



ELSEVIER

Contents lists available at ScienceDirect

SSM - Population Health

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

Article

“I wanted a skeleton ... they brought a prince”: A qualitative investigation of factors mediating the implementation of a Performance Based Incentive program in Malawi

Shannon A. McMahon^{a,*}, Adamson S. Muula^b, Manuela De Allegri^a^a Institute of Public Health at Heidelberg University, Im Neuenheimer Feld 324, 69120 Heidelberg, Germany^b College of Medicine at University of Malawi, Malawi

ARTICLE INFO

Keywords:

Malawi
 Performance-based incentives
 Performance-based financing
 Process evaluation
 Qualitative research
 Implementation research

ABSTRACT

While several evaluations have examined the extent to which performance based financing (PBF) programs induce changes in the quantity and quality of health services provided, less is known about the process of implementing PBF. We conducted a process evaluation of a PBF intervention in Malawi that focused on understanding moderators of program implementation. Informed by a seminal theory of implementation, we first created a timeline and taxonomy of key events in the program lifeline and then undertook 25 in-depth interviews with stakeholders including implementers, central-level ministry officials and district-level health staff. While seven “moderator categories” emerged in this study, two categories (program complexity and quality of delivery) proved especially crucial in terms of moderating implementation and sparking adaptations. Complexity refers primarily to the manner in which PBF requires that those implementing the program have business acumen and forecasting skills, which are often beyond the purview of a clinician’s training and thus proved challenging. Regarding quality of delivery, the program struggled to issue rewards in a timely and adequate manner, which proved highly problematic as it undermined a bedrock feature of PBF. Adaptations and adaptability refers here to a program’s ability to make changes; the program proved rigid in several respects although nimble in terms of adjusting the verification process (upon noticing revengeful behaviors in peer verification). This PBF program is unique in several respects and findings cannot be generalized to all PBF programs. Nevertheless, process evaluations that draw from or expand upon existing implementation theories can allow researchers to better disentangle complex programming. We hope that more process evaluations, which track both core elements and necessary adaptations of PBF implementation, can further advance understandings of why PBF implementation functions or fails within a given setting, thereby enhancing implementers’ abilities to replicate facilitators and bypass barriers.

1. Introduction

Performance-based financing (PBF) refers to a range of interventions within a health system that link financial and/or material rewards to the attainment of predefined quantity and/or quality outputs, verified on a regular basis (Eichler et al., 2013; Fritsche, Soeters, & Meessen, 2014; Musgrove, 2011). While straightforward, this definition belies the complexity of implementing PBF (Renmans, Holvoet, Orach, & Criel, 2016). Like other interventions aimed at producing system-wide changes, PBF necessitates an overhaul of longstanding norms; it involves rearranging how health facilities are financed and supplied, and how key actors such as providers, patients and regulators view their roles within the wider health system. As an illustrative example, the

distribution of supplies to facilities has historically been based on pre-determined metrics; facilities receive a quota of materials based on the size of a facility’s catchment area, and facility-based staff engage minimally in determining when, whether and which new equipment is received. In contrast to this, under PBF, program implementers and health administrators outline a package of materials and equipment deemed necessary to improve health outcomes or the working environment. Following this, implementers outline a series of health-related outputs (such as births in a facility or vaccines administered) and agree that once a goal or target is reached, supplies, equipment and/or salary bonuses be disbursed. This shift is referred to as moving from input to output-based financing within the PBF literature. An intended goal in undertaking such a shift is that those working within a health

* Corresponding author.

E-mail address: mcmahon@uni-heidelberg.de (S.A. McMahon).

facility feel in control of not only their existing environment, but also their ability to monitor progress toward a target and to enact behavioral changes (Renmans et al., 2016).

Despite the inherent complexity of such an intervention, the PBF community has devoted relatively limited research energy to implementation and process-related research. With notable exceptions (Renmans et al., 2016; Gergen et al., 2016; Bertone, Lagarde, & Witter, 2016; Bhatnagar and George, 2016; Wilhelm, Brenner, Muula, & De Allegri, 2016; Turcotte-Tremblay, Gali-Gali, De Allegri, & Ridde, 2017; Ssengooba, McPake, & Palmer, 2012), most studies and at least one systematic review (Witter, Fretheim, Kessy, & Lindahl, 2012) related to PBF have emphasized outcome-related research; in this respect PBF has proved promising but not without setbacks. PBF has been observed to produce changes in utilization (sparking increases in institutional deliveries (Bonfrer, Van de Poel, & Van Doorslaer, 2014), antenatal visits (Bonfrer et al., 2014), and preventative care visits (Basinga et al., 2011) for example), but not always in quality. It has been described as a “success story” in Rwanda (Shroff, Bigdeli, & Meessen, 2017), but has shown mixed results in other settings such as Tanzania, Benin and Cameroon (Paul, Albert, & Bisala, 2018). In considering program performance, researchers, donors and relevant stakeholders have begun seeking to unpack the “black box” (Renmans et al., 2016) of implementation. Did a given country’s PBF program fail or succeed on its own merits and due to its own attributes, or was the program never implemented as intended? The latter scenario represents a Type III error¹ - conclusions are formulated about a program without first ascertaining if implementation reflected program design (Hasson, 2010). Beyond discussions about implementation fidelity, as years of PBF programming unfold, it has become clear that a PBF program cannot be seamlessly transplanted from one setting to another and expected to achieve the same results (Shroff et al., 2017). Analyzing how a program is implemented, how factors moderate the implementation process and examining which elements are inserted or removed from a program to be responsive to on-the-ground realities is referred to as conducting a process evaluation (Renmans et al., 2016; Ssengooba et al., 2012; Hasson, 2010; Waweru, Goodman, Kedenge, Tsofa, & Molyneux, 2016; Magrath and Nichter, 2012; Witter et al., 2013).

Process evaluations, by definition, examine essential aspects of a program in an effort to define what a program is and how it has been delivered to its intended audience (Rossi, Lipsey, & Freeman, 2003). This definition is intertwined with the concept of implementation fidelity, which focuses on a program’s content, coverage, frequency and duration to assess whether “the active ingredients of the intervention have been received by the participants as often and for as long as was planned” (Hasson, 2010). Carroll et al. (2007) developed a seminal conceptual framework for fidelity, which emphasized that beyond examining adherence to core components - it is also necessary to assess moderators. Moderators are factors that “influence or moderate the degree of fidelity with which an intervention is implemented” (Carroll et al. 2007). Carroll (2007) emphasized the following moderators: participant responsiveness, the comprehensiveness of a policy (or program) description, inclusion of program facilitators (in terms of monitoring, feedback, and training), and the quality of program delivery. Hasson (2010) expanded upon this framework by adding two further moderators of implementation: recruitment procedures and context. Taken collectively, these moderators involve asking questions such as: How much did participants in a program feel engaged? Was the program well articulated with clear goals and recommendations? Did a program have adequate monitoring systems in place? Was the program delivered in a way that it could achieve what was intended? What procedures were used to attract participants? What social, historical or

political factors affected implementation? (Hasson, 2010; Carroll et al., 2007)

Carroll and then Hasson’s fidelity frameworks have informed studies addressing a range of health issues in high-and low-income settings, from care for the elderly and frail in Sweden (Hasson, Blomberg, & Dunér, 2012) to malaria control in Burkina Faso (Ridde, Druetz, Poppy, Kouanda, & Haddad, 2013). While we are aware of several studies that have looked at process elements of PBF programming in a holistic sense (Renmans et al., 2016) or within individual countries (Benin, (Antony, Bertone, & Barthes, 2017; Paul, Sossouhounto, & Eclou, 2014) Burkina Faso (Turcotte-Tremblay et al., 2017), Malawi, (Wilhelm et al., 2016) Sierra Leone, (Bertone and Meessen, 2013) Tanzania, (Chimhutu, Tjomsland, Songstad, Mrisho, & Moland, 2015)) we are not aware of published PBF-focused studies that have drawn in or expanded upon this framework to guide the collection, analysis and interpretation of data.

In this article, we present qualitative findings from a process evaluation of early experiences of a PBF program in Malawi, which draws upon and expands Carroll (Carroll et al., 2007) and then Hasson’s (Hasson, 2010) framework. We place a particular emphasis on program moderators, as we seek to understand the implementation process and to learn how and why a given program diverted from its initial blueprint. The findings are intended to inform ongoing discussions regarding future performance-based interventions in Malawi and similar settings. However, the incorporation and expansion of a widely-used fidelity framework may have applications to interventions that are similarly directed toward producing broad changes within the health system.

2. Methods

2.1. Study setting

Malawi is ranked 173 of 188 countries in terms of human development (UNDP, 2015), and has the world’s fourth lowest GDP per capita (World Bank, 2015). Infant and under-5 mortality rates are 42 and 63 per 1000 live births, respectively (NSONMa, 2017). For every 100,000 live births, an estimated 439 women die (NSONMa, 2017). HIV Prevalence is 10.6% (Measure, 2010), making Malawi among ten countries with the highest prevalence in the world. While the country met several of its Millennium Development Goals (MDG) targets including those related to child mortality (MDG 4) and HIV and AIDS (MDG 6), other targets were not met including those related to maternal mortality (MDG 5) (Countdown, 2015; UN, 2015).

Malawi’s health system uses an input based health financing approach; resources such as infrastructure, staff, drugs and equipment are distributed according to population, number of facilities and existing resources (WHO, 2015). While management of human resources is retained by the Ministry of Health, District Health Offices serve as a hub of service delivery, coordinating services such as routine operations, drug budgets and the management of health centers (Chriwa, 2013). Decentralization has been the guiding approach for nearly two decades, but progress toward decentralization has been “mixed” (Chriwa, 2013), and plagued by challenges such as “weak coordination of decentralization at the national level, underfunding of district implementation plans and high staff turnover within the health sector” (WHO, 2015). Calls have been made to better clarify roles and responsibilities at central versus district levels, and to better empower actors within zonal, district and community levels in terms of policy formation and health activity planning (WHO, 2015). Actors within Malawi’s peripheral facilities including health centers and dispensaries control relatively few resources.

Health services are provided by both public and private facilities (both for-profit and not-for-profit). Across public sector facilities (which represent the focus of this research), the availability of basic amenities varies: 63% have regular electricity, 91% have an improved water

¹ A type I error entails incorrectly rejecting the null hypothesis. A type II error entails failing to reject a false null hypothesis.

source and 22% have a client latrine (Ministry of Health Malawi II, 2014). Health facility staff have high client loads due to inadequate worker-to-patient ratios (0.02 physicians and 0.34 nursing and midwifery personnel per 1000 population) (Chimwaza, Chipeta, & Ngwira, 2014). Not surprisingly, health staff report feeling underpaid, inadequately supervised, subjected to poor working conditions and limited in terms of career and educational opportunities (Bradley and McAuliffe, 2009; McAuliffe, Manafa, Maseko, Bowie, & White, 2009). While national estimates suggest that hospitals typically have one doctor, one pharmacist and five technicians (Ministry of Health Malawi II, 2014), researchers highlight that this estimate masks two considerations: first, more than half of all doctors work in four of the country's central hospitals and a handful of district hospitals; second, doctors are routinely tasked with administrative duties that restrict time for clinical care (Chimwaza et al., 2014).

2.2. The SSDI-PBI program in Malawi

At least two performance based, health financing interventions have been recently introduced in the country including the “Results Based Financing for Maternal and Newborn Health (Brenner, Muula, & Robyn, 2014)” (RBF4MNH) and the “Support for Service Delivery Integration-Performance Based Incentives (McMahon, Brenner, & Lohmann, 2016)” (SSDI-PBI).

SSDI-PBI is the focus of this process evaluation. While global literature favors the term PBF, in this program the term PBI is more often employed (including in the name of the program). For purposes of this paper, we use the term PBF unless naming the SSDI-PBI program specifically. The SSDI-PBI program was designed and implemented by two SSDI sectors, SSDI-Systems (led by Abt Associates) and SSDI-Services (led by Jhpiego), respectively, and in partnership with the Ministry of Health.² Funded by USAID, the program was rolled out in 17 facilities across three districts (Chitipa, Nkhotakota and Mangochi) from 2014 through January 2017. Facilities that received the intervention were non-randomly selected based on criteria related to equipment, infrastructure and personnel. SSDI-PBI rewards were comprised of a combination of quantity and quality scores, with community scores serving as a source of potential bonus payments. Quantity, or utilization, indicators focused on increasing total service provision counts across the maternal health continuum of care, newborn and child health, and HIV and AIDS care and treatment (see Table 1). Quality indicators emphasized improvements in the broader facility environment and in the nature of how care was provided across 13 service areas (see Table 2). Strategies to attain targets (and to consider how rewards would be used) were outlined via business plans that were devised by staff within health facilities with support from program implementers. Rewards were paid to facilities upon achievement of set targets, but the rewards could only be used toward facility improvements and could not be partially redistributed in the form of performance bonuses to individual health workers; this marks a stark departure from common arrangements under other performance-based schemes (Fritsche et al., 2014). Program implementers worked with facility-based staff to ensure that rewards issued aligned with the ability to meet future quality and quantity indicators, with the intention that this would create a positive feedback loop.

Along with not distributing monetary incentives, SSDI-PBI did not allow facilities themselves to procure goods, materials or equipment; this role was retained by the program implementer, Jhpiego. This is also a deviation from most PBF programs.

Further details of the SSDI-PBI program are outlined in McMahon et al. (2016).

² In this paper, for purposes of clarity Abt and Jhpiego will be referred to as program “Designers” and program “Implementers,” respectively.

Table 1
Quantity indicators used in SSDI-PBI.

1.	Number of pregnant women starting antenatal care during the first trimester
2.	Number of women completing the four antenatal care visits
3.	Number of pregnant women receiving at least two doses of intermittent preventive therapy
4.	Number of births attended by skilled birth attendants (doctor, nurse or midwife)
5.	Number of 1-year-old children fully immunized
6.	Number of HIV-positive pregnant women who were initiated on antiretroviral therapy
7.	Number of HIV/AIDS cases screened for Tuberculosis
8.	Number of children receiving Vitamin A supplementation
9.	Number of clients counseled for family planning
10.	Number of couples tested for HIV during HIV testing and counseling services
11.	Number of infants born by HIV positive mothers tested for HIV
12.	Number of women who receive postnatal care after delivery by skilled health workers within seven days
13.	Number of pregnant women attending antenatal care receiving iron supplementation

Table 2
Quality dimensions assessed in SSDI-PBI.

1.	General activities
2.	Follow-up assessment and HMIS
3.	Hygiene, environment, and sterilization
4.	Outpatient and inpatient consultation
5.	Maternity ward
6.	Antenatal consultation
7.	Family planning
8.	Vaccination and monitoring of newborns
9.	HIV/AIDS control
10.	Tuberculosis
11.	Laboratory
12.	Minor surgery
13.	Drug and commodity management

2.3. Study methods

We conducted a process evaluation based on theories drawn from the literature and on discussions with stakeholders. Two main frameworks or theories informed this work: fidelity of implementation (Hasson, 2010; Carroll et al., 2007) and expectancy theory (Ssenooba et al., 2012; Lawler, 1988; Lawler and Suttle, 1973). As highlighted earlier, we sought to build on Carroll's - and then Hasson's - framework, which focuses primarily on moderating factors rather than a quantification of fidelity, but also worked on incorporating more pointedly participant expectations. In her work, Hasson hints at this dimension via the inclusion of “participant responsiveness” as a moderator. However, by participant responsiveness Hasson refers to a level of enthusiasm that participants feel toward an intervention with questions emphasizing whether participants felt engaged in and satisfied with an intervention. Expectation theory extends this notion of responsiveness by considering not just enthusiasm but an individual's internal perspectives of a program, and how these perspectives interact with external dynamics (such as the flow of rewards, and supportive supervision). Drawing on this theory inspires questions that we viewed as essential to understanding the program because they relate to how participants perceive: their internal capacity to engage in a performance-based intervention; trade-offs comparing extra work inherent to a program versus potential or actual gains; whether or how fairness in terms of reward issuance is respected; and the potential for an intervention to transform the working environment and the health system in the shorter and longer terms (Lawler, 1988; Lawler and Suttle, 1973).

Data collection included 25 qualitative interviews with a purposely selected sample of key informants involved in the implementation of SSDI-PBI and working at central, district and facility levels, and a

Table 3

Respondent table.

SSDI-PBI staff (incl. Abt, Jhpiego)	8
Ministry of Health	5
USAID	1
Health providers and professionals at district level (incl. District Medical Officers, District Nursing Officers and PBI coordinators)	8
Health provider at primary level	1
Community leaders	2
Total	25

review of all manuals and operational guidelines affiliated with the intervention. In the Fall of 2015, the research and implementing teams created a calendar of events outlining core components and key moments in the life of the program, looking at whether and/or when events that were supposed to happen actually happened. This outline of events was then cross-checked with program documents. The process of creating and crosschecking a program timeline informed the in-depth interviews (IDs), which were conducted for four weeks in the Spring of 2016, and focused on factors that affected program implementation.

The lead author conducted 19 interviews and trained two research assistants (each of whom possessed previous experience with qualitative research) to conduct the remaining six interviews. Interviews lasted 65 minutes on average with a range of 40 to 120 minutes. Most respondents (n=18) were male. See Table 3 for a breakdown of respondents. Interviews were recorded and transcribed by the research team. Analysis was thematic, applying deductive and inductive approaches; deductive in that questions used in interview tools guided coding, and inductive because as new themes or codes emerged these were added to the codebook (Braun and Clarke, 2006). Finally, while the approach and tools were based on constructs that underpin fidelity of implementation and theories of expectation, data was collected in an open-ended manner to allow for the discovery of new themes related to unintended and unexpected experiences within the program.

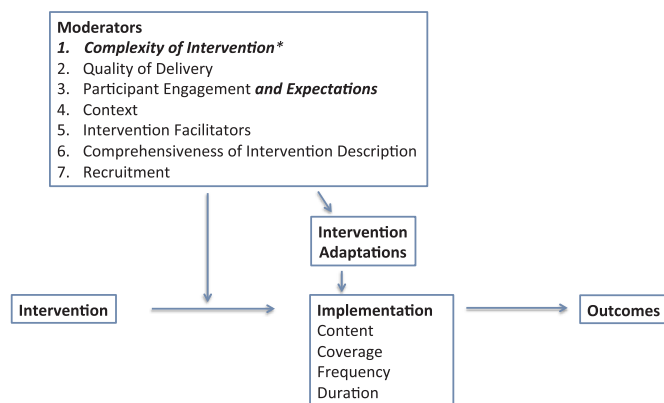


Fig. 2. An implementation framework adapted to a performance-based financing program.

*Carroll 2007 listed this as a component of Comprehensiveness. We see it as a discrete entity unto itself.

3. Findings

In this section, we first describe the program, outlining key events in the program lifeline. Following this, we present each of the moderators and the adaptations that affected the fidelity with which SSDI-PBI was implemented. To guide the discussion, we present a framework which draws upon and expands Carroll (Carroll et al., 2007) and then Hasson’s (Hasson, 2010) frameworks (see Fig. 2).

3.1. Execution and timing of key intervention components

Fig. 1 outlines key events in the program lifeline, distinguishing activities/events that were (a) unexpected and undesirable (from a program standpoint), (b) undertaken with delay, (c) inserted ad hoc due to a perceived need, (d) undertaken in a different format and (e)

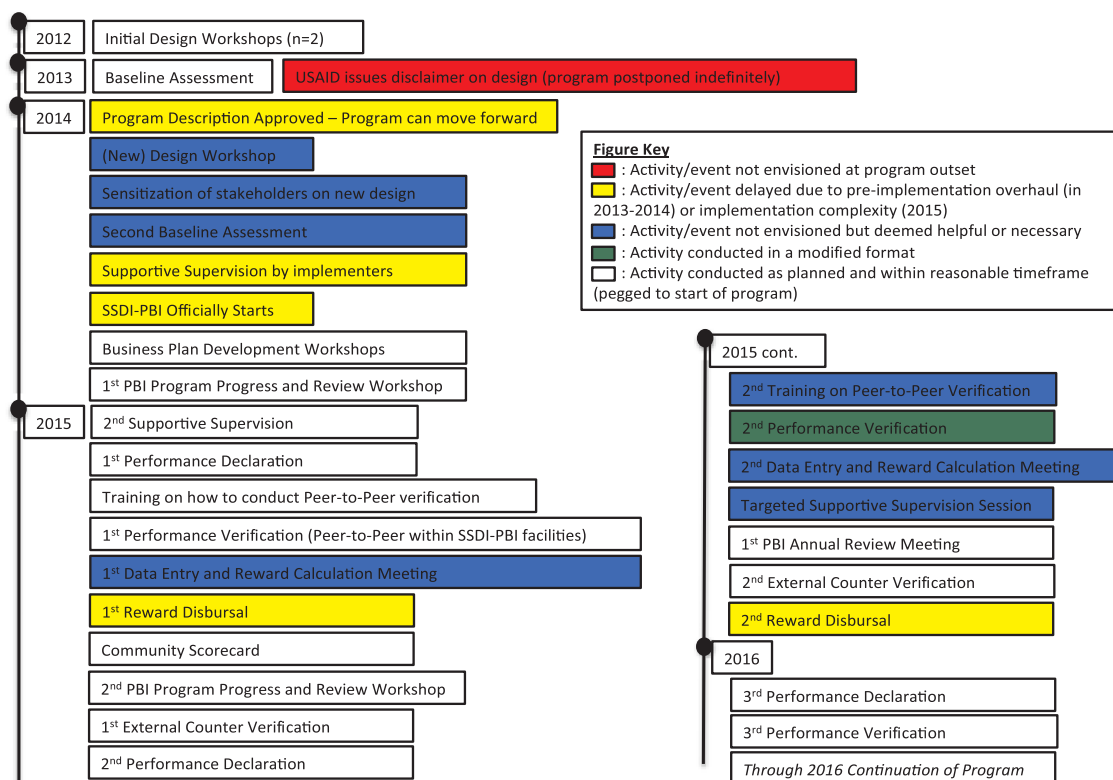


Fig. 1. Timeline for SSDI-PBI 2013–2016.

undertaken in a timely manner as pegged to program commencement. Produced in collaboration with key staff and crosschecked with program documents, the Figure's content served as a starting point for interviews on the program as a whole and how it was moderated. While not the focus of this paper, the timeline and content of the figure highlights that after a (pre-implementation) overhaul of the program design (wherein the provision of bonuses to individual providers was eliminated), SSDI was later executed in a manner that incorporated its most critical elements: program sensitization, business plan development, performance declarations, performance verifications, reward disbursements, and supportive supervision. Rather than examining the program's implementation fidelity per se, results herein focus on moderators – how factors affected the degree to which those engaged could implement the program - and adaptations that were made in light of moderators.

3.2. Program moderator - Complexity

Across respondents, SSDI-PBI was described as well-designed, well-articulated yet nevertheless exceptionally complex. Unlike many interventions with which respondents had experience, the program was described as transcending several dimensions of clinical or medical care. One implementer described how “other interventions get to sit together in one office” and focus on a given health domain (HIV or childhood immunizations or community-based health promotion), but SSDI-PBI extended across these (and other) categories. The program also placed demands on individuals devoted to financing, supply chain management, construction and procurement – all offices that do not typically engage closely in health interventions. District-level staff triangulated this point, stating that they had never encountered a program like SSDI-PBI. Nearly all respondents agreed that the concept of drafting business plans and being attuned to data monitoring was conceptually difficult for clinicians to grasp. One district medical officer (DMO) described the situation as, “Providers are more inclined to do patient care than to do data management. They had no concept ... of what it means to look at data ... and make plans based on data.”

The forced delay in the start of the program (from April 2013 to September 2014), and the change in the program's components, necessitated further rounds of program sensitization at several levels (from facilities up to the Ministry of Health) regarding the model of the new scheme and the rationale behind the elimination of individual bonuses. This process of altering an already complex design and then re-sensitizing stakeholders on the redesign, compressed the time within which the program could be executed, which in turn meant that program components and concepts had less time to “gel” among all involved.

Some timing delays were not the result of a delayed program start, but reflect under-estimation of the complex nature of the program and the amount of time necessary for facility staff and implementers to undertake a new task. For instance, the implementation team expected that supportive supervision activities (wherein each facility is visited by program implementers in order to ensure that the facility is functional in key respects) would last four hours. Each visit ultimately lasted approximately seven hours. Implementers expected to receive each facility's performance declaration forms by early March of 2015, but these trickled in about one week later. Such delays were often attributed to “realities of a new initiative”. At later periods, such activities were typically conducted on time and in a more fluid manner.

3.3. Program moderator – Quality of delivery

In terms of understanding whether SSDI-PBI program components were delivered in a manner that facilitated program success, the procurement of goods proved highly problematic. Upon meeting targets, facilities were meant to receive equipment, infrastructure or materials. The delayed delivery of such goods was mentioned in every interview.

Implementers highlighted that they underestimated the complexity of procuring goods:

“We're not procurement specialists. We didn't even have a procurement officer dedicated to PBI, not even an *officer*, we thought we would use the systems of Jhpiego. Nobody had seen how big this would turn out ... Now we have an engineer and procurement person *just* for PBI.... Have you ever heard of a health intervention hiring an engineer? An electrician?” - Implementer

Procuring goods was challenging for implementers for several reasons including the requirements that goods be procured via a tender (per standard protocols of the implementer), and that they meet safety and environmental standards of the donor (USAID). At the facility level, respondents described how goods (uniforms, generators, cloth wrappers, dustbins) would arrive late, be of substandard quality or not meet their specifications, which was frustrating in itself but more so problematic because facilities were penalized (or scored down) for goods that were ordered though not present – in other words, goods that had been included in their earliest business plan but did not arrive. Similarly, cash for outreach and meetings as stipulated in business plans were described as arriving late to facilities, preventing providers from proceeding with planned activities, which had been put in place in order to meet targets. The experience of being penalized because an ordered item or cash had not been delivered (and facilities could do nothing to expedite the process) was among the most negative facets of the SSDI-PBI program according to respondents. The situation was described as undermining facility and community motivation and autonomy, and breeding resentment and distrust. Implementers and Ministry officials were intimately aware of this problem, but consistently maintained that facilities could not be rewarded for goods that were not present during inspections. This rigidity further de-motivated district and facility-level staff.

3.4. Program moderator – Participant engagement, experiences and expectations

Implementers (including MoH officials) described how their expectations of the program's potential sank upon learning that monetary incentives (in the form of salary bonuses) would not be provided. As one implementer said, “I was at home ... thinking, ‘What am I going to do, am I not going to fail?’ I don't like failing.” Another implementer described the dread he felt when he needed to re-sensitize district-level staff on the new award arrangement. District-level staff (including facility-based SSDI-PBI coordinators) described feeling less concerned about the change. One DMO said the change improved his impression of the program as he had less concerns about staff leaving facilities and taking “all of the investment” with them when they rotated to another facility or left the health sector, which “won't happen if the investment is a motorcycle.” Another program coordinator said the equipment-focused iteration was more sustainable:

Imagine the program stops... tomorrow ... if people were receiving money (instead of equipment) we would have to say, ‘Keep doing what you're doing, but now we're not going to pay you anything.’ ... It just wouldn't work. – SSDI-PBI coordinator

Respondents described feeling hopeful but also overwhelmed by the program. At district and facility levels, all respondents described the components of SSDI-PBI as conceptually comprehensible yet nevertheless daunting. When successful, the process of devising business plans, gathering and assessing data about a series of health outputs, and procuring goods for structural improvements was deeply rewarding and invigorating. Yet oftentimes, facility-based staff felt that the ability to meet indicators related to quantity and quality targets was difficult (or impossible) to operationalize. Participants described feeling poised to fail in terms of meeting targets that they viewed as beyond their control, such as those that involved encouraging community members to

come to antenatal care (ANC) with their partners, or that relied on medical supplies that were in short supply (pregnancy tests, and ferrous sulfate with folic acid (FEFOL) tablets), or that required identifying HIV-positive populations. On this last point, one district medical officer (DMO) said,

We have been testing everyone (for HIV). I mean like *a lot* of people, but none ... are positive. They're all negative! ... You don't want to force people to have HIV. But it's also bad when you look through a day of testing and have found nobody. – DMO district 1

Respondents at district and facility levels said the situation of being evaluated on factors perceived as beyond one's control was frustrating and challenged sincere buy-in to enact the program. Implementers said it was too early to begin shifting targets, but this would be worth reviewing in the longer term.

Respondents across all tiers of the health system also mentioned that SSDI-PBI seemed prohibitively expensive. Several respondents said the Ministry would not be capable of continuing such a program without tremendous external support; this made it difficult to galvanize and sustain a long-term vision for the program.

3.5. Other moderators – Context, facilitators, comprehensiveness of program description, recruitment

While infeasible to detail the remaining four moderators in detail, we briefly highlight findings related to context, facilitators, comprehensiveness and recruitment below.

Contextual factors that affected the program include those inherent to Malawi's health system (high staff turnover, critical shortages in human resources for health, and a largely centralized health system), economic and political factors (sporadic (and sometimes acute) fuel shortages), currency inflation and the fact that SSDI-PBI was implemented immediately after the emergence of a political scandal (Malawi's "cashgate" scandal, 2014), and finally the existence of another health financing program, called RBF4MNH, which preceded the onset of SSDI-PBI and which (unlike SSDI-PBI) entailed cash bonuses to providers. RBF4MNH was ongoing and well known by those engaged with the health system, leading many respondents to sense a familiarity with the concept of PBF and to draw comparisons between SSDI-PBI and RBF4MNH.

In terms of facilitators, factors mentioned most often as underscoring success in terms of maintaining fidelity to the SSDI-PBI program cycle are largely rooted in the successful execution of essential PBF program ingredients. The program fostered changes in attitudes and behaviors across multiple levels and then benefitted from those changes in a positively reinforcing loop that bolstered fidelity. Respondents described engaging more intimately with catchment communities (who then provided added supports including sweat equity to meet targets), with other providers (who could galvanize support around common indicators and goals), and with other facilities (who, upon understanding that the program was not zero-sum, began sharing tips and tricks to meet targets). This facilitated the acquisition of rewards, which could be used to enhance service provision. Other facilitators included the management style of a program implementer whose presence and demeanor was lauded by those within districts and facilities, routine meetings between implementers and facilities, and an active line of communication in terms of discussing questions or concerns about the program. Perhaps the most powerful facilitator was the receipt of goods. While there were often procurement delays, the eventual arrival of goods astounded facility staff who proudly displayed their "PBI rewards" during the course of interviews (motorcycles, uniforms, computers, bed nets, lawnmowers, blood pressure machines and curtains) and detailed how the acquisition of products helped them attract more clients, and receive better quality scores.

In terms of comprehensiveness, several respondents from both central and district levels applauded the level of detail within the

program's operations manual. The manual was edited over an extended period (rooted largely in a donor-induced delay to program start) via consultative meetings across funders, implementers (including the MoH), and external consultants. The manual served as a reference among implementers during program implementation.

In terms of recruitment, the selection of facilities to participate was done non-randomly with decisions made collectively by implementers with DMOs. The implementation team sought to include only facilities with higher staffing capabilities and/or to re-assign staff to intervention facilities if they appeared to have human resource demands. In some cases, this sparked tensions with DMOs who felt that the implementer was overstepping their role in the health system; as one DMO said, "It is not their place to tell me how to do my job". District-and facility-level staff also described how some facilities appeared to have an added advantage at the program's outset by virtue of their pre-existing infrastructure and personnel; one facility-based manager said that compared to other facilities, it was harder for his staff to focus on output indicators and quality scores when their roof kept blowing away. An implementer described human resource competency challenges in tertiary facilities, "...you have medical assistants and nurse clinicians whose qualifications do not entail... competency in a situational analysis ... they cannot assess performance and (devise solutions)." For an overview of lessons learned across program modifiers see [Table 4](#).

3.6. Program adaptability and adaptations

In terms of adaptability, respondents mentioned that several adaptations that may have proved beneficial could nevertheless not be made (namely related to procurement). At the district and facility level, this fostered a sense that the program could not be adapted in a manner that would bolster autonomy and be more sensitive to on-the-ground priorities. This sentiment was perhaps best encapsulated by one DMO who said,

"... let me tell you, I wanted them to bring us a skeleton. A skeleton and then together we would put on some flesh. Build something together. But they came from Lilongwe and they brought a prince. He could not be touched, nothing could be changed or altered". – DMO district 2

Implementers described how their hands were often tied due to regulations and operating protocols of either the donor or of their organization(s). Facility-level respondents repeatedly argued that the nature of indicators and measurement of performance could be unrealistic or inappropriate when essential medical supplies (such as pregnancy tests) were out of stock, or when indicators were devised using "old, bad data". Several providers also described how quality indicators were too rigidly interpreted and could not be adjusted. A DMO described how the maternity ward was missing bed sheets on one bed during an inspection (the sheets were drying on a clothes line), but because each bed was not covered, the facility was penalized:

"It feels like... there is no flexibility... No understanding. If I'm at home, and I want to make a meal I need water, a pot, some fire and some food. These are the major things. This program is penalizing us because we don't have salt. The meal is there but that salt is missing. ... I don't need salt to eat a meal." – DMO district 1

Program implementers described several adaptations undertaken particularly at or around the program outset. Because of a condensed timeline outlined above, implementers undertook several sensitization tasks themselves rather than outsourcing the work. For example, rather than being able to conduct a training of trainers – wherein SSDI-PBI program staff would train trainers who would then provide on-the-job training to facility-based providers on the concept and progress of PBF – implementing staff collapsed this activity into district-wide trainings and led all trainings themselves.

Most adaptations in relation to the PBF cycle involved a change in

Table 4
Program moderators and fidelity.

Moderators	How a moderator influenced fidelity
1. Complexity of Intervention	Program implementers underestimated the amount of technical support necessary to get the program up and running. Concepts inherent to PBF (data monitoring and business plan development) proved challenging to master among many clinically inclined colleagues.
2. Quality of Delivery	Delays in the delivery of equipment, infrastructure and materials undermined the ability of facilities to meet targets and successfully progress through the program cycle.
3. Participant Engagement and Expectations	Changes in the program design (eliminating redistribution of incentives as bonuses to health workers) led several respondents to fear that provider motivation (and thus program implementation) would suffer. Ultimately, a delayed receipt of goods (see #2), and a rigid approach to targets were more substantive barriers to participant buy-in.
4. Context	Malawi's health system is characterized by high staff turnover and shortages in human resources, both of which challenged the introduction and implementation of the program. Training new employees on a complex program is especially difficult. Other factors such as fuel shortages and currency inflation also impeded implementation.
5. Intervention Facilitators	The (eventual) receipt of goods astounded and delighted health facility staff, who saw the goods as tangible proof that the program rewarded effort. This fostered buy-in and compelled providers to redouble their efforts to reach targets.
6. Comprehensiveness of Intervention Description	Respondents across levels said program manuals were detailed, clear and served as a reference throughout implementation.
7. Recruitment	The decision on which facilities would be chosen for the intervention sparked some conflict. Several respondents also highlighted that participating facilities had vastly different capabilities in terms of infrastructure and the nature and number of staff; lower-functioning facilities often found it difficult to absorb and undertake PBF.

the verification process (the process wherein reported data is checked for accuracy.) An important, early change involved the removal of verification by peers who were also part of the SSDI-PBI program. Despite being instructed otherwise, respondents said peer reviewers initially viewed the program as a zero-sum game; if they downgraded other facilities, more funds would be made available for their own facility's gain. This created cycles of punishment. In response, implementers cancelled peer verification and instead brought in verifiers who were also health providers, but not engaged in SSDI-PBI. Implementers also eventually increased the amount of resources (personnel and vehicles) to conduct quality verifications as they noticed that when the process was conducted in a step-wise fashion, some facilities would alert others to "quickly clean and tidy everything" as preparation for the visit.

In order to address program challenges implementers expanded their staff by hiring a procurement officer, an engineer and program assistants in each district (to assist in management of resources/finances at district level). Implementers also created a "Results Declaration and Reward Meeting" where representatives from facilities came together and data on facility performance was entered into software and simultaneously projected on a wall. This was done to quell misgivings about how funds were distributed, emphasize that the program is not a zero-sum game, "enhance transparency," create "a bit of healthy competition" and clarify questions and concerns among participating facilities collectively. Finally, to expedite the procurement of items, the implementer changed an NGO-wide procurement protocol to increase the local expenditure limit for purchases from \$5,000 to \$25,000.

4. Discussion

While there is ample literature on the impact produced by PBF programs, there has been relatively less focus on implementation processes and on how PBF is perceived among those enacting the intervention (Antony, 2017; Ogundeji, 2016; Bodson, 2018). This inhibits a more nuanced understanding of why PBF programs function or fail within a given setting, and it impedes programmers' abilities to replicate facilitators and bypass barriers. This study sought to highlight process-related findings from a health financing intervention and to assess whether and/or how seminal frameworks and theories from complementary fields could be used to better organize, understand and examine the implementation of a complex health financing intervention, such as the SSDI-PBI program.

We found that our research process (which began with building and crosschecking a program timeline, creating a taxonomy to arrange timeline events, and undertaking interviews that were guided by the

timeline but also drew upon existing frameworks and theories to formulate questions) allowed us to gather a comprehensive yet manageable body of knowledge. We note that other theories and models could also provide an organizing framework for such research (such as the Social-Ecological Model, which focuses on dynamic relationships across personal and environmental tiers (McLeroy, Bibeau, Steckler, & Glanz, 1988), and Walt and Gilson's policy analysis triangle framework (Walt & Gilson, 1994), which incorporates context, actors, process and content concepts in analyzing policies).

While PBF literature is largely quantitative and outcome-based, researchers have recently begun qualitatively examining discrepancies between what is expected during implementation of a PBF intervention versus what is actually taking place. These studies have placed particular emphasis on understanding verification efforts (the process of confirming reported data via engagement with communities, peers, implementers or contracting agencies). Verification is a cornerstone of PBF, but work in Burkina Faso (Turcotte-Tremblay et al., 2017), Uganda (Ssenooba et al., 2012), Benin (Antony et al., 2017) and our own study found that verification efforts are labor intensive, costly and can foment unintended negative consequences. A recent study in Benin highlighted a promising verification approach that is less resource intensive and relies on counter-verification via community level networks (Paul, Dramé, and Kashala, 2018). We learned that peer verifiers viewed the verification process as a zero-sum game, and thus sought to exact revenge on one another during peer verification thereby forcing an overhaul of the process. In Burkina Faso, community verification had the unintended (negative) effect of compromising patient confidentiality, sparking marital strife, and fostering fears within communities about retribution from facilities (Turcotte-Tremblay et al., 2017). Also in Burkina Faso, as well as Benin and Uganda, those who conducted verifications reported feeling overworked and/or dissatisfied with their compensation, which ultimately undercut the flow and fluidity of the broader program (Turcotte-Tremblay et al., 2017; Ssenooba et al., 2012; Antony et al., 2017).

The notion that effort be linked to reward in a timely feedback loop is the bedrock on which PBF is based. Despite this essential program feature, several studies have shown that implementation challenges can compromise the loop. In Benin, the time necessary to complete verification compromised time available for other aspects of the PBF cycle (providing feedback to facility staff, improving service delivery and awarding bonus payments) (Antony et al., 2017). In Sierra Leone and Nigeria, delays of a year or more in delivering payment led to a delinking of the relationship between effort and reward (Bertone et al., 2016; Bhatnagar and George, 2016). As described in another health financing intervention in Malawi (Wilhelm et al., 2016), our study found that several respondents had unfavorable experiences in terms of

the timing and nature of reward issuance. Despite knowing that timely procurement and disbursal of goods is a key feature of the SSDI-PBI program, implementers were unable to create a fluid, reliable supply chain in advance of program outset and/or to adapt performance metrics upon realizing that the disbursal system was flawed. Challenges related to reward issuance in Malawi stem from weaknesses across several tiers (limited human resource capacity, legally-mandated restrictions on how donor funds can be spent, internal standard operating procedures on how NGOs undertake purchasing, a poor road network, spikes in fuel costs etc). Nevertheless, we urge that programmers in this and similar settings prepare for this challenge within their forecasts.

Along with challenges related to verification and delays in terms of the timing or nature of reward issuance, we also found that in some respects the SSDI-PBI program overestimated existing technical and institutional capacities, which is echoed in studies from Uganda (Sengooba et al., 2012) and Tanzania (Chimhutu, Lindkvist, and Lange, 2014). As a health financing reform, it is expected that such an intervention will transcend many domains of the health system and that it will require human resource capacity beyond clinical expertise. Nevertheless, several respondents in our study described feeling surprised or underprepared regarding the business-management aspects of such an intervention and/or they noticed that the skills inherent to clinical care did not readily translate to an ability to harness data and conduct forecasting. Future efforts to undertake programs such as SSDI-PBI in Malawi may benefit from placing more emphasis on this facet of the program during sensitizations and trainings.

5. Limitations and opportunities for future research

This research would have been strengthened by the employment a longitudinal design. This study was commissioned after the program had begun, which limited the research team's ability to track developments and changes over time, thereby limiting recall bias. Furthermore, we highlight that our respondents were high-ranking within the health system (almost exclusively based in District Hospitals) and in terms of implementation (almost exclusively based at the Central level), which leads our findings to be weighted toward high-end implementation matters rather than issues that may be acute among community-based implementers or primary-care providers. Furthermore, we emphasize that this PBF program is unique; it was funded by the United States government, and was thus beholden to US government regulations and restrictions. Furthermore, unlike most PBF programs, this program did not allow redistribution of incentives as performance bonuses to individual health workers. In this sense, the rigidity or restricted adaptability of the program described herein may be linked to PBF programs generally, but it may also be linked to donor-imposed regulations, or some combination of the two. Regardless, the receipt of direct incentives affects expectations and motivations of providers, which in turn affects program fidelity. Finally, while we present our findings as organized via moderating factors, we caution that the creation of discrete moderator domains masks a more nuanced consideration of the ways in which moderators affect and are affected by one another. We hope that future research can better guide not only the mechanics of undertaking process evaluation research in this field, but also outline how moderators interrelate, and, possibly, how some moderators may be more (or less) important in a given program setting.

6. Conclusions

This study is among the few to focus on the implementation process of a performance-based program. While we identified and described several moderators of implementation, we view issues regarding the timely issuance of rewards and the need to consider existing technical capacities as essential considerations in terms of gauging how and where to place more pointed emphasis in similar programming within this and similar settings. As health financing reforms are introduced or

expanded into new settings, it becomes increasingly important that implementers and other stakeholders at a minimum remain attentive to program complexity, anticipate the nature and timing of obstacles and, more ideally, are prepared to recognize and then capitalize on promising features. We hope that research such as our own expands such discourse.

Ethical approval statement

This study is original, has not been submitted for review to other journals, has been approved by all authors, and the study has received ethical approval from the College of Medicine in Malawi (ethical approval code: P.09/15/1803) and the University of Heidelberg in Germany (ethical approval code: S-385/2015). The authors declare no competing interests.

Acknowledgements

The authors would like to thank the research team engaged in the broader evaluation including: Stephan Brenner, Rachel P. Chase, Julia Lohmann, Christopher Makwero, Don P. Mathanga and Aleksandra Torbica. We also thank data collectors who supported this research: Paul Mphemo and Ibrahim Sadala. We thank Valéry Ridde for his thoughtful review of our modified framework. This research project was made possible through Translating Research into Action, TRAction, which was funded by the United States Agency for International Development (USAID) under cooperative agreement number No. GHS-A-00-09-00015-00. The project team included the University of Heidelberg as prime recipient and the University of Malawi at the College of Medicine in Malawi as a sub-recipient. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Finally, we acknowledge financial support by the Deutsche Forschungsgemeinschaft within the funding programme Open Access Publishing, by the Baden-Württemberg Ministry of Science, Research and the Arts and by Ruprecht-Karls-Universität Heidelberg.

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 Services Research*, 17(1), 204.
- Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L., Sturdy, J., & Vermeersch, C. M. (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.
- Bertone, M. P., Lagarde, M., & Witter, S. (2016). Performance-based financing in the context of the complex remuneration of health workers: Findings from a mixed-method study in rural Sierra Leone. *BMC Health Services Research*, 16(1), 286.
- 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.
- Bhatnagar, A., & George, A. S. (2016). Motivating health workers up to a limit: Partial effects of performance-based financing on working environments in Nigeria. *Health Policy and Planning*, czw002.
- 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(1), 4.
- 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.
- Bradley, S., & McAuliffe, E. (2009). Mid-level providers in emergency obstetric and newborn health care: Factors affecting their performance and retention within the Malawian health system. *Human Resources for Health*, 7(1), 1.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Brenner, S., Muula, A. S., Robyn, P. J., et al. (2014). 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 Services Research*, 14(1), 180.
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implement Sci*, 2(1), 40.
- 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(1), 1.
- Chimhutu, V., Tjomsland, M., Songstad, N. G., Mrisho, M., & Moland, K. M. (2015). Introducing payment for performance in the health sector of Tanzania—the policy process. *Globalization and Health*, 11(1), 1.
- Chimwaza, W., Chipeta, E., Ngwira, A., et al. (2014). What makes staff consider leaving the health service in Malawi? *Human Resources for Health*, 12(1), 1.
- Chriwa M. (2013). Knowledge synthesis on Malawi health system: Literature review. <http://www.ndr.mw:8080/xmlui/bitstream/handle/123456789/904/Knowledge_synthesis_on_Malawi_health_system_literature_review.pdf?sequence=1>. Accessed May 29, 2017.
- Countdown to 2015, A Decade of Tracking Progress for Maternal, Newborn and Child Survival. Malawi Country Report. 2015; <<http://www.countdown2015mnch.org/country-profiles/malawi>> . Accessed 25 October 2016.
- 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), S36.
- Fritsche, G. B., Soeters, R., & Meessen, B. (2014). *Performance-based financing toolkit*. World Bank Publications.
- Gergen, J., Josephson, E., Coe, M., Ski, S., Madhavan, S., & Bauhoff, S. (2016). *How do performance-based financing programs incorporate quality of care? An overview of 32 programs in 29 countries*. Washington DC: ThinkWell, University Research Co., U.S. Agency for International Development, Center for Global Development11.
- Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implementation Science*, 5(1), 1.
- Hasson, H., Blomberg, S., & Dunér, A. (2012). Fidelity and moderating factors in complex interventions: A case study of a continuum of care program for frail elderly people in health and social care. *Implementation Science*, 7(23), 1–11.
- Lawler, E. (1988). Pay for performance – Making it work. *Personnel*, 65(10), 68–71.
- Lawler, E. E., & Suttle, J. L. (1973). Expectancy theory and job behavior. *Organizational Behavior and Human Performance*, 9(3), 482–503.
- Magrath, P., & Nichter, M. (2012). Paying for performance and the social relations of health care provision: An anthropological perspective. *Social Science Medicine*, 75(10), 1778–1785.
- Malawi's "cashgate" scandal (2014). The \$32m heist. *The Economist*. London.
- McAuliffe, E., Manafa, O., Maseko, F., Bowie, C., & White, E. (2009). Understanding job satisfaction amongst mid-level cadres in Malawi: The contribution of organisational justice. *Reproductive Health Matters*, 17(33), 80–90.
- McLeroy, K. R., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education Quarterly*, 15(4), 351–377.
- McMahon, S. A., Brenner, S., Lohmann, J., et al. (2016). Evaluating complex health financing interventions: Using mixed methods to inform further implementation of a novel PBI intervention in rural Malawi. *BMC Health Services Research*, 16(1), 414.
- Measure DHS (2010). Malawi Demographic and Health Survey (MDHS) - HIV Prevalence. MDHS 2010; <<https://dhsprogram.com/pubs/pdf/HF34/HF34.pdf>> . Accessed 17 November 2016.
- Ministry of Health Malawi II (2014). *Malawi Service Provision Assessment (MSPA) 2013-14*. Lilongwe, Malawi, and Rockville, MD, USA2013.
- Musgrove, P. (2011). *Financial and other rewards for good performance or results: A guided tour of concepts and terms and a short glossary*. Washington, DC: World Bank1–12.
- NSONMa I. (2017). 2015-16 Demographic and Health Survey Key Findings. In: I. UNA (ed.) Zomba, Malawi, and Rockville, Maryland.
- Ogundeji, Y. K., Jackson, C., Sheldon, T., Olubajo, O., & Ihebuzor, N. (2016). Pay for performance in Nigeria: The influence of context and implementation on results. *Health policy and planning*, 31(8), 955–963.
- Paul, E., Albert, L., Bisala, B. N. S., et al. (2018). Performance-based financing in low-income and middle-income countries: Isn't it time for a rethink? *BMJ Global Health*, 3(1), e000664.
- Paul, E., Dramé, M. L., Kashala, J.-P., et al. (2018). Performance-based financing to strengthen the health system in Benin: Challenging the mainstream approach. *International Journal of Health Policy and Management*, 7(1), 35.
- 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 of Health Policy and Management*, 3, 207–214.
- 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 and Planning*, czw045.
- Ridde, V., Druetz, T., Poppy, S., Kouanda, S., & Haddad, S. (2013). Implementation fidelity of the national malaria control program in Burkina Faso. *PLoS One*, 8(7), e69865.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A systematic approach*. Sage publications.
- Shroff, Z. C., Bigdeli, M., & Meessen, B. (2017). From scheme to system (part 2): Findings from ten countries on the policy evolution of results-based financing in health systems. *Health Systems Reform*, 3(2), 137–147.
- Sengooba, 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.
- 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.
- UN (2015). Millennium Development Goals - Malawi; <<http://www.mw.one.un.org/millennium-development-goals/>>. Accessed 17 November 2016.
- UNDP (2015). Human Development Index and its components. *Human Development Reports*; <<http://hdr.undp.org/en/composite/HDI>>. Accessed 17 November 2016.
- Walt, G., & Gilson, L. (1994). Reforming the health sector in developing countries: The central role of policy analysis. *Health Policy and Planning*, 9(4), 353–370.
- Waweru, E., Goodman, C., Kedenge, S., Tsofa, B., & Molyneux, S. (2016). Tracking implementation and (un) intended consequences: A process evaluation of an innovative peripheral health facility financing mechanism in Kenya. *Health Policy and Planning*, 31(2), 137–147.
- WHO (2015). The African Health Observatory - Malawi. *The African Health Observatory*; <http://www.who.afro.int/profiles_information/index.php/Malawi:Index>. Accessed 29 May 2017.
- Wilhelm, D. J., Brenner, S., Muula, A. S., & De Allegri, M. (2016). A qualitative study assessing the acceptability and adoption of implementing a results based financing intervention to improve maternal and neonatal health in Malawi. *BMC Health Services Research*, 16(1), 398.
- 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. *Cochrane Database System Review*, 2(2), CD007899.
- Witter, S., Toonen, J., Meessen, B., Kagubare, J., Fritsche, G., & Vaughan, K. (2013). Performance-based financing as a health system reform: Mapping the key dimensions for monitoring and evaluation. *BMC Health Services Research*, 13(1), 367.
- World Bank (2015). GDP per capita *World Bank Open Data*; <http://data.worldbank.org/indicator/NY.GDP.PCAP.CD?Order=wbapi_data_value_2013+wbapi_data_value&sort=asc&year_high_desc=false>. Accessed 17 November 2016.