2. METHODOLOGY

Approach : a useful and participative evaluation :

For this evaluation, we decided to use a mixed evaluation approached: utilization-focussed and participative.

One of the most well-known evaluator in the world, M.Q. Patton[5], said: "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 why we propose to adopt the Utilization-focused Evaluation strategy for the review.

The objective of our mandate is to answer the needs identified by the ToR in terms of lessons learned and best practices (I prefer "better" practices) to improve the actual AMI intervention and to identify future strategic priorities in the future. To attain this objective we will use a participative approach. Participatory evaluation provides for active involvement in the evaluation process of those with a stake in the program. Listening to and learning from program beneficiaries, field staff, and other stakeholders who know why a program is or is not working is critical to making improvements[6].

During all phases of the process, the role of the consultant is be sure that the four evaluation standards are respected: utility, feasibility, propriety and accuracy[7]. In this participative approach, mechanisms are developed to help program staff learn from both the successes and problems encountered in implementing the activities in order to improve the program in future. The evaluation coordinator's role is not only to structure and facilitate each step in the evaluation process but also to contribute as a full member of the evaluation team. In addition, the 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[8].

Evaluation aim : a process evaluation

During the first working meeting with the headquarters manager in Paris and then with the medical coordinator for the Afghan Program in Kabul, we clarified the mandate (in accordance with the Statement of Work, see appendices) in detail and tried to reach a better understanding of how the evaluation results would be used. What purpose will they serve? Will it be possible to implement changes in the AMI program following the evaluation? Who, how and in how much time? This, in view of the fact that standardized recipe approaches had not worked.

Through interaction between Paris and Kabul, evaluators and stakeholders were able to negotiate the question of evaluation (see next section) according to the fourth-generation evaluation

[5] Patton, M.Q., Utilization-Focused Evaluation. 3rd ed. 1997, Thousand Oaks-London-New Delhi: Sage Publications. 431p

[6] L'expérience d'une démarche pluraliste dans un pays en guerre : l'Afghanistan, *Canadian Journal of Program Evaluation*, 2003, vol 18, n°1, 25-48

[7] Joint Committee, 1994

[8] The evaluation process in Afghanistan is an adapted process from : Aubel, J., 1999, Participatory Program Evaluation Manual, Involving Program Stakeholders in the Evaluation Process Second Edition, Child Survival Technical Support Project-Catholic Relief Services-USAID

approach[9]. Due to resource and time constraints, it was impossible for any one evaluation to answer everyone's questions or attend to all possible issues raised. However, in a Utilization-focused Evaluation the stakeholders can participate. This does not only occur at the community level, but it is applicable at all levels, enabling people to reflect on the evaluation results and how they will be used. Therefore, after this first discussion and negotiation, 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 service delivery mechanisms, their management practices and the linkages among these[10].

An evaluation team: to be participative

The participatory evaluation process began with an evaluation planning workshop held in Kabul from 29th to 31st October (see photos). As the team was not prepared in advance as we had initially planned for in the evaluation proposal (sent two weeks prior to our departure[11]), we were delayed a couple of days in Kabul before to being able to organize it. In the present Afghani case, we established an evaluation team composed of the following 6 people: general medical coordinator, nutrition coordinator, medical coordinator for the Eastern Zone, two midwives and an administrator. This evaluation team was balanced in terms of gender, location and professional status.

Name	Date of birth	Gender	Profession	Place of duty
Fauzia Raouf	1953	F	Midwife	Metherlam Hospital
Sahibullah Shakir	1966	M	General medical coordinator	Kabul AMI Office
Abdul Zaher	1958	M	Administrator	Metherlam AMI Office
Mujeeburrahman Shirzad	1974	M	Physician (Eastern zone coordinator)	Eastern zone AMI Office
Zermina Arian		F	Health Educator	Jubul Saraj Clinic (Parwan)
Sylvie Goossens	1975	F	Physician (nutrition coordinator)	Kabul AMI Office

Table 1 : List of evaluation team members

The purpose of the first workshop was to build consensus around the aim of the evaluation; to refine the scope of work and clarify roles and responsibilities of the evaluation team and facilitator; to review the schedule, logistical arrangements, and agenda; and to train participants in basic data collection and analysis. Assisted by the facilitator, participants identified the evaluation questions they wanted answered. Participants then selected appropriate methods and developed data-gathering instruments and analysis plans needed to answer the questions. Some of the participants already had some knowledge on evaluation and for them this workshop represented a form of revision[12].

[9] Guba, E. G., & Lincoln, Y. A. (1987). *The countenances of fourth-generation evaluation : description, judgment and negociation*. In D. J. Palumbo (Ed.), The politics of program evaluation (Vol. 15, pp. 202-234). Newbury Park, Beverly Hills, London: Sage Publications.

[10] Contandriopoulos, A.-P., Champagne, F., Denis, J.-L., & Avargues, M.-C. (2000). L'évaluation dans le domaine de la santé : concepts et méthodes. *Revue d'épidémiologie et de Santé Publique, 48*, 517-539.

[11] Ridde, V, Final Evaluation Proposal, Draft, October 15, 2004, 7p.

[12] Ridde, V, Seeds against malnutrition in Afghanistan: an experience in participative performance evaluation training, in *Evaluation Encyclopaedia*, Mathison, S., 2004, Sage Pub, In press

The evaluation planning workshop was held in Kabul over three days and the aims were set out as follows :

General aim:
To involve project stakeholders in developing the evaluation methodology

Specific aims:
1) To define concepts and basic notions in evaluation
2) To explain the different types of approach in program evaluation (i.e participative and utilization focussed evaluation)
3) To describe the logic model approach
4) To review the AMI/DGrelex logic model
5) To define the main types of program evaluation
6) To define the evaluation questions for the AMI/DGrelex project vs current context and utilization of evaluation results
7) To identify from whom/what source information should be collected for each evaluation question (i.e selection of case studies and people)
8) To describe data collection techniques which can be used in health and nutrition projects
9) To identify the most appropriate data collection technique/s for each evaluation question and analysis techniques
10) To develop evaluation data collection instruments

Figure 1 : Evaluation planning workshop aims

As the evaluation was implemented during Ramadan, lack of time meant that the set of evaluation tools could not be fully developed in Kabul, so an extra day had to be organized in the field (Metherlam).

Assessment prior to evaluation and selection of topics

During this workshop we assess whether or not the AMI program is ready for evaluation. Evaluators have a means of deciding whether a program is ready for evaluation. During the assessment, calls for the early evaluation are made, in collaboration with people working on the programs, in order to ascertain whether its objectives are adequately defined and its results verifiable. To do this assessment evaluators used the Logical Framework Approach (LFA[13]).

LFA is an analytical, presentational and management tool which can help us to : analyze the current situation during project preparation, establish a logical hierarchy of means by which objectives will be reached, identify potential risks, establish how outputs and outcomes might best be monitored and evaluated, present a summary of the project in a standard format; and monitor and review projects during implementation. The evaluation team was trained to understand the purpose of a LF and what the different types of evaluation are with the help of the following model.

[13] Sartorius, R.H., The logical framework approach to project design and management. *Evaluation practice*, 1991. 12(2): p. 139-147.

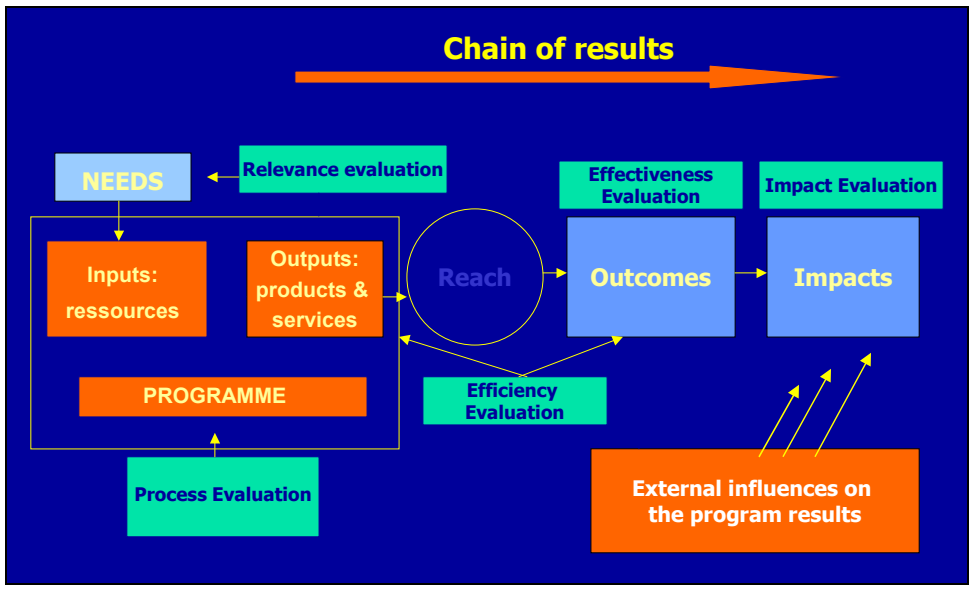

Figure 2 : A logic model and different types of evaluation

The evaluation team first reviewed the current LF of the AMI/DG relex program. For most of the team, it was the first time that they saw the LF with its activities and objectives.

After this, it was necessary for the evaluators to study the LF of the next program financed by the European Union. Indeed, since we had decided to carry out an evaluation of the implementation process of the program, it was necessary to select the relevant fields of activity (Topics) to be evaluated.

So that the lessons learned be useful to improve the program developed in the following months, it was necessary to choose some common activities. For example, the problem of the TFC was not selected because it is clear that AMI will have no program in this area. The consultant had selected some topics. The evaluation team first checked that there were no missing topics regarding the LF of the AMI/DGrelex program. Then, we organized a vote regarding the importance (1= not important, 5= very important), in terms of evaluation (rather than implementation) of the 16 topics selected.

The original results are the following:

	TOPICS	A	B	C	D	E	F	Mean
1	HIS	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	CHW/TBA	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**

Table 2 : Original vote on topics for evaluation

Following the vote, we organized a discussion on the results and tried to reach a consensus in terms of the topics for evaluation. Different criteria were used to reach this consensus, such as: the availability of data, AMI's capacity to take decisions, the time constraint etc. This is why we decided to forego the HIS topic since AMI is obliged to follow the national policy on that. We also decided, as requested by some participants, to reformulate the topic regarding access to care for the poorest.

The following 6 topics were selected for evaluation by the three evaluation groups. People were grouped in virtue of their ability to find data and their knowledge about the topic. For example, it was impossible to assign men to the women's health evaluation topic.

Women's health	**Fauzia**
Health education	**Zermina**
Exemption schemes	**Zaher**
Management of health facilities	**Mujib**
Curative care	**Shahibullah**
Training	**Sylvie**

Table 3 : Evaluation groups and topics

Program/Topics evaluation question

A process evaluation is an evaluation of the internal dynamics of implementing organizations, their program instruments, their service delivery mechanisms, their management practices, and the linkages among these. Each evaluation group developed a number of evaluation questions for each topic. A maximum of three questions could be answered during the evaluation but each team could start by choosing more than three. Then, the consultant selected the three most important (or feasible) questions and the evaluation team agreed on the choice. The role of the consultant, as in other phases of the evaluation process, is 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 evaluation question, the group had to identify what kind of data they would need to collect (quantitative/qualitative) and where and from whom they would be able to find it.

TOPICS	EVALUATION QUESTIONS	QUANT/QUALI	WHERE/WHOM
Women's health	•Were deliveries carried out in proper conditions in hospital?	Quantitative	Hospital, Midwife, Women,
	•Did the pregnant women go for ANC/PNC to the HF and TBA?	Qualitative/ quantitative	Registration book, clinic, hospital, TBA report, Midwife, pregnant women, ANC card, AMI health passport, TBA
	•Did the women use FP?	Qualitative/ quantitative	Hospital, midwife, clinic, director of clinic, village, women, register book, FP card, gynecologist
Health education	•Was the HE session organized properly?	Qualitative/ quantitative	patients, hospital, clinics, Health educators, director of clinic, nurses, registration book, documents
	•Were the topics chosen according to the time of year?	Qualitative/ quantitative	patients, registration books, clinic, hospital, village, health educator, documents
Exemption schemes	•What was the degree of community participation in the system ?	Qualitative	Hospital, clinics, villages; member of HC, directors HF, villagers, poorest people
	•How were people who qualify for exemption identified?	Qualitative	Hospital, clinics, villages; doctors, members of HC, HF staff, poorest people
	•What were the difficulties in implementing the schemes?	Qualitative	Hospital, clinics, medical staff, director HF, director MoH
Management of health facilities	•Were HMT/HC meetings regularly organized?	Quantitative/ Qualitative	Hospital, clinics, villages, director oh HF, HMT/HC members, meetings registration book, reports
	• Was cost recovery implemented according to AMI standards	Quantitative/ Qualitative	Hospital, clinics, registration book, ticket book, pharmacy register, prescription, income generation sheet, medical staff, director of HF, director of MOH, beneficiaries
	•Were the activities regularly supervised by the AMI coordination team	Quantitative/ Qualitative	clinics, hospital, staff, monthly reports, general medical coordinator
Curative care	•Did the HW follow AMI's treatment protocol?	Qualitative	hospital, clinic, doctor, nurses, midwives, library, drug survey
	•Did AMI have a proper system for follow-up of patients?	Quantitative/ Qualitative	hospital, clinics, villages, doctors, midwives, patients, villagers, documents
	•Were patients satisfied with the curative care services?	qualitative	hospital, clinics, patients, villagers
Training	•Did AMI have a relevant program (curriculum, methodology, plan) for training of each HW category?	Quantitative/ Qualitative	hospital, clinics, coordination team, HW, documents
	•Did the HW use the archives of training session?	Qualitative	hospital, HW, library manager, documents

Table 4 : Evaluation question, type of data and place/person

The answers to these questions enable the consultant and the evaluation team to identify better practices, in this specific context only, to improve the AMI program. This means the identification of new ideas or lessons learned about effective program activities developed and implemented in the field that have been shown to produce positive outcomes.

Method strategy

Evaluation strategy : case studies

In the evaluation itself, due to time, logistics and safety[14] constraints, we studied only one (1) in-depth case. Case studies are particularly useful for understanding a program in depth[15]. These case-studies of AMI interventions were selected in the field in terms of location, in collaboration with the stakeholders and bearing in mind safety constraints. Cases were also selected in terms of their ability to help us answer the evaluation questions. Cases were rich in information, in the sense that a great deal could be learned from these examples of AMI interventions. This is why we decided to organize this evaluation in the Laghman province where AMI has been working since 1996 and supports one hospital and three clinics. During the case studies, the consultant and stakeholders used multiple sources of evidence as outlined below.

Evaluation tools :

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 quantitative as well as qualitative methods. Quantitative methods are useful for getting broad descriptions of a situation, how it has changed or measured impacts. Qualitative methods are useful for understanding the reasons for events described in an evaluation. For the final evaluation we chose to focus on both methods. Through the use of simple data collection and analysis techniques all the program staff were actively involved and had to develop basic data collection skills. The consultant checked all evaluation tools and worked with each group to ensure that they fulfilled standards of quality. To answer the evaluation questions the evaluation team used the following evaluation tools :

Archival Data and Documentation Review
Archival data already exists. This data is usually inexpensive and may be fairly easy to obtain. However, we had little choice in the data format since it had previously been collected by someone else for other purposes. Existing records from different AMI departments were used as a data source. Record reviews usually involve counting the frequency of different operations, programs... In this category, we studied program proposals, monthly reports, evaluation reports, accounting reports etc.

Focus groups
Focus groups are typically used for collecting background information on a subject, creating new ideas and hypotheses, assessing how a program is working, or helping to interpret the results from other data sources. The focus group interview generally involves 6 to 12 individuals who discuss a particular topic under the direction of a moderator, who promotes interaction and assures that the discussion, remain on the topic of interest (see photo). Focus groups can provide a quick and inexpensive way to collect information from a group (as opposed to a one-on-one interview), allow for clarification of responses, obtain more in-depth information, and create easy-to-understand results. However, since focus groups use only a small number of people, they may not accurately represent the larger population.

Unstructured Interviews
Similar to a focus group, but with just one person, an unstructured interview is designed to obtain very rich and detailed information by using a set of open-ended questions (see photo). The interviewer guides the participant through the questions, but allows the conversation to flow

[14] We were allowed to travel in Laghman province from Kakass clinic up to Metherlam hospital only
[15] Yin, R. K. (1994), Case Study Research Design and Method. London, New Delhi, Sage Publications

naturally, encouraging the participant to answer in his or her own words. The interviewer often will ask follow-up questions to clarify responses or get more information. It takes a great deal of skill to conduct an unstructured interview and analyze the data. It is important to define criteria that determine who will be interviewed and the evaluation team accordingly identified the people to be interviewed.

Observation:
While an activity is going on, an observer records what he sees either using a checklist or by taking descriptive notes. The observation can include information on the setting (the actors, context, and surroundings); the actions and behavior of the actors; and what people say, including direct quotations. In the field (2 case studies) where activities were still being carried out by AMI, the evaluation team collected some information using this method.

Workshops :
At the end of the field case-study (analysis and recommendations workshop) and at the end of the final evaluation (lessons learned workshop), a workshop is organized in the presence of all the stakeholders. The aim of these workshops will be to share the current and partial knowledge of the evaluation team regarding the implementation processes of the AMI projects. Participants in these workshops will again have the opportunity to give their own input regarding the project, thereby correcting any misunderstandings on the part of the evaluator team.

The list of the tools used by each evaluation group for each topics is shown in the following table.

	Interview	**Focus group**	**Observation**	**Document**	**Questionnaire**
Women's health	2	3 (30)	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)**

Table 5 : Instruments and number of participants

Most of the people came from Laghman but as Metherlam is the reference base for Kunar, we took the opportunity of the visit of 3 people from AMI health facilities in Kunar to interview them.

So, in addition to the people observed during this evaluation, 205 people had the opportunity to express their thoughts and possible concerns regarding the implementation of the AMI program in Afghanistan. Of those 205 people, we had a gender balanced approach: 105 women and 100 men.

As each group was composed of two people, during interviews and focus groups, one person took notes whilst the other conducted the interview. The following principles of note-taking served as a guide: 1) notes should be recorded in the first person, 2) key words and ideas should be recorded, 3) original, descriptive phrases or sayings should be recorded word for word as quotations, 4) information should be recorded exactly as it is heard and not "filtered" based on interviewers' ideas or values 5) as many notes should be taken as possible, 6) in group interviews the various opinions in the group should be recorded.

Data analysis, results and lessons learned :
Once the data had been gathered, a participatory approaches to its analysis and interpretation helped participants to build a common body of knowledge. The consultant led the evaluation group to carry

out their own analysis but was always present to ensure that the quality of the analysis was of the right level..

The daily qualitative data analysis process was structured around the interview questions asked of each category of interviewees. A simplified approach to content analysis[16] based on a series of five steps was used by each group.

Step 1: *Re-read the interview questions*. One-by-one the interview questions should be read to the group. This allows the team members to recall the focus of each interview question.

Step 2: *Read the interview notes.* The note-taker/s should read aloud the responses, found in the notes for each question. If there are more than one set of notes, each set of notes should be read.

Step 3: *Discuss the responses.* The team leader asks the group to discuss the information included in the notes, to share other comments made by the interviewees that may not have been written down, to clarify exactly what the interviewees were saying.

Step 4: *Categorize the responses and summarize findings.* Together the group identifies the categories of responses in the information collected and summarizes the findings in a concise fashion. The example below illustrates a summary of the findings for one interview question.

Step 5. *Identify unclear or missing information.* A last step in the discussion of each interview question is for the group to determine whether there is missing or unclear information that should be further investigated in subsequent interviews.

Figure 3 : Qualitative data analysis process

Once the analysis is complete, the facilitator worked with the evaluation team during the last workshop in the field 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.

Methodological constraints

Before presenting our results for each topic, it is important to clarify the different constraints at play in this evaluation. Three types of limitations were identified as constraints: logistical, methodological, program.

1. *Logistical* : due to time constraints we were obliged to collect the data during a maximum of 6 days and the team was not prepared before the consultant's arrival. As this was the final evaluation, we were also obliged to wait until the program ended so that the time of evaluation could coincide with Ramadan. For security reasons we were also obliged to focus our evaluation only in the Laghman province. This constraint could limit the generalization, for the whole AMI program, of the

[16] Aubel, J., 1999, ibid

results and recommendations reached. In addition, also due to security reasons, the Country Director decided to ask the consultant and one of the evaluation team members (the only expatriate) to leave Laghman province before the end of the data collection and analysis stage. Therefore, the group dynamic was broken during 4 days (i.e 30% of the total evaluation time).

2. *Methodological* : in certain places, it was difficult to ask the head of the village or the director of the health facilities not to be involved in the focus group, for example. This problem could have an impact of the capacity of the other participants to give free answers to the questions asked by the evaluation team. Some evaluators were involved in the program under evaluation and this could have an impact of their objectivity, but we organized each team in a balanced way in order to compensate for this potential liklihood.

3. *Program* : At the time of this evaluation, AMI had some funding difficulties. Most of the health facilities staff was not being paid anymore by AMI and some of them (in Kakass, Laokar) were aware that AMI was no longer going to support their clinics. This constraint could be problematic in terms of staff willingness to have a discussion or to give unbiased, honest answers.