CrossMark

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, 7101 Avenue du Parc, Room 3060, Montreal, QC H3N 1X9, Canada
[b] School of Public Health, University of Montreal, 7101 Avenue du Parc, Montreal, QC H3N 1X9, 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, Im Neuenheimer Feld 130, 69120 Heidelberg, Germany

ARTICLE INFO

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.

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

* Corresponding author. University of Montreal Public Health Research Institute, 7101 Avenue du Parc, Room 3060, Montreal, QC H3N 1X9, Canada.
E-mail address: anne-marie.turcotte-tremblay@umontreal.ca (A.-M. Turcotte-Tremblay).

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 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. Fig. 1 illustrates our theoretical framework.
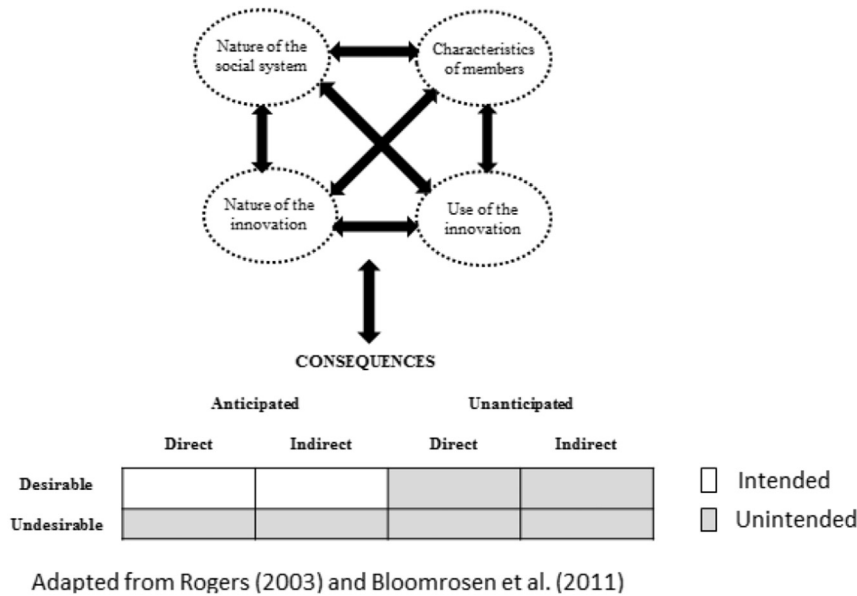
Fig. 1. Theoretical framework.

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 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 *Comités de gestion* (COGES — healthcare facility management committees). 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 statuses, 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 about 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 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 1 provides a breakdown of the data collected for each case and across cases.

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 2, 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.

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

Table 1
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

Table 2
Classification of unintended consequences.

| | 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

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 the local association. 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 (2500 CFA francs) per patient found, but the local association deducted overhead fees ($0.81—1.62 USD/patient; 500—1000 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 family 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! (…) [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 non-participant 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 3 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 for each patient found.

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)

Table 3
Characteristics of the social system and its members that made it difficult to find patients and obtain valid data.

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)

"*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; 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) indexing 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.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.socscimed.2017.09.007.

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.1365-3156.2009.02430.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 santé. 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), 207–214. 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/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.