

VERIFICATION IN RESULTS-BASED FINANCING FOR HEALTH

DISCUSSION PAPER

DECEMBER 2016

Petra Vergeer
Anna Heard
Erik Josephson
Lisa Fleisher



WORLD BANK GROUP
Health, Nutrition & Population

VERIFICATION IN RESULTS-BASED FINANCING FOR HEALTH

*Summary of findings and recommendations from a cross-
case analysis*

Petra Vergeer, Anna Heard, Erik Josephson, Lisa Fleisher

December 2016

Health, Nutrition and Population (HNP) Discussion Paper

This series is produced by the Health, Nutrition, and Population Global Practice of the World Bank. The papers in this series aim to provide a vehicle for publishing preliminary results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character.

The World Bank does not guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

For information regarding the HNP Discussion Paper Series, please contact the Editor, Martin Lutalo at mlutalo@worldbank.org or Erika Yanick at eyanick@worldbank.org.

RIGHTS AND PERMISSIONS

The material in this work is subject to copyright. Because The World Bank encourages dissemination of its knowledge, this work may be reproduced, in whole or in part, for noncommercial purposes as long as full attribution to this work is given.

Any queries on rights and licenses, including subsidiary rights, should be addressed to World Bank Publications, The World Bank Group, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2625; e-mail: pubrights@worldbank.org.

© 2016 The International Bank for Reconstruction and Development / The World Bank

1818 H Street, NW Washington, DC 20433

All rights reserved.

Health, Nutrition and Population (HNP) Discussion Paper

Verification in Results-Based Financing for Health *Summary of findings and recommendations from a cross-case analysis*

Petra Vergeer^a Anna Heard^b Erik Josephson^c Lisa Fleisher^d

^a Health, Nutrition and Population, World Bank, Washington DC, United States

^b Independent consultant, Washington DC, United States

^c Independent consultant, Washington DC, United States

^d Independent consultant, Washington DC, United States

Abstract: Despite the increasing popularity of Results Based Financing, there is little evidence or documentation of different verification strategies and how strategies relate to the verification results. Documentation of implementation processes including those pertaining to verification of outputs/results is lacking in World Bank-financed RBF projects in the health sector. The overall objective of this cross-case analysis is to expand knowledge about verification processes and practices to address the design and implementation needs of RBF projects. This study adds to available knowledge by comparing the characteristics of verification strategies as well as available data on costs (using level of effort as a proxy), savings, and verification results to date in six countries: Afghanistan, Argentina, Burundi, Panama, Rwanda, and the UK. These case studies were purposively selected to explore a number of factors, including: how a variety of results are verified; how the verification strategy is being implemented at different levels in the health system; and the implications of having different types of actors (that is, third-party versus internal verifiers) involved in the verification process. In this cross-case analysis, the discussion of similarities and differences in verification methods across the six cases as well as the analysis of findings is guided by a conceptual framework developed for this study. This study presents seventeen key findings, and nine recommendations.

Keywords: verification, results-based financing, performance-based financing

Disclaimer: The findings, interpretations and conclusions expressed in the paper are entirely those of the authors, and do not represent the views of the World Bank, its Executive Directors, or the countries they represent.

Correspondence Details: Petra Vergeer, World Bank, 1818 H Street NW, Washington DC, USA, pvergeer@worldbank.org

Table of Contents

RIGHTS AND PERMISSIONS	II
ACKNOWLEDGMENTS	V
PREFACE.....	VI
INTRODUCTION.....	ERROR! BOOKMARK NOT DEFINED.
RECOMMENDATION 1. VERIFICATION STRATEGIES SHOULD BE DYNAMIC, NOT STATIC, AND USE A RISK-BASED APPROACH	10
RECOMMENDATION 2. ANALYZE AND USE DATA AVAILABLE FROM VERIFICATION AND COUNTER-VERIFICATION.....	11
RECOMMENDATION 3. CONSIDER CONTEXT TO DETERMINE TO WHAT EXTENT FUNCTIONS NEED TO BE SEPARATED	12
RECOMMENDATION 4. LEARN MORE ABOUT THE BEST WAYS TO MEASURE AND PAY FOR QUALITY	13
RECOMMENDATION 5. PROTECT PATIENT CONFIDENTIALITY.....	14
RECOMMENDATION 6. CONSIDER HOW INDICATORS WILL BE VERIFIED DURING THE DESIGN PHASE ...	14
RECOMMENDATION 7. AS NEW RBF PROGRAMS EMERGE, COUNTRIES SHOULD BE AWARE OF AND UNDERSTAND THE CONSEQUENCES INVOLVED WITH DIFFERENT VERIFICATION APPROACHES.....	15
RECOMMENDATION 8. DOCUMENT RBF DESIGN DECISIONS.....	16
RECOMMENDATION 9. ADDITIONAL RESEARCH IS NEEDED	16
REFERENCES.....	17
<i>Case studies</i>	17
<i>Other literature</i>	18

ACKNOWLEDGMENTS

The content of this cross-case analysis is based on six individual country case studies. The case study on Afghanistan was co-authored by Cheryl Cashin, Lisa Fleisher and Tawab Hashemi. The case study on Argentina was co-authored by Alfredo Perazzo and Erik Josephson. The case study on Panama was co-authored by Alfredo Perazzo, Carmen Carpio and Renzo Sotomayor. The case study on Burundi was authored by Adrien Renaud and the one on Rwanda was co-authored by Adrien Renaud and Jean-Paul Semasaka. The case study on the United Kingdom was co-authored by Cheryl Cashin and Petra Vergeer. We are grateful to the case study authors as well as to colleagues in the World Bank for feedback received on earlier versions of this paper. The authors also would like to thank the World Bank for publishing this report as an HNP Discussion Paper.

PREFACE

This discussion paper is the executive summary of a cross-case analysis which draws from six individual country case studies. The full cross-case analysis will be published as either a discussion paper or book.

SUMMARY

Results-based financing (RBF) has become increasingly popular as governments with limited resources respond to pressures to improve health service outcomes and the effectiveness and efficiency of health service delivery (Witter 2012, Canavan et al. 2008). At the core of most countries' responses is an effort to increase accountability (Brinkerhoff 2004). RBF increases accountability through a system of rewards and punishments (Jack 2003, England 2000, Mills and Broomberg 1998) and results-driven approaches. Verification, defined as the first order substantiation of results paid for in RBF – whether coverage rates or quantities of patients seen, quality of services provided or patient satisfaction – is critical in assuring that these RBF rewards and sanctions are justly allotted. Verifying results can include ensuring the consistency of routinely collected data, directly observing the conditions of service delivery and care, and conducting patient surveys either at the facility or contacting patients later at home. Ensuring a country has the relevant systems in place to implement and conduct robust verification is of interest to donors and governments which are sensitive to the potential for “over-payments” based on inflated service reporting.

Despite the increasing popularity of RBF, there is little evidence or documentation of different verification strategies and how strategies relate to the verification results (Naimoli and Vergeer 2010). Documentation of implementation processes including those pertaining to verification of outputs/results is lacking in World Bank-financed RBF projects in the health sector (Brenzel et al. 2009). In demand-side schemes, such as conditional cash transfers (CCTs), verification processes are less documented than other aspects of CCT design and implementation (Fiszbein and Schady 2009). One study has explored the trade-offs between different verification strategies used in RBF programs in sub-Saharan Africa (Ergo and Paina 2012), but there has been little consideration for the influence of the context on verification strategies and limited learning of the effect of verification characteristics on data accuracy and costs.

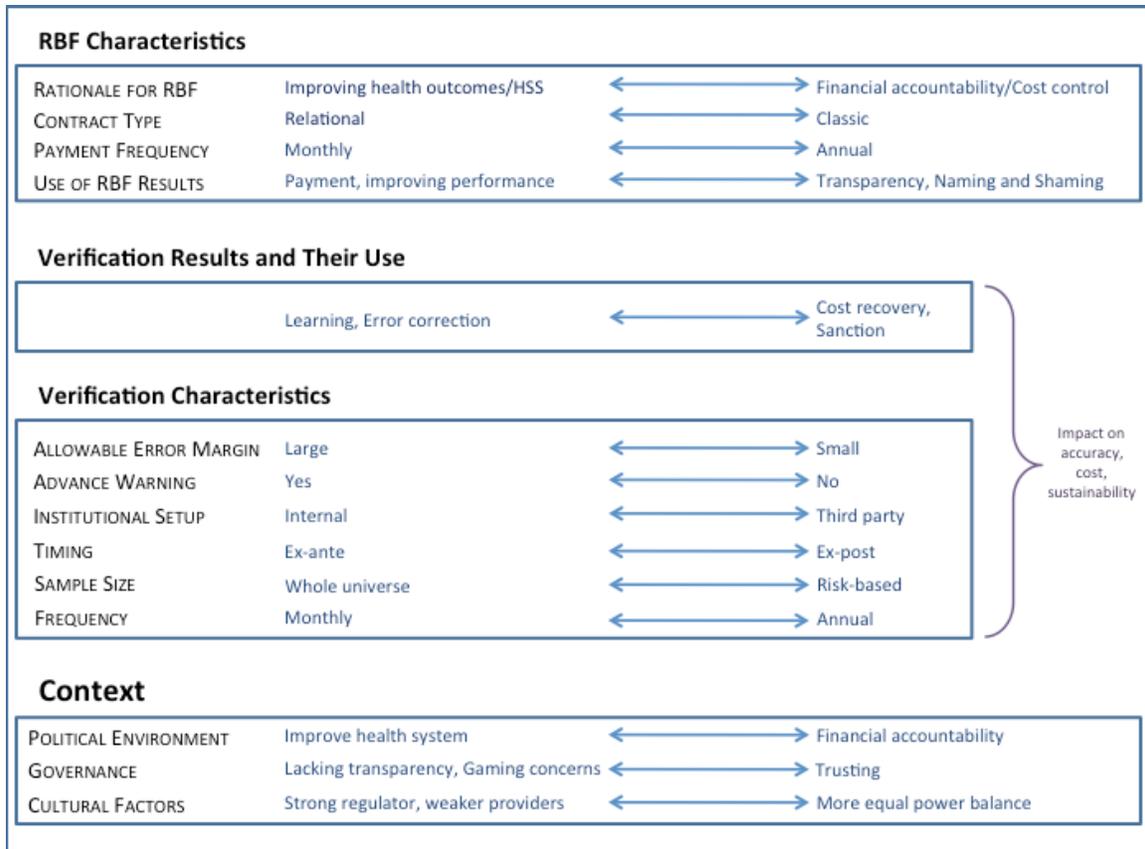
The overall objective of this cross-case analysis is to expand knowledge about verification processes and practices to address the design and implementation needs of RBF projects.

This study adds to available knowledge by comparing the characteristics of verification strategies as well as available data on costs (using level of effort as a proxy), savings, and verification results to date in six countries: Afghanistan, Argentina, Burundi, Panama, Rwanda, and the UK.¹ The six case study countries have considerable variability in context and experience with RBF. As a result, each RBF program encompasses a variety of institutional arrangements, foci, levels of implementation, and results purchased. Implementing RBF at different levels of the health system will create different principal-agent problems and the design and implementation of verification strategies are therefore highly heterogeneous across the case countries.

These case studies were purposively selected to explore a number of factors, including: how a variety of results are verified; how the verification strategy is being implemented at different levels in the health system; and the implications of having different types of actors (that is, third-party versus internal verifiers) involved in the verification process. In this cross-case analysis, the discussion of similarities and differences in verification methods across the six cases as well as the analysis of findings is guided by a conceptual framework developed for this study.

The development of the conceptual framework was informed by a review of the literature on the theory of verification and related topics. The principal-agent problem helps to explain why and how verification is used to increase information, accountability, and motivation. The literature also suggests that the incentives for accountability that are included in the verification strategy, how verification results are used, and the structure of verification, are important to ensure that agents are held accountable throughout the verification process.

¹ Individual case studies, published as six separate HNP Discussion Papers were written on each of these countries (see Annex for references, and download links)



Source: Authors

The conceptual framework uses a spectrum for each area of focus: context and RBF characteristics, verification characteristics and the use of verification results.

It is important to note that this framework is not intended to be *normative*: this paper is not suggesting that verification *should* be structured one way or another based on the context. Rather, the framework is *descriptive* with respect to how RBF characteristics, and more relevant to this paper, verification characteristics, and the use of verification results differ. This then allows for consideration of potential implications on data accuracy, cost and sustainability. As we will argue in this paper, these considerations and decisions are critical throughout the implementation of RBF and its verification strategies and the conceptual framework may be helpful for this.

Key findings from across the six case studies are:

FINDINGS RELATED TO CONTEXT AND RBF CHARACTERISTICS

Finding 1. Implementation of verification is highly dependent on context and RBF characteristics

Finding 2. Verification strategies have changed over time along the spectrum of the conceptual framework

Finding 3. Use of e-data and civil registration systems reduces error rates and cost of verification

FINDINGS RELATED TO VERIFICATION RESULTS AND THEIR USE

Finding 4. Error rates decline after the first few years of verification

Finding 5. Verification and counter-verification databases and reporting present disaggregated data at the level of the contracted party except in Burundi

Finding 6. Verification has spillover effects

Finding 7. Verification identifies both under- and –over-reporting

FINDINGS RELATED TO VERIFICATION CHARACTERISTICS

Sampling Strategy

Finding 8. Sampling strategy has consequences for cost

Finding 9. Patient tracing confirms patient existence and receipt of service; however it is time consuming

Finding 10. Indicators with high patient volume and complexity have high error rates

Allowable Error Margin

Finding 11. Errors not met with sanctions will persist

Advance Warning

Finding 12. All countries except Afghanistan give warning for verification visits

Ex Ante vs. Ex Post Verification

Finding 13. Ex ante verification is used in all cases

Institutional Setup

Finding 14. Varied levels of separation of functions are observed in the countries, influencing (counter-) verification design

Finding 15. Only one country systematically counter-verifies quality and shows high discrepancies between verification and counter-verification results

Finding 16. Patient confidentiality affects verification accuracy and in turn is affected by verification

FINDINGS RELATED TO CONSEQUENCES OF VERIFICATION CHARACTERISTICS FOR COST, ACCURACY, AND SUSTAINABILITY

Finding 17. Verification characteristics have consequences for cost, accuracy, and sustainability

Nine recommendations emerge from the cross-case findings:

RECOMMENDATION 1. VERIFICATION STRATEGIES SHOULD BE DYNAMIC, NOT STATIC, AND USE A RISK-BASED APPROACH

A verification strategy should not be static but change based on needs. **As context changes, RBF programs evolve and face challenges during implementation, the associated verification strategy (and its characteristics) also will need to change.** For example, it became evident that the rationale for RBF influences what is done with the verification results (finding 1): countries which started out more focused on health system strengthening and support, like Burundi and Rwanda, used the verification results mostly for error correction and learning, while countries that were more focused on financial accountability, like Argentina, Afghanistan and the UK, used it mostly for sanctioning or cost-recovery purposes. Over time though, countries made a lateral shift along the spectrum: with Burundi intensifying its efforts to sanction while Argentina moved towards using verification in a more supportive manner. The magnitude and direction of the shift will be determined by a country's culture and context. It is therefore important to assess periodically whether verification strategies are still aligned with contextual needs.

Verification was found to be an evolving process (finding 2). Changes were made to verification strategies in response to specific challenges that arose, such as making payment for verification visits performance-based due to delays experienced in counter-verification in Rwanda, including a stronger focus on improving quality in Burundi, and an enhanced focus on the number of services provided in Afghanistan.

The country experiences with RBF are also thought to influence verification characteristics: for example, the actor providing the funds seems to influence whether verification and/or counter-verification are carried out by internal actors or by a third party (finding 1): In the UK, verification was put in place following the desire by the government to have much more transparency and a general shift in the political climate, demanding greater financial accountability. In Burundi, the government preferred internal verification but the actors providing external funding required that third party counter-verification be implemented. In Afghanistan, third-party verification was also implemented based partly on the requirements of the funder. A robust sampling strategy was developed and implemented in Afghanistan to verify the quantity of services delivered by facilities and received by the community. Concerns about gaming and sub-optimal performance by contracted parties motivated the large sample size and frequency of patient tracing. It is recognized that such considerations may change over time, especially in view of possible learning from the contracted party and verifiers, enhancing the level of trust of the financiers in the RBF system.

Across most countries there was a pattern of error rates related to reported quantity of services and beneficiary enrollment declining after the first few years of verification implementation with some learning apparent from pilots or phased implementation (finding 4). For example, in Argentina, quick improvements (within 2 years) were seen in the number of errors reported with beneficiary enrollment from more than 20 percent to less than 1 percent, while third party verification of declarations made by provinces on tracer indicators showed a more gradual decline in the rejection rate from 25 percent in 2005 to 5 percent in 2012. The idea of a learning curve is supported by the fact that Phase 2 provinces in Argentina have initial error rates for beneficiary enrollment records (around 0.60 percent) in line with provinces that had already been in the program for a few years.

In Afghanistan, improvements were also seen, with errors in reporting in the HMIS reducing from 17 percent to around 5 percent in two years' time. Patient tracing also showed a reduction in patients that cannot be found (from 33 percent to 5 percent in two years). In Rwanda, the percentage of service indicators which were erroneous (either over- or under-reported) did not change dramatically. However, the size of the error for over-reporting declined substantially (from over 100 percent to around 7 percent) in one year's time. After two years of implementation, patient tracing confirmed the existence of almost 97 percent of the patients, all of which received the service they were referred for by the CHW (Ministry of Health, Rwanda, 2012). In Burundi, counter-verification shows less than 1 percent difference in results against those found by verification at

health center level and 31 percent difference at hospital level (the latter can in part be explained by lack of standardized registers), implying verification is effective in identifying errors and ensuring the results paid for are accurate. In view of the learning curve occurring with verification, it is useful to amend the verification strategies along the way. Piloting certain verification strategies on a smaller scale may also show some benefits as the learning can be taken on when further rolling out the RBF program, as was seen in Argentina.

A risk-based sampling approach, in which a sub-set of contracted parties (such as facilities) or indicators are selected to be verified because of their expected or observed status as outliers relative to province- or national-level averages, should be considered as long as a credible threat of verification remains for all contracted parties. A risk-based approach, as used in both Argentina and the UK, may be more cost-effective and sustainable over time (finding 8). Criteria to select providers can include those that have higher volumes of services and/or have indicator trends that differ significantly from national averages. Criteria to select indicators can include those with high volumes of services and those which are more complex to fulfill, as both were found more prone to error, (finding 10). It makes sense to also include those indicators that have more value attached to them, even though indicators with higher associated payments did not necessarily demonstrate higher rates of over-reporting in Burundi (the only country that collected information on this). All these criteria are also applicable for determining the number of patients to be traced. Interim steps to a fully risk-based approach can include reducing the sample size and/or frequency of verification, and making the acceptable level of error smaller with time. Starting out by providing advance warning of verification visits and progressing to unannounced visits is also worth considering. Should a risk-based approach be pursued, the terms of reference and contract of any third party involved should include a fixed contract with a variable component. This type of contract mechanism will enable the implementation of a risk-based approach. Lump-sum contracts, like those currently used in Argentina and Panama, promote having as low a sample as possible.

RECOMMENDATION 2. ANALYZE AND USE DATA AVAILABLE FROM VERIFICATION AND COUNTER-VERIFICATION

Verification strategies facilitated the strengthening of already existing HMIS systems in all cases (finding 6). Broadly, the availability of accurate data and use of that data improved as RBF programs matured. There is some evidence that contracted parties are now more attentive to how data registries are managed, perhaps in recognition that because incentive payments are tied to the accuracy of the data extracted from the HMIS, quality and accuracy of data entry is critical. In Burundi, the existence of the verification system showed positive effects on the completeness of the HMIS.

Verification identified both under- and over-reporting (finding 7): for example, in Burundi there was about 20 percent over-reporting and about 10 percent underreporting. In Rwanda, the verification system also resulted in a major improvement of data accuracy for paid indicators (identifying 24 percent over-reporting and 28 percent under-reporting), and it is likely that the accuracy for indicators that are not paid, which are verified during the same process, improved too. Importantly, the observed improvement in data is not restricted to low-income country cases. The UK has also experienced more accurate data in the HMIS as well as greater use of HMIS data by contracted parties.

Recognizing the cost and level of effort verification requires, it would be advisable to leverage the information gleaned from it as much as possible. In most cases, both verification and counter-verification results are used **to adjust payments to contracted parties** to correct for over- or under-reporting of quantity results above the pre-agreed error threshold. Adjustments are also made in respect to quality checklist scores in some countries (UK, Burundi and Afghanistan). In Argentina 11.5 percent of the capitation amount was reduced for adjustments related to tracer indicators following verification, and 2.2 percent for fines. In Panama the results of verification are also used for adjustments with the payment being reduced dependent on the level of coverage for each indicator (adjustments totaled 2.8 percent of the capitation amount in 2011). In addition, the

data should be used **to apply the sanctions** as outlined in the RBF manual and/or contract. Not applying these sanctions may undermine the deterrent function of verification (finding 11) consequently encouraging gaming and may also diminish the level of trust in the RBF program in general. Afghanistan, which applied all-or-nothing payment for indicators with error rates, saw a significant reduction in error rates within two years. In Burundi, an explanation for the discrepancy between declared and verified indicators (31 percent error in health centres and 38 percent in hospitals) after two years of RBF implementation may be that the erroneous indicator declarations are not sanctioned sufficiently. Applying such sanctions may be easier when verification is performed solely by a third party.

In some countries, however, very little is done with verification results (that is, error reporting). Other than reporting back to the provider, and in nearly all cases, the results are not distributed beyond the reviewed contracted party. In the UK however, the results are published. It is vital for countries to analyze and use the data from verification and counter-verification. **Reviewing error trends can help inform a risk-based approach and can help identify those providers and areas that may need more support to improve reporting and/or improve their performance.** For that reason, it is imperative that the data can be disaggregated to the level of the contracted party (finding 5) as it is otherwise difficult to analyze the extent of misreporting and identify contracted parties needing additional support or sanctions, as in Burundi.

Lastly, analyzing verification and counter-verification data is useful as it can help determine whether a country should change its verification strategy, in terms of the use of the verification findings (for example, applying sanctions) or in the verification characteristics (for example, frequency, sample size, error margin).

RECOMMENDATION 3. CONSIDER CONTEXT TO DETERMINE TO WHAT EXTENT FUNCTIONS NEED TO BE SEPARATED

As RBF payments are linked to reported results, the contracted party has an incentive to over-report in order to obtain a larger RBF payment. For that reason, a separate actor verifies the results (finding 14). In Afghanistan, the Ministry of Public Health purchases from NGOs the services that are provided in their health facility. These same NGOs subcontract (and thus are also a purchaser of the services delivered by) the health facilities, which they are responsible for. This could create a conflict of interest for the NGO as over-reporting of results would lead to increased revenue from RBF for the NGO. As was noted earlier, donors had concerns about the accuracy of the results reported by the NGOs. Having a party independent of the contracted party verify these results, as is the case in Afghanistan, helps maintain a level of trust in the system.

Maintaining a separation between the purchaser² and the contracted party is recommended in RBF, particularly in countries where competition can be created. If there is no separation, the purchaser is unlikely to end the contract in the event of disappointing results or fraudulent reporting. Creating competition can help improve health outcomes if the best party is awarded the contract to provide results.

Having a third party counter-verify the results of the verifier is critical in cases where the verifier and the contracted party are closely related. In Burundi and Rwanda, for example, all aspects are first internally verified by an actor close to the contracted party, namely the provincial verification committees and health facilities respectively, which may create an incentive to over-report. In these situations, considerations for the structure of counter-verification (frequency, independence, sampling strategy) are important to prevent potential conflicts of interest.

² Purchaser refers to the actor who enters into a contract with the contracted party and signals the payer to pay the contracted party for results.

Countries also need to carefully consider their context to determine if it is appropriate to entrust the supervision and verification functions to one actor given the experiences with verifying quality. Health authorities that are hierarchically responsible for the contracted health centers may have an incentive to find high quality while peers assessing hospital quality (such as setups where doctors from a hospital in a neighboring district perform the quality verification) would like equal treatment when they are evaluated: this may result in either under- or over-scoring as the potential exists to get even for poor ratings. Supportive supervision by internal actors can play an important role to improve RBF data and performance, without necessarily needing to incorporate the verification function. Supervision can be used in a supportive/incentivizing manner that improves results for both the purchaser and the contracted party, such as a health facility. Information collected by a third party can be used concurrently by both the purchaser and the supervisor to identify problems so that improvements can be made, thus benefitting both.

At the same time, if the purchaser is the payer as well as the verifier, it has an incentive to be overly strict in verification to save money. Third party counter-verification and independent financial audits of the purchaser are therefore needed. For beneficiary enrollment in Argentina, and for coverage and service provision in Panama, the use of an independent third party to counter-verify the results helps mediate these conflicting incentives. Particularly in a federal context like Argentina, a third party is essential to establish the level of error, and ensure trust in results by both the purchaser (national level) and contracted party (provincial level). In the UK, verification of results is carried out by the purchaser, counter-verification occurs in a small sample and independent audits of the purchaser are undertaken.

The separation of functions with a split between the purchaser and contracted party as well as the verifier can benefit the relationships between the different actors involved in RBF and prevent any (appearance of) conflict of interest. Using a third party organization for verifying results linked to quantity and quality should not automatically be equated with higher costs as local organizations can be used or established. In fact, this actor could continue to play a role if the provider payment mechanism evolves, for example into a health insurance agency whereby a purchaser-provider split is created. A third party cannot have any hierarchical links with either the purchaser or the contracted party, which does not preclude the third party from being a public institution even if the purchaser, payer and contracted party are also public institutions. Using a public institution as third party verifier can therefore be considered for sustainability purposes, but this depends in large part on governance, the independence of the civil service and the rule of law.

RECOMMENDATION 4. LEARN MORE ABOUT THE BEST WAYS TO MEASURE AND PAY FOR QUALITY

It is vital that RBF programs do not merely pay for quantity of services provided (whether through the use of targets or a fee for service mechanism) but also assure quality of services provided. Argentina and Panama's review of paper-based clinical records remains a proxy as the verification processes are unable to objectively determine whether clinical records accurately describe the care provided to the patient. This raises the question of measurement validity and whether we measure, and pay for, what we intend in RBF. The UK uses a similar process to verify adherence to clinical guidelines but given that electronic medical records are used for so many related purposes (for example, drug prescriptions), they are less likely to be prone to gaming. **Pending the introduction of reliable electronic medical records, such as those in the UK, it remains challenging to objectively measure and verify quality in a way that is both valid and reliable.**

In Burundi quality assessment through a checklist initially appears to have led to improvements in provider behavior but over time improvements may become less significant as providers are judged by the same measurement each time. In addition, opportunities for gaming may increase as the providers learn how to improve their score without making (lasting) changes. Burundi is the only country which systematically counter-verifies quality of service delivery (finding 15), and the results were substantially different from the quality scores established by internal actors during verification (systematically overestimated in 79 percent of counter-verified health centers and in 84 percent of counter-verified hospitals with a magnitude of 11 percent and 17 percent respectively). This

conclusion must be nuanced because the verification and counter-verification were carried out more than a month apart. It may also be linked to the tools used to measure quality which may lack reliability, in that they do not provide stable and consistent results when repeated over time. **Regular changes to the quality checklist and unannounced assessment visits may help mitigate risks and ensure continuous improvement of performance.**

All countries, except for Afghanistan, provide advance warning of verification visits (finding 12). The reason most often given for providing advance warning is to ensure that contracted parties or providers have their paperwork prepared and that staff is present to meet with verifiers. There is unfortunately no data available to support any hypotheses in regard to whether providing advance warning has any effect on reporting of results. This paper proposes that a study be done to test the effects of providing advance warning. The authors of this paper suspect that providing advance warning may, particularly in regard to quality of services provided, allow providers to game the system.

Measuring quality at health facilities is not necessarily sufficient as patient perceptions significantly influence people's decision-making in utilizing health care providers. Panama and Burundi both measure patient satisfaction. In the UK, patient experience data used to be collected but this was no longer the case in 2013, for reasons not known to the authors. However a study on the role of community-based organizations in verification of performance-based financing schemes in Burundi found that, while the use of community organizations helped relay messages about patient satisfaction to the providers, action was not necessarily taken to act on the information (Falisse, Meessen, Ndayishimiye and Bossuyt, 2012).

Further analytical work is required to determine how best to measure and verify paying for quality in RBF.

RECOMMENDATION 5. PROTECT PATIENT CONFIDENTIALITY

Patient confidentiality must be protected, both in electronic and paper-based systems. In the case of electronic checks of records, such as practiced in the UK, confidentiality has been assured as the records have individual-specific identity numbers while personal information, such as name and place of residence of the patient, are not available to the verifiers.

In countries where verification of receipt of services and/or quality of care can only be determined by looking at actual patient records and/or contacting the patient, patient confidentiality plays a role. Patient confidentiality was found to be both limiting to, and negatively affected by, verification (finding 16). Concerns for patient confidentiality resulted in the exclusion of certain indicators from verification: for example in Burundi where indicators related to HIV, TB and family planning clients were excluded from patient tracing for confidentiality reasons. At the same time, indicators for which services are not verified are clearly open to gaming.

In Afghanistan the community monitors, and in Burundi the local organizations (for example, local development associations, women's associations), who carry out household verification visits are drawn from the same community as those being visited. While patients are asked to provide consent at the beginning of the household visit, they are not asked to provide consent prior to being visited in the first place. In Afghanistan respondents also mentioned of the public identification of visited households. In contexts where patient tracing requires actual contact with patients, their confidentiality must be warranted.

RECOMMENDATION 6. CONSIDER HOW INDICATORS WILL BE VERIFIED DURING THE DESIGN PHASE

Quantity indicators should be simple as complex indicators are more likely to be measured with error. While there might be a desire to include comprehensive and complex indicators in RBF, having multiple compliance criteria within a single indicator (such as completely vaccinated children in Burundi, or tracking healthy children between 1 and 6 years old in Argentina) appeared to result in higher error rates. Such indicators require that the provider keep track not only of each child coming in for check-up or vaccination but also their age or the number of vaccinations for each child which may be up to five vaccinations given at different times. In settings where patient charts

are not well kept or are well kept but are paper-based, requesting that a provider keep track of complex combinations of criteria for a single indicator may be counter-productive: results will likely include errors, not helping either the provider or the purchaser.

RECOMMENDATION 7. AS NEW RBF PROGRAMS EMERGE, COUNTRIES SHOULD BE AWARE OF AND UNDERSTAND THE CONSEQUENCES INVOLVED WITH DIFFERENT VERIFICATION APPROACHES

The dynamic nature of verification inhibits the potential to define best practices or classify country programs as blueprint models for future efforts. **New RBF programs should therefore think through the various verification characteristics and their consequences to decide what is most appropriate for their context.** The conceptual framework developed for this cross-case analysis may be of use to facilitate this process. Each of the characteristics outlined in the conceptual model (allowable error margin, advance warning, independence, frequency of verification and sample size) has **implications for costs incurred, accuracy of data obtained as well as for sustainability.**

There are important justifications for allocating funds to verification. Most obviously, “savings” are achieved through countering over-reporting. In the cases reviewed such savings appeared small (finding 17): Payment adjustments were in the range of less than 1 percent to 3 percent of gross payments in Burundi, Rwanda and the UK, with Argentina a little higher at 5.9 percent (including fines). However, it is important to remember that verification may have an unobserved deterrent effect which is difficult to value. That is to say, if verification were not carried out, there would likely be substantial misreporting (given that it already occurs despite verification) and might cost significantly more in false payments and reputational damage to the RBF program than the cost of verification. Hence, in countries in which the total budget for RBF is large (Argentina, UK) there are likely to be economies of scale in verification and a consequently lower percentage of cost of verification relative to the output budget. Burundi also noted a decline from 25-30 percent to 16 percent in verification costs since the nationwide scale up. Reducing technical assistance as capacity is being built is also likely to create further reductions.

The main cost drivers of verification include personnel, transport, and technology (communication, computers). Of course these input costs for verification (for example, the salary of a verifier/assessor, transport costs) can vary widely from country to country. Countries considering implementing verification as part of new RBF programs should **think through the budget implications when deciding on the characteristics of verification strategies:** whether there is one level of verification or two (verification and counter-verification), whether verification (or counter-verification) is conducted by a third party, whether the verification is focused primarily on electronic records, whether there are field visits involved, how frequently verification (or counter-verification) is carried out, and what the size of the sample is (or whether the whole universe of facilities/households is included).

It has not been possible for this study to obtain and analyze full and complete cost information. However, information on level of effort (measured by person-hours) was calculated. **The personnel hours needed to conduct verification depends almost entirely on the sampling strategy used** (finding 7). In countries in which verification is relatively frequent (quarterly in Afghanistan, monthly in Burundi and Rwanda), where a large sample or all contracted parties are visited (all facilities verified in Burundi and all CHWs in Rwanda) or where geographical access can be problematic (Afghanistan in particular) the level of effort is much higher. In countries using a risk-based sampling strategy, such as Argentina and the UK, the level of effort will be lower as it uses smaller samples facility visits (Argentina) or infrequent facility visits (UK).

Verification of quantity was also seen to benefit significantly from electronic records and IT systems in countries like Argentina, Panama, and the UK (finding 3). Electronic records and IT systems can help to reduce possible errors and create costs savings on verification as facility visits or patient tracing (as used in Afghanistan and Burundi) can become obsolete or made less frequent. Patient tracing was found to be a time-consuming exercise: patient interviews generally last between 15 and 45 minutes, and finding patients in contexts where street names and house numbers are rare is cumbersome. Understandably, geographic conditions of a country will also

influence costs. Exploring how technologies, such as mobile phones, could be applied to verification may also lead to potential cost savings. In Panama comparing the beneficiary register to the national civil registration database ensures that the basis for capitation is accurate.

Costs should not be the only consideration. As **decisions on the verification characteristics outlined in the conceptual model (allowable error margin, advanced warning, independence, frequency of verification and sample size) also have implications for accuracy of data obtained and for sustainability**. For example, a verification approach which is more on the left-hand side of the spectrum with verification implemented by internal actors (that is, government) has led to ownership and supportive capacity building. It also has likely helped manage costs of the verification processes. However, the consequence could be compromised data accuracy, for example if there are concerns about possible conflict of interest with the verifier being too closely linked to the contracted party to objectively carry out their role, which could in turn have implications for the perceived legitimacy of the program.

All countries showed that verification strategies evolved and shifted along the spectrum of the conceptual framework. The cases did not provide evidence whether changes made to the focus of the verification strategy would be more or less accepted by stakeholders and more or less easy to implement if the initial focus is on the left-hand side (and changes over time to the right-hand side) or if the initial focus is on the right-hand side (and changes over time to the left-hand side). A country therefore may benefit from **considering whether the country's culture and context will make it easier and/or more effective to shift from the right-hand side of or vice versa**. It would also be useful to study the benefits of shifting along the spectrum.

RECOMMENDATION 8. DOCUMENT RBF DESIGN DECISIONS

The research conducted to write the six individual case studies on which this cross-case analysis is based revealed what is anecdotally well known by many researchers who seek to learn from past experiences, that is, that documentation of project design and decisions taken during initial implementation is frequently non-existent. Both the fallibility of human memory and the frequent turnover of personnel mean that this effectively results in knowledge lost. Yet it is very important to learn from implementation and to know what decisions were taken and why so that countries can avoid pitfalls experienced by others.

It is highly recommended that the logic and reasoning behind RBF design, including verification, be properly documented as well as the decisions made after the design phase which alter or amend the RBF implementation materially.

RECOMMENDATION 9. ADDITIONAL RESEARCH IS NEEDED

To date, an increasing number of rigorously-designed impact evaluations have focused on providing information about the effect of various types of RBF programs on health services in several countries. However, more needs to be learned about verification in RBF both through impact evaluations and qualitative research. Additional research is needed in the following areas: (i) the costs, savings, and cost-effectiveness of different verification and counter-verification strategies; (ii) how technologies, such as mobile phones, could be applied to verification and their potential for cost savings; (iii) how to best ensure that patient confidentiality is protected in both paper-based and electronically-oriented systems; (iv) relative ease and acceptability of shifting along the spectrum of the conceptual framework in one direction or another and its benefits of doing so; (v) how best to measure, verify, and pay for quality in RBF; (vi) the cost-effectiveness of patient tracing and potential other ways of achieving the same goals.

REFERENCES

Case studies

Cashin, C and Petra Vergeer. 2013. "Verification in Results-based Financing (RBF): The Case of United Kingdom." HNP Discussion Paper. Washington, DC. The World Bank.

<https://www.rbfhealth.org/sites/rbf/files/Verification%20in%20Results-Based%20Financing%20-%20The%20Case%20of%20the%20United%20Kingdom.pdf>

Cashin, C., Lisa Fleisher and Tawab Hashemi. 2015. "Verification of Performance in Results-based Financing (RBF): The Case of Afghanistan." HNP Discussion Paper. Washington, DC. The World Bank.

<https://www.rbfhealth.org/sites/rbf/files/Verification%20of%20Performance%20in%20RBF%20Afghanistan.pdf>

Perrazzo, A., and Erik Josephson, 2014. "Verification of Performance in Results-based Financing (RBF): The Case of Argentina." HNP Discussion Paper. Washington, DC. The World Bank.

https://www.rbfhealth.org/sites/rbf/files/Verification_Argentina_Final_March%202015.pdf

Perrazzo, A., Carmen Carpio and Renzo Sotomayor, 2015. "Verification of Performance in Results-based Financing (RBF): The Case of Panama's Health Protection for Vulnerable Populations (PSPV) Program." HNP Discussion Paper. Washington, DC. The World Bank.

<https://www.rbfhealth.org/sites/rbf/files/Verification%20Case%20Study%20Panama.pdf>

Renaud, A. 2013. "Verification of Performance in Results-based Financing (RBF): The Case of Burundi." HNP Discussion Paper. Washington, DC. The World Bank.

https://www.rbfhealth.org/sites/rbf/files/Verification%20of%20Performance%20in%20Results-Based%20Financing%20-%20The%20Case%20of%20Burundi_1.pdf

Renaud, A. and Jean-Paul Semasaka. 2014. "Verification of Performance in Results-based Financing (RBF): The Case of Community and Demand-side RBF in Rwanda." HNP Discussion Paper. Washington, DC. The World Bank.

https://www.rbfhealth.org/sites/rbf/files/Verification%20of%20Performance%20in%20RBF%20-%20The%20Case%20of%20Community%20and%20Demand-Side%20RBF%20in%20Rwanda_0.pdf

Other literature

Aidt, Toke. 2011. The causes of corruption. CESifo DICE Report 2/2011. Online. http://www.econ.cam.ac.uk/faculty/aidt/papers/web/DICE_corruption.pdf. Accessed May 5, 2013.

Brenzel L et al. 2009. Taking Stock: World Bank Experience with Results-Based Financing (RBF) For Health. Washington D.C.: The World Bank.

Brinkerhoff D. W. 2004. "Accountability and Health Systems: Toward Conceptual Clarity and Policy Relevance." *Health Policy and Planning*. 19(6): 371-379.

England R. 2000. Contracting and Performance Management in the Health Sector. London, England: DFID Health Systems Resource Centre.

Ergo, Alex and Ligia Paina. August 2012. Verification in Performance-Based Incentive Schemes. Bethesda MD: Health Systems 20/20, Abt Associates Inc.

Hood, C. 1991. Public Management for all Seasons, in *Public Administration*, Vol. 69, pp. 3-19.

Jack W. 2003. "Contracting for Health Services: An Evaluation of Recent Reforms in Nicaragua." *Health Policy and Planning*. 18(2): 195-204.

Klitgaard, Robert. 1988. Controlling Corruption. Los Angeles and Berkeley, CA: University of California Press.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*. 115(2): 200-249.

Rose-Ackerman, Susan. 2005. "The Challenge of Poor Governance and Corruption." *Revista Direitogv* Special Issue 1, November: 207-266.

Stiglitz J. 1989. Principal and Agent. In Eatwell J, Milgate M, and Newman P (eds). *The New Palgrave: Allocation, Information and Markets*. London: The Macmillan Press.

Vian, Taryn. 2007. "Review of Corruption in the Health Sector: Theory, Methods and Interventions." *Health Policy and Planning* (23): 83-94.

Despite the increasing popularity of Results Based Financing, there is little evidence or documentation of different verification strategies and how strategies relate to the verification results. Documentation of implementation processes including those pertaining to verification of outputs/results is lacking in World Bank-financed RBF projects in the health sector. The overall objective of this cross-case analysis is to expand knowledge about verification processes and practices to address the design and implementation needs of RBF projects. This study adds to available knowledge by comparing the characteristics of verification strategies as well as available data on costs (using level of effort as a proxy), savings, and verification results to date in six countries: Afghanistan, Argentina, Burundi, Panama, Rwanda, and the UK. These case studies were purposively selected to explore a number of factors, including: how a variety of results are verified; how the verification strategy is being implemented at different levels in the health system; and the implications of having different types of actors (that is, third-party versus internal verifiers) involved in the verification process. In this cross-case analysis, the discussion of similarities and differences in verification methods across the six cases as well as the analysis of findings is guided by a conceptual framework developed for this study. This study presents seventeen key findings, and nine recommendations.

ABOUT THIS SERIES:

This series is produced by the Health, Nutrition, and Population Global Practice of the World Bank. The papers in this series aim to provide a vehicle for publishing preliminary results on HNP topics to encourage discussion and debate. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations or to members of its Board of Executive Directors or the countries they represent. Citation and the use of material presented in this series should take into account this provisional character. For free copies of papers in this series please contact the individual author/s whose name appears on the paper. Enquiries about the series and submissions should be made directly to the Editor Martin Lutalo (mlutalo@worldbank.org) or HNP Advisory Service (healthpop@worldbank.org, tel 202 473-2256).

For more information, see also www.worldbank.org/hnppublications.



1818 H Street, NW
Washington, DC USA 20433

Telephone: 202 473 1000
Facsimile: 202 477 6391
Internet: www.worldbank.org
E-mail: feedback@worldbank.org