2014 to December 2007 for performance.	Includes subsidies and administration costs. No detailed description of the included and omitted costs. Health outcomes were not included.	1) insufficient use of pre-intervention data; 2) the lack of a control group without PBF during the entire time period; and 3) the possibility that other interventions (e.g., mutual health organizations, sensitization campaigns) influenced the results.	- None declared - 5 of 6 authors affiliated with organisations involved in the implementation
Sabri (2007)	No	Comparison of three different PBF programs.	2002 to 2006	No clear description of the included and omitted costs. Limited data on healthcare services and health outcomes.	1) the lack of information on the methodology used; and 2) the absence of links between the costs and outcomes.	- None declared - At least one author employed by an organisation involved in the implementation
Soeters (2006)	No	Comparison of PBF in early vs later stages.	2003 to 2005	Lack information on how PBF affects total health expenditures (only focuses on out-of-pocket health spending).	1) absence of pre-intervention data; 2) absence of a control group;) possibility that other interventions occurring simultaneously reduced catastrophic user fee payments.	- None declared - Authors worked for an organisation involved in the implementation

Synthesis of results and appraisal

The section below presents our overall assessment of the strength of the evidence, using the list of questions we adapted from Drummond et al. [10].

1. Was a clear relation between costs and consequences demonstrated empirically?

None of the included studies were classified as full economic evaluations that make clear connexions between the PBF costs and healthcare services and/or health (Type I). In other words, none of the studies included cost-effectiveness analyses, cost-utility analyses or cost-benefit analyses. For this reason, we classified the 7 studies as partial economic evaluations (Type II), as. they described the costs and consequences of PBF initiatives without making a clear connection between the two. It is important to note that full economic evaluations are necessary to evaluate whether PBF provides good value for money in LMICs because they are more methodologically sound than partial economic evaluations [10].

2. Which types of designs were used to assess the effectiveness of PBF?

An intervention that is not effective cannot provide good value for money. Therefore, we examined the designs that were used to assess the effectiveness of PBF in the included studies. Of the seven articles, only one study reported using a randomized control trial to assess the consequences of PBF [16]. However, Witter and colleagues [14] have identified problems with the allocation to the treatment and control-groups for this study. It appears that some districts were found to have existing pay for performance schemes, requiring the allocation to be adjusted in a non-random way. This study found that the intervention group had an increase in institutional deliveries and preventive care visits, compared to the control group. However, there were no improvement in the number of women receiving any prenatal care; the number of women completing four or more prenatal visits; and the number of children receiving full immunisation schedules[16]. The other articles included in this review adopted a variety of observational designs, for instance, relying on difference-in-difference estimates, time series and trend analyses. The majority of studies did not use

---

[16] It should be noted that a recent study published using the same data found that there no measurable difference in estimated probability of reporting illness with diarrhea, fever or acute respiratory infections between the intervention and comparison groups [17].

pre-intervention data in their analyses. Potential biases and mitigated results limit our confidence in the effectiveness of PBF programs, as presented in the studies.

3. Were different types of interventions considered as alternatives?

Economic evaluations require the comparison of two alternatives to identify which is more efficient [2]. Most studies in this review compared the implementation of PBF to the status quo. System-strengthening alternatives to improve the motivation of healthcare workers or service delivery were not used as comparators. Potential alternatives that could have been considered to test whether PBF provides the best value for money include: other funding mechanisms; monitoring (without financial incentives); providing performance feedback; training health workers; increasing leadership skills; encouraging collaboration; and fostering a culture that promotes trust and the intrinsic value of work [18]. In addition, more studies should attempt to tease apart the incentive effect from the resource effect. Only one study included in this systematic review increased the budgets of the PBF intervention and control groups by the same amount [16].

4. Were the costs (or inputs) and consequences measured longitudinally to examine change over time?

The seven articles examined the impact of PBF programs over different time periods. Gok & Altmdag (2014)'s study ranges from 2001-2008; Bowser et al. (2014) and Sabri et al. (2007)'s study cover a four year time period; and Zeng et al. (2013), Basinga et al. (2011), Rusa et al. (2009), and Soeters et al. (2006) report change over a two year period.

From the studies in this review, little is known about how the relation between PBF costs and outcomes in LMICs evolves in the long term.

5. Were all important costs (or inputs) and consequences considered?

The studies did not provide a detailed description of the costs that were included or omitted. The studies mostly examined the immediate/direct financial costs and effects of the interventions. Authors generally did not attempt or were not able to quantify all the different types of costs and inputs (e.g., time and funds invested to monitor the delivery of health services, time spent filling out forms). Only aggregated costs were presented.

Overall, important effects on health outcomes and unintended consequences (e.g., reduction of healthcare services not rewarded financially) were not sufficiently considered.

6. Were the studies conducted in different countries and contexts)?

The seven articles were conducted in only 5 LMICs. Table 4 presents the number of articles, the region and the income level for each of these countries. Many regions and countries currently implementing PBF are not represented in these studies [19]. Moreover, some countries like Rwanda are characterised by unique political contexts and demographic situations, limiting the generalizability of results to other countries.

Table 4: Countries classified according to region and income level

Country	**Number of articles**	**Region**	**Income level**
Rwanda	3	Sub-Saharan Africa	Low-income economy
Belize	1	Latin America and the Caribbean	Upper middle-income economy
Haiti	1	Latin America and the Caribbean	Low-income economy
Afghanistan	1	South Asia	Upper middle-income economy
Turkey	1	Europe and Central Asia	Upper middle-income economy

*This classification is based on World Bank criteria.

7. Did the authors report potential conflicts of interest?

Six out of 7 articles had at least one author that was or had been affiliated with an organisation involved in the implementation of PBF, thereby resulting in a potential conflict of interest. The interpretation of data or presentation of information may have been influenced by their personal or financial relationship with other people or organizations. Interestingly, only one author explicitly reported having been employed by an organization involved in the implementation of PBF as a potential conflict of interest [20].

Summary of the assessment

Only seven articles fit out inclusion criteria. Overall, the evidence of economic evaluations of PBF is weak for the following reasons: (1) none of the studies were full economic evaluations; (2) only one study used a randomized controlled trial, but issues with the randomization procedure were reported; (3) important alternative interventions to strengthen the capacities of the healthcare system have not been used as a comparator; (4) few studies examined the costs and consequences of PBF in the long term; (5) important costs and consequences were omitted from the evaluation; (6) very few LMICs are represented in the literature, despite wide implementation in these countries and (7) most articles had at least one author that was affiliated with an organisation involved in the implementation of PBF, thereby resulting in a potential conflict of interest.

DISCUSSION

This systematic review highlights a lack of strong empirical evidence that supports the idea that PBF increases value for money in LMICs. This result is consistent with past findings [4], [5], [11], [14]. For example, a Cochrane review addressing the effectiveness of PBF in LMICs found that the current evidence base is too weak to draw general conclusions about the effectiveness of PBF in LMICs. Without reasonable reliable effectiveness-estimates, cost-effectiveness estimates cannot be calculated. Thus, it would have been surprising if this review had concluded differently.

The added value of this review is threefold. First, replications of past reviews are useful to validate results and find articles that might have been overlooked. Second, past reviews only included studies published up to 2011-2012. An update was therefore warranted, especially considering the rapid implementation of PBF in LMICs and the large number of studies that have published on PBF since then. Third, this is the first literature review with a search strategy that specifically targeted articles on the efficiency of PBF in LMICs. Thus, the current review has a different focus than past reviews, providing a collection of economic evaluations of PBF in LMICs that were not previously identified. For example, 6 of the 7 studies in this systematic review were not included in the Cochrane review. Three of the studies were published after the Cochrane authors conducted their search [20]–[22]. The three other studies included in this systematic review, but not in the Cochrane review, were published and available in time to be considered [23]–[25]. However, they were not included and are not mentioned under "excluded studies" in the Cochrane review. Consequently, our systematic review may be useful to inform researchers and decision-makers specifically concerned with optimizing value for money in LMICs.

The reasons why so few PBF economic evaluations have been conducted in LMICs, despite wide implementation, is worth exploring. First, PBF is a complex intervention that targets multiple services. It is therefore difficult to evaluate the impact of PBF on health. Economic evaluations on this topic require complex modelling because diverse people and many conditions are affected. Second, it is difficult to obtain good quality cost data in LMICs because the information is not easily accessible. Last, international partners occasionally resist sharing their costs, usually substantial at start up. Promoting transparency may be useful to facilitate economic evaluations on PBF.

Strengths and limitations

While systematic reviews can take years to complete, this review was conducted within a few months to respond to timely concerns about whether PBF provides the best value for money in LMICs. The time frame usually required for producing systematic reviews has been found to be inappropriate for local policy makers that have urgent decisions to make [26]. This issue was highlighted by a decision-maker in Haiti, who widely shared an e-poster on the current results, sadly claiming that "long publication delays would eliminate the important benefits of this review" (personal communication, June 13, 2015). Despite its rapidity, this review adheres to the core principles of systematic reviews in order to avoid bias and ensure rigor. A detailed description of the methods used was provided to promote methodological transparency, and to facilitate replication.

Our review has limitations. First, the studies varied in methodological quality and study characteristics. These differences made it difficult to adequately compare the results of the articles included in our systematic review. Second, as in the case with most reviews, our review might have suffered from publication bias. Sponsors of inefficient PBF programs may have blocked publishing to protect their interests [4]. Last, as with any review, we may have missed some relevant information during the selection and data extraction process.

Future directions

Future researchers and evaluators should attempt to make a direct relation between costs and consequences of PBF in order to draw conclusions about whether this financing option represents good value for money. There is a need to adopt stronger designs and to consider the long-term implications of these programs on costs and health outcomes. In addition, future studies should compare PBF to promising alternative interventions that aim to strengthen the healthcare system. It would also be beneficial to analyze the literature around PBF in LMICs using Drummond and Jefferson (1996)'s 38 defined quality criteria, as seen in Emmert et al. (2012)'s systematic review, in order to generate an average quality score for each article.

During our search, it has come to our attention that at least two economic evaluations of PBF are currently being conducted in LMICs. Borghi et al. [27] published a protocol on the evaluation of a P4P program in Tanzania. Using a controlled before and after study, the authors aim to measure the cost-effectiveness of the P4P program. Moreover, two economic evaluations are being conducted on PBF initiatives in Malawi [28]. Together, these studies should contribute to the evidence on the efficiency of PBF in LMICs.

CONCLUSION

In contexts of limited resources such as LMICs, it is essential that funders and decision-makers aim to optimize the value obtained from the money invested in healthcare services, in order to address the pressing health needs of the population. Some stakeholders have proposed PBF as a promising avenue. However, this review has demonstrated that there is a lack of empirical evidence to support the claim that PBF represents value for money. We still do not know if, given the same amount of resources, PBF buys more healthcare services or health than the status quo or other interventions. Full economic evaluations of PBF are needed to truly inform decision-makers in LMICs on how to make better use of limited resources to improve population health.

COMPETING INTERESTS

VR and MD are co-researchers on the baseline study of an impact evaluation of PBF in Burkina Faso but they have not received any salary from the funder (World Bank) for this activity. VR's participation fees for the PBF workshop in Argentina in 2014 were supported by the World Bank. The other authors have no conflict of interests regarding the publication of this paper.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Emmanuel Guindon for his comments on an earlier version of this manuscript. The authors would also like to thank the health economic experts for their feedback, namely on the search results.

FUNDING

AMTT received a training bursary from the Canadian Institutes of Health Research (CIHR). JS received a training bursary from the Fonds de recherche du Québec - Santé (FRQS). VR holds a CIHR-funded Research Chair in Applied Public Health. Sponsors did not have a role in the study design; the collection, analysis and interpretation of data; the writing of the report; and the decision to submit the article for publication.

REFERENCES

[1] World Health Organization, « World Health Report. Health Systems Financing. The Path to Universal Coverage. », 2010.

[2] J. Hurley, *Health Economics*, First Edition. Toronto: McGraw-Hill Ryerson, 2010.

[3] G. B. Fritsche, R. Soeters, et B. Meessen, *Performance-Based Financing ToolKit*. Washington, DC: World Bank, 2014.

[4] M. Emmert, F. Eijkenaar, H. Kemter, A. S. Esslinger, et O. Schöffski, « Economic evaluation of pay-for-performance in health care: a systematic review », *Eur. J. Health Econ. HEPAC Health Econ. Prev. Care*, vol. 13, nº 6, p. 755-767, déc. 2012.

[5] R. Meacock, S. R. Kristensen, et M. Sutton, « The cost-effectiveness of using financial incentives to improve provider quality: a framework and application », *Health Econ.*, vol. 23, nº 1, p. 1-13, janv. 2014.

[6] P. Marschall et S. Flessa, « Efficiency of primary care in rural Burkina Faso. A two-stage DEA analysis », *Health Econ. Rev.*, vol. 1, nº 1, p. 5, 2011.

[7] V. Ridde, A.-M. Turcotte-Tremblay, A. Souares, J. Lohmann, D. Zombré, J. L. Koulidiati, M. Yaogo, H. Hien, M. Hunt, S. Zongo, et M. D. Allegri, « Protocol for the Process Evaluation of Interventions Combining Performance-Based Financing with Health Equity in Burkina Faso », *Implement. Sci.*, vol. 9, nº 1, p. 149, oct. 2014.

[8] PRISMA, « PRISMA. Transparent reporting of systematic reviews and meta-analyses. » [En ligne]. Disponible à: http://www.prisma-statement.org/.

[9] World Bank, « Country and Lending Groups | Data », 2016. [En ligne]. Disponible à: http://data.worldbank.org/about/country-and-lending-groups. [Consulté le: 29-juin-2014].

[10] M. F. Drummond, M. J. Sculpher, G. W. Torrance, B. J. O'Brien, et G. L. Stoddart, *Methods for the economic evaluation of health care programmes*, Oxford University Press. Oxford, 2005.

[11] F. Eijkenaar, M. Emmert, M. Scheppach, et O. Schöffski, « Effects of pay for performance in health care: a systematic review of systematic reviews », *Health Policy Amst. Neth.*, vol. 110, nº 2-3, p. 115-130, mai 2013.

[12] G. Flodgren, M. P. Eccles, S. Shepperd, A. Scott, E. Parmelli, et F. R. Beyer, « An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes », *Cochrane Database Syst. Rev.*, nº 7, p. CD009255, 2011.

[13] A. C. Gorter, P. Ir, et B. Meessen, « Evidence Review. Results-Based Financing of Maternal and Newborn Health Care in Low- and Lower-Middle-Income Countries », German Federal Ministry for Economic Cooperation & Development, 2013.

[14] S. Witter, A. Fretheim, F. L. Kessy, et A. K. Lindahl, « Paying for Performance to Improve the Delivery of Health Interventions in Low- and Middle-Income Countries (Review) », *Cochrane Database Syst. Rev.*, vol. 2, p. 1-82, 2013.

[15] M. Rethlefsen, H. H. Murad, et E. Livingston, « Engaging Medical Librarians to Improve the Quality of Review Articles », *The Journal of the America Medical Association*, 2014.

[16] P. Basinga, P. J. Gertler, A. Binagwaho, A. L. Soucat, J. Sturdy, et C. M. Vermeersch, « Effect on Maternal and Child Health Services in Rwanda of Payment to Primary Health-Care Providers for Performance: An Impact Evaluation », *The Lancet*, vol. 377, nº 9775, p. 1421-1428, 2011.

[17] M. P. Skiles, S. L. Curtis, P. Basinga, G. Angeles, et H. Thirumurthy, « The effect of performance-based financing on illness, care-seeking and treatment among children: an impact evaluation in Rwanda », *BMC Health Serv. Res.*, vol. 15, sept. 2015.

[18] V. Ridde, « Building trust or buying results? », *The Lancet*, vol. 366, nº 9498, p. 1692, nov. 2005.

[19] World Bank, « Results-Based financing for Health », 2014. [En ligne]. Disponible à: https://www.rbfhealth.org/project/our-projects.

[20] W. Zeng, M. Cros, K. D. Wright, et D. S. Shepard, « Impact of performance-based financing on primary health care services in Haiti », *Health Policy Plan.*, vol. 28, nº 6, p. 596-605, sept. 2013.

[21] D. M. Bowser, R. Figueroa, L. Natiq, et A. Okunogbe, « A preliminary assessment of financial stability, efficiency, health systems and health outcomes using performance-based contracts in Belize », *Glob. Public Health*, vol. 8, nº 9, p. 1063-1074, 2013.

[22] M. S. Gok et E. Altındağ, « Analysis of the cost and efficiency relationship: experience in the Turkish pay for performance system », *Eur. J. Health Econ. HEPAC Health Econ. Prev. Care*, 2015.

[23] R. Soeters, C. Habineza, et P. B. Peerenboom, « Performance-based financing and changing the district health system: experience from Rwanda. », *Bull. World Health Organ.*, vol. 84, nº 11, p. 884-889, nov. 2006.

[24] B. Sabri, S. Siddiqi, A. M. Ahmed, F. K. Kakar, et J. Perrot, « Towards sustainable delivery of health services in Afghanistan: options for the future », *World Hosp. Health Serv. Off. J. Int. Hosp. Fed.*, vol. 43, nº 3, p. 10-16, 2007.

[25] L. Rusa, J. de D. Ngirabega, W. Janssen, S. Van Bastelaere, D. Porignon, et W. Vandenbulcke, « Performance-based financing for better quality of services in Rwandan health centres: 3-year experience », *Trop. Med. Int. Health TM IH*, vol. 14, nº 7, p. 830-837, juill. 2009.

[26] S. Oliver et K. Dickson, « Policy-relevant systematic reviews to strengthen health systems: models and mechanisms to support their production », *Evidence and Policy*, nov-2015.

[27] J. Borghi, I. Mayumana, I. Mashasi, P. Binyaruka, E. Patouillard, I. Njau, O. Maestad, S. Abdulla, et M. Mamdani, « Protocol for the evaluation of a pay for performance programme in Pwani region in Tanzania: a controlled before and after study », *Implement. Sci. IS*, vol. 8, p. 80, 2013.

[28] S. Brenner, A. S. Muula, P. J. Robyn, T. Bärnighausen, M. Sarker, D. P. Mathanga, T. Bossert, et M. D. Allegri, « Design of an impact evaluation using a mixed methods model – an explanatory assessment of the effects of results-based financing mechanisms on maternal healthcare services in Malawi », *BMC Health Serv. Res.*, vol. 14, nº 1, p. 180, avr. 2014.

APPENDIX 1 : LITERATURE SEARCH HISTORY

Appendix 1.1 : PubMed Search History

#1: 2014/06/20; N=2079

Search (((("pay for performance" OR "P4P" OR "PFP" OR "pay for value" OR "pay for quality" OR "payment for quality" OR "value-based purchasing" OR ("financial incentive*" AND quality) OR ("monetary incentive*" AND quality) OR (bonus AND quality) OR (reward* AND quality) OR "performance-based payment" OR "performance-based reimbursement" OR "performance-based contracting" OR "performance-based pay" OR "performance-based financing" OR "results-based financing" OR "output-based payment" OR "incentive reimbursement" OR "incentive program" OR "quality-based purchasing" OR "quality incentive" OR "Quality Improvement/economics"[Mesh] OR "Reimbursement, Incentive/economics"[Mesh] OR "Financing, Government/methods"[Mesh] OR "Quality Assurance, Health Care/economics"[Mesh]))) AND ((French[Language]) OR English[Language])) AND ("2012/01/01"[Date - Publication] : "2014/06/20"[Date - Publication])

#2: 2014/06/20; N=157 204

Search (((("program evaluation*" OR "economic evaluation*" OR "financial analysis" OR "cost*" OR "profit" OR "efficiency" OR "efficient" OR "return on investment" OR "ROI" OR "rate of return" OR "net present value" OR "benefit-cost ratio" OR "business case" OR "economic case" OR "social case" OR "quality-adjusted life years" OR "qaly*" OR "Costs and Cost Analysis/economics"[mesh] OR "Cost Control/economics"[mesh] OR "Cost Savings/economics"[mesh] OR "Cost-Benefit Analysis/economics"[mesh] OR "Program Evaluation/economics"[mesh] OR "Health Services Research/economics"[mesh] OR "Utilization Review/economics"[mesh] OR "Efficiency, Organizational/economics"[mesh]))) AND ((French[Language]) OR English[Language])) AND ("2012/01/01"[Date - Publication] : "2014/06/20"[Date - Publication])

#3: 2014/06/20; N= 605

Search #1 AND #2

Appendix 1.2 Econlit Search History

#1: 2014/06/20; N= 10,030

AB (("pay for performance" OR "p4p" OR "pfp" OR "pay for value" OR "pay for quality" OR "payment for quality" OR "value-based purchasing" OR ("financial incentive*" AND "quality") OR ("monetary incentive*" AND "quality") OR ("bonus" AND "quality") OR ("reward*" AND "quality") OR "performance-based pay*" OR "performance-based reimbursement" OR "performance-based contracting" OR "output-based payment" OR "incentive reimbursement" OR "incentive program" OR "quality-based purchasing" OR "quality incentive" OR "performance-based financing" OR "results-based financing")) or TI (("pay for performance" OR "p4p" OR "pfp" OR "pay for value" OR "pay* for quality" OR "value-based purchasing" OR ("financial incentive*" AND "quality") OR ("monetary incentive*" AND "quality") OR ("bonus" AND "quality") OR ("reward*" AND "quality") OR "performance-based payment" OR "performance-based reimbursement" OR "performance-based contracting" OR "performance-based pay*" OR "output-based payment" OR "incentive reimbursement" OR "incentive program" OR "quality-based purchasing" OR "quality incentive" OR "performance-based financing" OR "results-based financing")) or SU (("Analysis of Health Care Markets" OR " Health: Government Policy; Regulation; Public Health" or "National Government Expenditures and Health " or "Economic Development: Human Resources; Human Development; Income Distribution; Migration"))

Dates: January 2012 to July 2014

#2: 2014/06/20; N= 24,137

AB (("program evaluation*" OR "economic evaluation*" OR "financial analysis" OR "saving*" OR "cost*" OR "profit" OR "efficiency" OR "efficient" OR "return on investment" OR "roi" OR "rate of return" OR "net present value" OR "business case" OR "economic case" OR "social case" OR "quality-adjusted life years" OR "qaly*")) or TI (("program evaluation*" OR "economic evaluation*" OR "financial analysis" OR "saving*" OR "cost*" OR "profit" OR "efficiency" OR "efficient" OR "return on investment" OR "roi" OR "rate of return" OR "net present value" OR "benefit-cost ratio" OR "business case" OR "economic case" OR "social case" OR "quality-adjusted life years" OR "qaly*"))

Dates: January 2012 to July 2014

#3: 2014/06/20; N= 1,924

#1 AND #2

Appendix 1.3 Google Scholar Search History

#1: 2014/06/25; N= 91

("P4P" OR "performance-based financing" OR "results-based financing") AND ("economic evaluation")

("2012"[Date - Publication] : "2014"[Date - Publication])

Appendix 1.4 Google Search History

#1: 2014/06/25; N= 19

("P4P" OR "performance-based financing" OR "results-based financing") AND ("economic evaluation")

("2012/01/01"[Date - Publication] : "2014/06/20"[Date - Publication])

APPENDIX 2: ABSTRACTS OF INCLUDED STUDIES

Appendix 2.1: Bowser et al., 2014

Over the last 10 years, Belize has implemented a National Health Insurance (NHI) program that uses performance-based contracts with both public and private facilities to improve financial sustainability, efficiency and service provision. Data were collected at the facility, district and national levels in order to assess trends in financial sustainability, efficiency payments, yearend bonuses and health system and health outcomes. A difference-indifference approach was used to assess the difference in technical efficiency between private and public facilities. The results show that per capita spending on services provided by the NHI program has decreased over the period 2006–2009 from BZ$177 to BZ$136. The private sector had achieved higher levels of technical efficiency, but lower percentages of efficiency and year-end bonus payments. Districts with contracts through the NHI program showed greater improvements in facility births, nurse density, reducing maternal mortality, diabetes deaths and morbidity from bronchitis, emphysema and asthma than districts without contracts over the period 2006–2010. This preliminary assessment of Belize's pay-for-performance system provides some positive results, however further research is needed to use the lessons learned from Belize to implement similar reforms in other systems.

Appendix 2.2: Gok and Altmdag (2014)

This paper analyzes the effects of the pay for performance (PFP) system on the efficiencies of public and private hospitals in Turkey. In order to evaluate these effects, we examine the relationship between hospital efficiency and health care costs in Turkey, and addressed the impact of the PFP system on the efficiencies of public and private hospitals. In an effort to analyze the efficiencies of public and private hospitals, this study used data envelopment analysis. The Malmquist Productivity Index is also used to analyze the patterns of efficiency change for the study years from 2001 to 2008. This study shows that health care costs and hospital efficiency are negatively correlated for private hospitals, while they are positively correlated for public hospitals. In other words, increased health care costs might reduce efficiency in private hospitals in contrast to public hospitals. Our findings also indicate that average efficiencies of public hospitals tend to increase, particularly during the implementation period ofPFP system. The efficiency trend of private hospitals, conversely, decreased in the latter periods of the PFP system. Suggestions for improvement are provided to the health care policy makers regarding the impact of health care reforms on public and private hospitals.

Appendix 2.3: Zeng et al., 2013

To strengthen Haiti's primary health care (PHC) system, the country first piloted performance-based financing (PBF) in 1999 and subsequently expanded the approach to most internationally funded non-government organizations. PBF complements support (training and technical assistance). This study evaluates (a) the separate impact of PBF and international support on PHC's service delivery; (b) the combined impact of PBF and technical assistance on PHC's service delivery; and (c) the costs of PBF implementation in Haiti. To minimize the risk of facilities neglecting potential non-incentivized services, the incentivized indicators were randomly chosen at the end of each year. We obtained quantities of key services from four departments for 217 health centres (15 with PBF and 202 without) from 2008 through 2010, computed quarterly growth rates and analysed the results using a difference-in-differences approach by comparing the growth of incentivized and non-incentivized services between PBF and non-PBF facilities. To interpret the statistical analyses, we also interviewed staff in four facilities. Whereas international support added 39% to base costs of PHC, incentive payments added only 6%. Support alone increased the quantities of PHC services over 3 years by 35% (2.7%/quarter). However, support plus incentives increased these amounts by 87% over 3 years (5.7%/

quarter) compared with facilities with neither input. Incentives alone was associated with a net 39% increase over this period, and more than doubled the growth of services (P<0.05). Interview findings found no adverse impacts and, in fact, indicated beneficial impacts on quality. Incentives proved to be a relatively inexpensive, well accepted and very effective complement to support, suggesting that a small amount of money, strategically used, can substantially improve PHC. Haiti's experience, after more than a decade of use, indicates that incentives are an effective tool to strengthen PHC.

Appendix 2.4: Basinga et al., 2011

Background: Evidence about the best methods with which to accelerate programs towards achieving the Millennium Development Goals is urgently needed. We assessed the effect of performance-based payment of health-care providers (payment for performance; P4P) on use and quality of child and maternal care services in health-care facilities in Rwanda.

Methods: 166 facilities were randomly assigned at the district level either to begin P4P funding between June, 2006, and October, 2006 (intervention group; n=80), or to continue with the traditional input-based funding until 23 months after study baseline (control group; n=86). Randomisation was done by coin toss. We surveyed facilities and 2158 households at baseline and after 23 months. The main outcome measures were prenatal care visits and institutional deliveries, quality of prenatal care, and child preventive care visits and immunisation. We isolated the incentive effect from the resource eff ect by increasing comparison facilities' input-based budgets by the average P4P payments made to the treatment facilities. We estimated a multivariate regression specification of the difference-in-difference model in which an individual's outcome is regressed against a dummy variable, indicating whether the facility received P4P that year, a facility-fixed effect, a year indicator, and a series of individual and household characteristics.

Findings: Our model estimated that facilities in the intervention group had a 23% increase in the number of institutional deliveries and increases in the number of preventive care visits by children aged 23 months or younger (56%) and aged between 24 months and 59 months (132%). No improvements were seen in the number of women completing four prenatal care visits or of children receiving full immunisation schedules. The authors also estimated an increase of 0·157 standard deviations (95% CI 0·026–0·289) in prenatal quality as measured by compliance with Rwandan prenatal care clinical practice guidelines.

Interpretation: The P4P scheme in Rwanda had the greatest effect on those services that had the highest payment rates and needed the least effort from the service provider. P4P financial performance incentives can improve both the use and quality of maternal and child health services, and could be a useful intervention to accelerate progress towards Millennium Development Goals for maternal and child health.

Appendix 2.5: Rusa et al., 2009

In 2005, the Ministry of Health in Rwanda, with the support of the Belgian Technical Cooperation, launched a strategy of performance-based financing (PBF) in a group of 74 health centres (HCs), covering 2-m inhabitants. In 2006, PBF was extended to an additional group of 85 HCs, thus reaching 3.8-m inhabitants. This study evaluates the effect of PBF on HC performance from 2005 to 2007. Composite indicators for measuring quantity and quality of services were developed and evaluated through monthly formative supervisions by qualified and well-trained district supervisors. The strategy was based on a fixed fee per quality-approved service. The entire budget spent on the implementation of PBF amounted to \$0.25 ⁄ cap ⁄ year, of which \$0.20 ⁄ cap ⁄ year for subsidies and an estimated \$0.05 ⁄ cap ⁄ year for administration, supervision and training. A positive effect on utilization rates was only seen for activities that were previously less well organized; in this case, growth monitoring services and institutional deliveries. The quality of services, defined as the compliance rate with national and international norms, rose considerably for all services in both groups. A sustained level of quality between 80% and 95% was reached within 18 months in the first group. A similar result was reached in the second group in 8 months.

Appendix 2.6: Sabri et al., 2007

Disruption caused by decades of war and civil strife in Afghanistan has led many international and national nongovernmental organizations (NGOs) to assume responsibility for the delivery of health services through contracts with donor agencies. Recently the Afghan Government has pursued the policy of contracting for a basic package of health services (BPHS) supported by funds from three major donors – the World Bank, the United States Agency for International Development (USAID) and the European Commission. With the gradual strengthening of the public health ministry, options for the future include pursuing the contracting option or increasing public provision of health services.

Should contracting with NGOs be pursued, a clear strategy is required that includes developing accreditation instruments, better contracting mechanisms and a system for monitoring and evaluating the entire process. Should the government opt for an increasing role, problems to be solved include securing the transition to public provision, obtaining guarantees that appropriate financing will be provided and reconfiguration of the public health delivery system. Large-scale contracting with the private for-profit sector cannot be recommended at this stage, although this option could be explored via subcontracting by larger NGOs or smallscale trial contracts initiated by the public health ministry. Irrespective of the option chosen, an important challenge remaining is the recalcitrant problem of high out-of-pocket payments.

Sustainable delivery of health services in Afghanistan can only be achieved with a clear national strategy in which all stakeholders have roles to play in the financing, regulation and delivery of services.

Appendix 2.7: Soeters et al., 2006

Evidence from low-income Asian countries shows that performance-based financing (as a specific form of contracting) can improve health service delivery more successfully than traditional input financing mechanisms. We report a field experience from Rwanda demonstrating that performance-based financing is a feasible strategy in sub-Saharan Africa too. Performance-based financing requires at least one new actor, an independent well equipped fundholder organization in the district health system separating the purchasing, service delivery as well as regulatory roles of local health authorities from the technical role of contract negotiation and fund disbursement. In Rwanda, local community groups, through patient surveys, verified the performance of health facilities and monitored consumer satisfaction. A precondition for the success of performance-based financing is that authorities must respect the autonomous management of health facilities competing for public subsidies. These changes are an opportunity to redistribute roles within the health district in a more transparent and efficient fashion.

#	Authors (year)	Reason for exclusion*
1	Agee & Gates (2012)	Not focused on LMIC
2	Allen, Nobel & Burton (2012)	Not focused on LMIC
3	Awoonor-Williams (2013)	Main focus not on PBF intervention
4	Baral (2012)	Not an economic evaluation
5	Bernstein (2014)	Not an empirical study
6	Blecker (2014)	Not focused on LMIC
7	Blumenthal, Song, Jena & Ferris (2013)	Not focused on LMIC
8	Broughton et al. (2013)	Main focus not on PBF intervention
9	Chee (2003)	Main focus not on PBF intervention
10	Cheng, Lee & Chen (2012)	Not focused on LMIC
11	Eichler et al. (2013)	Not an empirical study
12	Falisse, Meessen, Ndayishimiye & Brossuyt (2012)	Not an economic evaluation
13	Gerber-Grote & Windeler (2014)	Main focus not on PBF intervention
14	Ginsburg (2013)	Not focused on LMIC
15	Greene, Hibbard & Overton (2014)	Not focused on LMIC
16	Higgs,Stammer, Roth & Balster (2013)	Not an empirical study
17	Himmelstein, Ariely & Woolhandler (2014)	Not focused on LMIC
18	Holcombe (2014)	Not focused on LMIC
19	Hupp (2014)	Not focused on LMIC
20	Ireland, Paul & Dujardin (2011)	Not an empirical study
21	Jeong (2012)	Not an economic evaluation
22	Johnson & Higgins (2014)	Not an empirical study
23	Karash (2013)	Not an empirical study
24	Lee, Cheng, Chen & Lai (2010)	Not focused on LMIC
25	Lorincz, Lawson & Long (2013)	Not focused on LMIC
26	Maynard (2011)	Not focused on LMIC
27	McMahon & Chopra (2012)	Not focused on LMIC
28	Moore & DeBuono (2013)	Not focused on LMIC
29	Peabody et al. (2010)	Main focus not on PBF intervention
30	Rajkumar, Conway & Tavenner (2014)	Not focused on LMIC
31	Ran, Luo, Wu, Yao & Feng (2013)	Not an economic evaluation
32	Robeznieks (2012)	Not an economic evaluation
33	Rosenau, Lal & Lako (2012)	Not focused on LMIC
34	Ryan (2013)	Not focused on LMIC

35	Saronga et al. (2014)	Main focus not on PBF intervention
36	Swensen, Dilling, Mc Carty, Bolton & Harper (2013)	Not focused on LMIC
37	Tan, Pwu, Chen & Yang (2014)	Not focused on LMIC
38	Tummers, Schrijvers & Visser-Meily (2013)	Not focused on LMIC
39	VanLare, Blum & Conway (2012)	Not focused on LMIC
40	Wilson (2013)	Not focused on LMIC
41	Wranik (2012)	Not focused on LMIC
42	Yip et al. (2014)	Not an economic evaluation
43	Zeng, Rwiyereka, Amico, Avila-Figueroa & Shepard (2014)	Not an economic evaluation

* Articles that were excluded often did not meet more than one inclusion criteria. However, only one reason for exclusion is mentioned in the table above.

References for excluded studies

Agee, M. D., & Gates, Z. (2012). Lessons from Game Theory about Healthcare System Price Inflation. *Applied Health Economics and Health Policy*, *11*(1), 45–51. http://doi.org/10.1007/s40258-012-0003-z

Allen, H., Nobel, J. J., & Burton, W. N. (2012). Making health care reform work: where physician and employer interests converge. *JAMA*, *308*(23), 2465–2466. http://doi.org/10.1001/jama.2012.45428

Awoonor-Williams, J. K., Bawah, A. A., Nyonator, F. K., Asuru, R., Oduro, A., Ofosu, A., & Phillips, J. F. (2013). The Ghana essential health interventions program: a plausibility trial of the impact of health systems strengthening on maternal & child survival. *BMC Health Services Research*, *13 Suppl 2*, S3.

Baral, G. (2012). An assessment of the safe delivery incentive program at a tertiary level hospital in Nepal. *Journal of Nepal Health Research Council*, *10*(21), 118–124.

Bernstein, J. (2014). Not by bread alone: shortcomings of the pay-for-performance approach. *Clinical Orthopaedics and Related Research*, *472*(2), 405–409. http://doi.org/10.1007/s11999-013-3417-5

Blecker, E. (2014). The challenges in achieving successful P4P programs. *Findings Brief: Health Care Financing & Organization*, *42*(2), 1–3.

Blumenthal, D. M., Song, Z., Jena, A. B., & Ferris, T. G. (2013). Guidance for structuring team-based incentives in healthcare. *The American Journal of Managed Care*, *19*(2), e64–70.

Broughton, E., Saley, Z., Boucar, M., Alagane, D., Hill, K., Marafa, A., … Sani, K. (2013). Cost-effectiveness of a quality improvement collaborative for obstetric and newborn care in Niger. *International Journal of Health Care Quality Assurance*, *26*(3), 250–261. http://doi.org/10.1108/09526861311311436

Chee, G. (2003). Using financing to motivate a for-profit health care provider to deliver family planning services: Is it a cost-effective intervention? A study of AAR health services in Kenya. *The International Journal of Health Planning and Management*, pp. 205 – 220.

Cheng, S.-H., Lee, T.-T., & Chen, C.-C. (2012). A longitudinal examination of a pay-for-performance program for diabetes care: evidence from a natural experiment. *Medical Care*, *50*(2), 109–116. http://doi.org/10.1097/MLR.0b013e31822d5d36

Eichler, R., Agarwal, K., Askew, I., Iriarte, E., Morgan, L., & Watson, J. (2013). Performance-based incentives to improve health status of mothers and newborns: what does the evidence show? *Journal of Health, Population, and Nutrition*, *31*(4 Suppl 2), 36–47.

Falisse, J.-B., Meessen, B., Ndayishimiye, J., & Brossuyt, M. (2012). Community participation and voice mechanismes under performance-based financing schemes in Burundi. *Tropical Medicine and International Health*, pp. 674–682.

Gerber-Grote, A., & Windeler, J. (2014). What is the contribution of health economic evaluations to decision-making in health care? Experiences from 7 selected countries. *Zeitschrift Für Evidenz, Fortbildung Und Qualität Im Gesundheitswesen*, *108*(7), 358–359. http://doi.org/10.1016/j.zefq.2014.08.018

Ginsburg, P. B. (2013). Achieving Health Care Cost Containment Through Provider Payment Reform That Engages Patients And Providers. *Health Affairs*, *32*(5), 929–934. http://doi.org/10.1377/hlthaff.2012.1007

Greene, J., Hibbard, J. H., & Overton, V. (2014). A Case Study of a Team-Based, Quality-Focused Compensation Model for Primary Care Providers. *Medical Care Research and Review*, *71*(3), 207–223. http://doi.org/10.1177/1077558713506749

Higgs, E. S., Stammer, E., Roth, R., & Balster, R. L. (2013). Evidence acquisition and evaluation for evidence summit on enhancing provision and use of maternal health services through financial incentives. *Journal of Health, Population, and Nutrition*, *31*(4 Suppl 2), 23–35.

Himmelstein, D. U., Ariely, D., & Woolhandler, S. (2014). Pay-for-performance: toxic to quality? Insights from behavioral economics. *International Journal of Health Services: Planning, Administration, Evaluation*, *44*(2), 203–214.

Holcombe, R. F. (2014). Cancer clinical research: return on investment in the era of value-based purchasing. *Journal of Oncology Practice / American Society of Clinical Oncology*, *10*(5), 327–328. http://doi.org/10.1200/JOP.2014.001416

Hupp, J. R. (2014). The value of improving value. *Journal of Oral and Maxillofacial Surgery: Official Journal of the American Association of Oral and Maxillofacial Surgeons*, *72*(5), 843–845. http://doi.org/10.1016/j.joms.2014.03.001

Ireland, M., Paul, E., & Dujardin, B. (2011). Can performance-based financing be used to reform health systems in developing countries? *Bull World Health Organ*, pp. 695–698.

Jeong, H.-S. (2012). Designing an Effective Pay-for-performance System in the Korean National Health Insurance, Designing an Effective Pay-for-performance System in the Korean National Health Insurance. *Journal of Preventive Medicine and Public Health, Journal of Preventive Medicine and Public Health*, *45*(3), 127–136. http://doi.org/2012.45.3.127

Johnson, J., & Higgins, A. (2014). Evaluating the fair market value of pay for performance. *Healthcare Financial Management: Journal of the Healthcare Financial Management Association*, *68*(4), 80–84.

Karash, J. A. (2013). Investing in value-based health care. *Hospitals & Health Networks / AHA*, *87*(5), 54–58.

Lee, T.-T., Cheng, S.-H., Chen, C.-C., & Lai, M.-S. (2010). A pay-for-performance program for diabetes care in Taiwan: a preliminary assessment. *The American Journal of Managed Care*, *16*(1), 65–69.

Lorincz, I. S., Lawson, B. C. T., & Long, J. A. (2013). Provider and patient directed financial incentives to improve care and outcomes for patients with diabetes. *Current Diabetes Reports*, *13*(2), 188–195. http://doi.org/10.1007/s11892-012-0353-9

Maynard, A. D. (2011). The powers and pitfalls of payment for performance - Maynard - 2011 - Health Economics - Wiley Online Library. *Health Economics*, *21*(1). Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/hec.1810/abstract

McMahon, L. F., & Chopra, V. (2012). Health care cost and value: the way forward. *JAMA*, *307*(7), 671–672. http://doi.org/10.1001/jama.2012.136

Moore, L. G., & DeBuono, B. (2013). Total Cost of Care: A Discipline That Leads to Better Care. *Journal of Ambulatory Care Management*, *36*(3), 193–198. http://doi.org/10.1097/JAC.0b013e3182955b4b

Peabody, J. W., Florentino, J., Shimkhada, R., Solon, O., & Quimbo, S. (2010). Quality variation and its impact on costs and satisfaction: evidence from the QIDS study. *Medical Care*, *48*(1), 25–30. http://doi.org/10.1097/MLR.0b013e3181bd47b2

Rajkumar, R., Conway, P. H., & Tavenner, M. (2014). CMS--engaging multiple payers in payment reform. *JAMA*, *311*(19), 1967–1968. http://doi.org/10.1001/jama.2014.3703

Ran, L., Luo, K., Wu, Y., Yao, L., & Feng, Y. (2013). An analysis of China’s physician salary payment system. *Journal of Huazhong University of Science and Technology. Medical Sciences = Hua Zhong Ke Ji Da Xue Xue Bao. Yi Xue Ying De Wen Ban = Huazhong Keji Daxue Xuebao. Yixue Yingdewen Ban*, *33*(2), 309–314. http://doi.org/10.1007/s11596-013-1116-9

Robeznieks, A. (2012). Looking at value. NQF backs effort focusing on resource use, costs. *Modern Healthcare*, *42*(6), 10.

Rosenau, P. V., Lal, L. S., & Lako, C. (2012). Managing pay for performance: aligning social science research with budget predictability. *Journal of Healthcare Management / American College of Healthcare Executives*, *57*(6), 391–404; discussion 404–405.

Ryan, A. M. (2013). Will value-based purchasing increase disparities in care? *The New England Journal of Medicine*, *369*(26), 2472–2474. http://doi.org/10.1056/NEJMp1312654

Saronga, H. P., Duysburgh, E., Massawe, S., Dalaba, M. A., Savadogo, G., Tonchev, P., … Loukanova, S. (2014). Efficiency of antenatal care and childbirth services in selected primary health care facilities in rural Tanzania: a cross-sectional study. *BMC Health Services Research*, *14*, 96. http://doi.org/10.1186/1472-6963-14-96

Swensen, S. J., Dilling, J. A., Mc Carty, P. M., Bolton, J. W., & Harper, C. M. (2013). The business case for health-care quality improvement. *Journal of Patient Safety*, *9*(1), 44–52. http://doi.org/10.1097/PTS.0b013e3182753e33

Tan, E. C.-H., Pwu, R.-F., Chen, D.-R., & Yang, M.-C. (2014). Is a diabetes pay-for-performance program cost-effective under the National Health Insurance in Taiwan? *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *23*(2), 687–696. http://doi.org/10.1007/s11136-013-0502-x

Tummers, J. F. M. M., Schrijvers, A. J. P., & Visser-Meily, J. M. A. (2013). A qualitative study of stakeholder views on the effects of provider payment on cooperation, quality of care and cost-containment in integrated stroke care. *BMC Health Services Research*, *13*, 127. http://doi.org/10.1186/1472-6963-13-127

VanLare, J. M., Blum, J. D., & Conway, P. H. (2012). Linking performance with payment: implementing the Physician Value-Based Payment Modifier. *JAMA*, *308*(20), 2089–2090. http://doi.org/10.1001/jama.2012.14834

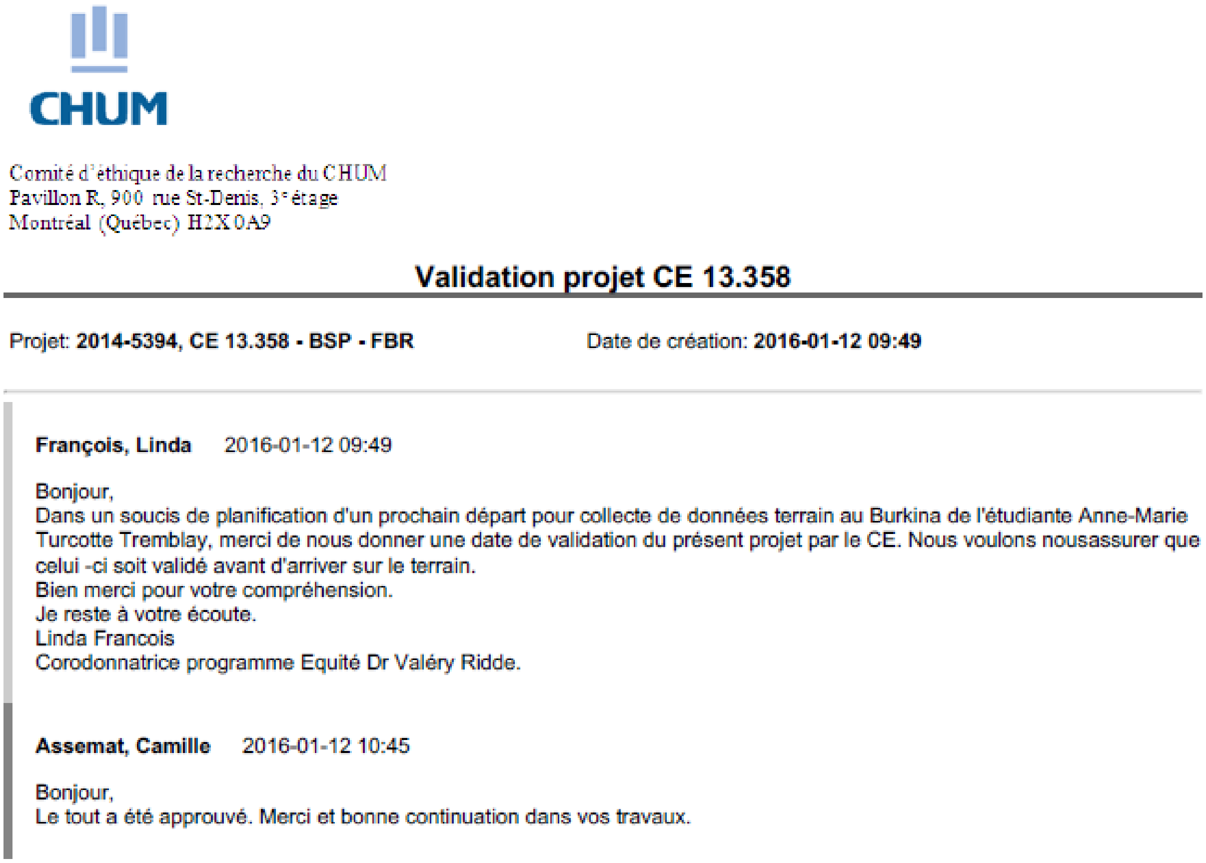
Wilson, K. J. (2013). Pay-for-performance in health care: what can we learn from international experience? *Quality Management in Health Care*, *22*(1), 2–15. http://doi.org/10.1097/QMH.0b013e31827dea50

Wranik, D. (2012). Healthcare policy tools as determinants of health-system efficiency: evidence from the OECD. *Health Economics, Policy, and Law*, *7*(2), 197–226. http://doi.org/10.1017/S1744133111000211

Yip, W., Powell-Jackson, T., Chen, W., Hu, M., Fe, E., Hu, M., … Hsiao, W. C. (2014). Capitation combined with pay-for-performance improves antibiotic prescribing practices in rural China. *Health Affairs (Project Hope)*, *33*(3), 502–510. http://doi.org/10.1377/hlthaff.2013.0702

Zeng, W., Rwiyereka, A. K., Amico, P. R., Avila-Figueroa, C., & Shepard, D. S. (2014). Efficiency of HIV/AIDS health centers and effect of community-based health insurance and performance-based financing on HIV/AIDS service delivery in Rwanda. *The American Journal of Tropical Medicine and Hygiene*, *90*(4), 740–746. http://doi.org/10.4269/ajtmh.12-0697

Appendix B. Approval from the University of Montreal Hospital Research Center



CHUM

Comité d'éthique de la recherche du CHUM
Pavillon R, 900 rue St-Denis, 3ᵉ étage
Montréal (Québec) H2X 0A9

Validation projet CE 13.358

Projet: **2014-5394, CE 13.358 - BSP - FBR** Date de création: **2016-01-12 09:49**

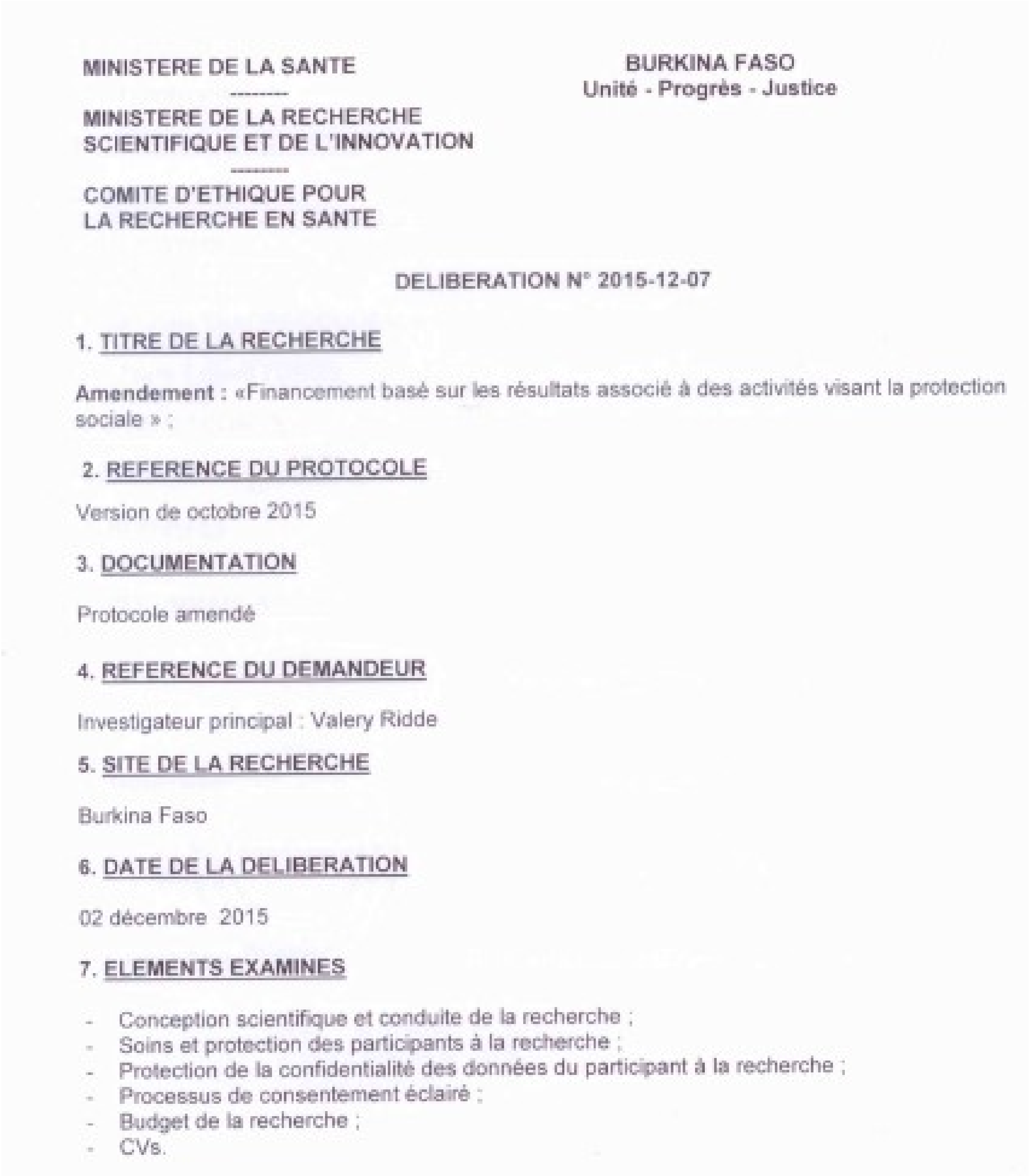
François, Linda 2016-01-12 09:49

Bonjour,
Dans un soucis de planification d'un prochain départ pour collecte de données terrain au Burkina de l'étudiante Anne-Marie Turcotte Tremblay, merci de nous donner une date de validation du présent projet par le CE. Nous voulons nousassurer que celui -ci soit validé avant d'arriver sur le terrain.
Bien merci pour votre compréhension.
Je reste à votre écoute.
Linda Francois
Corodonnatrice programme Equité Dr Valéry Ridde.

Assemat, Camille 2016-01-12 10:45

Bonjour,
Le tout a été approuvé. Merci et bonne continuation dans vos travaux.

Appendix C. Approval from the Ethics Committee in Burkina Faso



MINISTERE DE LA SANTE

MINISTERE DE LA RECHERCHE
SCIENTIFIQUE ET DE L'INNOVATION

COMITE D'ETHIQUE POUR
LA RECHERCHE EN SANTE

BURKINA FASO
Unité - Progrès - Justice

DELIBERATION N° 2015-12-07

1. **TITRE DE LA RECHERCHE**

Amendement : «Financement basé sur les résultats associé à des activités visant la protection sociale » ;

2. **REFERENCE DU PROTOCOLE**

Version de octobre 2015

3. **DOCUMENTATION**

Protocole amendé

4. **REFERENCE DU DEMANDEUR**

Investigateur principal : Valery Ridde

5. **SITE DE LA RECHERCHE**

Burkina Faso

6. **DATE DE LA DELIBERATION**

02 décembre 2015

7. **ELEMENTS EXAMINES**

- Conception scientifique et conduite de la recherche ;
- Soins et protection des participants à la recherche ;
- Protection de la confidentialité des données du participant à la recherche ;
- Processus de consentement éclairé ;
- Budget de la recherche ;
- CVs.

8. OBSERVATIONS

Pas d'observation

9. MEMBRES AYANT SIEGE

Dr Séni KOUANDA ;
Dr Djénéba SANON/OUEDRAOGO ;
Dr Ida SAWADOGO;
Dr Maminata TRAORE;
Mr Mathias SAM;
Dr Maminata TRAORE/COULIBALY;
Dr Lincoln Théophile TAPSOBA. ;
Dr R Frank Edgard ZONGO.

10. AVIS DU COMITE

Favorable pour la poursuite de l'étude

11. RESERVES

12. RECOMMANDATIONS

Ouagadougou, le 02 décembre 2015

Le Rapporteur

Rapporteur
CERS

[redacted]

Dr R. Frank Edgar ZONGO

Le Président P/I

Président

[redacted]

Dr Germaine L. MINOUNGOU-COMPAORE

Appendix D. Information and consent form

Formulaire d'information et de consentement

Titre: Financement basé sur les résultats et protection sociale au Burkina Faso

Objectifs du formulaire de consentement

Nous sollicitons votre participation à un projet de recherche affilié à l'Université de Montréal au Canada et à l'Association Action Gouvernance Intégration Renforcement (AGIR) au Burkina Faso. Avant d’accepter de participer, veuillez prendre le temps de comprendre et de considérer les renseignements qui suivent. Nous vous invitons à poser toutes les questions que vous jugerez utiles.

Nature et déroulement du projet de recherche

Pour améliorer les soins de santé au Burkina Faso, les autorités ont mis en œuvre une nouvelle intervention qui associe le financement basé sur les résultats à des interventions de protection sociale. L'objectif de cette recherche est d'étudier la manière dont l'intervention a été mise en oeuvre et les divers changements que cela a pu engendrer. Cette collecte de données s'inscrit dans le cadre d'un programme de recherche sur le processus d'implantation du financement basé sur les résultats.

Pour les participants, ce volet de l'étude peut impliquer des entrevues individuelles ou des entretiens de groupes. Durant environ une heure, nous allons vous poser des questions sur la mise en oeuvre et les changements dus à l'intervention. Nous allons également effectuer de l'observation dans les milieux reliés à l'intervention. De l'information sera collectée sur les soignants, l'infrastructure des formations sanitaires, la disponibilité des médicaments, la gestion des établissements de santé, la rémunération du personnel, les recettes des formations sanitaires et l'exemption de paiement des soins des indigents. Le comité d’éthique de la recherche du Burkina Faso et du Centre de recherche de l'Université de Montréal ont approuvé ce projet de recherche. Ce projet de recherche est financé par le Centre de recherches pour le développement international (CRDI) du Canada et l'Institut de recherche en santé du Canada (IRSC).

Risques et bénéfices associés à la recherche

Votre participation à cette recherche ne devrait vous causer aucun préjudice. Vous pourriez ressentir de l'inconfort lors de l'entrevue tel que de la gêne, de l’anxiété, de la fatigue, du stress et de la frustration suite au temps consacré à la recherche. Vous ne retirerez aucun bénéfice personnel de votre participation à cette recherche. Les participants ne recevront pas de compensation.

Il n’y a aucun avantage ou bénéfice direct pour le participant. Les résultats contribueront à l’avancement des connaissances. Les résultats de l’étude serviront à entreprendre des actions en vue d’améliorer la qualité et l’utilisation des services de santé maternelle et infantile au Burkina Faso.

Confidentialité

Les participants ne seront pas identifiables lors de la diffusion des résultats. Les renseignements recueillis demeureront strictement confidentiels dans les limites prévues par la loi. Afin de préserver la confidentialité, vous ne serez identifié que par un code de participant. La clé du code reliant votre nom à votre dossier de recherche sera conservée par les membres de l'équipe de recherche. Les données seront utilisées dans le but de répondre aux objectifs de la recherche seulement. Les données (enregistrements audio, transcription, etc.) seront conservées au Canada. Les données seront détruites après une période de 7 ans en supprimant tous les fichiers.

Dans le cas d'entretien de groupe, la confidentialité des échanges dépend de l'engagement réciproque des participants à ne pas divulguer l'identité des autres participants et la nature des échanges aux personnes n'ayant pas participé à la rencontre.

Participation volontaire

Vous êtes libre d'accepter ou de refuser de participer à ce projet de recherche. Vous pouvez vous retirer de cette étude à n'importe quel moment, sans avoir à donner de raison et sans conséquence pour vous. Vous n'avez qu'à en informer une personne-ressource de l'équipe de recherche, et ce, par simple avis verbal. En cas de retrait, vous pouvez demander la destruction des données ou du matériel vous concernant. Cependant, il sera impossible de retirer vos données ou votre matériel des analyses menées une fois ces dernières publiées ou diffusées.

Personnes-ressources

Si vous avez des questions ou si vous voulez vous retirer de l'étude, n'hésitez pas à communiquer avec les membres de l'équipe de recherche:

Coordination au Burkina Faso

Paul-André Somé
Association Action Gouvernance Intégration Renforcement (AGIR)
14 PB 254 Ouagadougou 14, Burkina Faso

Chercheur candidat au doctorat
Anne-Marie Turcotte-Tremblay
Institut de recherche en santé publique de l'Université de Montréal (IRSPUM)
7101 avenue du Parc, bureau 3060
Montréal, Québec, Canada H3N 1X9

Chercheur principal
Valéry Ridde
Chercheur et professeur agrégé
Institut de recherche en santé publique de l'Université de Montréal
7101 avenue du Parc
Montréal, Québec, Canada H3N 1X9

Pour toute préoccupation sur vos droits ou sur les responsabilités des chercheurs concernant ce projet, vous pouvez contacter le conseiller en éthique du Comité d’éthique de la recherche en santé par courriel à l’adresse ceres@umontreal.ca ou par téléphone au 00-1-514-343-6111, poste 2604 ou consulter le site http://recherche.umontreal.ca/participants. Toute plainte concernant cette recherche peut être adressée à l’ombudsman de l’Université de Montréal, au numéro de téléphone 00-1-514-343-2100 ou à l’adresse courriel ombudsman@umontreal.ca. L’ombudsman accepte les appels à frais virés.

Consentement du participant

Je comprends que je peux prendre mon temps pour réfléchir avant de donner mon accord ou non à participer à la recherche.

Je peux poser des questions à l’équipe de recherche et exiger des réponses satisfaisantes.

Je comprends qu’en participant à ce projet de recherche, je ne renonce à aucun de mes droits ni ne dégage les chercheurs de leurs responsabilités.

J’ai pris connaissance du présent formulaire d’information et de consentement et j’accepte de participer au projet de recherche.

____________________________	____________________________
Prénom et nom du participant	Signature du participant
	____________________________
	Date :

Engagement du chercheur

J’ai expliqué les conditions de participation à la recherche au participant. J’ai répondu au meilleur de ma connaissance aux questions posées et je me suis assuré de la compréhension du participant. Je m’engage, avec l’équipe de recherche, à respecter ce qui a été convenu au présent formulaire d’information et de consentement.

____________________________	____________________________
Prénom et nom du chercheur	Signature du chercheur
	____________________________
	Date :

Appendix E. List of reviewer's comments for Article 2 – 2nd round of revisions

Related to the abstract

1. Improve the following:
    - Methods: what kinds of interviews were done? (were they in depth interviews, key informant interviews?)
    - When were the interviews done?
    - In which districts in Burkina Faso did you conduct the study?
    - How many participants were interviewed?
    - Results: Are they qualitative results? Then indicate whether majority of the respondents provided the results or all respondents said what was described.

Related to the background section

2. There are also related studies in Ghana that the authors can use.

Related to the framework

3. What is the source of the conceptual framework (Figure 1)?

Related to the methods

4. Provide names of the districts and briefly explain why the study was done in those districts. Features of the districts should be described e.g number of health facilities, number of health staff, why was PBF done in those districts etc.
5. What did the pre-pilot PBF test entail?
6. You didn't tell us much about why the intervention was modified and expanded to 12 districts. How was it modified? Why was it expanded to 12 districts? This should be brief.
7. Why was a multiple case study done?
8. What kind of providers were involved? What kind of patients?
9. Indicate how many different types participants were interviewed in the text.

10. Were they key informant interview guides or in-depth interview guides?

11. What kind of quantitative data was it? What indicators were collated?

12. How did you process the recorded interviews?

13. Provide the full meaning of QDA miner.

14. In the text, interpret the other indicators presented in the table describing patients (e.g., age group)

15. In your analysis, you should also talk about using quotes from study participants

Related to the results

16. Explain the consequences before showing the table 3.
17. Reorganise the table 3
18. Regarding filling out registers retrospectively, was it all of them? A few of them or majority
19. The findings are extensive. Is it possible to highlight the key consequences?
20. The findings would be more interesting if authors could let us know whether participants talking about the consequences were a majority, few or all participants said what was described. This could then be followed by a few quotes.

Related to the discussion

21. Reorganise the discussion a bit. What were the key results? What is the significance of the results? How does it fit in with what else is known about the topic?/interpret results in light of what is already known about the subject

22. Was it a mixed method study?

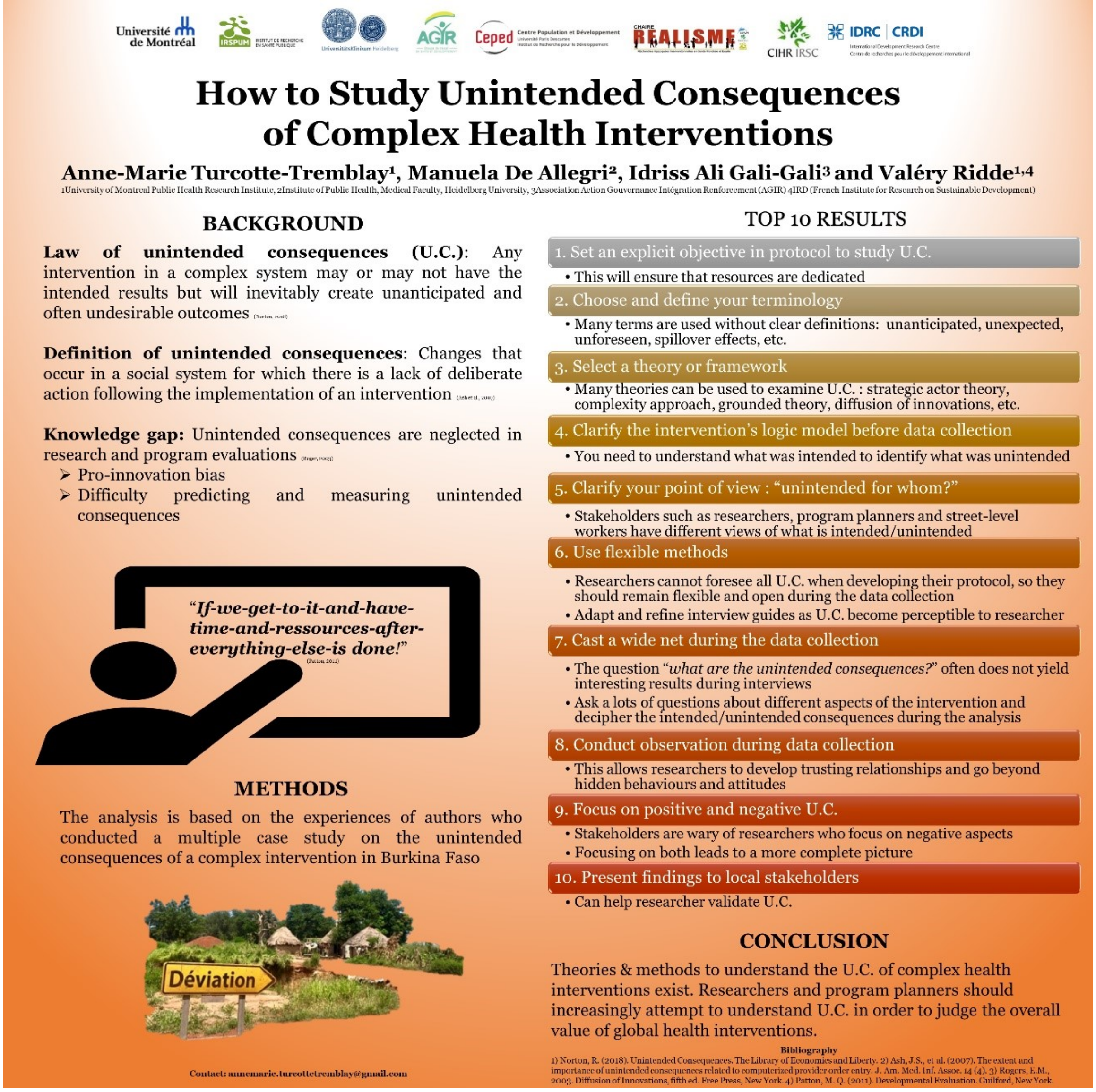
Related to the conclusion

23. In the conclusion, what kind of complex interventions?
24. The index tool should not be part of the final paper.

Related to the appendices

25. In appendix 4, what kind of interviews?

Appendix F. Examples of posters presented in scientific conferences



How to Study Unintended Consequences of Complex Health Interventions

Anne-Marie Turcotte-Tremblay[1], Manuela De Allegri[2], Idriss Ali Gali-Gali[3] and Valéry Ridde[1,4]

[1]University of Montreal Public Health Research Institute, [2]Institute of Public Health, Medical Faculty, Heidelberg University, [3]Association Action Gouvernance Intégration Renforcement (AGIR) [4]IRD (French Institute for Research on Sustainable Development)

BACKGROUND

Law of unintended consequences (U.C.): Any intervention in a complex system may or may not have the intended results but will inevitably create unanticipated and often undesirable outcomes (Norton, 2018)

Definition of unintended consequences: Changes that occur in a social system for which there is a lack of deliberate action following the implementation of an intervention (Ash et al., 2007)

Knowledge gap: Unintended consequences are neglected in research and program evaluations (Roger, 2003)

- Pro-innovation bias
- Difficulty predicting and measuring unintended consequences

> ***"If-we-get-to-it-and-have-time-and-ressources-after-everything-else-is done!"***
> (Patton, 2011)

METHODS

The analysis is based on the experiences of authors who conducted a multiple case study on the unintended consequences of a complex intervention in Burkina Faso

Contact: annemarie.turcottetremblay@gmail.com

TOP 10 RESULTS

1. Set an explicit objective in protocol to study U.C.
   - This will ensure that resources are dedicated
2. Choose and define your terminology
   - Many terms are used without clear definitions: unanticipated, unexpected, unforeseen, spillover effects, etc.
3. Select a theory or framework
   - Many theories can be used to examine U.C. : strategic actor theory, complexity approach, grounded theory, diffusion of innovations, etc.
4. Clarify the intervention's logic model before data collection
   - You need to understand what was intended to identify what was unintended
5. Clarify your point of view : "unintended for whom?"
   - Stakeholders such as researchers, program planners and street-level workers have different views of what is intended/unintended
6. Use flexible methods
   - Researchers cannot foresee all U.C. when developing their protocol, so they should remain flexible and open during the data collection
   - Adapt and refine interview guides as U.C. become perceptible to researcher
7. Cast a wide net during the data collection
   - The question "*what are the unintended consequences?*" often does not yield interesting results during interviews
   - Ask a lots of questions about different aspects of the intervention and decipher the intended/unintended consequences during the analysis
8. Conduct observation during data collection
   - This allows researchers to develop trusting relationships and go beyond hidden behaviours and attitudes
9. Focus on positive and negative U.C.
   - Stakeholders are wary of researchers who focus on negative aspects
   - Focusing on both leads to a more complete picture
10. Present findings to local stakeholders
    - Can help researcher validate U.C.

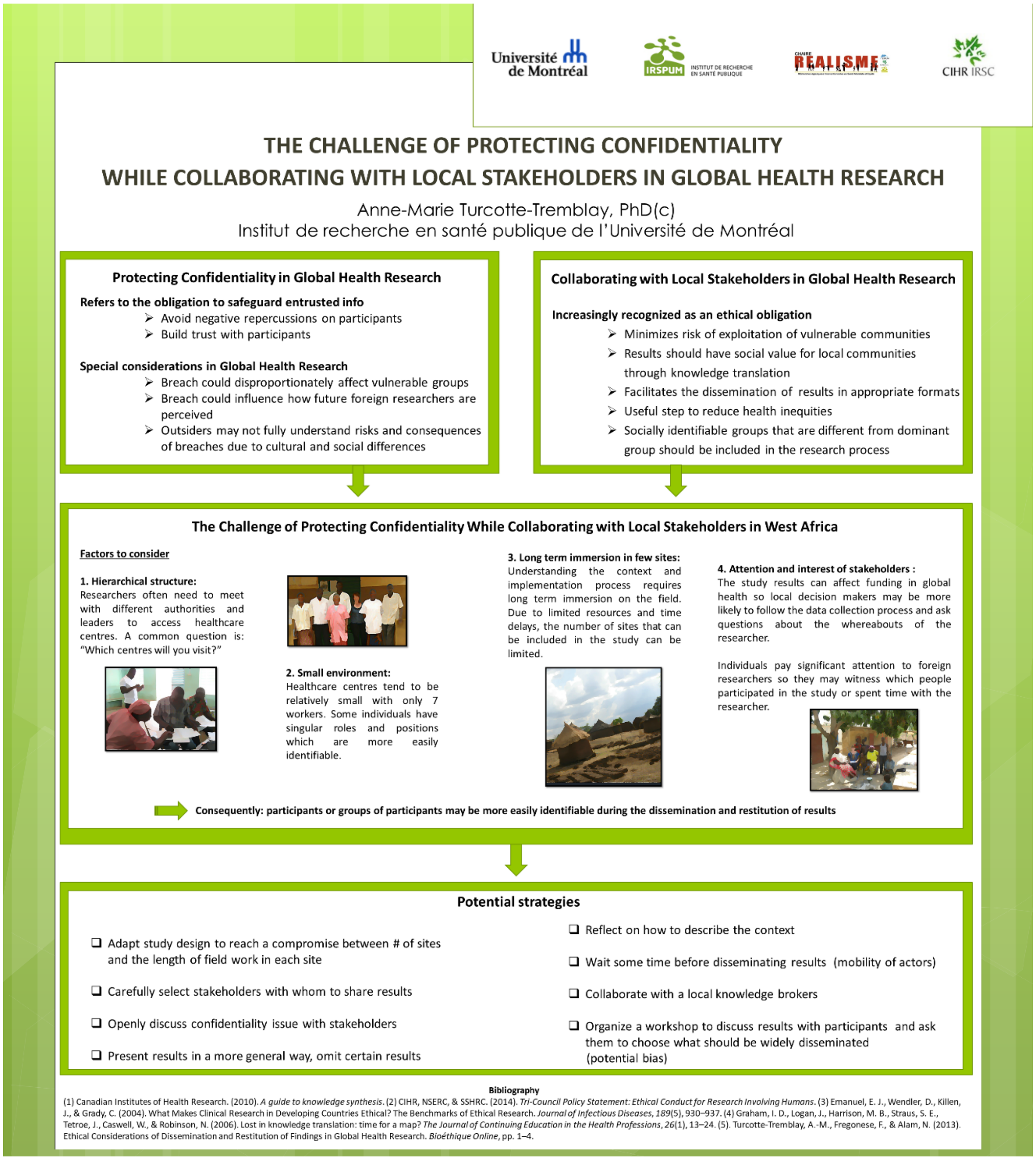
CONCLUSION

Theories & methods to understand the U.C. of complex health interventions exist. Researchers and program planners should increasingly attempt to understand U.C. in order to judge the overall value of global health interventions.

Bibliography

1) Norton, R. (2018). Unintended Consequences. The Library of Economics and Liberty. 2) Ash, J.S., et al. (2007). The extent and importance of unintended consequences related to computerized provider order entry. J. Am. Med. Inf. Assoc. 14 (4). 3) Rogers, E.M., 2003. Diffusion of Innovations, fifth ed. Free Press, New York. 4) Patton, M. Q. (2011). Developmental Evaluation. Guilford, New York.

Reference: Turcotte-Tremblay, A.-M., De Allegri, M., Ridde, V., (2018, October 8–12). *How to study unintended consequences of complex health interventions* [Poster presentation]. 5th Global Symposium on Health Systems Research. Health Systems Global, Liverpool, UK.



Université de Montréal IRSPUM INSTITUT DE RECHERCHE EN SANTÉ PUBLIQUE CHAIRE REALISME CIHR IRSC

THE CHALLENGE OF PROTECTING CONFIDENTIALITY WHILE COLLABORATING WITH LOCAL STAKEHOLDERS IN GLOBAL HEALTH RESEARCH

Anne-Marie Turcotte-Tremblay, PhD(c)
Institut de recherche en santé publique de l'Université de Montréal

Protecting Confidentiality in Global Health Research

Refers to the obligation to safeguard entrusted info

- Avoid negative repercussions on participants
- Build trust with participants

Special considerations in Global Health Research

- Breach could disproportionately affect vulnerable groups
- Breach could influence how future foreign researchers are perceived
- Outsiders may not fully understand risks and consequences of breaches due to cultural and social differences

Collaborating with Local Stakeholders in Global Health Research

Increasingly recognized as an ethical obligation

- Minimizes risk of exploitation of vulnerable communities
- Results should have social value for local communities through knowledge translation
- Facilitates the dissemination of results in appropriate formats
- Useful step to reduce health inequities
- Socially identifiable groups that are different from dominant group should be included in the research process

The Challenge of Protecting Confidentiality While Collaborating with Local Stakeholders in West Africa

Factors to consider

1. Hierarchical structure:
Researchers often need to meet with different authorities and leaders to access healthcare centres. A common question is: "Which centres will you visit?"

2. Small environment:
Healthcare centres tend to be relatively small with only 7 workers. Some individuals have singular roles and positions which are more easily identifiable.

3. Long term immersion in few sites:
Understanding the context and implementation process requires long term immersion on the field. Due to limited resources and time delays, the number of sites that can be included in the study can be limited.

4. Attention and interest of stakeholders :
The study results can affect funding in global health so local decision makers may be more likely to follow the data collection process and ask questions about the whereabouts of the researcher.

Individuals pay significant attention to foreign researchers so they may witness which people participated in the study or spent time with the researcher.

Consequently: participants or groups of participants may be more easily identifiable during the dissemination and restitution of results

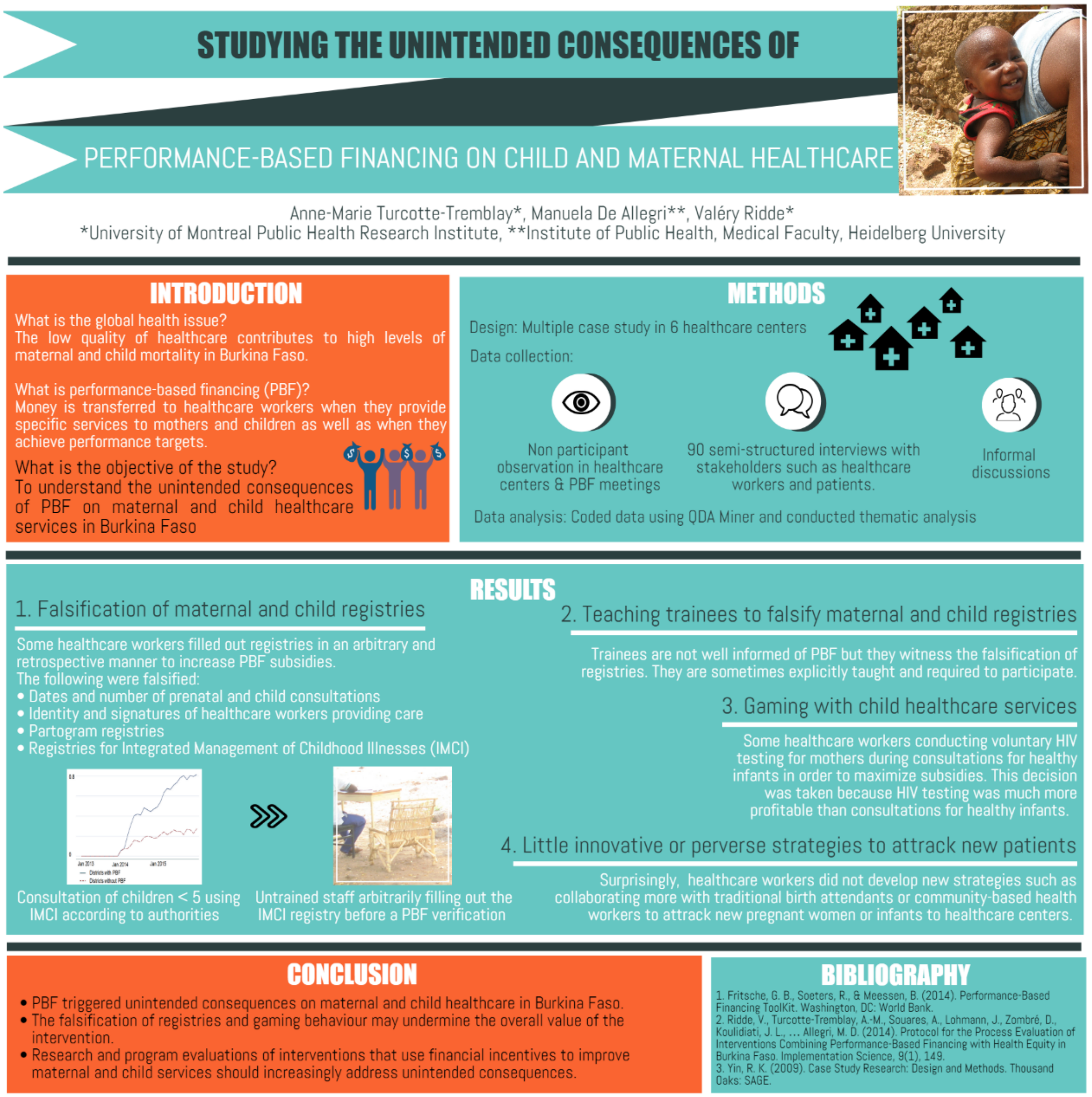
Potential strategies

- ☐ Adapt study design to reach a compromise between # of sites and the length of field work in each site
- ☐ Carefully select stakeholders with whom to share results
- ☐ Openly discuss confidentiality issue with stakeholders
- ☐ Present results in a more general way, omit certain results
- ☐ Reflect on how to describe the context
- ☐ Wait some time before disseminating results (mobility of actors)
- ☐ Collaborate with a local knowledge brokers
- ☐ Organize a workshop to discuss results with participants and ask them to choose what should be widely disseminated (potential bias)

Bibliography

(1) Canadian Institutes of Health Research. (2010). *A guide to knowledge synthesis*. (2) CIHR, NSERC, & SSHRC. (2014). *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans*. (3) Emanuel, E. J., Wendler, D., Killen, J., & Grady, C. (2004). What Makes Clinical Research in Developing Countries Ethical? The Benchmarks of Ethical Research. *Journal of Infectious Diseases*, *189*(5), 930–937. (4) Graham, I. D., Logan, J., Harrison, M. B., Straus, S. E., Tetroe, J., Caswell, W., & Robinson, N. (2006). Lost in knowledge translation: time for a map? *The Journal of Continuing Education in the Health Professions*, *26*(1), 13–24. (5). Turcotte-Tremblay, A.-M., Fregonese, F., & Alam, N. (2013). Ethical Considerations of Dissemination and Restitution of Findings in Global Health Research. *Bioéthique Online*, pp. 1–4.

Reference: Turcotte-Tremblay, A.-M. (2016, November 11). *The challenge of protecting confidentiality while collaborating with local stakeholders in global health research* [Poster presentation]. Conference on Ethical Challenges and Issues in Global Health Research. Québec Population Health Research Network, Montréal, Canada.



STUDYING THE UNINTENDED CONSEQUENCES OF PERFORMANCE-BASED FINANCING ON CHILD AND MATERNAL HEALTHCARE

Anne-Marie Turcotte-Tremblay*, Manuela De Allegri**, Valéry Ridde*
*University of Montreal Public Health Research Institute, **Institute of Public Health, Medical Faculty, Heidelberg University

INTRODUCTION

What is the global health issue?
The low quality of healthcare contributes to high levels of maternal and child mortality in Burkina Faso.

What is performance-based financing (PBF)?
Money is transferred to healthcare workers when they provide specific services to mothers and children as well as when they achieve performance targets.

What is the objective of the study?
To understand the unintended consequences of PBF on maternal and child healthcare services in Burkina Faso

METHODS

Design: Multiple case study in 6 healthcare centers

Data collection:

Non participant observation in healthcare centers & PBF meetings

90 semi-structured interviews with stakeholders such as healthcare workers and patients.

Informal discussions

Data analysis: Coded data using QDA Miner and conducted thematic analysis

RESULTS

1. Falsification of maternal and child registries

Some healthcare workers filled out registries in an arbitrary and retrospective manner to increase PBF subsidies.
The following were falsified:
- Dates and number of prenatal and child consultations
- Identity and signatures of healthcare workers providing care
- Partogram registries
- Registries for Integrated Management of Childhood Illnesses (IMCI)

Consultation of children < 5 using IMCI according to authorities

Untrained staff arbitrarily filling out the IMCI registry before a PBF verification

2. Teaching trainees to falsify maternal and child registries

Trainees are not well informed of PBF but they witness the falsification of registries. They are sometimes explicitly taught and required to participate.

3. Gaming with child healthcare services

Some healthcare workers conducting voluntary HIV testing for mothers during consultations for healthy infants in order to maximize subsidies. This decision was taken because HIV testing was much more profitable than consultations for healthy infants.

4. Little innovative or perverse strategies to attrack new patients

Surprisingly, healthcare workers did not develop new strategies such as collaborating more with traditional birth attendants or community-based health workers to attract new pregnant women or infants to healthcare centers.

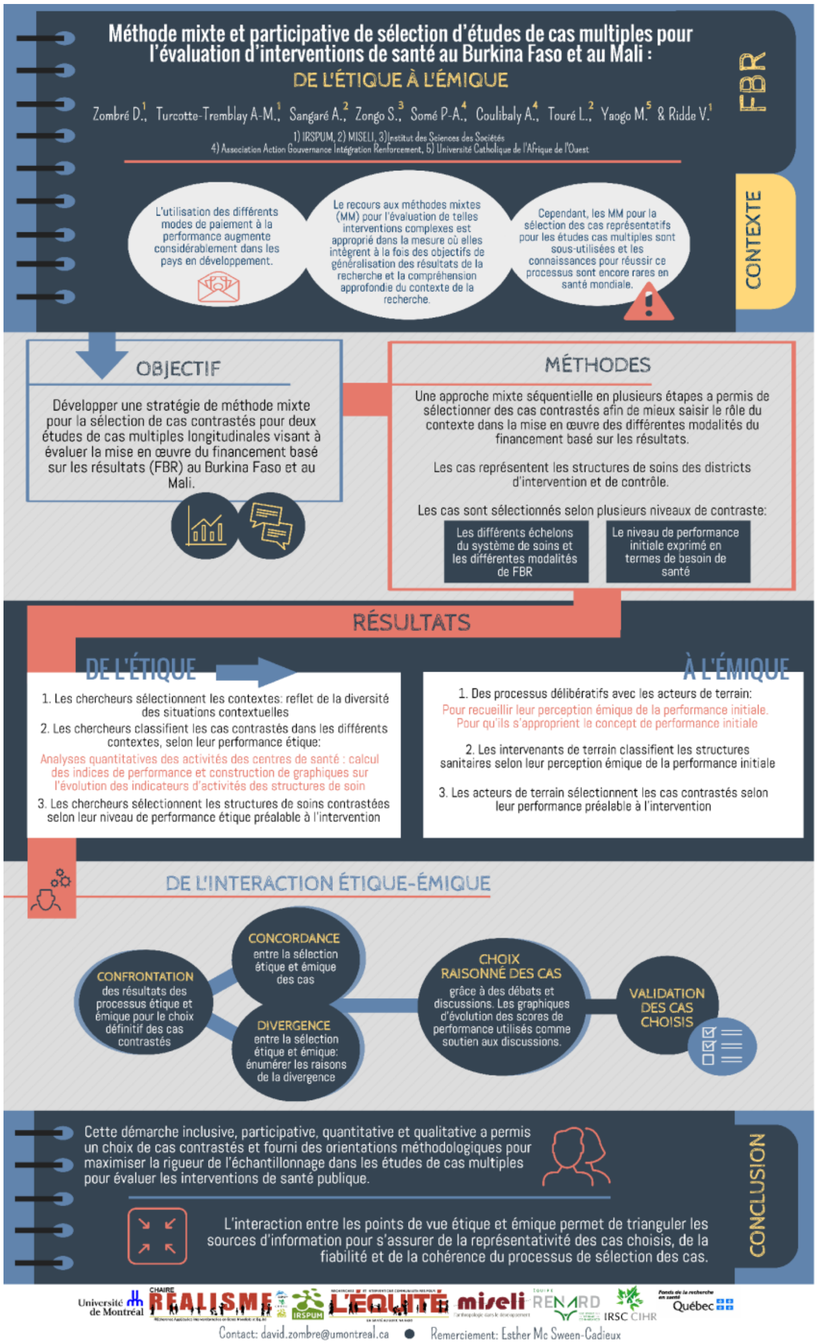
CONCLUSION

- PBF triggered unintended consequences on maternal and child healthcare in Burkina Faso.
- The falsification of registries and gaming behaviour may undermine the overall value of the intervention.
- Research and program evaluations of interventions that use financial incentives to improve maternal and child services should increasingly address unintended consequences.

BIBLIOGRAPHY

1. Fritsche, G. B., Soeters, R., & Meessen, B. (2014). Performance-Based Financing ToolKit. Washington, DC: World Bank.
2. Ridde, V., Turcotte-Tremblay, A.-M., Souares, A., Lohmann, J., Zombré, D., Koulidiati, J. L., … Allegri, M. D. (2014). Protocol for the Process Evaluation of Interventions Combining Performance-Based Financing with Health Equity in Burkina Faso. Implementation Science, 9(1), 149.
3. Yin, R. K. (2009). Case Study Research: Design and Methods. Thousand Oaks: SAGE.

Reference: Turcotte-Tremblay, A.-M., De Allegri, M. & Ridde, V. (2016, November 3–4). *Studying the unintended consequences of performance-based financing on maternal and child healthcare in Burkina Faso* [Poster presentation]. Partnerships, Policy and Progress: Healthy Lives for Women, Children and Adolescents Worldwide. Canadian Network for Maternal, Newborn and Child Health, Montreal, Canada.



Méthode mixte et participative de sélection d'études de cas multiples pour l'évaluation d'interventions de santé au Burkina Faso et au Mali :

DE L'ÉTIQUE À L'ÉMIQUE

Zombré D.[1], Turcotte-Tremblay A-M.[1], Sangaré A.[2] Zongo S.[3] Somé P-A.[4] Coulibaly A.[4] Touré L.[2] Yaogo M.[5] & Ridde V.[1]

1) IRSPUM, 2) MISELI, 3) Institut des Sciences des Sociétés
4) Association Action Gouvernance Intégration Renforcement, 5) Université Catholique de l'Afrique de l'Ouest

FBR

CONTEXTE

L'utilisation des différents modes de paiement à la performance augmente considérablement dans les pays en développement.

Le recours aux méthodes mixtes (MM) pour l'évaluation de telles interventions complexes est approprié dans la mesure où elles intègrent à la fois des objectifs de généralisation des résultats de la recherche et la compréhension approfondie du contexte de la recherche.

Cependant, les MM pour la sélection des cas représentatifs pour les études cas multiples sont sous-utilisées et les connaissances pour réussir ce processus sont encore rares en santé mondiale.

OBJECTIF

Développer une stratégie de méthode mixte pour la sélection de cas contrastés pour deux études de cas multiples longitudinales visant à évaluer la mise en œuvre du financement basé sur les résultats (FBR) au Burkina Faso et au Mali.

MÉTHODES

Une approche mixte séquentielle en plusieurs étapes a permis de sélectionner des cas contrastés afin de mieux saisir le rôle du contexte dans la mise en œuvre des différentes modalités du financement basé sur les résultats.

Les cas représentent les structures de soins des districts d'intervention et de contrôle.

Les cas sont sélectionnés selon plusieurs niveaux de contraste:

- Les différents échelons du système de soins et les différentes modalités de FBR
- Le niveau de performance initiale exprimé en termes de besoin de santé

RÉSULTATS

DE L'ÉTIQUE

1. Les chercheurs sélectionnent les contextes: reflet de la diversité des situations contextuelles
2. Les chercheurs classifient les cas contrastés dans les différents contextes, selon leur performance étique:
   Analyses quantitatives des activités des centres de santé : calcul des indices de performance et construction de graphiques sur l'évolution des indicateurs d'activités des structures de soin
3. Les chercheurs sélectionnent les structures de soins contrastées selon leur niveau de performance étique préalable à l'intervention

À L'ÉMIQUE

1. Des processus délibératifs avec les acteurs de terrain:
   Pour recueillir leur perception émique de la performance initiale.
   Pour qu'ils s'approprient le concept de performance initiale
2. Les intervenants de terrain classifient les structures sanitaires selon leur perception émique de la performance initiale
3. Les acteurs de terrain sélectionnent les cas contrastés selon leur performance préalable à l'intervention

DE L'INTERACTION ÉTIQUE-ÉMIQUE

CONFRONTATION des résultats des processus étique et émique pour le choix définitif des cas contrastés

CONCORDANCE entre la sélection étique et émique des cas

DIVERGENCE entre la sélection étique et émique: énumérer les raisons de la divergence

CHOIX RAISONNÉ DES CAS grâce à des débats et discussions. Les graphiques d'évolution des scores de performance utilisés comme soutien aux discussions.

VALIDATION DES CAS CHOISIS

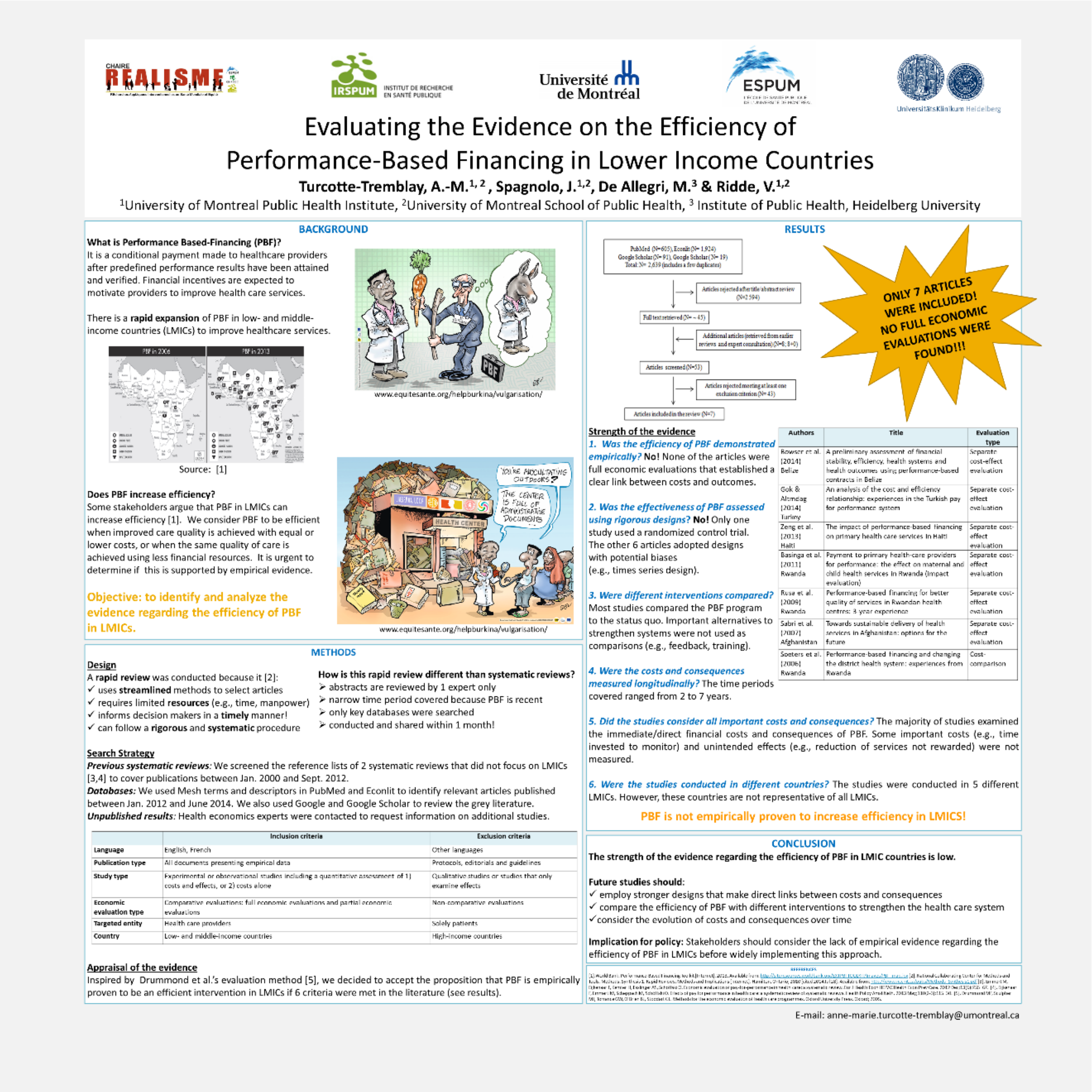
CONCLUSION

Cette démarche inclusive, participative, quantitative et qualitative a permis un choix de cas contrastés et fourni des orientations méthodologiques pour maximiser la rigueur de l'échantillonnage dans les études de cas multiples pour évaluer les interventions de santé publique.

L'interaction entre les points de vue étique et émique permet de trianguler les sources d'information pour s'assurer de la représentativité des cas choisis, de la fiabilité et de la cohérence du processus de sélection des cas.

Contact: david.zombre@umontreal.ca ● Remerciement: Esther Mc Sween-Cadieux

Reference: Zombré, D. Turcotte-Tremblay, A.-M., Zongo, S., Somé, P.-A., Sangaré, A., Coulibaly, A., Touré, L., Yaogo, M., & Ridde, V. (2016, May 9–13). Méthode mixte et participative de sélection d'études de cas multiples pour l'évaluation d'interventions de santé au Burkina Faso et au Mali : de l'étique à l'émique [Poster presentation]. 84th Conference of the Association francophone pour le savoir (Acfas), Montréal, Canada.



Evaluating the Evidence on the Efficiency of Performance-Based Financing in Lower Income Countries

Turcotte-Tremblay, A.-M.[1, 2], Spagnolo, J.[1,2], De Allegri, M.[3] & Ridde, V.[1,2]

[1]University of Montreal Public Health Institute, [2]University of Montreal School of Public Health, [3] Institute of Public Health, Heidelberg University

BACKGROUND

What is Performance Based-Financing (PBF)?
It is a conditional payment made to healthcare providers after predefined performance results have been attained and verified. Financial incentives are expected to motivate providers to improve health care services.

There is a **rapid expansion** of PBF in low- and middle-income countries (LMICs) to improve healthcare services.

Source: [1]

www.equitesante.org/helpburkina/vulgarisation/

Does PBF increase efficiency?
Some stakeholders argue that PBF in LMICs can increase efficiency [1]. We consider PBF to be efficient when improved care quality is achieved with equal or lower costs, or when the same quality of care is achieved using less financial resources. It is urgent to determine if this is supported by empirical evidence.

Objective: to identify and analyze the evidence regarding the efficiency of PBF in LMICs.

www.equitesante.org/helpburkina/vulgarisation/

METHODS

Design
A **rapid review** was conducted because it [2]:
- ✓ uses **streamlined** methods to select articles
- ✓ requires limited **resources** (e.g., time, manpower)
- ✓ informs decision makers in a **timely** manner!
- ✓ can follow a **rigorous** and **systematic** procedure

How is this rapid review different than systematic reviews?
- ➢ abstracts are reviewed by 1 expert only
- ➢ narrow time period covered because PBF is recent
- ➢ only key databases were searched
- ➢ conducted and shared within 1 month!

Search Strategy
Previous systematic reviews: We screened the reference lists of 2 systematic reviews that did not focus on LMICs [3,4] to cover publications between Jan. 2000 and Sept. 2012.
Databases: We used Mesh terms and descriptors in PubMed and Econlit to identify relevant articles published between Jan. 2012 and June 2014. We also used Google and Google Scholar to review the grey literature.
Unpublished results: Health economics experts were contacted to request information on additional studies.

	Inclusion criteria	Exclusion criteria
Language	English, French	Other languages
Publication type	All documents presenting empirical data	Protocols, editorials and guidelines
Study type	Experimental or observational studies including a quantitative assessment of 1) costs and effects, or 2) costs alone	Qualitative studies or studies that only examine effects
Economic evaluation type	Comparative evaluations: full economic evaluations and partial economic evaluations	Non-comparative evaluations
Targeted entity	Health care providers	Solely patients
Country	Low- and middle-income countries	High-income countries

Appraisal of the evidence
Inspired by Drummond et al.'s evaluation method [5], we decided to accept the proposition that PBF is empirically proven to be an efficient intervention in LMICs if 6 criteria were met in the literature (see results).

RESULTS

PubMed (N=605), Econlit (N= 1,924)
Google Scholar (N= 91), Google Scholar (N= 19)
Total: N= 2,639 (includes a few duplicates)

→ Articles rejected after title/abstract review (N=2 594)

Full text retrieved (N= ~ 45)

← Additional articles (retrieved from earlier reviews and expert consultation) (N=8; 8+0)

Articles screened (N=53)

→ Articles rejected meeting at least one exclusion criterion (N= 43)

Articles included in the review (N=7)

ONLY 7 ARTICLES WERE INCLUDED! NO FULL ECONOMIC EVALUATIONS WERE FOUND!!!

Strength of the evidence

1. Was the efficiency of PBF demonstrated empirically? **No!** None of the articles were full economic evaluations that established a clear link between costs and outcomes.

2. Was the effectiveness of PBF assessed using rigorous designs? **No!** Only one study used a randomized control trial. The other 6 articles adopted designs with potential biases (e.g., times series design).

3. Were different interventions compared? Most studies compared the PBF program to the status quo. Important alternatives to strengthen systems were not used as comparisons (e.g., feedback, training).

4. Were the costs and consequences measured longitudinally? The time periods covered ranged from 2 to 7 years.

Authors	Title	Evaluation type
Bowser et al. [2014] Belize	A preliminary assessment of financial stability, efficiency, health systems and health outcomes using performance-based contracts in Belize	Separate cost-effect evaluation
Gok & Altındag [2014] Turkey	An analysis of the cost and efficiency relationship: experiences in the Turkish pay for performance system	Separate cost-effect evaluation
Zeng et al. [2013] Haiti	The impact of performance-based financing on primary health care services in Haiti	Separate cost-effect evaluation
Basinga et al. [2011] Rwanda	Payment to primary health-care providers for performance: the effect on maternal and child health services in Rwanda (impact evaluation)	Separate cost-effect evaluation
Rusa et al. [2009] Rwanda	Performance-based financing for better quality of services in Rwandan health centres: 3-year experience	Separate cost-effect evaluation
Sabri et al. [2007] Afghanistan	Towards sustainable delivery of health services in Afghanistan: options for the future	Separate cost-effect evaluation
Soeters et al. [2006] Rwanda	Performance-based financing and changing the district health system: experiences from Rwanda	Cost-comparison

5. Did the studies consider all important costs and consequences? The majority of studies examined the immediate/direct financial costs and consequences of PBF. Some important costs (e.g., time invested to monitor) and unintended effects (e.g., reduction of services not rewarded) were not measured.

6. Were the studies conducted in different countries? The studies were conducted in 5 different LMICs. However, these countries are not representative of all LMICs.

PBF is not empirically proven to increase efficiency in LMICS!

CONCLUSION

The strength of the evidence regarding the efficiency of PBF in LMIC countries is low.

Future studies should:
- ✓ employ stronger designs that make direct links between costs and consequences
- ✓ compare the efficiency of PBF with different interventions to strengthen the health care system
- ✓ consider the evolution of costs and consequences over time

Implication for policy: Stakeholders should consider the lack of empirical evidence regarding the efficiency of PBF in LMICs before widely implementing this approach.

REFERENCES

E-mail: anne-marie.turcotte-tremblay@umontreal.ca

Reference: Turcotte-Tremblay, A.-M., Spagnolo, J., De Allegri, M., Ridde, V. (2015, May 24–27). *Evaluating the evidence on the efficiency of performance-based financing in lower income countries* [Poster presentation]. 36th Evaluation Conference of the Canadian Society for Evaluation, Montreal, Canada.