

Contribution analysis as an evaluation strategy in the context of a sector-wide approach: Performance-based health financing in Rwanda

Authors:

Martin Noltze¹
Felix Gaisbauer¹
Thomas Schwedersky¹
Stefanie Krapp¹

Affiliations:

¹German Institute for Development Evaluation (DEval), Germany

Correspondence to:

Martin Noltze

Email:

martin.noltze@deval.org

Postal address:

German Institute for Development Evaluation, Tulpenfeld 7, Bonn 53113, Germany

Dates:

Received: 30 Apr. 2014
Accepted: 18 Sept. 2014
Published: 18 Dec. 2014

How to cite this article:

Noltze, M., Gaisbauer, F., Schwedersky, T. & Krapp, S., 2014, 'Contribution analysis as an evaluation strategy in the context of a sector-wide approach: Performance-based health financing in Rwanda', *African Evaluation Journal* 2(1), Art. #81, 8 pages. <http://dx.doi.org/10.4102/aej.v2i1.81>

Note:

The paper was presented at the 7th African Evaluation Association (AfrEA) conference held in Yaounde, Cameroon, 1–5 March 2014.

Copyright:

© 2014. The Authors.
Licensee: AOSIS OpenJournals. This work is licensed under the Creative Commons Attribution License.

Read online:

Scan this QR code with your smart phone or mobile device to read online.

Sector-wide approaches (SWAPs) emerged as a response to donor fragmentation and non-adjusted and parallel programming. In the health sector, SWAPs have received considerable support by the international donor community due to their potential to reduce inefficiencies through alignment to common procedures and hence to increase development effectiveness. Evaluating development cooperation in the context of a SWAP, however, translates into methodological challenges for evaluators who have to disentangle the cumulative effects in strongly donor-aligned, complex sector environments. In this article the authors discussed the application of a methodological strategy for evaluating development interventions in complex settings – for example in the context of a SWAP – and reflected the suitability of the approach. The authors conducted a contribution analysis, a theory-based approach to evaluation, and exemplified the approach for an intervention of performance-based financing for Rwandan health workers supported by the Rwanda-German cooperation. The findings suggested that the Rwandan system of performance-based financing increased service orientation and outputs of health professionals, but also indicated that negative motivational side effects and resource constraints are real. With regard to the methodological approach, the authors conclude that contribution analysis has a high potential to evaluate development cooperation in the context of a SWAP due to its high flexibility to use different data collection tools and its capability to assess risks and rival explanations. Challenges can be identified with regard to the efficiency of the evaluation strategy and a remaining trade-off between scope and causal strength of evidence.

Les approches sectorielles (SWAP, Sector-Wide Approach) ont émergé comme réponse à la fragmentation des donateurs et à la programmation non-actualisée et parallèle. Dans le secteur de la santé, les approches sectorielles ont reçu un appui considérable de la communauté internationale des donateurs en raison de leur potentiel à réduire les inefficacités par l'alignement de procédures communes et donc d'augmenter l'efficacité du développement. Évaluer la coopération de développement dans le contexte d'une approche sectorielle se traduit toutefois par des difficultés méthodologiques pour les évaluateurs, qui doivent démêler les effets cumulatifs dans des environnements complexes et fortement alignés sur les donateurs. Dans cet article, les auteurs ont discuté de l'application d'une stratégie méthodologique pour l'évaluation des interventions de développement dans des contextes complexes, par exemple dans le contexte d'une approche sectorielle, et ont réfléchi sur la pertinence de l'approche. Les auteurs ont mené une analyse de contribution, une approche de l'évaluation basée sur la théorie, et illustré l'approche d'une intervention de financement basé sur les résultats auprès d'agents de santé rwandais soutenue par la coopération rwan-do-allemande. Les résultats suggèrent que le système rwandais de financement basé sur les résultats a amélioré l'orientation et les prestations de service des professionnels de la santé, mais ils ont également indiqué que les effets secondaires négatifs sur la motivation et que les contraintes de ressources étaient réels. En ce qui concerne l'approche méthodologique, les auteurs concluent que l'analyse de contribution a un fort potentiel d'évaluation de la coopération au développement dans le contexte d'une approche sectorielle, en raison de sa grande flexibilité à utiliser différents outils de collecte de données et sa capacité à évaluer les risques et les explications rivales. Les défis peuvent être identifiés eu égard à l'efficacité de la stratégie d'évaluation et d'un compromis persistant entre la portée et la force de causalité des preuves.

Introduction

Sector-wide approaches (SWAp) were introduced as a response to donor fragmentation and non-adjusted and parallel programming (Peters, Paina & Schleimann 2013). In a SWAp funds by development partners contribute directly to developing and implementing a defined sector policy under a government authority (Cassels 1997). Key principles of a SWAp are that the partner government leads and takes ownership of the sector programme and that there is a shared effort by development partners to support the programme (Cassels 1997; Peters *et al.* 2013; Vaillancourt 2009). Ideally, different aid modalities are combined, including joint financing and technical assistance.

Since the early 1990s SWAp have evolved in the health sector of developing countries and received considerable support by the international donor community due to their potential to reduce inefficiencies through alignment to common procedures and hence to increase development effectiveness (Cassels 1997; Vaillancourt 2009). Whilst this led some scholars to declare SWAp as a panacea for development and also influenced the policy process initiated by the 2005 Paris Declaration on Aid Effectiveness, recent experiences from Africa provide a less positive assessment with regard to development effectiveness of interventions embedded in the context of a health SWAp (Walford 2007). Moreover, new global health initiatives introduced to target specific diseases in the course of the health-related Millennium Development Goals (MDGs), such as the Global Fund for AIDS, TB and Malaria (GFATM) or the President's Emergency Plan for HIV and/or AIDS (PEPFAR), are often found to bypass policies and management processes of partner countries (Peters *et al.* 2013).

In spite of a wide strand of literature about the SWAp model itself, there is still limited evidence from empirical case studies investigating the capacities and performance of interventions embedded in a SWAp environment and the factors that determine their success (or failure). The authors argue that this also stems from a lack of methodological approaches suitable to disentangle the cumulative efforts of multiple development partners, influencing factors and alternative explanations (cf. Leeuw & Vaessen 2009; Pawson & Tilley 1997) for evaluating highly aligned development interventions in complex sector environments for the purpose of learning and (mutual) accountability.

In this article, the authors propose contribution analysis, a theory-based evaluation approach, to manage the methodological challenges using the Rwandan-German cooperation on performance-based financing as an example. To date, the applied methodological approach has not, to the knowledge of the authors, been used for development evaluation in the context of a (health) SWAp. The rest of the article is organised as follows: the next section outlines the approach and specific features of the Rwandan-German cooperation in health. The subsequent section introduces the methodological strategy, a roadmap for its implementation

and the results. Finally, the authors discuss some methodological limitations and present conclusions.

The Rwandan-German cooperation in health

Rwanda has made remarkable development progress since the dramatic genocide in the mid-1990s, which severely set back the country's social and economic development (Binagwaho *et al.* 2014). Although strong economic growth and pro-poor development together with large investments into the public health system brought Rwanda on track to achieve most of the health-related MDG targets before 2015 (International Health Partnership + related initiatives 2011), Rwanda remains one of the poorest and most aid-dependent countries in the world as it struggles with persistent poverty and economic inequality (Abbott & Rwirahira 2012). As a result, the Rwandan health sector is still highly dependent on foreign support which, in turn, manifests in a multitude of donors whose support has to be managed.

Looking back on more than 30 years of development cooperation in health, the Rwandan-German cooperation evolved along the Paris Declaration paradigm and was intensified by aligning it to the umbrella of a SWAp and incorporating joint financing modalities since 2007. In compliance with a national policy on the division of labour of development partners, introduced in 2011 (cf. GoR & MINECOFIN 2013), Germany phased out its support at the end of 2012.

In the course of the health SWAp, major development partners, including Belgium, the United States of America, the United Kingdom, Switzerland, Luxembourg and Germany, aligned their interventions to the health sector strategic plans and the overall national Economic Development and Poverty Reduction Strategy of Rwanda. The process was accompanied by a national health policy, a strategy which followed the policy process initiated by the Paris Declaration (GoR 2005). At the national level, donor coordination was managed in a development partner group, a health sector working group which coordinated the SWAp and about 30 subordinate technical working groups (GoR & MoH 2012). At the decentralised level, development partners were involved in joint action development forums. Besides various forms of technical assistance, joint financing modalities such as sector budget support and basket funding were established.

Within the SWAp structure, the Rwandan-German strategy focused on health system strengthening by promoting primary health care. The cooperation followed a multi-level approach by operating on all health system levels: from support to the technical working groups at the national level, over cooperation with health district administrations to the support of community-based initiatives. Support to the SWAp was provided to improve coordination and cooperation between the Rwandan government (GoR) and development partners. Joint financing modalities were used to increase ownership of the Rwandan Ministry of Health

(MoH) and to improve the capacities of public financial management. These outcomes were envisaged to enhance the quality of care and to contribute to a health system that better responds to the needs of the population, especially the poor. Besides support to the SWAp itself and in coordination with the MoH as well as other development partners, Germany concentrated its support on three components including (1) human resource development, (2) sexual and reproductive health and (3) health financing.

In this article, the authors exemplarily investigate the support by German Development Cooperation (GDC) to one selected intervention of the health financing component: performance-based financing (PBF). PBF is a flagship amongst the Rwandan health sector reforms designed to overcome low-quality services and has sparked special international interest (Logie, Rowson & Ndagije 2008). PBF can be defined as the transfer of conditional financial support on the basis of performance targets (Eichler & Levine 2009). Rwanda piloted PBF approaches in 2002 and rolled out a nation-wide program in 2005 (cf. Basinga *et al.* 2011; Soeters, Musango & Meessen 2005). Whilst PBF encompasses all care levels, in the following the authors only refer to the administrative procedure at the level of hospitals (cf. Rusa *et al.* 2009): PBF budgets for hospitals are set prospectively, based on an annual value of about \$600 per bed. Each quarter, a team from a peer hospital assesses the hospital's performance according to a checklist of over 50 indicators. The percentage of the prospective budget disbursed as incentive is contingent on the degree of indicator achievement. The district hospital decides on how to use these performance fees. Following a multi-level approach, GDC support to PBF included technical advice at the national level, and direct financial assistance to one specific hospital, Ruhengeri in the Musanze district, between 2007 and 2011.

Methodological approach

In this section, the authors present their methodological approach for evaluating the effectiveness of development interventions in the context of a health SWAp in Rwanda. Firstly, the nature and suitability of contribution analysis as a methodological strategy for this evaluation are introduced. Secondly, Mayne's (2011; 2012) steps to contribution analysis are outlined and applied using the example of the Rwandan-German PBF intervention. Moreover, the authors also account for suggestions made by Delahais and Toulemonde (2012).

Rationale for applying contribution analysis

Adequately assessing aid effectiveness for a highly integrated program with multiple interventions that are operating at different levels in a multi-donor SWAp environment is a methodological challenge. In the case of this study, applying a (quasi-)experimental approach proved impossible because neither the Rwandan-German programme nor parts of it have been implemented following randomisation (cf. Shadish, Cook & Campbell 2002). In addition, the scarcity of available

data (e.g. lack of baseline information, absence of program-specific monitoring) does not allow eliminating threats to internal validity (Connell & Kubisch 1998; Leeuw & Vaessen 2009; Shadish *et al.* 2002). For such complex situations, Stern *et al.* (2012) point to a trade-off between scope of a programme and strength of causal evidence.

Given these preconditions for evaluating interventions in a complex SWAp environment, the authors chose to adopt a contribution analysis as the evaluation strategy. Contribution analysis is a theory-based evaluation approach that has been developed for scenarios when attribution analysis is not feasible (Delahais & Toulemonde 2012; Mayne 2011; 2012). Contribution analysis has recently raised attention: the journal *Evaluation* published a special issue on this matter in 2012. The methodology seeks to answer whether and to what extent the intervention made a noticeable contribution to an observed result under some given influencing factors (Mayne 2012). Thereby contribution analysis puts stronger focus on causal mechanisms and the interplay of influencing factors than the (quasi-)experimental rationale (Leeuw 2012). This emphasis makes contribution analysis a suitable strategy for taking stock of the variety of experiences gained in the course of the Rwandan-German cooperation in a SWAp environment.

Steps of conducting a contribution analysis

The conceptual backbone to contribution analysis is examining and testing the theory of change of an intervention against logic and evidence. Causality is inferred if the intervention is based on a reasoned theory of change, has been implemented as planned and can be corroborated by evidence whilst rival explanations are accounted for (Mayne 2012). Thereby, contribution analysis follows six key steps (cf. Mayne 2012:272): (1) set out the cause-effect issue to be addressed, (2) develop the postulated theory of change and risks to it, including rival explanations, (3) gather the existing evidence on the theory of change, (4) assemble and assess the contribution claim and challenges to it, (5) seek out additional evidence and (6) revise and strengthen the contribution story.

The basic unit of analysis is the causal link: the postulated causal mechanism linking two elements of a results chain. Accordingly, an evaluation would question whether an intended change did (or did not) occur, due (or not due) to the intervention's contribution (cf. Delahais & Toulemonde 2012:291). The robustness of a causal claim depends on the items of evidence gathered from existing studies, secondary data analysis or analysis of newly collected primary data, as well as their strength of evidence and their triangulation. Finally, a contribution story presents one or more contribution claims as a coherent narrative of (a branch of) the theory of change and its evidence base.

Step 1: Set out the cause-effect issue to be addressed

Addressing the cause-effect issue is central for identifying useful evaluation questions. The authors included stakeholders from ministries in Rwanda and Germany, the

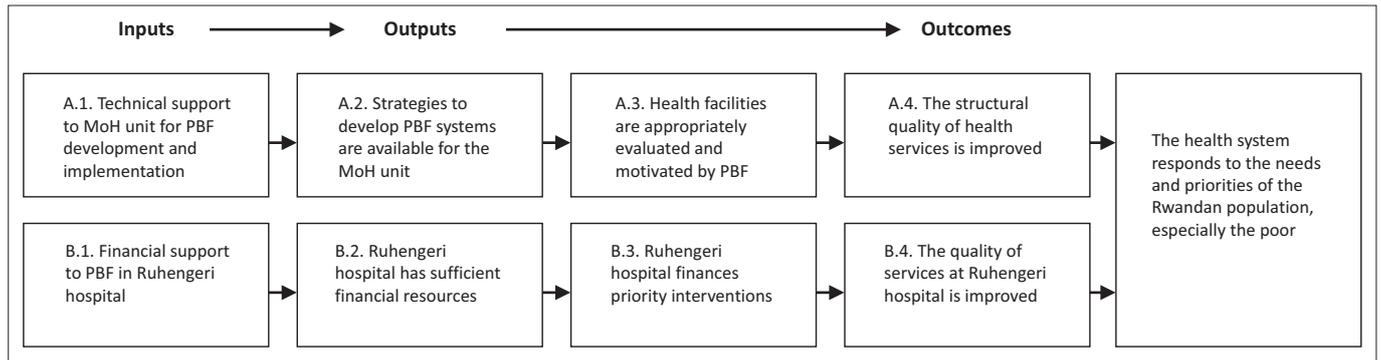


FIGURE 1: Theory of change for PBF within the Rwandan-German cooperation in health.

TABLE 1: Assumption, risks and rival explanations (for causal link A.3. → A.4.).

Causal link	A.3. → A.4. (see Figure 1)
Narrative	The appropriate evaluation and motivation of health facilities by PBF leads to an improvement of the structural quality of health services.
Main assumptions	Low motivation of health professionals is the main impediment to high-quality services. Financial incentives conditional on pre-defined performance achievements can overcome low motivation of individual health workers.
Risks	Financial incentives (extrinsic motivation) may crowd out intrinsic motivation (Gorter <i>et al.</i> 2013; Grittner 2013). Over-focus on incentivised indicator ('gaming') at the expense of other relevant indicators including the distortion of information to score higher on indicators (Kalk <i>et al.</i> 2010; Oxman & Fretheim 2009; Paul 2009).
Rival explanation	Overall budget increase: between 2000 and 2006, Rwanda experienced a four-fold increase of total health expenditure (GoR & MoH 2008).

German implementing agencies, partner institutions and academia to agree on a common object of the evaluation. Based on a shared understanding of the problem statement, policy relevance, resource inputs and feasibility of the study, the thematic focus of the evaluation was selected and shaped. Finally, the stakeholders agreed to focus the purpose of the study on learning, accountability and strategic management. Thus, this first step was used to put principles of utilisation-focused evaluation into practice (cf. Patton 2011).

Step 2: Develop the postulated theory of change and risks to it, including rival explanations

For developing the theory of change, planning documents for the programme were screened using an analytical framework on inputs, processes, outputs, outcomes and impact levels as orientation. Simultaneously, assumptions, risks and rival explanations were collected from planning documents, critical reasoning, experience and both academic and grey literature in order to verify the causal links as reconstructed in the theory of change.

Figure 1 presents a schematic overview of two strands of the theory of change for PBF. The first strand (A.1. to A.4.) represents the programme's logic at central level that is intervening at the responsible unit at the Rwandan MoH. According to the intervention logic, advice on policy and technical issues to MoH (input A.1.) leads to suitable PBF strategies (output A.2.). The implementation of these strategies at the level of health facilities leads to proper performance assessments which trigger a higher motivation of staff (output A.3.) and finally lead to an improvement in the structural quality of health services (outcome A.4.). The second strand (B.1. to B.4.) represents the program's efforts on the decentralised level, where GDC was solely in charge of paying the incentives for the district hospital of Ruhengeri, which should lead – adhering to the same

logic – to an improvement of service quality at the hospital. The programme's activities on both the central and decentralised levels should finally contribute to a more pro-poor responsive Rwandan health system.

For the purpose of this article, the authors exemplify the steps for contribution analysis in detail by referring to a single causal link: the relationship between output A.3. and outcome A.4., in other words the causal mechanism explaining how performance monitoring leads to improved health services by increasing the health workers' motivation. As suggested by Mayne (2012), assumptions, risks and rival explanations related to this link were identified (see Table 1). Afterwards, the results section again refers to the two causal strands of PBF as a whole.

Two core assumptions underlie the PBF evaluation system: (1) low motivation of health professionals is the main impediment to high-quality services – as opposed to other resource constraints – which (2) can be overcome through (financial) incentives conditional on pre-defined performance achievements by individual health workers.

A literature review yielded two main risks related to PBF: financial incentives, by definition extrinsically motivating, (1) may crowd out intrinsic motivation and (2) can lead to gaming, or manipulating the system to achieve a desired outcome. If applicable, these risks could result in unintended behaviours such as skimping on quality or focusing on incentivised indicators at the expense of other relevant services (Gorter, Ir & Meessen 2013; Grittner 2013; Oxman & Fretheim 2009).

Whilst there are several rival explanations that PBF did not (or did not as intended) contribute to the structural quality

TABLE 2: Data collection tools.

Number	Data collection tool	Health system level [†]	Population
1	Literature review	National and decentralised	Unspecified.
2	Survey with follow-up in-depth interviews	Decentralised	German technical assistants.
3	Interviews, mostly semi-structured in-depth interviews	National and decentralised	Project and programme staff and counterparts; key informants of health SWAp; other development partners, non-governmental organisations.
4	Comparative case study with in-depth interviews and focus group discussions	Decentralised	District health staff at all district care levels, community health workers and end users of health care services.

[†] The national level represents the Ministry of Health and national administration system; the decentralised level refers to the health districts.

of health services in Rwanda, the authors highlight only a single prominent alternative explanation in this article for the sake of brevity. One might argue that improvements in health service quality stem from an increased budget allocation for health which trickled down to health facilities and finally improved service quality. This contextual rival explanation is a priori valid as Rwanda experienced a four-fold increase of overall health expenditure between 2000 and 2006, which overlapped the time period of German support to the Rwandan PBF initiative (GoR & MoH 2008).

Step 2 allows the evaluator to assess the inherent logic of causal links according to plausibility, one criterion for assessing a theory of change suggested by Connell and Kubisch (1998). The authors found that the programme's planning documents did not plausibly outline how output A.3. ('facilities are appropriately evaluated and motivated by PBF system') leads to outcome A.4. ('the structural quality of health services is improved'). At this stage, a first literature review did not lead to the conclusion that improvements in the quality of services are strictly attributable to external incentives under PBF (cf., e.g., Gorter *et al.* 2013; Ireland, Paul & Dujardin 2011). This step of contribution analysis therefore is useful to identify formal flaws in theories of change and can help in patching them in evaluations and for future interventions. In this example, policymakers and implementing agencies should more explicitly outline how performance assessments in the Rwandan context supposedly trigger motivational mechanisms and behaviour change.

Step 3 and 4: Gather existing evidence on the theory of change and assemble and assess the contribution claim and challenges to it

On the basis of the elaborated assumptions, risks and alternative explanations, and a first literature review to assess the plausibility of the theory of change, the authors crafted evidence analysis tables for each branch of the theory of change in accordance with Delahais and Toulemonde (2012), concentrating on information confirming or refuting a causal link as outlined in the theory of change. This stocktaking revealed that assembling the contribution claim, as recommended by Mayne (2011), based on the available items of evidence would only allow causal inferences of limited robustness. Given the necessity to collect further (primary) data on PBF in Rwanda, the authors directly went to empirical data collection (step 5). It should be noted that Delahais and Toulemonde (2012) found as well that collecting primary and secondary data in parallel may better fit the time requirements of real-world evaluations.

Steps 5: Seek out additional evidence

Structuring the available items of evidence along the theory of change allowed formally assessment of the extent to which single links had already been covered by evidence and which needed to be investigated in more detail. The authors follow Mayne (2012) and White and Phillips (2012) that contribution analysis rather provides a more analytical strategy for scrutinising causal links of a theory of change than a strict design or even proposal for specific methods. As a corollary, evaluators should feel encouraged to choose from the toolbox of social science research methods whatever fits their purpose for investigating the nature of a causal link.

In this example, and based on the gaps in the items of evidence as derived from step 3, the authors designed the following data collection tools for assessing the causal strands for PBF (see Table 2): (1) a comprehensive literature review, (2) a survey with follow-up in-depth interviews with former technical assistants, (3) semi-structured interviews with key informants including Rwandan officials, German programme staff and other development partners and (4) a comparative case study of four district health systems, including two districts where the programme did not operate at decentralised level, to collect cross-sectional data with regard to the administrative structure and service provision. The latter comprised in-depth interviews and focus group discussions with different populations.

The authors incorporated questions on the causal link to be tested in the data collection tools. As a result, they gathered a multitude of items of evidence from different methods and sources, both of a quantitative and qualitative nature and at national as well as decentralised level. The authors organised and analysed all items of evidence (literature, interviews, reports, etc.) using MAXQDA 11, a qualitative data analysis software program.

Step 6: Revise and strengthen the contribution story

This step encompasses building a more credible contribution story and reassessing its strengths and weaknesses. Due to time constraints, the authors did not follow the iterative process between step 4 through 6 as suggested by Mayne (2012) and hence went directly to developing the contribution claims and composing the contribution story. The contribution claim for the exemplary causal link discussed in this article, as well as (parts of) the contribution story for the overarching PBF strand, are part of the following section on the results.

TABLE 3: Evidence analysis table (for causal link A.3.→A.4.).

Item of evidence	Type of source	Convergence with theory of change	Type of causal mechanism	Strength of evidence [†]
Increase in use and quality of incentivised services	Interviews; literature; comparative case study	Confirming	Intended	High
Increase in motivation, performance, responsibility and participation of health professionals who receive performance-based financing	Literature; interviews; comparative case study	Confirming	Intended	Moderate
PBF created feedback loop and hence better mutual communication between management and staff	Interviews	Confirming	Intended	Low
'Gaming' occurred (manipulation of medical records)	Literature; interviews	Refuting	Unintended	Moderate
Indicators are not consistent with staff members' self-assessment, hence the narrative accompanying the indicators does not allow to infer 'structural improvements'	External evaluation (case study in one hospital)	Refuting	Unintended	Low
Lack of motivation due to delayed payments of salaries or generally low salary level	Comparative case study	Refuting	Unintended	Moderate
Hospital management and staff report an insufficient number of health professionals, high turnover rates and insufficient equipment and materials	Comparative case study	Refuting and confirming different aspects	Unintended	Moderate

[†], The strength of causal evidence is based on the type of sources that confirm or refute the convergence with the theory of change. Low = Limited sources that support this item of evidence; Moderate = Some relevant and reliable evidences are available; High = Sufficient and strong evidence available.

Results

In this section, the effectiveness of GDC support to the Rwandan strategy to overcome low-quality services in the health sector by investigating GDC interventions focusing on PBF is assessed. For this the authors examine in detail the extent to which a single causal link examined in the previous section contributes to the overall contribution story for the two strands of the PBF programme theory of change. The causal link A.3.→A.4., nevertheless, still serves as an example for an evidence analysis table, as outlined in Table 3. Special emphasis is placed on the programme's effectiveness towards increased pro-poor responsiveness of the Rwandan health system.

The self-reported German ambition regarding PBF was to put stronger emphasis on assessing service quality with the PBF indicator system. For this, GDC actively participated through SWAp mechanisms in technical working groups at the national level. Other contributors were the GoR followed by the Belgian cooperation, the World Bank, and non-governmental organisations (World Health Organization 2009).

Regarding decentralised support, GDC contributed to the training of PBF evaluators and data managers, participated in peer reviews of hospitals and disbursed about €0.5 million as performance incentives to the supported district hospital Ruhengeri. Over the time period of interest, Ruhengeri hospital has made far-reaching administrative rearrangements. Whilst the PBF evaluation system was not well understood and accepted by hospital staff at its introduction in 2008 (Deutsche Gesellschaft Für Technische Zusammenarbeit 2008), the comparative case study sheds light on the most recent perceptions: focus group discussants at hospitals and resource persons in the districts (including Ruhengeri) acknowledged the alignment of PBF to the district performance contract system. The respondents often perceived the PBF indicator system as too rigid and at times unreachable and questioned whether the indicators validly measure the defined targets. Nevertheless, health staff self-report higher motivation for their work due to the PBF system.

These self-assessments do not, however, allow determining whether this effect is attributable to setting incentives rather than to a general increase in money.

The analysis of the items of evidence also sheds light on the formulated assumptions and risks as presented above (see Table 1). Evidence on the assumption of low motivation of health professionals as the main impediment to high-quality services is mixed: several interviews indicate that intrinsic motivation might be low amongst health workers, whilst secondary interview data by Paul (2009) suggest that Rwandan health workers abide by a high work ethos. Whilst this inconclusiveness impedes a final effectiveness assessment for this aspect of the causal link, it underscores the importance of making explicit and testing one's assumptions about how a programme works. Regarding the programme, it becomes obvious that the relationship between motivation and performance is conditional on more than (the postulated) monetary rewards. This, in turn, lends further plausibility to the assumption that negative side effects occurred. Regarding the identified risks, negative effects of PBF in Ruhengeri such as crowding out and gaming have been consistently reported and researched (Kalk 2011; Kalk, Paul & Grabosch 2010; Paul 2009) and pose a threat to improving the structural quality of health services, usually under-researched (Ireland *et al.* 2011). Whilst these experiences in Ruhengeri also led GDC to more strongly focus its support on other interventions, it is unclear if these effects prevail today. The nationwide PBF roll-out has been conducted in a standardised format, thus lending plausibility to the assumption that other district hospitals have faced similar problems.

Indubitably, it can be concluded from interviews, reports and studies that PBF has induced positive changes: quality and quantity of health services have increased (Sherry, Bauhoff & Mohanan 2013), health staff self-report higher personal performance, less absenteeism and improved internal communication and supervision (Kalk *et al.* 2010; Paul 2009). According to the PBF indicator score, the overall performance of Ruhengeri hospital rose from 78% in 2008 to 84% in 2010 (data provided by GDC).

Next to the improvement of the structural quality of health service, it is assumed that PBF increases capacity to manage for results and autonomy at the provider level. End users report in focus group discussions to have been treated with more respect at health facilities since the introduction of PBF. Improved client satisfaction can be seen as an indicator for service responsiveness (Kruk & Freedman 2008). Moreover, health professionals report that due to improved supervision, the management can indeed tailor the facilities' services more closely to the clients' needs. The same key informants voiced additional major risks: firstly, PBF might exacerbate inequities between health providers because the inflexible indicator system will favour well-capacitated over resource-poor facilities in the long run; secondly, most respondents considered that their work environment suffered from problems not tackled by PBF, such as high turnover rates and lack of equipment. In particular, the raised affordability of health services due to the introduction of a nation-wide health insurance scheme has markedly increased their workload and the strain on hospital capacities, so that other resource constraints may become the bottleneck to high-quality health services.

The finding that hospital management and staff report an insufficient number of health professionals, high turnover rates and insufficient equipment and materials again illustrates that insufficient motivation is not the sole impediment to high-quality health services, thus questioning one of the main assumptions underlying PBF. On the other hand, the recurrent theme of a poorly capacitated work environment somewhat invalidates the rival explanation that an upward trend in Rwanda's total health expenditure has trickled down to improve health services on facility level to such an extent that it would override the positive changes induced by PBF.

To conclude, the evidence confirms that PBF does increase service orientation and outputs of health professionals, but also suggests that negative motivational side effects and resource constraints are real and constrain the intervention's overall effectiveness.

Discussion

On the basis of the presented results, in this section the authors discuss their experience with contribution analysis as a methodological strategy and its strengths and weaknesses in assessing the effectiveness of development interventions in a complex programme environment.

The authors applied contribution analysis in order to formally manage the high complexity surrounding the interventions under evaluation when attribution is not possible. To this end, they firstly reconstructed the programme's theory of change, before the main analysis concentrated on single causal links by using evidence analysis tables (cf. Delahais & Toulemonde 2012). By doing so the authors ended up with a multitude of causal links to be tested. With regard to the evaluation of similar interventions, two fundamental problems arise: firstly, every additional link under examination increases the resource demands for the overall

study. This necessarily leads to the above-mentioned trade-off between scope and causal strength of evidence (Stern *et al.* 2012). Only deliberate decisions regarding the scope of an evaluation by the evaluation stakeholders and team, taking into account the principle of evaluation efficiency, can thus ensure a utilisation-focus of similarly complex evaluations like that of a highly integrated programme-based approach in this article. Secondly, and related to the first problem, programmes operating in a SWAp environment reshape the attribution problem. Rather than an attribution gap, the attribution problem is arguably better framed as gradient or continuum: in the case that support is strongly aligned to national priorities and well harmonised amongst development partners, even tracing more proximal effects back to the partners' deliverables becomes increasingly difficult, let alone attributing the more highly aggregated outcome levels to the interventions. A resulting lesson learnt was that sufficient resources should be devoted to explicitly formulating and collecting other influencing factors and rival explanations questioning the programme's contribution. The authors allocated considerable time and resources to this step when reconstructing the theory of change and are confident they have gathered sufficient evidence on the mechanisms of the PBF intervention to solidly back the conclusions drawn in this article. For future complex evaluations, however, the authors see an untapped potential for more strictly testing theories of change by placing more emphasis on rival explanations and invalidating their postulated influence and putting more focus on validating the theory of change with key stakeholders.

Conclusion

In this article the authors analysed the effectiveness of a development intervention in the context of a health SWAp in Rwanda applying contribution analysis as the overarching evaluation strategy. They conclude that contribution analysis is suitable to provide orientation for the analysis of single interventions in a complex sector environment when the programme is strongly aligned to and integrated into the policy framework of the partner country. As such, contribution analysis was capable to also consider alternative explanations which are prevalent in a sector environment with multiple donor activities in place. Finally, the evaluation strategy was flexible enough to incorporate different data collection tools and analytical methods to test different causal links in a systematic manner.

Challenges, however, were found with regard to the efficiency of the investigation as each additional link under examination increases the resource demands for the overall study. This necessarily leads to a trade-off between scope and causal strength of evidence. Moreover, for similar complex evaluations, the authors see an untapped potential for more strictly testing theories of change by placing more emphasis on rival explanations and invalidating their postulated influence and putting more focus on validating the theory of change with key stakeholders.

The evidence gathered confirms that the Rwandan system of PBF does increase service orientation and outputs of health

professionals, but also suggests that negative motivational side effects are real and resource constraints limit the overall effectiveness. GDC's contribution to PBF is thus rated as moderate. For further probing the potential of PBF in Rwanda, the aspects of effectiveness and efficiency would merit further methodologically rigorous research, especially with regard to ruling out alternative explanations such as a motivational increase by raising health professionals' salaries without performance orientation and appraising the cost-effectiveness of the laborious verification and reporting process. Future PBF interventions should be coupled more strongly with capacity development for the supervising and managerial staff and, at best, with interventions improving the physical work environment. Otherwise the goals and indicators set are perceived as out of reach and might lead to demotivation or faked reporting on goal achievement.

Acknowledgements

The authors of this article would like to thank all interviewees, the stakeholders in Rwanda and Germany, the Rwandan Ministry of Health, the German Embassy, the German Federal Ministry for Economic Cooperation and Development and the implementing agencies for their explicit and valuable support of the evaluation.

Competing interests

The authors declare that they have no financial or personal relationship(s) that may have inappropriately influenced them in writing this article.

Authors' contributions

S.K. (German Institute for Development Evaluation, DEval) and T.S. (DEval) supervised the evaluation process. All authors were involved in the development of the evaluation design. T.S., F.G. (DEval) and M.N. (DEval) applied the evaluation strategy, collected the data and conducted the analysis. M.N. drafted the manuscript, F.G. drafted the passages on the empirical example and provided further comments. S.K. and T.S. commented on and supplemented the draft. All authors read and approved the final manuscript.

References

- Abbott, P. & Rwirahira, J., 2012, 'Against the odds: Achieving the MDGs in Rwanda', Rwanda public observatory report number 3, Institute of Policy Analysis and Research-Rwanda, Kigali, Rwanda.
- Basinga, P., Gertler, P., Binagwaho, A., Soucat, A., Sturdy, J. & Vermeersch, C., 2011, 'Effect on maternal and child health services in Rwanda of payment to primary healthcare providers for performance: An impact evaluation', *The Lancet* 377, 1421–1428. [http://dx.doi.org/10.1016/S0140-6736\(11\)60177-3](http://dx.doi.org/10.1016/S0140-6736(11)60177-3)
- Binagwaho, A., Farmer, P.E., Nsanzimana, S., Karema, C., Gasana, M., Ndirabaga, J.D. et al., 2014, 'Rwanda 20 years on: Investing in life', *The Lancet* 384, 371–375. [http://dx.doi.org/10.1016/S0140-6736\(14\)60574-2](http://dx.doi.org/10.1016/S0140-6736(14)60574-2)
- Cassels, A., 1997, *A guide to sector-wide approaches for health development: Concepts, issues and working arrangements*, World Health Organization, Geneva.
- Connell, J.P. & Kubisch, A.C., 1998, 'Applying a theory of change approach to the evaluation of comprehensive community initiatives: Progress, prospects, and problems', in K.F. Anderson (ed.), *New approaches to evaluating community initiatives*, The Aspen Institute, Washington, DC.
- Delahais, T. & Toulemonde, J., 2012, 'Applying contribution analysis: Lessons from five years of practice', *Evaluation* 18, 281–293. <http://dx.doi.org/10.1177/1356389012450810>
- Deutsche Gesellschaft Für Technische Zusammenarbeit (GTZ), 2008, 'Evaluation dans le cadre de l'approche contractuelle à l'hôpital de Ruhengeri. Rapport de l'évaluation. Version finale', external evaluation commissioned by GTZ, GTZ, Eschborn.
- Eichler, R. & Levine, R., 2009, 'Money into health', in R. Eichler, R. Levine & Performance-Based Incentives Working Group (eds.), *Performance incentives for global health: Potential and pitfalls*, Center for Global Development, Washington, DC.
- Government of Rwanda (GoR), 2005, *Health sector policy*, GoR, Kigali.
- GoR & Ministry of Finance and Economic Planning (MINECOFIN), 2013, *Division of labor in Rwanda*, GoR, Kigali.
- GoR & Ministry of Health (MoH), 2008, *National health accounts Rwanda 2006 with HIV/AIDS, Malaria, and reproductive health subaccounts*, GoR/MoH, Kigali.
- GoR & MoH, 2012, *Third health sector strategic plan 2012–2018*, GoR/MoH, Kigali, Rwanda.
- 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', study commissioned and funded by the German Federal Ministry for Economic Cooperation and Development through the sector project PROFILE at Deutsche Gesellschaft für Internationale Zusammenarbeit, Bonn & Eschborn, Germany.
- Gritner, A.M., 2013, 'Result-based financing: Evidence from performance-based financing in the health sector', discussion paper 6, Deutsches Institut für Entwicklungspolitik, Bonn.
- International Health Partnership + related initiatives (IHP+), 2011. *Health in Rwanda is improving! Mid-term review (MTR) of the Rwanda second health sector strategic plan (HSSP II, July 2009 – June 2012)*, IHP+, Kigali, Rwanda.
- Ireland, M., Paul, E. & Dujardin, B., 2011, 'Can performance-based financing be used to reform health systems in developing countries?', *Bulletin of the World Health Organization* 89, 695–698. <http://dx.doi.org/10.2471/BLT.11.87379>
- Kalk, A., 2011, 'The costs of performance-based financing', *Bulletin of the World Health Organization* 89, 319. <http://dx.doi.org/10.2471/BLT.11.087247>
- Kalk, A., Paul, F.A. & Grabosch, E., 2010, 'Paying for performance in Rwanda: Does it pay off?', *Tropical Medicine and International Health* 15, 182–190. <http://dx.doi.org/10.1111/j.1365-3156.2009.02430.x>
- Kruk, M.E. & Freedman, L.P., 2008, 'Assessing health system performance in developing countries: A review of the literature', *Health Policy* 85, 263–276. <http://dx.doi.org/10.1016/j.healthpol.2007.09.003>
- Leeuw, F. & Vaessen, J., 2009, *Impact evaluations and development: NONIE guidance on impact evaluation*, The Network of Networks on Impact Evaluation, Washington.
- Leeuw, F.L., 2012, 'Linking theory-based evaluation and contribution analysis: Three problems and a few solutions', *Evaluation* 18, 348–363. <http://dx.doi.org/10.1177/1356389012452051>
- Logie, D.E., Rowson, M. & Ndagije, F., 2008, 'Innovations in Rwanda's health system: Looking to the future', *The Lancet* 372, 256–261. [http://dx.doi.org/10.1016/S0140-6736\(08\)60962-9](http://dx.doi.org/10.1016/S0140-6736(08)60962-9)
- Mayne, J., 2011, 'Contribution analysis: Addressing cause and effect', in K. Forss, M. Marra & R. Schwartz (eds.), *Evaluating the complex*, pp. 53–96, Transaction Publishers, New Brunswick, NJ.
- Mayne, J., 2012, 'Contribution analysis: Coming of age?', *Evaluation* 18, 270–280. <http://dx.doi.org/10.1177/1356389012451663>
- Oxman, A.D. & Fretheim, A., 2009, 'Can paying for results help to achieve the Millennium Development Goals? Overview of the effectiveness of results-based financing', *Journal of Evidence-Based Medicine* 2, 70–83. <http://dx.doi.org/10.1111/j.1756-5391.2009.01020.x>
- Patton, M.Q., 2011, *Essentials of utilization-focused evaluation*, Sage, Thousand Oaks, CA.
- Paul, F., 2009, 'Health worker motivation and the role of performance based finance systems in Africa: A qualitative study on health worker motivation and the Rwandan performance based finance initiative in district hospitals', London School of Economics and Political Science working paper series 08-96, Development Studies Institute, London.
- Pawson, R. & Tilley, N., 1997, *Realistic evaluation*, Sage, London.
- Peters, D.H., Paina, L. & Schleimann, F., 2013, 'Sector-wide approaches (SWAps) in health: What have we learned?', *Health Policy and Planning* 28, 884–890. <http://dx.doi.org/10.1093/heapol/czs128>
- Rusa, L., Schneiderman, M., Fritsche, G. & Musango, L., 2009, *Rwanda: Performance-based financing in the public sector*, Center for Global Development, Washington, DC.
- Shadish, W., Cook, T. & Campbell, D., 2002, *Experimental and quasi-experimental designs for generalized causal inference*, Houghton Mifflin Company, New York, NY.
- Sherry, T.B., Bauhoff, S. & Mohanan, M., 2013, 'Paying for performance in health care: Results from randomized roll-out of Rwanda's national program', Economic Research Initiatives at Duke working paper no. 136, Sanford School of Public Policy, Duke University, Durham, NC.
- Soeters, R., Musango, L. & Meessen, B., 2005, *Comparison of two output based schemes in Butare and Cyangugu provinces with two control provinces in Rwanda*. Global Partnership on Output-Based Aid/World Bank/Ministry of Health, Butare, Rwanda/Antwerp, Belgium/The Hague, the Netherlands.
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. & Befani, B., 2012, 'Broadening the range of designs and methods for impact evaluations', report of a study commissioned by the Department for International Development (DFID), working paper 38, DFID, London, UK/Glasgow, Scotland.
- Vaillancourt, D., 2009, 'Do health sector-wide approaches achieve results? Emerging evidence and lessons from six countries: Bangladesh, Ghana, Kyrgyz Republic, Malawi, Nepal, Tanzania', IEG working paper 4/2009, Independent Evaluation Group, Washington, DC.
- Walford, V., 2007, *A review of health sector wide approaches in Africa*, HILSP Institute, London.
- White, H. & Phillips, D., 2012, 'Addressing attribution of cause and effect in small n impact evaluations: Towards an integrated framework', working paper 15, International Initiative for Impact Evaluation (3ie), New Delhi, India.
- World Health Organization, 2009, *Republic of Rwanda. Health Financing Systems Review 2008. Options for universal coverage*, World Health Organization, Geneva.