G Model
EPP-910; No. of Pages 9
ARTICLE IN PRESS

Evaluation and Program Planning xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Evaluation and Program Planning

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

Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan

Valéry Ridde [a,b,*], Sylvie Goossens [c], Sahibullah Shakir [d]

[a] Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Département de Médecine Sociale et Préventive, Université de Montréal, 3875, rue Saint-Urbain, Montréal, QC, Canada H2W 1V1
[b] Institut de Recherche des Sciences de la santé (IRSS) du CNRST, Burkina Faso., Canada
[c] Médecins Sans Frontières, France
[d] Ministry of Public Health in Afghanistan, Kabul, Afghanistan

ARTICLE INFO

Article history:
Received 21 June 2008
Accepted 1 November 2010

Keywords:
Afghanistan
Participative evaluation
Consultancy
Utilization

ABSTRACT

In humanitarian and post-conflict settings, evaluations are rarely participative due to constraints such as limited time, resources or expertise. External evaluators control most of the process while stakeholders are the most consulted but are rarely participants. Yet, we believe that not all evaluators are comfortable with that sort of practice and that it is possible to involve stakeholders in the evaluation process in the field. The objective of this paper is to demonstrate, that even in a post-conflict setting such as Afghanistan in 2003, a short term consultancy allowed for the adoption of a collaborative and useful evaluation approach. After the description of the whole process in order to facilitate its reproduction by humanitarian evaluators, we discuss strengths and weaknesses of the process and highlight some positives and negatives factors affecting the use of evaluation finding.

If this is true that in general there is a "*lack of reports of participatory approaches to evaluation in health promotion*" (Springett, 2001: 91), this is more evident in post-conflict and humanitarian action settings. Despite "*a move toward* [...] *participatory evaluation approaches*" (ALNAP, 2004: 33) for the past few years, the consultation and participation component of the annual meta-evaluation of humanitarian evaluation action remain poorly rated. The percentage of satisfactory results for this component during the annual meta-evaluation done by the Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP) varied from 13% to a maximum of 20% during the last four years (203 reports from 2000 to 2004) (ALNAP, 2004). We believe that part of the explanation of this poor utilization of participatory approach in the humanitarian field, most of the time, is due to many contextual constraints as limited time, resources, or expertise. Another explanation is stated by ALNAP "*although participative approaches are widespread in the development sector, there is a lack of tools and guidance specifically oriented to the humanitarian environment*" (http://www.alnap.org/themes/participation.htm). This is why evaluations in humanitarian and post-conflict settings are mostly done through a directive approach where external evaluators are controlling most of the process. Stakeholders, and even more beneficiaries, are consulted but they are rarely participants. One of the consequences of this rigid humanitarian evaluation process is the poor level of result utilization which is validated by a recent survey (ALNAP, 2006). Nevertheless, we would like to document that this is not always the case and that it is not impossible to involve stakeholders in the evaluation process on the field. The objective of this paper is twofold, as two of the current encouraging trends in humanitarian evaluation are a move towards more participatory approaches and an increased focused on evaluation use (Feinstein & Beck, 2006). First, using an empirical case study, we will attempt to demonstrate that, even in a post-conflict setting such as Afghanistan in 2003 with the utilization of a short term consultancy of 3 weeks, it was possible to adopt a collaborative and utilization-focused evaluation approach. Case studies concerning the evaluation process and it's impact in terms of result's utilization in that particular context (post-conflict country, short-term evaluation, participative approach) are rare (Maclure, 2006). We need more empirical studies to document the range of practices in order to improve knowledge in that particular field and be able to test that approach. We believe this to be the foundation, as we have done elsewhere (Ridde, 2006a, 2006b), of a theoretical framework on participative evaluation. Second, we will draw some lessons three years after this process, particularly in terms of the use of findings.

* Corresponding author at: Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Département de Médecine Sociale et Préventive, Université de Montréal, Canada.
Tel.: +1 514 890 8000x15928; fax: +1 514 412 7108.
E-mail addresses: valery.ridde@umontreal.ca (V. Ridde), sylviegoossens@hotmail.com (S. Goossens), smohammadalam@yahoo.com (S. Shakir).

0149-7189/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.evalprogplan.2010.11.007

This is an important task due to the fact that this participative approach is clearly based on the use dimension of evaluation (Alkin, 2004; Patton & LaBossière, 2009). Making recommendations and follow up their implementation are part of the collaborative evaluator's tasks (Iriti, Bickel, & Nelson, 2005).

1. Evaluation practice in international context

According the specialists in the history of program evaluation, we are in the fourth generation of evaluation (Guba & Lincoln, 1987). After measurement, description then judgement, this fourth generation sees emerging evaluators who have to use their competences in negotiation. Contrary to evaluation in humanitarian action, this type of pluralist approach has been pioneered by NGOs to evaluate development intervention (Cracknell, 1996; Estrella, 2000; Ridde, 2003). It is true that one of the bases of this approach, which stresses active and not fictitious participation of the social actors as well as the reinforcement of their capacities, approaches the concept of "empowerment and conscientization", addressed by Paolo Friere, which underlie the majority of the interventions of these international NGOs (Ridde, 2006a, 2006b). In the field of the international public health, the eighties witnesses the development of participative approaches in evaluation (or rapid assessment procedures), borrowing much from anthropology in order to mitigate the deficiencies of a bio-medical approach extensively adopted by the health professionals (Desclaux, 1992; Ridde, Delormier, & Gaudreau, 2007; Scrimshaw & Gleason, 1992).

There is no simple or universally recognized definition of participative and collaborative evaluation and one can affirm that two streams exist (Cousins & Whitmore, 1998; Ridde, 2006a, 2006b). On one hand, there is the approach known as "*practical participatory evaluation*", which makes it possible for the organizations to make decisions or to solve particular problems. On other hand, there is the "*transformative participatory evaluation*" approach, where the participation of the actors aims at social change by empowering of people and stakeholders.

Following the creation of various organizations and national groups devoted to program evaluation, primarily in developed countries, the International Organization for Cooperation in Evaluation was launched in 2003 (Feinstein & Beck, 2006). More concretely, multiple guides were recently produced in order to support NGOs in their evaluation practice (Aubel, 1999; IFAD, 2002). Much of them are accessible for free on Internet. A myriad of Internet sites exists and thus gives access to multiple resources. The U.S. Agency for International Development (USAID) has just launched a website, so has Care International. In addition, many training sessions are proposed, like the one of the Carleton University in Canada in coordination with the World Bank's Operations Evaluation Department: the International Program for Development Evaluation Training (IPDET). This stream is not exclusively North-American. Already in 1991, the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) proposed guiding principles for the evaluation of the development aid (OCDE, 1991). These principles were largely adopted later by the European Union and its office of emergency aid, ECHO. Since 1994 in France, funds for the promotion of the preliminary studies, cross-sectional studies and evaluations (F3E), have been put in place. This body, co-managed by NGOs and the French Government, places some extremely useful methodological guidelines on the Internet (F3E, 2000). Based in London, ALNAP is an international interagency forum working since 1997 to improve learning, accountability and quality across the Humanitarian Sector. ALNAP also proposes some training and tools. Additionally, in the United Kingdom, the Humanitarian Policy Group recently published a review of the actions undertaken to evaluate the humanitarian interventions (Hofmann et al., 2004).

In addition to this evaluation practice presentation in the field of humanitarian action, we must say that the large majority of evaluations are carried out by consultants over a short period. This paper does not deal with evaluation research in post-conflict countries, since we focus on evaluation consultancy practices. If there are some initiatives to set up monitoring and evaluation systems throughout project implementation, most of the time NGOs recruit external evaluators by the means of invitation to tender. The evaluations are thus carried out by individuals or teams which spend a few days to one month on the field to judge the merit of the programs or to improve them. We know that the depth of the participation is one of the key elements of the collaborative evaluations (Cousins & Whitmore, 1998). However, to call upon external consultants who spend little time on the projects is not an ideal context to support the use of a collaborative approach. The political and sociological context in a post-conflict country is a second type of constraint for the use of an evaluation approach. Let us now explain where this collaborative context was implemented in order to not be accused to "*systematically ignoring the views and perspectives of primary stakeholders*" (Feinstein & Beck, 2006).

2. Afghanistan evaluation context

With an estimated population of about 23 millions inhabitants, Afghanistan occupies a region of striking geographic diversity in Central Asia. About 75% of Afghanistan's inhabitants are rural agriculturalists. Conflict has lasted about 24 years in parts of the country, marked by periods of heavy fighting, the loss of nearly 1.5 million lives, and the displacement of some 8 million persons. After more than 20 years of conflict and important economic decline, chances for development in Afghanistan are impaired by the worsening health condition of the population. The health system is adversely affected by major infrastructure and resource problems. In addition, the pre-war in human resource capacity has been eroded and there is scarcity of personal with managerial and technical skills throughout the country. There is also a lack of training and a lack of public health expertise for all health staff and doctors are generally not able to deal with the most urgent problems at community level. Indeed, medical facilities and personnel are very few in number and are primarily found in the Capital. In rural areas, NGOs are in charge of the large majority of the health facilities. However, access to health services remains appalling for rural populations because of limited public transport, cultural constraints that limit the access to health care for women, high illiteracy levels with lack of knowledge about health care, few hardtop and rural roads and absence of telecommunications (Ridde, Dalil, Wahidi, & Barlett, 2007). Afghanistan, qualified at the time being as a Fragile State (Palmer, Strong, Wali, & Sondorp, 2006), is not yet safe and secure, tensions still running high in most parts of the country. Moreover, there are signs of nascent problems, notably harassment of the International Community by Government authorities and the potential return to violence in some areas. Current insecurity and political instability will obviously constrain the pace and geographic scope for extending health services. Intense ethnic rivalries and local conflicts have undermined trust in public and government institutions and will remain a challenge in years to come.

3. The evaluated program

Aide Médicale Internationale (AMI), is implementing programs in Afghanistan since the early eighties, undertaking different activities focusing on the rehabilitation of Health Care Structures

Please cite this article in press as: Ridde, V., et al. Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan. *Evaluation and Program Planning* (2010), doi:10.1016/j.evalprogplan.2010.11.007

and on Medical Training for Health Care Workers. Between1985 and 1993 AMI ran a training program (Medical Training for Afghans) in Peshawar (Pakistan), and provided the 115 graduated students with medical kits to start their activities inside Afghanistan. In 1996, AMI started to support health facilities in three provinces (Kunar, Logar, and Laghman) in the Eastern Region of Afghanistan (Ridde, Bonhoure, Goossens, Shakir, & Shirzad, 2004).

The program under evaluation started on October 1st 2001 that ran for 2 years. This multi-prong health program was funded by the European Union (EU). The general objectives of this program were to improve the quality of services and to improve access to health care for the most vulnerable groups in the target areas of the project, especially women. To reach these objectives, AMI provided financial, technical and logistical support to implement the following activities in three provincial hospitals and six clinics as well as in the surrounding communities: (i) train the medical and administrative staff, (ii) supply the facilities with necessary medications and equipment to treat patients, (iii) maintain the buildings in proper conditions and add new constructions where necessary, (iv) train community health workers and organize information meetings in communities, and (v) edit, publish and distribute a quarterly distance-learning magazine.

4. Practical participative evaluation process

The evaluation run in November 2003 was the final formative evaluation of the program and was part of the general terms of the agreement signed between the NGO and the EU. The objective of this evaluation was to answer the needs in terms of lessons learned and best practices in order to improve the actual AMI intervention and to identify future strategic priorities for the reconstruction of the Afghan Health System.

Because of resource and time constraints, it was impossible to answer all the questions or cover all possible issues raised. Therefore, after some e-mail and phone discussions and negotiations between Paris (NGO Headquarters), Québec (External Evaluator) and Kabul (Program Field), evaluators and stakeholders, it was decided that the general purpose of the evaluation was the process of the AMI program implementation. A process evaluation is an evaluation of the internal dynamics of implementing organizations, their program instruments, their services delivery mechanisms, their management practices and the linkages among these.

As Patton said during the first conference of the African Evaluation Association in 1999: "No matter how rigorous the methods of data collection, design and reporting are in evaluation, if it does not get used it is a bad evaluation". This is the reason why it was proposed to adopt a mixed evaluation approach: utilization-focused and practical participative evaluation strategy. Practical participatory evaluations provide active involvement in the evaluation process of all the stakeholders of the program (Cousins & Whitmore, 1998). This does not only occur at the community level, but it is applicable at all levels, enabling people to consider the evaluation results and the way they will be used. Listening to and learning from program beneficiaries, field staff and other stakeholders to understand why a program is or is not working, is critical to make improvements, as we have done previously in Afghanistan (Ridde & Shakir, 2005). In this collaborative approach, mechanisms are developed to help program staff learn from both successes and problems encountered in implementing the activities in order to improve the program in the future. In addition, the practical participatory approach constitutes a learning experience for the program stakeholders who are involved; it reinforces their skills in program evaluation and increases their understanding of their own program strategy, its strengths and weaknesses.

The rest of the paper will document that, even in a short-term and post-conflict context, it is possible and even more useful to adopt a collaborative approach. Most of the evaluation processes used in Afghanistan were an adapted process from the guideline provided by Aubel (1999). Many others offered participatory evaluation guideline. But this guide was the most useful because it was developed and adapted for an international development context, even if it was never implemented in a post-conflict context for humanitarian action.

4.1. Step 1: Build and train an evaluation team

The participatory evaluation process began with the establishment of an evaluation team composed by 6 members and balanced in terms of gender, professional status and location of professional activity: 1 general medical coordinator (Male, Kabul), 1 nutrition coordinator (Female, Kabul), 1 medical coordinator (Male, Eastern region), 2 midwifes (Female, Eastern and Central region) and 1 administrator (Male, Eastern region).

To define concepts and basic notions in program evaluation

To explain the different types of approach in program evaluation

To describe the logic model approach of a program and to review the program logic model

To define the main types of program evaluation

To select the topics to evaluate in priority (versus current context and utilization of evaluation results) for the project

To define for each topic 2 or 3 evaluation questions

To identify the source from which the information should be collected to answer the evaluation questions

To describe and select the most appropriate data collection technique for each evaluation question

To identify data analysis techniques to analyze the collected data

To develop evaluation data collection instruments

Fig. 1. Evaluation planning workshop objectives.

Please cite this article in press as: Ridde, V., et al. Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan. *Evaluation and Program Planning* (2010), doi:10.1016/j.evalprogplan.2010.11.007

Table 1
Original vote on topics for evaluation.

	Topics/persons	Evaluators vote						
		A	B	C	D	E	F	Mean
1	Health Information System	5	5	5	5	5	3	4.67
2	Women's health	5	4	3	3	5	5	4.17
3	Health education	3	4	4	5	3	4	3.83
4	Access to care for the poorest	5	5	5	2	2	3	3.67
5	Management of health facilities	5	2	5	2	4	3	3.50
6	Curative care	2	3	3	3	5	5	3.50
7	Training	5	5	5	2	1	1	3.17
8	Nutrition	5	1	1	4	4	4	3.17
9	Drugs	5	3	3	3	3	2	3.17
10	Supervision/monitoring	3	3	4	1	3	4	3.00
11	Community health worker and traditional birth attended	4	2	4	2	2	2	2.67
12	Infrastructure	4	4	1	2	3	2	2.67
13	Sustainability/cost recovery	4	5	1	2	2	2	2.67
14	Community participation	5	3	1	1	1	3	2.33
15	Publication	3	2	1	3	2	1	2.00
16	Assessment	5	1	1	1	2	1	1.83

A three day evaluation planning workshop was then organized for this team. The main purposes of this workshop were (i) to review some concepts and basic notions of evaluation (some of the participants having already some knowledge about evaluation, the workshop represented a form of revision (e.g. Ridde, 2004)), (ii) to build consensus around the aim of the evaluation, (iii) to train and involve the evaluation team in developing the evaluation methodology, as required by the participatory evaluation strategy. The specific purposes of the planning workshop are presented in the Fig. 1.

4.2. Step 2: Evaluability assessment and selection of topics

During the evaluation planning workshop, as it is advised for the evaluability assessment stage, it was first assessed whether or not the program was ready for evaluation. Therefore, evaluators used the Logical Framework Approach (LFA), in order to ascertain whether the objectives of the program are adequately defined and its results verifiable. The evaluation team was thus trained to understand the purpose of a LFA and the place of intervention of the different existing types of evaluation in this chain. The team first reviewed the current LF of the program. Since it was decided to carry out an evaluation of the implementation process of this program, it was necessary to select the relevant fields of activity (topics) to be evaluated. However, in order for the lessons learned to be useful to improve the program in the future, it was important to choose activities that were common between the past program and the next one, financed by the EU. Therefore, the team also studied the LF of the future EU program and selected a total of 16 common activities. After that, a vote was organized regarding the importance, in terms of evaluation rather than implementation, of the 16 topics selected (Table 1).

Following the vote, a discussion was organized on the results and participants tried to reach a consensus in terms of the choice of the 6 main topics for evaluation. Different criteria were used to reach this consensus, such as the availability of data, NGO's capacity to take decisions, the time constraints, etc... Finally, the following 6 topics were selected for evaluation by the evaluation team: Women's health, Health education, Exemption schemes, Management of health facilities, Curative care and Training.

4.3. Step 3: Evaluation question topics

The six team members were then divided in three groups, each group being responsible for the development of two topics, in accordance with their ability to find data and their knowledge about the topic. For each topic, a list of evaluation questions was developed. Due to time constraints, a maximum of three questions could only be answered during the evaluation, the consultant selected the three most important (and feasible in the context) evaluation questions from this list, with the agreement of the

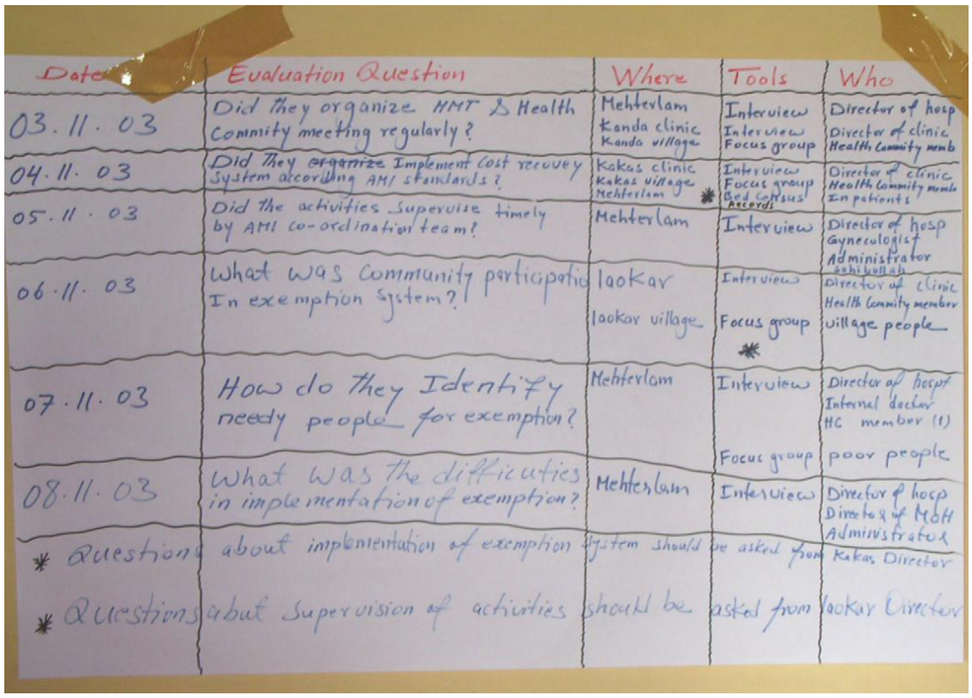

Fig. 2. Picture of the evaluation plan and evaluation question.

Please cite this article in press as: Ridde, V., et al. Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan. *Evaluation and Program Planning* (2010), doi:10.1016/j.evalprogplan.2010.11.007

evaluation team. The role of the consultant, as in other phases of the evaluation process, was both to structure the task for the group and to actively contribute to the development of evaluation questions based on insights from the fieldwork and on their own experience with other programs.

For each selected evaluation question, the group had to identify the type of data they would need to collect (quantitative or qualitative) and where and from whom they would be able to find it (Fig. 2).

The answers to these questions enabled the evaluation team to identify better practices, new ideas or lessons learned from actual program activities developed and implemented in the field that have been shown to produce actual outcomes.

4.4. Step 4: Evaluation tools and data analysis

The next step consisted in the selection of appropriate methods and the development of data gathering instruments needed to answer the questions. The validity of evaluation results depends in large part on the adequacy and reliability of the data. Hence, it is important to use different sources of data collected through mixed methods, both quantitative and qualitative methods (Pluye et al., 2009). The consultant checked all evaluation tools and worked with each group to ensure that they fulfilled standards of quality. Four types of methods were used: archival data and documentation review; focus groups discussion (FGD), unstructured interviews and observation. Contrary to some textbooks on program evaluation, the purpose of this paper is not to describe them but to illustrate that during a short period of time it is possible to use various methods and involve many stakeholders. A total of 205 people (105 woman and 100 men) had the opportunity to express their thoughts and possible concerns regarding the implementation of the programs in Afghanistan, through those four evaluation tools (Table 2).

Once the data was gathered, a participatory approach to its analysis and interpretation helped participants build a common body of knowledge. The consultant led the evaluation group in carrying out their own analysis, and was always present to ensure the quality of the analysis. As it was the first time for the evaluation team to interview people (in FGD or unstructured interviews) and analyze data, the consultant provided them a simple guideline adapted for Aubel (1999) (Table 3).

Once the analysis of each group was completed, the facilitator worked with the evaluation team to reach a consensus on findings, conclusions and recommendations. Developing a common understanding of the results, on the basis of empirical evidence, became the cornerstone for the group's commitment to an action plan. By focusing the evaluation exercise on developing the lessons learned from program implementation, the program stakeholders could analyze past problems and successes more openly.

5. Action plan and recommendations implementation

5.1. Step 5: Lessons learnt workshop and implementation process

Often, there is a lack of continuity between evaluation results and their use in program planning (Cousins, 2004; Ridde, 2007). In the case of the evaluation of an ongoing program, the recommendations should be directly applied to the program. When the incorporating of findings is left for "later", it mostly never happens. To overcome this problem, a one-day "lessons learned workshop" was organized in order to share findings and recommendations with the major stakeholders of the program in the eastern zone and to determine how to implement them in future programs. The workshop was an opportunity for the evaluation team to share current and partial knowledge regarding the program and its implementation difficulties, and for the participants (around 40) to give their own input regarding the project and thereby to correct any misunderstandings on the part of the evaluation team. At the end of the workshop, in order to start the process of development of an action plan and to increase the possibility of appropriation, the participants were divided into six work groups. Each group was instructed to develop a draft of an action plan based on the evaluation findings and lesson learned from one of the six topics and to specify, for each of the "lessons learned", what, where and when they will act and finally who will run the projected tasks.

Another important aspect of any evaluation is that the findings should be shared with all the program collaborators and transmitted back to the communities where the data were collected. Everyone involved in the program should not only be

Table 2
Evaluation tools and participants.

Evaluation topics	Interview	FGD	Observation	Document	Questionnaire
Women's health	2	3 (30[a])	1	1	
Health education	1	4 (32)	4		
Exemption schemes	6	3 (10)			
Management of health facilities	6	2 (13)			1 (32)
Curative care	3	3 (27)		1	1 (10)
Training	4	1 (8)	1		1 (21)
Total	**22**	**16 (120)**	**6**	**2**	**3 (63)**

[a] Number of participants.

Table 3
Guideline for Interviewers.

Note-taking	Data analysis
– Record notes in the first person – Record key words and ideas – Record original, descriptive phrases or sayings word-for-word as quotations – Record information exactly as it is heard and not "filter" it based on interviewers' ideas or values – Take as many notes as possible – In group interviews, record the various opinions of the group members	– Re-read the interview questions to allow the team members to recall the focus of each interview question – Read the interview notes – Discuss the information included in the notes, share other comments made by the interviewees that may not have been written down, clarify exactly what the interviewees where saying for each question – Categorize the responses in the collected information and summarize the findings in a concise fashion – Identify unclear or missing information that should be further investigated in subsequent interviews

Please cite this article in press as: Ridde, V., et al. Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan. *Evaluation and Program Planning* (2010), doi:10.1016/j.evalprogplan.2010.11.007

Table 4
Factors affecting evaluation findings use.

	Human factors	Evaluation factors	Context factors
Positives factors	– Motivated evaluation team members – Motivated medical director at the headquarters – General population openness vis-à-vis the evaluation process and NGO staff	– Practical participative evaluation approach – Evaluation team from within the program – Team training/building and Follow-up – Selection of useful evaluation question for/by stakeholders – Lessons learnt and action plan workshop – Follow-up recommendations workshop – A known consultant with negotiation and didactic skills, aware of the culture and the context – Recommendations accompanied by action plan	– Past evaluation done for the same NGO – No specific orientation for the evaluation requested by donors – Context focused program and evaluation – Known NGO and NGO staff at the local level
Negatives factors	– Staff turn over at the headquarters and in the field – Expatriate-lead program – Power relation between men/female, medical/non medical – Evaluation training depth	– Depth and regional focus of the data analysis – No recommendations options or alternative for each topics	– Staff attraction by projects concurrency – Difficulties to recruit expatriate in a post-conflict country – NGO Donors dependency vs. project – Absence of NGO evaluation policy – Contextual constraints vs. women work – News tasks due to national health policy without past NGO experiences/expertise

informed of the important lessons learned, but should also have an opportunity to discuss the results. This is why it was also suggested that a presentation of the findings should undertaken in the district were the data was collected.

After this participative process and "lessons learned" workshop, the consultant came back to Canada, and a steering committee was established in order to decide which actions should be implemented. The report writing process was also participative as each of the six topic teams has to produce his own chapter under the consultant's coordination. This committee included country director, general medical coordinator, general administrator, expatriate medical coordinator, hospital director and expatriate physician, the general medical coordinator also having the role of coordinator.

The steering committee named a total of six persons working in the field as team leaders, each one being responsible for one of the six topics of the evaluation. These six team leaders had to organize some work groups with other stakeholders on the field in order to elaborate a precise and detailed action plan and to present it to the steering committee after two months. In the middle of this period, a meeting was held with the steering committee and the team leaders in order to see if they understood the objective of their work and to ensure that their work process is exact.

When the team leaders were ready with their action plan, a half day workshop was organized for each one of them in order to share the elaborated action plan with the steering committee. The proposed actions were discussed and when needed, a second day of discussion was planned after further investigations in the field. In the light of these final workshops, the evaluation steering committee made necessary decisions in order to implement the most feasible actions.

6. Implementation of recommendations

In order to follow up the implementation of recommendations made to ameliorate the program, two workshops were organized in 2005. The first workshop, which took place in January, stressed the recommendations of the 2003 evaluation. It was found that about 40% of the recommendations were applied, 10% were in the process of being implemented and for 50% of the recommendations, none had been implemented, 14 months after the evaluation had been conducted. In addition, new recommendations had been formulated. The analysis of the results of this first workshop revealed that the failure to implement certain recommendations can be attributed to the fact that their execution was not the sole responsibility of the NGO. In other words, most of the suggestions that could be directly put in application by the NGO without having to involve other external parties were followed. If this was not the case, it is because the context was not conducive, for example the difficulty in recruiting female personnel in health centers or conducting discussion within communities, and most certainly the most difficult of all, circumstances related to the internal organizations of hospitals which the NGOs managed. Therefore, contextual elements are essential to the feasibility of the recommendations. Utilization recommendations practices are highly "context-dependent" (Patton, 2001). This observation holds true as the general context in Afghanistan changed since the end of the evaluation. During the evaluation period NGOs had a great deal of liberty in conducting their programs, currently it is no longer the case. The reconstruction of Afghanistan and the formulation of new health policies require NGOs and donors to follow clear directives and precise objectives. In addition, a certain number of recommendations were not implemented as they involved building a partnership with new institutions, which sometimes takes long to develop.

In October of 2005, a state of affairs was established which in turn allowed for the follow up of the recommendations, two years after the end of this participative evaluation. The organization of the workshop at the beginning of the year certainly enhanced dynamism around changes to undertake. Thus, almost 60% of the recommendations were implemented in October 2005, 30% are in the process of being implemented and 10% are not carried out. Major concerns remain the same. Although constraints are of an essentially systematic and contextual nature they can be related to the obligation to collaborate. The change in context incited NGOs to work in partnership with state service providers and other organizations, constraints related to negotiations to be undertaken proved important. An additional element which contributes to the difficulties in implementing recommendations is the turnover of personnel. This turnover, which is in expatriate personnel (in host countries or the NGO headquarters) or national hires, provokes difficulties in the follow up of recommendations; this will be further discussed in this paper.

7. Discussion

In a context where the evaluator only has a few days to undertake the evaluation, which is usually the norm and not the exception in evaluation of development intervention and humanitarian action, what can be learnt concerning findings use? In fact

Please cite this article in press as: Ridde, V., et al. Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan. *Evaluation and Program Planning* (2010), doi:10.1016/j.evalprogplan.2010.11.007

since our perspective is one of an evaluative process that maximizes the chances of findings use, it is this last point that we reflect upon. This reflection can be structured by the propositions of Alkin and Taut (2003) concerning findings use but we have to precise two things. Firstly, they involve reflection concerning the instrumental use of evaluation results and not their symbolic or conceptual use. Secondly, we will focus on finding use and not process use as we have already dealt elsewhere with the latter (Ridde & Shakir, 2005). We will handle the factors that seem to have contributed to the amelioration of findings use. Alkin and Taut (2003) suggest three types of factors that influence findings use: human, evaluation and contextual factors. A summary of empirical elements that seem to have favoured the findings use or non-use is presented in Table 4.

The analysis of this data and the data presented in preceding pages leads us to think that the three influence factors are interdependent. In fact, evaluative factors were favourable to findings use, but human contextual factors were certainly less favourable. In other words, if the evaluators could have some control over the evaluative process, they could only have marginal control over human and contextual factors. This constitutes the limit of evaluative exercises undertaken by external evaluators who do not necessarily have contact with the organization in question once their report is submitted. This case study demonstrates that the choice of the evaluative process, even when it is under taken with a clear and deliberate perspective favourable to findings use, cannot resolve all the difficulties related to the usefulness of evaluations. We have already demonstrated the importance of context during evaluations in a humanitarian work context (Ridde, 2003). Beyond these findings, it is useful to revisit the force and weakness of the process. A summary of process lessons learnt is provided in Table 5.

But we could first focus on the presence of two major strengths: the utilization of a gradual participative approach at the same time of a staff evaluation skills building.

7.1. Gradual participation

We have to say that the depth of participation during that evaluation must be understood as a continuum in terms of participative evaluation since the year 2001 with the same NGO. The first author of the paper, as an independent consultant, has done an evaluation each year since 2001. Implicitly, he decided to use an evaluation model which employed approaches nearer to the ideal-type of the practical type of participative model (Weaver & Cousins, 2005). The goal was to gradually reinforce competences and knowledge of the NGO stakeholders in terms of evaluation and try to favor the institutionalization of these activities. Although in the context of international development, NGOs have been first to mainly apply this type of pluralist approach (Cracknell, 1996), this NGO was not truly accustomed to such a process. The context of permanent war during more than 20 years in Afghanistan obliging the NGO to work in substitution of the Government and without much participation of the communities in decision making, is one of the explanations to the lack of use of such an approach and the lack of implementation of some recommendations. It should be noted that the implementation of the participative approach for the first time in 2001 proceeded in parallel with the will of the NGO to give a wider role to the local populations in the management of health centers. It also came at the time of first attempts to establish Health Management Communities. Also, the gradual approach with regards to participation is justified by the gradual evolution of the context passing from a situation of war with the presence of Taliban (2001) to a situation of post-conflict and rebuilding of the State (2003). In another paper we illustrate the depth of the participation in the three evaluation processes and how, gradually, we used the appropriate evaluation model according to the context and the NGO wishes (Ridde & Shakir, 2005).

7.2. Evaluation skills

We can say that this whole evaluation process done by an evaluation team from the NGO was a perfect approach to develop their evaluation skills in all the evaluation areas, from the evaluability assessment to the data analysis and action plan formulation phase (Ridde, Fournier, Banza, Tourigny, & Ouedraogo, 2009). It is also clear that skills to participate in the whole process were increased, for some, partly due to the capacity building process done over the past two years. Some of them were able in 2004 to use evaluation techniques (focus group and bed census) during an assessment of the NGO cost-recovery schemes done in some provinces of the country. According to a new contract between the NGO and the EU, the NGO was obliged in 2004 and 2006 to implement a baseline/end line survey on the health status of the population in the targeted clusters of the new project (pre-post test). The NGO recruited two different expatriates specifically for this task in 2004 and then in 2006. The first one in 2004 was, not surprisingly, one of the six members of the 2003 participatory evaluation team. This was a good opportunity for her to use some of the knowledge that she had acquired. Therefore in three other settings, the baseline survey process was in the hand of the local staff. The medical coordinator, who was one of the six team members, delivered the training for the surveyors. The good quality of the 2006 end line surveys could be understood as an interesting unintended impact of this global approach (Morell, 2010). The report of that survey and the following workshop findings (September 2006) were also perfect moments for old and new (due to the staff turn over) program stakeholders to be aware of the non-implemented 2003 recommendations.

In term of weakness, we have to note, in addition to Table 5, logistical and methodological constraints that could have undermined the validity of the findings.

The evaluation team was not prepared before the arrival of the consultant. For this reason, some time was wasted and primary data had to be collected in the field for a maximum of 6 days. Moreover, as it was a final evaluation, the NGO was obliged to wait until the end of the program, which coincided with Ramadan (holy month for Muslims) and thus a reduction of the working hours for the local staff. For security reasons, the evaluation team was obliged to focus on an evaluation only in one of the three provinces

Table 5
Summary of strength and weakness of the evaluation process.

Strength	Weakness
Focus on immediate evaluation results utilization for intended users	Limited time for and depth/rigors of analysis
Relevance of evaluation questions	Limited training of stakeholders
Increase the appropriation of evaluation results	Co-evaluator distance
Reinforce actor's competencies and motivations	Limited evaluation topic and regional focus
Arose program change through action plan workshop	Team work and "US Army" due to repatriation
NGO and program context adapted	Men/women power relation

Please cite this article in press as: Ridde, V., et al. Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan. *Evaluation and Program Planning* (2010), doi:10.1016/j.evalprogplan.2010.11.007

where the program was implemented. This constraint limited the external validity of the findings and recommendations for the whole NGO program. The lessons learnt during this evaluation where more formulated in terms of "better practice" according to the context than "best practices", following the critique done by Patton (2001). In addition, for security reasons ("team work and US army" in Table 5), the Country Director decided to ask the consultant and the other expatriate team members to leave the province before the end of the data collection and analysis stage. Security consideration is one the major constraints in humanitarian evaluation (Feinstein & Beck, 2006). Therefore, the group dynamic was broken during a total of 4 days, representing some 30% of the total evaluation time.

During the data collection, it was sometimes difficult to ask the head of village or the director of the health facility not to be involved in focus groups. However, his presence had probably an impact on the capacity of the other participants to give free answers to the questions asked by the evaluation team. But as we have seen, we used more than just one data collection method and we collected information from a broad perspective, trying to increase the triangulation of data and perspective of the evaluation. At the time of the evaluation, the NGO had some funding difficulties; most of the staff of the health facilities were not being paid anymore and some of them were aware that the NGO was no longer going to support their health facility. This constraint could have been problematic in terms of staff willingness to have a discussion or to give unbiased or honest answers. Moreover, some evaluators were involved in the program under evaluation. This could have had an impact on their "objectivity", but each team was organized in a balanced way in order to compensate this likelihood, the consultant was involved in asking questions and insuring that findings were consistent and most importantly, findings were discussed and validated by stakeholders during the final workshop.

8. Conclusion

This case study show that using a participative approach in a humanitarian and a post-conflict context over a short term is feasible, such as in "stable countries" like Burkina Faso, Senegal or Kenya (Aubel, 1999; Holte-McKenzie, Forde, & Theobald, 2006). The process was documented in accordance with the needs updated by the ALNAP *tools and guidance,* in order for consultants to undertake it. It should be mentioned that the selection of the people participating in the process was undertaken from a practical perspective ("intended use by intended users" (Patton, 2008)) and not an empowerment perspective. This is why affected populations were consulted but did not participate. This could be the next humanitarian challenge. We equally learn that a participative approach is primordial to the utilization of the evaluation results. This case study seems to be an empirical confirmation of the fact that participation and leadership are the two use-promoting factors most frequently mentioned by independent/external humanitarian evaluators (ALNAP, 2006). Nonetheless, we have demonstrated that a participative approach is necessary but not sufficient. If the context and the orientations of the potential users are not favourable to the implementation of the recommendations, the users do not necessarily win. In this case, we are obligated to launch the debate concerning the role of evaluators. Till what lengths should evaluators go in following up the recommendations resulting from evaluations they have undertaken? Should they limit themselves to making recommendations, suggest action plans, or are they uniquely "rigorous result producing machines"? Three quarters of humanitarian evaluators are rarely or never involved in structured post-evaluation activities (ALNAP, 2006). Therefore it is the social responsibility of the humanitarian action evaluators that should be revisited (Ridde, 2007, 2009).

Acknowledgements

Thanks is expressed to everyone met in Afghanistan, specifically the evaluation team members and AMI staff: F. Raouf, Z. Arian, A. Zaher, M. Shirzad, T. Bhatia Theuss, S. Robin, M. Bévalot. The evaluation was funded by the European Union (DG Relex). Thanks to Lara Arjan for the translation support. V. Ridde is a *Canadian Institutes for Health Research* (CIHR) *New Investigator*.

References

Alkin, M. C. (Ed.). (2004). *Evaluation Roots. Tracing Theorists'Views and Influences*. Thousand Oaks: Sage.
Alkin, M. C., & Taut, S. M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation, 29*, 1–12.
ALNAP. (2004). *ALNAP review of humanitarian action in 2003. Field level learning*. London: Overseas Developpement Institute.
ALNAP (2006). *Results from survey questions for independent/external evaluators*. London: ALNAP. http://www.alnap.org/pdfs/external_evaluators_survey_results.pdf Accessed 24.10.06.
Aubel, J. (1999). *Participatory program evaluation manual. involving program stakeholders in the evaluation process*. Calverton, Maryland: Child Survival Technical Support Project and Catholic Relief Services.
Cousins, J. B. (2004). Commentary: minimizing evaluation misuse as principled practice. *American Journal of Evaluation, 25*(3), 391–397.
Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. In Whitmore, E. (Ed.). *Understanding and practicing participatory evaluation* (Vol. 80, pp. 5–23). Jossey-Bass Publishers.
Cracknell, B. E. (1996). *Evaluating development aid. issues, problems and solutions*. New Delhi, Thousand Oaks, London: Sage Publications.
Desclaux, A. (1992). Le "RAP" et les méthodes anthropologiques rapides en santé publique. *Cahier Santé, 2*, 300–306.
Estrella, M. (2000). *Learning from change: Issues and experiences in participatory monitoring and evaluation*. Ottawa: Intermediate Technology Publications, International Development Research Centre.
F3E (2000). Le Guide Evaluation "L'évaluation, un outil au service de l'action". http://f3e.asso.fr/methodo/guides.htm.
Feinstein, O., & Beck, T. (2006). Evaluation of development interventions and humanitarian action. In I. F. Shaw, J. C. Greene, & M. M. Mark (Eds.), *Handbook of evaluation. policies, programs and practices* (pp. 536–557). London, Thousand Oaks, New Delhi: Sage Publications Ltd.
Guba, E. G., & Lincoln, Y. A. (1987). The countenances of fourth-generation evaluation: Description, judgment and negociation. In Palumbo, D. J. (Ed.). *The politics of program evaluation* (Vol. 15, pp. 202–234). Newbury Park, Beverly Hills, London: Sage Publications.
Hofmann, C.-A., Roberts, L., Shoham, J., & Harvey, P. (2004). *Measuring the impact of humanitarian aid. A review of current practice*. London: Humanitarian Policy Group HPG Research Report No 17.
Holte-McKenzie, M., Forde, S., & Theobald, S. (2006). Development of a participatory monitoring and evaluation strategy. *Evaluation and Program Planning, 29*, 365–376.
IFAD. (2002). *Managing for impact in rural development. a guide for project M&E*. Roma: International Fund for Agricultural Development.
Iriti, J. E., Bickel, W. E., & Nelson, C. A. (2005). Using Recommendations in Evaluation. A Decision-Making Framework for Evaluators. *American Journal of Evaluation, 26*(4), 464–479.
Maclure, R. (2006). Pragmatism or transformation? Participatory evaluation of a humanitarian education project in Sierra Leone. *Canadian Journal of Program Evaluation, 21*(1), 107–129.
Morell, J. (2010). *Evaluating in the face of uncertainty*. New York: Guilford Press.
OCDE. (1991). *Development assistance committee. Principles for evaluation of development assistance*. Paris: OCDE/DAC.
Palmer, N., Strong, L., Wali, A., & Sondorp, E. (2006). Contracting out health services in fragile states. *British Medical Journal, 332*(7543), 718–721.
Patton, M. Q. (2001). Evaluation, knowledge management, best practices, and high quality lessons learned. *American Journal of Evaluation, 22*(3), 329–336.
Patton, M. Q. (2008). *Utilization – Focused evaluation* (4th ed.). Thousand Oaks, London, New Delhi: Sage Publications.
Patton, M. Q., & LaBossière, F. (2009). L'évaluation axée sur l'utilisation. In V. Ridde & C. Dagenais (Eds.), *Approches et pratiques de l'évaluation de programme* (pp. 143–158). Montréal: Presses de l'Université de Montréal.
Pluye, P., Nadeau, L., Gagnon, M.-P., Grad, R., Johnson-Lafleur, J., & Griffiths, F. (2009). Les méthodes mixtes pour l'évaluation des programmes. In V. Ridde & C. Dagenais (Eds.), *Approches et pratiques en évaluation de programme* (pp. 123–141). Montréal: Presses de l'Université de Montréal.
Ridde, V. (2003). L'expérience d'une démarche pluraliste dans un pays en guerre: l'Afghanistan. *Canadian Journal of Program Evaluation, 18*(1), 25–48.
Ridde, V. (2004). Seeds against malnutrition in Afghanistan: An experience in participative performance evaluation training. In S. Mathison (Ed.), *Encyclopedia of evaluation* (pp. 433–434). Thousand Oaks: Sage Publication.

Please cite this article in press as: Ridde, V., et al. Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan. *Evaluation and Program Planning* (2010), doi:10.1016/j.evalprogplan.2010.11.007

Ridde, V. (2006a). Introduction to the thematic segment of 11 articles. Programmes communautaires et innovations méthodologiques: Participation, accompagnement et empowerment. *Canadian Journal of Program Evaluation, 21*(3), 133–136 [Special Issue].

Ridde, V. (2006b). Suggestions d'amélioration d'un cadre conceptuel de l'évaluation participative. *Canadian Journal of Program Evaluation, 21*(2), 1–23.

Ridde, V. (2007). Are program evaluators judges and/or knowledge brokers? *Journal of Epidemiology and Community Health, 61*, 1020.

Ridde, V. (2009). Knowledge transfer and the university system's functioning: Need for change. *Global Health Promotion, 16*(3), 3–5.

Ridde, V., & Shakir, S. (2005). Evaluation capacity building and humanitarian organization. *Journal of Multidisciplinary Evaluation, 3*, 78–112.

Ridde, V., Bonhoure, P., Goossens, S., Shakir, S., & Shirzad, M. (2004). User fees and hospital health care utilization in Afghanistan: Lessons learned from a participative evaluation. *Journal of Afghanistan Studies, 2*, 93–109.

Ridde, V., Dalil, S., Wahidi, S., & Barlett, L. (2007a). Health promotion in Afghanistan. In M. O'Neill, A. Pederson, I. Rootman, & S. Dupéré (Eds.), *Health promotion in Canada: Critical perspectives* (pp. 259–260). Toronto: Canadian Scholars Press Inc.

Ridde, V., Delormier, T., & Gaudreau, L. (2007b). Evaluation of empowerment and effectiveness: universal concepts? In McQueen, D. V., & Jones, C. M. (Eds.), *Global perspectives on health promotion effectiveness* (Vol. I, pp. 389–403). New York: Springer Science & Business Media.

Ridde, V., Fournier, P., Banza, B., Tourigny, C., & Ouedraogo, D. (2009). Programme evaluation training for health professionals in francophone Africa: Process, competence acquisition and use. *Human Resource for Health, 7*(1), 3 doi:10.1186/1478-4491-7-3 [January 15].

Scrimshaw, N. S., & Gleason, G. R. (1992). *Rapid assessment procedures, qualitative methodologies for planning and evaluation of health related programmes*. Boston: International Nutrition Foundation for Developing Countries.

Springett, J. (2001). Participatory approaches to evaluation in health promotion. In I. Rootman, M. Goodstadt, B. Hyndman, D. V. McQueen, L. Potvin, J. Springett, & E. Ziglio (Eds.), *Evaluation in health promotion: Principles and perspectives* (pp. 83–105). WHO Regional Publications. European Series, No. 92.

Weaver, L., & Cousins, J. B. (2005). Unpacking the participatory process. *Journal of Multidisciplinary Evaluation, 1*, 19–40.

Valéry Ridde Ph.D., has been doing research and evaluation on equity and access to healthcare systems in Africa for more than 10 years. He is a researcher in global health and program evaluation at the University of Montreal/CRCHUM (Canada) and an associate researcher at the Institut de Recherche en Science de la Santé of the CNRST in Burkina Faso. His research portfolio is organized in Mali, Niger, Burkina Faso and Benin.

Please cite this article in press as: Ridde, V., et al. Short-term consultancy and collaborative evaluation in a post-conflict and humanitarian setting: Lessons from Afghanistan. *Evaluation and Program Planning* (2010), doi:10.1016/j.evalprogplan.2010.11.007