



**Implications of team building and reporting  
on the implementation of impact evaluation in a post-conflict setting**

**1. Background**

**Intervention:** The Democratic Republic of Congo (DRC) has been implementing several Results-Based Financing (RBF) schemes throughout the country. One of them, financed by the World Bank has been implemented in the Haut Katanga province, located about two thousand miles away from the capital Kinshasa, since June 2010. Compared to other areas of DRC, Haut Katanga benefits from better road infrastructure and is less afflicted by instability. The Haut Katanga intervention is subject to an impact evaluation (IE).

**Data collection:** The baseline survey for the IE was conducted in September-November 2009. It consisted in five questionnaires: household, health facility, health facility staff and exiting patients, and community questionnaires. The questionnaires and CS-Pro data entry program were based on World Bank templates, with additional modules. Baseline data cleaning and analysis were completed in June 2011, almost two years after initial data collection. The follow-up survey is planned for the end of 2012.

**2. Country Experience**

The difficult context of program and IE implementation created challenging conditions to start with. In addition, the relatively long time span between the design of the IE, the preparation and implementation of the baseline survey and its analysis posed various challenges, which were all related to team turnover and communication within the team.

**External out-of-control factors lied in challenging geopolitical and practical context:** After suffering an extremely violent war until 2003, DRC still remains under permanent threat of conflict and violence. Long lasting rivalries and armed conflicts have had several consequences that challenged the implementation of both the RBF program and the IE. On the implementation side, instability and recent elections created input shortages and delays in non-RBF and RBF payments to staff. This threatened the intervention itself and the rationale that RBF payments incentivize workers because they top off regular salaries. Therefore, it also threatened the possibility of an impact evaluation if the intervention turned out not to be the one initially intended. On the impact evaluation side, the quality and security of the roads and conflicts remaining in certain areas of the country posed serious security issues for the implementation of the baseline survey and endangered the integrity of the data collected.

**Institutional arrangements of the project created layers in the project and IE management chain:** The Haut Katanga RBF pilot intervention was embedded within the Health Sector Rehabilitation and

Support Project (HSRSP) as one of the numerous measures aiming to improve the health system in DRC. The national Program Implementation Unit (PIU) of the project was therefore only remotely involved in the design and implementation of the pilot intervention. Instead, an international Non-Governmental Organization (NGO) was contracted as the implementing agency to support the operations of the PIU in Haut Katanga. This additional institution involved in the project did not create ideal basis for coordination between the IE team, project team and Government officials.

***Internally, the initially scattered team with little presence on the ground and reporting impacted IE design and baseline survey preparation:*** At the early stages of the impact evaluation and for a significant portion of baseline data collection, the Principal Investigator (PI) of the IE team and the project Task Team Leader (TTL) were not based in DRC, until a new health specialist TTL was posted in Kinshasa in January 2010 to manage the project and the IE. Before that, international consultants were in charge of leading the sampling and randomization on the one hand, and the preparation and implementation of data collection on the other hand, with only a few missions on the ground and no contacts with the PIU and the World Bank Team in DRC. The predominance of consultants within the team and the lack of presence full time in DRC led to a concentration of their feedback and deliverables in the hands of the Task Team Leader, with no proper feedback loop within the team. Inadequate communication within the IE team, and disconnect between IE team and project team had unintended consequences. The sample from the baseline survey ended up being different from the sample used for the randomization of treatment and comparison groups, partly because the randomization was done one year after the baseline survey, and partly because the information on the sampling frame and the primary sampling units (PSUs) was not shared properly within the team. In addition, the unique identifiers used for the IE did not match those used within the program itself for RBF payments, and no document existed as to how to match the two identifiers.

***Additional communication bottlenecks arose between country authorities, project team and IE team before and during program implementation:*** Outside from the team, the communication between provincial authorities and the project and IE teams was also lacking during the initial stages of the intervention. Therefore, the project and the IE teams could not use monitoring data to address potential challenges in project implementation. As a result, while RBF payments were not disbursed properly and the comparison group benefited from higher payments than the treatment groups, the IE team did not identify this issue until a later stage in the IE.

***In the field, little presence on the ground led to insufficient training and supervision of the survey firm on fairly complicated instruments:*** Aside from the fact that no double data entry was required, whereas it is usually recommended, the IE team had well defined the terms of reference of the survey firm, and the quality of the local survey firm selected was fair given country capacity constraints. However, because of the team turnover and lack of presence on the ground, the survey firm lacked training, supervision, verification and quality control during fieldwork. This was coupled with the fact that survey instruments were numerous, relatively complex, long to administer and had not been fully tailored to the local context.

***Central data entry in an unstable country and the lack of training and oversight of the survey firm threatened data integrity:*** The IE team chose to enter data centrally in Lubumbashi. This created a significant delay between data collection and entry. In a post-conflict unstable country, with undermined road infrastructure and somewhat dangerous itineraries, teams could not go back to

facilities or households once data entry operators detected errors centrally, and the data incurred a greater risk of being lost before being transmitted to the central level. The data entry program was provided by a quality assurance firm working with the World Bank (Sistemas) and ran on CS-Pro software. At the data entry stage, the local survey firm turned out to have insufficient capacity and training on CS-Pro, and faced a lack of clarity as to who should provide data entry training and supervision. As a result, they could not necessarily address data entry issues as they came up. The absence of double data entry reinforced the lack of controls and safeguards to the quality of the data.

***Data cleaning procedures had not been properly anticipated:*** The IE team had not defined any protocol to address wrong data collection or entry beforehand. Data cleaning procedures following data entry were undefined. Team members on the ground did not benefit from sufficient training on statistical software (such as Stata) to properly address data discrepancies and document any change made to the data. This led to ad-hoc undocumented data cleaning. Consequently, the analysis team hired later on could not always understand what kind of cleaning had been done.

***Data ownership had not been defined ex-ante:*** Finally, once the raw data was available, the IE team and the survey firm encountered issues on the ownership of the data, as team members had not signed any Memorandum of Understanding on data ownership. This delayed the release of the data from the survey firm to the IE team.

***Ex-post corrective measures can mitigate the consequences of initial hands-off supervision, reporting bottlenecks and insufficient capacity building, starting with more team building:*** During and after the completion of baseline data collection, both the TTL and the PI rotated. Country counterparts and implementing agency staff also rotated during that period. This reinforced international consultants having to handle the IE on the ground with visits or stays in-country. After baseline data collection, consultants in charge of the sampling and data collection stayed on board from afar, however the new TTL of the project appointed a new PI and co-PI around April 2011 in order to analyze the baseline data and prepare for the follow-up survey.

***Appointing an evaluation coordinator based in the field full time builds cohesive communication within the IE team:*** The PI and co-PI were not based in DRC, but they appointed a research assistant (RA), based in Lubumbashi full time, to handle day to day activities of the IE and understand practical challenges to address in a post-conflict setting. The RA was the focal point of activities on the ground and gave regular feedback to IE team members off the ground. The RA was also in charge of sorting out as many past issues as possible, so the team could come up with corrective measures for the follow-up survey. For example, the RA went back to questionnaires and other documentation of what had been done during baseline data collection. She communicated with the survey firm on the difficulties faced and measures taken during data collection and entry.

***Local presence builds bridges between the IE team and the project team:*** The RA was also able to understand several challenges linked to program implementation, such as delays in RBF payments or input shortages, which ultimately would threaten the impact of the intervention if the project team remained unaware of it. Without such a presence on the ground, many issues related to program implementation would not have come up, and the impact evaluation would have become obsolete.

***Defining clear rules and corrective measures for survey implementation, data entry and cleaning ex-ante limits the chances of poor data quality at follow-up:*** Given the lack of clarity on past baseline data cleaning, the new team defined clear rules for baseline data cleaning and documented all changes made to the data. The team produced a baseline data report based on the World Bank Hub team template in July 2011. In preparation for the follow-up survey, the team revised the current questionnaires: many sections administered at baseline were not necessarily relevant for the program in DRC (for example, they included a great emphasis on consumption modules), and contributed to the complexity and length of the instruments. The questionnaires for follow-up should be more specific to the intervention and shorter. The team also conducted ex-post power calculations to assess whether the study had sufficient power at baseline to detect a change in the outcomes of interest. These calculations revealed the study might be at risk of under-power. Therefore the team came up with correctives measures for follow-up, including putting safeguards in place to increase data quality, and better targeting the outcomes to research questions.

***Focusing on communication among the IE team, and between IE team, project team and local stakeholders limits the risks of jeopardizing the IE:*** Finally and most importantly, the whole team emphasized communication between team members. The team makes a point in documenting any analysis conducted or measure taken to improve data quality at follow-up.

### **3. Lessons Learned**

- Local continued presence of IE team members is key to the success of the IE. Having at least one co-PI, evaluation coordinator or research assistant on the ground, especially during data collection activities, is an extremely valuable strategy. It fosters communication between actors on the ground (Government and research partners) and the IE team. It allows teams to understand the local context and practicality issues, to respond to those issues as they arise and therefore take corrective measures on time, including making adjustments to the IE design if the project evolves.
- A common issue encountered during impact evaluations is the lack of communication between IE and project teams. This can clearly jeopardize the validity of the IE. World Bank Task Team Leaders must make a point in bridging the information gap between both teams, and facilitate collaboration between operational and IE teams.
- Written reporting is key to address issues of team turnover, which is likely to occur during impact evaluations that take several years to be completed. Practically, teams can use web-based solutions if internet connections allow (e.g. Dropbox), or consider using cell phones for day-to-day reporting if need be. Documenting post field-work activities such as data cleaning is also strongly recommended (e.g. using do-files cataloguing each change made to the data, with systematic rules defined).
- It is important to clearly define the role of each team member. Terms of reference should include *all* activities a team member is expected to endorse, and reporting modalities to the TTL and/or the team. In addition to defining clear terms of reference, TTLs are in charge of ensuring each team member has the capacities to fulfill those terms of reference, and training them if not.
- Before fieldwork, spending a fair amount of time on understanding local context, preparing the survey, mapping sampling units, training the survey firm and building local capacity is a preparatory step worth spending the time.

During fieldwork, the team on the ground is strongly encouraged to supervise the survey firm and fieldworkers closely. Reporting error rates can be computed on a daily basis in order to take corrective measures on time if necessary. Teams could consider penalties to limit poor data collection or entry.

[www.worldbank.org/health/impactevaluationtoolkit](http://www.worldbank.org/health/impactevaluationtoolkit)

2012/03/08