

Meta-evaluation Stop the Ebola Disaster

SHO members' intervention 2014-2015

COMMISSIONED BY SHO NETHERLANDS

July 1, 2017
TFM Consult, the Netherlands

Meta-evaluation Stop the Ebola Disaster

SHO members' intervention 2014-2015

Contents:

1.	SUMMARY	2
2.	INTRODUCTION	3
3.	METHODOLOGY	4
4.	ASSESSMENT OF INDIVIDUAL SHO MEMBER EVALUATIONS	5
	A. CARE	5
	B. CORDAID	6
	C. ICCO	8
	D. OXFAM NOVIB	9
	E. RODE KRUIS NETHERLANDS (IFRC)	11
	F. SAVE THE CHILDREN (SC)	12
	G. STICHTING VLUCHTELING	14
	H. TERRE DES HOMMES	15
	I. UNICEF NETHERLANDS	17
5.	DISCUSSION AND COMPARISON WITH PREVIOUS META-EVALUATIONS	19
6.	GENERAL CONCLUSIONS AND RECOMMENDATIONS	21
7.	ANNEXES	23

1. Summary

Not included, see document 'Management response' for clarification

2. Introduction

On 5 November 2014 the Foundation Cooperating Relief Organizations (Samenwerkende Hulp Organisaties - SHO) started the action "Help victims of the Ebola disaster". This national action brought in a total of over €10.5 million. The collected funds were spread over 10 member organisations to take part in this action according to a pre-defined division key: CARE Netherlands, Cordaid, ICCO and Kerk in Actie, Netherlands Red Cross, Save the Children, Oxfam Novib, Stichting Vluchteling, Netherlands Terres des Hommes, UNICEF and World Vision. With the Giro555-financed funds these organizations aided in all or some of the three most affected countries, Sierra Leone, Liberia and Guinea in the period from 5 November 2014 to 31 October 2015 (see Table 1).

SHO MEMBER ORGANISATION	% OF THE TOTAL SHO FUNDING	TOTAL RECEIVED	COUNTRIES RECEIVING AID	MAIN AREAS OF ASSISTANCE
CARE NEDERLAND	6,04	€ 609.721	LIBERIA	HEALTH CARE, WATSAN AND LIVELIHOOD SUPPORT
CORDAID MENSEN IN NOOD	15,49	€ 1.564.358	LIBERIA, SIERRA LEONE, GUINNEE	HEALTH CARE, FOOD, LIVELIHOOD SUPPORT, SHELTER / PROTECTION
ICCO EN KERK IN ACTIE	9,79	€ 988.308	LIBERIA, SIERRA LEONE	FOOD, LIVELIHOOD SUPPORT AND SHELTER/PROTECTION
NEDERLANDSE RODE KRUIS	19,10	€ 1.929.587	LIBERIA, SIERRA LEONE, GUINNEE	HEALTH CARE
OXFAM NOVIB	16,66	€ 1.682.299	LIBERIA	HEALTH CARE, WATSAN AND LIVELIHOOD SUPPORT
SAVE THE CHILDREN	6,91	€ 697.893	SIERRA LEONE	EDUCATION
STICHTING VLUCHTELING	5,30	€ 535.588	SIERRA LEONE	HEALTH CARE
TERRE DES HOMMES	2,49	€ 251.374	GUINEA	HEALTH CARE & WATSAN
UNICEF NEDERLAND	16,41	€ 1.657.029	LIBERIA	HEALTH CARE, WATSAN, EDUCATION, CHILD PROTECTION
WORLD VISION	1,82	€ 183.888	SIERRA LEONE	HEALTH CARE
TOTAL	100,00	€ 10.100.045		

TABLE 1: OVERVIEW OF SHO FUNDS DIVISION AND AID DELIVERED BY THE PARTICIPATING SHO MEMBER ORGANISATIONS.¹

The SHO prescribes in its Organisation Chapter that it 'will carry out a meta-evaluation in any action undertaken, what is a quality check of the delivered interventions. The meta-evaluation covers both the emergency phase and the reconstruction phase. The participants are required to actively participate in the implementation of the meta-evaluation. The organizations that work in international networks of relief and rehabilitation programmes do not have to carry out own separate external evaluations, but are requested to provide as many relevant documents for the benefit of the meta-evaluation. The evaluations used for this meta-evaluation do not only need to relate to the SHO-resources, as has been indicated'.²

This document is the result of the meta-evaluation "Help victims of the Ebola disaster" action as described above. Of the evaluation reports nine out of ten organisations were assessed, as World Vision did not had to carry out an evaluation of its intervention (funds received were less than 250.000 Euro). The evaluation reports were used to

¹ Source: SHO Eindrapportage-Stop-Ebola-deel-1 + 2, March 2016

² Extract from the SHO Organisation Protocol 1.16, June 2015

verify the impact, effectiveness and general quality of the members' operations. The main purpose of this meta-evaluation was to assess the quality and validate the conclusions of the individual evaluation reports. The meta-evaluation focused thus on the quality of the evaluation reports. The quality of the implemented projects of SHO member organisations was not assessed by the meta-evaluation, as this has already been accounted for in the individual evaluation reports.³

3. Methodology

The meta-evaluation uses a desk-study to analyse the different evaluations as its main tool, by means of a predefined quality assessment grid for each evaluation. Prior to the desk-study the evaluator made himself familiar with the SHO Organisational Chapter, the SHO end-report of the Ebola disaster and the Terms of Reference of the actual meta-evaluation (see annexes and consulted documents). The Quality Assessment Grid was used as a framework for every evaluation report in assessing and analysing the various reports whereas scoring of the various criteria led to an overall score per assessed evaluation report. As Cordaid provided 3 evaluation reports and an internal assessment of their actions the scores of these individual reports are combined to a general score for Cordaid. The same applies to Unicef, where the management response to the evaluation was taken into account, but not separately assessed as such.

Initially it was foreseen to have interviews with representatives of the various SHO member organisations (see semi-structured interview template in Annexe 3), but due to constraints in time and unavailability of staff (summer holidays) this was not possible and completely skipped.

Other constraints encountered were the unavailability of some Terms of References (essential to assess the first criteria of the quality assessment grid). Eventually the ToR's of 8 evaluation reports were available and 5 were not. Another constraint was the timing, July being the holiday season in the Netherlands, where many representatives of participating member organisations were not available to provide timely information or feed-back.

³ Source: ToR Meta-evaluation Stop the Ebola Disaster

4. Assessment of individual SHO member evaluations

a. CARE

CARE received a sum of € 609,721 or 6,04% of the total EBOLA funds. The interventions were mainly geared at Health Care, Water and Sanitation (WatSan) and Livelihood support in Liberia. An external evaluation of the intervention was carried out in December 2015, and this evaluation was subject of the assessment within the current meta-evaluation as shown in the Quality assessment grid.

TABLE 2: EVALUATION OF CARE ACTIVITIES IN LIBERIA, DEC. 2015

	unacceptable	weak	good	excellent
1. Meeting needs				
2. Appropriate design				
3. Reliable data				
4. Sound analysis				
5. Valid findings				
6. Impartial conclusions				
7. Useful recommendations				
8. Clear report				
Overall assessment			2.5	

The evaluation focused specifically on the SHO funded activities and SHO is clearly mentioned as the funder, though no specifications are made towards the total amount of the project under evaluation.

1. The ToR are quite elaborated, with clear purposes and five of the DAC criteria (Effectiveness, Efficiency, Appropriateness, Impact and Sustainability/Connectivity) included. Furthermore the criterium Connectivity is mentioned and the evaluator added Gender aspects. All these criteria, as well as the Humanitarian principles are discussed in the report, though some very shortly (e.g. effectiveness) and some only in the conclusion section (e.g. coordination).
2. The design is reasonably well defined, though most of the information was to be found in annexes (workplan) rather than in the report itself. Methodological limitations are explained as well as gender aspects (e.g. the percentage of female participants to be included in the evaluation). Total numbers of persons interviewed or participating in group discussions are, however, lacking.
3. The data are obtained using various approaches, one of them being a brain-storming session with the project staff on process, implementation and outcomes. The data as presented are, however, a mere presentation of this brain-storming session, without critical discussing potential biases, reducing the information's reliability. Furthermore the data on achievements of targets are presumably derived from project reports and in the presentation of the DAC criteria data sources of information are neither presented.
4. The analysis of the data is hence considered weak as well, as no systematic cross-checking e.g. by triangulation is carried out on data that are already less reliably obtained. Assumptions are sometimes mentioned (e.g. *if indicators have been achieved it is assumed that resources have been put to good use*), but sometimes the interpretation of data are unclear for the evaluator (e.g. in analysing certain output indicators: access to drinking water, crop planting, awareness raising, distribution of NFIs, the outcome is that stigmatization of survivors persists).

5. The validity of the findings is also considered weak and sometimes even questionable (e.g. participation of men and women in a certain activity is reflected as equal to the demographic sex ration of the population). Finding on effectiveness are virtually absent at that section and the information presented on (cost-)efficiency, although laudable that an effort has been made, lack any reference to “humanitarian standards”.
6. The report lacks as well systematic conclusions on all assessed parts of the evaluation. The conclusion for instance on the above-mentioned cost-efficiency is not presented nor are the other DAC criteria concluded. Coordination, however, is determined here but the conclusion itself (*coordination was well organized and given a high priority*) is not substantiated apart from an annexe where the number of meetings CARE staff participated in is presented.
7. Recommendations are not separately presented but at the end of every conclusion. They are valid and most of them considered useful, though only rarely clearly addressed to the (level of) authority concerned.
8. The report is interesting to read and the information presented easily accessible. Graphs and tables would have enhanced the readability as well as a more structured presentation of the conclusions. The summary, not completely summarizing all aspects of the report, is in itself a good presentation of the most important aspects of the evaluation

This evaluation is fairly well done and reported if all the information of the annexes is included as well. Still this evaluation report gives the impression that time was very limited and allocation of some more days (now only 14 in total were allocated) could probably have improved the quality of the final report.

b. Cordaid

Cordaid received a sum of € 1.564.358 or 15,49% of the total EBOLA funds. The interventions were mainly geared at Health Care and to a lesser extent food, livelihood support, shelter and protection in Liberia, Guinea and Sierra Leone . External evaluations of the various interventions were carried out in December 2015 and March 2016. Apart from these evaluations, Cordaid carried out a Performance Based Financing assessment, a partly internal, partly external assessment of this specific intervention in Sierra Leone in December 2015. All these evaluations were included in the assessment of the current meta-evaluation as shown in the respective Quality assessment grids and in an aggregated overall Cordaid Quality assessment grid (see below).

TABLE 3: EVALUATIONS OF CORDAID PARTNERS' ACTIONS IN SIERRA LEONE, GUINEA AND LIBERIA

	Sierra Leone				Liberia				Guinea				Cordaid overall			
	unac	weal	good	exce	unac	weal	good	exce	unac	weal	good	exce	unac	weal	good	exce
1. Meeting needs																
2. Appropriate design																
3. Reliable data																
4. Sound analysis																
5. Valid findings																
6. Impartial conclusions																
7. Useful recommendations																
8. Clear report																
Overall assessment		2.1				2.4				1.7				2.1		

The evaluations of Sierra Leone, Liberia and Guinea used more or less the same ToRs. Reason why the meta-evaluation assessment criteria will be collectively presented. The Performance Based Financing assessment is, however, not following the normal evaluation structure, is more a learning review rather than an evaluation and will therefore be discussed separately. In none of the reports the SHO is mentioned as the (co-)funder of the programme.

1. All the three evaluation reports follow the ToR requirements, though they vary considerably in the way they do this. Most reports answer more or less the questions raised in the ToR, though none of them to the extent needed (e.g. SPHERE standards application are sometimes mentioned, but never to what extent they are followed, the Guinea report does not answer the impact questions). Although the ToR of Cordaid is clear and reasonably elaborated, shortcomings in e.g. the description of conclusions and efficiency had its effect directly on the way the – assumingly not very experienced – evaluation teams executed their jobs.
2. The differences in interpretation of the ToR and the individual quality differences of the evaluation teams is quite clearly demonstrated in the design of the various evaluations. Here the Guinea evaluation scores unacceptable low (no specification of the number or rational of interviews, of the collection of data from beneficiaries etc.) whereas the Liberian team scores good, although also some more detailing in their methodological approach would have been appreciated. The Sierra Leone methodology scored in between, but weak, as e.g. no rationale for the number of interviewees nor a confidence level of the sampling size was presented.
3. Data gathered are hence considered rather weak and not so reliable for the Sierra Leone and the Guinea evaluations (e.g. a percentage of adhering or disagreeing respondents is often presented, without specifying whether these are key informants, direct or indirect beneficiaries etc.), but sufficient reliable for the Liberia evaluation as here the data collection instruments and most of the sources are quoted as well when assessing the outputs of the action. In assessing the DAC criteria these sources are, however, less visible.
4. The data cross-checking, interpretation and systematic analysis of all three evaluations is considered weak. No clarification of assumptions is done nor a sound analysis of data against indicators is presented.
5. As a result the validity of the findings is also considered weak as a clear chain of reasoning is often lacking. The exception is here the Sierra Leone evaluation where it is mostly clear where findings stem from and where several validation meetings were held with key informants and project staff before the report was published.
6. As mentioned in 1. the ToR provides a rather limited guidance on what is expected in the conclusions section. Most reports are therefor (?) rather short and incomplete in providing the necessary concluding information. The Guinea report scores in this respect again unacceptable low, as only some strong points of the programme are memorized here.
7. More or less the same applies for the recommendations: they are not mentioned in the ToR Scope, though they are to be formulated in the outline of the expected report. The Guinea report provides recommendations in between the findings, with no summary at all, the Liberia recommendations are rather limited and mostly focusing on management, though the recommendations of the Sierra Leone report are considered applicable, useful and clearly addressed to the responsible actors.
8. All three reports are not the best readable, reliable and accessible. The Sierra Leone report spends lots of space on presentation of the sample size and the findings vis a vis outputs – better presented in graphs and tables and some in annexes, discusses relevance and impact of the project in the same paragraph and presents percentages in the summary of the results of the evaluation that are not explained, reflected or elaborated in the summarized section. The Guinea report is quite inconsistent in its set-up, starting with some organisational audit with quite some interesting but unsolicited findings, and then ending up with very incomplete results on the DAC criteria (e.g impact is presented as whether certain disputable criteria are achieved). No summary is provided. The Liberia evaluation report is the best readable of the three, but lacks useful graphical presentation of findings, is inconsistent and distorted in the discussion of project components in the DAC criteria and the conclusions are missing in the summary.

All in all these three reports demonstrate that even with a common ToR the quality of evaluations and their reports can vary considerably. Moreover they all score ‘weak’, what is pitiful as quite some time and efforts have been put in these evaluations. A reason for this below average score could be that individual evaluations were carried out in the

three affected countries, for sure limiting the available budget for the respective evaluations. One evaluation could be traced to have a budget of some 6000 €, a rather small amount for the quality work that is expected.

The Performance Based Financing Project of Cordaid itself in Sierra Leone was eating up around another 25% of the total SHO budget. The Review was carried out by the Ministry of Health, Cordaid itself and an external consultant. Findings show mostly considerable progress towards indicators, but as this is more a project report, no judgement can be made about reliability of data, validity of findings and impartiality of conclusions. Moreover, this project was only possible and actually started in the recovery phase, thus limiting the possibilities for comparison with other SHO Stop Ebola emergency projects as well.

c. ICCO

ICCO received a sum of € 988.308 or 9,79% of the total EBOLA funds. The interventions were mainly geared at Food, Livelihood support and shelter/protection and to a lesser extent at health care and education in Liberia and Sierra Leone. An external evaluation of the intervention in Liberia was carried out in December 2015, and, although it is not clear whether this evaluation covered only the SHO funded actions - SHO is nowhere mentioned - it was subject of the assessment within the current meta-evaluation as shown in the below Quality assessment grid.

TABLE 4: EVALUATION OF ICCO PARTNERS’ ACTION IN LIBERIA

	unacceptable	weak	good	excellent
1. Meeting needs				
2. Appropriate design				
3. Reliable data				
4. Sound analysis				
5. Valid findings				
6. Impartial conclusions				
7. Useful recommendations				
8. Clear report				
Overall assessment		2.1		

1. The ToR specifies that the evaluation concerns a project of some 400.000 euro, so probably this was all SHO money, although consistently ICCO as funder of the through Christian Aid Liberia executed actions is mentioned. The evaluation had the following two-fold purpose: a. To assess and validate the outcomes and impact of the response from the standpoint of beneficiaries and partners, and b. To capture and synthesize Christian Aid (CA) and partner’s learning from the response. In addition some specific questions were posed, but no further DAC criteria were mentioned. The evaluation responds to some of the questions, though not sufficiently and more emphasis is put on evaluating the outputs and management of the operation than on outcomes or impact.
2. The design is at large clearly described, though not specific enough e.g. the number of intended household surveys was mentioned, though no argumentation provided. Methodological limitations were not mentioned nor their consequences on conclusions or recommendations.
3. As far as could be assessed, the collected data seem rather reliable, as data collection tools seem to have been used correctly. Mostly the sources are quoted in the various findings, though potential biases are not discussed.
4. The analysis is also considered weak as data were not consistently analysed. Efficiency and effectiveness are not fully analysed (e.g. assessing timeliness as only criterium for efficiency), appropriateness of the distribution in

Food items is differently presented than that of non-food items and the analysis is not consistent (e.g. the unavailability of NF items is presented per district where the unavailability of Food items is presented per item). The cash transfer component was only analysed for effectiveness and appropriateness, whereas especially this component could have benefitted greatly from analysing efficiency and impact.

5. The findings are in itself valid, as far as output indicators are concerned, but when it comes to evidence based more outcome related findings or conclusions, a reasoning is often lacking or false (e.g. *the majority of the people clarify that the items received were in line with their expectation*, suggesting the appropriateness of the action is sufficient). Limitations to validity are not mentioned at all.
6. Conclusions are derived from the various quantitative and qualitative findings, though often not specifying from which findings they stem. Again most conclusion are dealing with outputs and management of the operation, rather than on outcomes and impact (e.g consistent field monitoring of IPs’ activities by Christian Aid (CA) - ICCO’s partner in the field - was weak, CA was more concerned with strategic level activities), without making critical conclusions on the results of the operation itself.
7. Hence the recommendations are not very useful either, as again they are mostly focusing on improving the actual implementation of the operation by the IPs, rather than giving recommendations how impact and e.g. sustainability of such emergency interventions could be improved. Some effort is done with respect to the cash transfers – *to deepen the impact of the intervention the amount of cash should be increased to an amount that can feasibly start a business* -, completely neglecting , however, the very beneficial effect cash transfers can have on e.g. independency and self-respect of beneficiaries. In fact the buying of food from the received cash was even considered a negative effect.
8. The report is not consistent or very clear in its set-up, e.g. the sequence of the findings differs over the various components (Food items, non-food items and cash transfers) of the intervention, the heading of Observing of humanitarian Standards is missing and the chapter is only presenting the results of some dubious questions, not really addressing the in itself important aspects of preserving humanitarian standards in cash transfers. In the summary the conclusions are missing and graphs and tables are not consistent, scattered around, taking up quite some space and not always very appropriately presenting data (e.g. one time a simple yes/no question is presented as pie-chart, sometimes as cones). This is all in all not contributing to the readability and reliability of the report.

The ICCO evaluation report of the by Christian Aid executed action in Liberia is only receiving a total score of 2.1 and considered not up to SHO standards.

d. Oxfam Novib

Oxfam Novib received a sum of € 1.682.299 or 16,66% of the total EBOLA funds. The interventions were mainly geared at Health Care, WatSan and Livelihood support in Liberia. An external evaluation of the intervention was carried out in August 2016, and this evaluation was subject of the assessment within the current meta-evaluation as shown in the Quality assessment grid.

TABLE 5: STOP DE EBOLA DISASTER, FINAL EVALUATION OF OXFAM, AUG 2016

	unacceptable	weak	good	excellent
1. Meeting needs				
2. Appropriate design				
3. Reliable data				
4. Sound analysis				
5. Valid findings				
6. Impartial conclusions				

7. Useful recommendations				
8. Clear report				
Overall assessment		2.4		

The report has a strong focus on outcomes / early impacts on programme level. The evaluation includes OXFAM's overall response to the Ebola Disaster in Liberia with specific attention to the cross-cutting support provided by SHO (14 %of the total funding).

1. The Terms of Reference are well elaborated and provide good description of the context and background of the programme. The evaluation focuses on 1) outcomes and impacts level during emergency and early recovery programme, 2) relevance and 3) documentation and sharing of findings, lessons learned and provide recommendations. The scope of the evaluation is clearly defined with focus on the specific WASH and EFSVL (*Emergency Food Security and Vulnerable Livelihoods*) activities funded by SHO. The evaluation includes some of the DAC-criteria and adequately meets the information needs.
2. The evaluation methodology is described with a strong focus on the collection methods of the quantitative data. Specific evaluation questions are not elaborated and outcome-indicators and areas related to the logical framework are not mentioned. The design could have included more information on the approach for the qualitative data (e.g. *specific respondents and key informants internally/externally*). Main limitations are mentioned (e.g. *absence of the programme staff, closure of schools, difficulty to get in touch with KI*), but the consequences of these limitations on conclusions and recommendations are not defined. The choice for the two areas to be evaluated (out of four) is not explained.
3. The obtained data seem reliable and are gathered by involving different beneficiaries and stakeholders. Different collection methods have been used (Survey, Focus Groups Discussions, KII). Nevertheless the data presented in the tables are difficult to interpret without additional information (e.g. number of respondents in relation to total number). Potential biases are not discussed. It is not clear if the information could be equally collected across the two regions. It seems that Key Informant Interviews at the strategic level (MoH, other NGOs) did not take place (at least they cannot be traced in the interviewees list). Only limited information from the OXFAM overall evaluation is included in the report.
4. The analysis is consistent in problem analysis and activities responding developments in the different phases of the programme (emergency – early recovery). The focus is on output level of WASH interventions and verification whether activities had been realized (e.g. # of constructed and rehabilitated water points, on quality and accessibility). Some aspects of behaviour changes and sustainability are included, but the report lacks a clear overview of the formulated outcome indicators, realized outcomes and underlying assumptions. Findings are presented according to the prescribed DAC-criteria indicating crucial factors and successful approaches. Relevance of the WASH, social mobilization activities & contact tracing are clearly elaborated, but information on the other activities is limited (cash transfers, supply of agricultural tools, seedlings). Analysis on efficiency is sufficiently described (including management and cost effectiveness). The chapter on effectiveness include detailed situational descriptions and background information (e.g. on Waste Management) that could have been introduced earlier in the report. In some cases information, analysis and findings are mixed up. Coordination and connectedness are sufficiently covered.
5. The findings seem valid and cross-checked by different sources and by own observations of the evaluators (field visits). Qualitative data are used to triangulate and explain quantitative results, but sources are not consistently mentioned. Findings are not always precise (e.g. *some of the hand washing facilities are no longer functional, it is reported*). Limitations to validity are not mentioned.

6. The conclusions seem impartial and based on the logic of analysis and findings. Conclusions on (adaptive) management are limited and the conclusion ‘*challenges still remain in sustaining the gains made through the project, particularly at the community level*’ could have been more elaborated.
7. Main recommendations are forward looking and at operational level (WASH activities and required follow-up). Lessons learned focus on strategic level providing good elements for future disaster management. Recommendations on management and sustainability (ownership, handover after the programme) could have been more elaborated and included.
8. The report is well written and informative. The structure and readability could have been improved with a comprehensive overview of formulated and obtained outcomes and separation of background information, analysis and findings. The sequence of activities could have been presented according to an ‘implementation timeline’ improving the logical flow of the report. A summary is included in the report but leaves out the key lessons learned and recommendations.

Overall assessment: this evaluation report is considered only partly meeting the SHO standards.

e. Rode Kruis Netherlands (IFRC)

Rode Kruis Nederland as member of the International Federation of Red Crosses (IFRC) received a sum of € 1.929.587 or 19,1% of the total EBOLA funds. These funds were transferred to the IFRC and deployed together with other funds in Sierra Leone, Liberia and Guinea. The interventions were mainly geared at Health Care in the three countries affected. An external real-time evaluation of the intervention was carried out in December 2015 and this evaluation was subject of the assessment within the current meta-evaluation as shown in the Quality assessment grid below, although the SHO contribution is not specifically mentioned nor the percentage of the total amount of the action evaluated. A rough estimation, based on total funds of CHF 41.1 million for Sierra Leone only, reveals, however, that the contribution of SHO in the total IFRC response has been around 2%.

TABLE 6: IFRC (RED CROSS) REAL-TIME EVALUATION IN GUINEA, SIERRA LEONE AND LIBERIA

	unacceptable	weak	good	excellent
1. Meeting needs				
2. Appropriate design				
3. Reliable data				
4. Sound analysis				
5. Valid findings				
6. Impartial conclusions				
7. Useful recommendations				
8. Clear report				
Overall assessment			3,0	

Notwithstanding this rather small part of the SHO contribution in the total IFRC response, the evaluation report is considered valid enough to be assessed in this meta-evaluation. The evaluation report covers all 3 countries, and though the findings, conclusions and recommendations are set out individually, the report is assessed in its entirety.

1. Unfortunately the ToR for this real-time evaluation were not available, so it is hard to assess whether all the information needs are adequately met. In all 3 country sections the DAC criteria appropriateness/ relevance, efficiency/ effectiveness and connectedness are, however, discussed, answering specific predefined questions

alongside a specific IFRC upholding policy section. As these criteria are very elaborately and consistently discussed, it is trusted that the information needs are sufficiently met.

2. The methodology chapter mainly describes the itinerary of the evaluation team and the way the qualitative data were cross-checked and analysed. A good overview of timing, number of interviews and their rationale and the specifics of this real-time evaluation are, however, missing and this is considered a pity, as the report itself is very thoroughly and professionally set-up.
3. The presented data are hence somewhat difficult to judge on their reliability. Data are only qualitative and no data collection tools such as interview formats are provided nor are the data presented in a systematized way. However, often the sources of information are quoted and the overall impression is that of a well-executed and professional evaluation, providing the trust that data are reliable enough to derive conclusions from.
4. In a real-time evaluation data cross-checking and systematic analysis is somewhat difficult, though here again the evaluation team deserves the benefits of the doubt, as the way data are interpreted and presented is trustworthy. Feedback from key-persons is included and apparently some verification by tri-angulation of received information e.g. by interviewing donor and peer organisation representatives is carried out.
5. The findings are considered sufficiently based on evidence, though again, this is not easily assessed in a real-time evaluation and given the wealth of the information provided.
6. The individual conclusions per country are rather limited and do not conclude all the findings presented (e.g. no conclusion on the efficiency of operations, whereas one of the countries spent only half of the budget at the time of evaluation). Yet they are valid and, although sometimes not completely impartial, very useful, especially given the timing of the evaluation – at the midst of the epidemic.
7. Although most of the recommendations stem directly from the findings and not from the condensed conclusions, they are excellent in their applicability and usefulness. Most of the recommendations are detailed and could be applied instantaneously and at various levels – up to the Secretary General of the IFRC.
8. This is a very elaborated report, readable and providing massive information especially for the implementing Red Crosses on the ground in a structured and comprehensive manner. However, some sections could benefit from more consistency (e.g. use of resources, financial performance) and for such a vast terrain to evaluate and to compare findings in between the three countries the use of graphics or comparison grids would have made the abundant information more accessible.

All in all this is a good evaluation of a very complex, at the time ongoing and difficult matter. The fact that this was a real-time evaluation of this scope is interesting, helpful, quite unique and laudable. However, it makes quality comparisons within the scope of this meta-evaluation with the other – end-term - evaluations somewhat problematic.

f. Save the Children (SC)

Save the Children Netherlands received a sum of € 697,893 or 6,04% of the total EBOLA funds. The interventions were mainly geared at Education in Sierra Leone. An external evaluation of the intervention was carried out in December 2015, and this evaluation was subject of the assessment within the current meta-evaluation as shown in the Quality assessment grid below.

TABLE 7: STOP DE EBOLA RAMP, FINAL EVALUATION OF SAVE THE CHILDREN-SIERRA LEONE, FEB 2016

	unacceptable	weak	good	excellent
1. Meeting needs				
2. Appropriate design				

3. <i>Reliable data</i>				
4. <i>Sound analysis</i>				
5. <i>Valid findings</i>				
6. <i>Impartial conclusions</i>				
7. <i>Useful recommendations</i>				
8. <i>Clear report</i>				
<i>Overall assessment</i>		1,8		

The assessment grid shows an evaluation report that is generally weak and not up to the SHO standards. For unexplained reasons SC chose to have a process evaluation rather than an impact evaluation, concentrating more on outputs than on outcomes, though the Purpose and Objectives of the evaluation clearly state to assess the performance of the project according to DAC evaluation criteria. SHO is visibly mentioned as funder of the project.

1. The report meets some of the information needs as expressed in the Purpose and Objectives of the Evaluation and the evaluation criteria and questions (a ToR for the Evaluation was not included), though especially information on adherence to the Core Humanitarian Standard (CHS) on Quality and Accountability is lacking. Answers to the key questions are often vague and incomplete (e.g. the question on meeting the needs of the most vulnerable) and not sufficiently based on hard evidence (see later).
2. The evaluation methodology is well described and clear. The methodology should provide enough information to adequately answer the questions – the household interviews provide as well unsolicited information about the impact of the project such as the use of the solar powered radios – though some criteria e.g. the reasoning for the actual number and the locations of focus group discussions and key informant interviews were lacking.
3. Hardly any quantitative data e.g. of the household survey are presented in a systematic way. Valuable, available information about effect and impact is therefore missing. Outcome indicators and data are completely lacking and the reliability of presented data is questionable e.g. the confusing presentation of number of direct beneficiaries (varying between 10.500, 47.000 and 114.000). Triangulation is not mentioned nor the reliability of quoted sources or the potential biases.
4. Consequently the analysis of the data is insufficiently appropriate. There is no evidence of cross-checking data nor are external factors e.g. the course of the Ebola epidemic itself and flooding sufficiently identified and their influence taken into account.
5. The findings are often sourced, but as the collected data themselves are not adequately presented it is very difficult to assess the validity of the findings. A clear chain of reasoning is missing most of the times and limitations to the validity aren't stated either.
6. Only very limited conclusions were presented, hardly critical, not covering all DAC criteria (e.g. coherence and coverage are not included). Furthermore the conclusions do not take into account all findings (e.g. effectiveness is only concluded in addressing the number of beneficiaries reached and the timeliness, whereas a wealth of other data was presented in the evaluation findings). Conclusions are neither based on explicit criteria nor benchmarks.
7. Hence the recommendations that stem from the conclusions are not very useful and even sometimes inappropriate (e.g. the recommendation to increase the number of learning materials or support to more teachers/trainers without the previous conclusion that these issues were hampering the success of the project).
8. The report is sufficiently readable and accessible to the intended user, though the many linguistic and factual unclaritys (e.g. the erroneous numbers of direct beneficiaries, the “potential to reduce the risk for girls’ teenage marriage and pregnancy”, in spite of information elsewhere in the report from MoHS that teenage pregnancy increased nationwide with 27% etc.) decreased the trustworthiness of the report.

All in all the report does not satisfy the criteria of a proper and well-executed evaluation.

g. Stichting Vluchteling

Stichting Vluchteling received a sum of € 535.588 or 5,30% of the total EBOLA funds. The interventions were through its partner IRC mainly geared at Health Care in Sierra Leone. An external evaluation of the intervention was carried out in September 2015, and this evaluation was subject of the assessment within the current meta-evaluation as shown in the Quality assessment grid.

TABLE 8: : STOP DE EBOLA DISASTER, FINAL EVALUATION OF STICHTING VLUCHTELING, SEPT 2015

	unacceptable	weak	good	excellent
1. Meeting needs				
2. Appropriate design				
3. Reliable data				
4. Sound analysis				
5. Valid findings				
6. Impartial conclusions				
7. Useful recommendations				
8. Clear report				
Overall assessment			2.8	

Main purpose of the evaluation was an independent systematic assessment of IRC's implementation of the project 'Ensuring Appropriate Health care for all suspected and confirmed Ebola cases in Bo Isolation Unit, Bo district, Sierra Leone'. The specific purpose was to assess and report on the obtained and expected results and an analysis of any deviation. The project is supported with SHO-funds (specific activities are described), and co-funded by DFID and ECHO is briefly mentioned in the report. The evaluation focussed on the key activities funded by SV/SHO.

1. The ToR is clearly formulated and includes information on the project and references to background documents. The ToR refers to the 8 DAC-criteria with elaborated sub-questions and an expanded evaluation methodology. The evaluation is process oriented, complying with the request in the ToR to focus on learning and adaptive capacity. The description of 'Evaluation objectives and methodology' (Annex 1) is not referring to the ToR (not included in the report), but the specific evaluation questions are quite systematically answered in the report.
2. The methodology is briefly described in the report with a strong focus on collection of qualitative data. The choice for the Appreciative Inquiry Approach could have been better explained could have included the limitations of the approach. The Appreciative Inquiry approach focuses on what went well and improvements, but less on what a different approach would have resulted in (*e.g. IRC took over the management of the hospital; could co-management have been a solution to create more ownership after the project period?*). Some limitations of the evaluation are included in the report though the consequences of these limitations are not further elaborated (*e.g. number of important players had left the country with no possibility to interview them remotely; what were the consequences, conclusions and recommendations?*). Only 5 DAC-criteria are covered in the report, Coverage, Connectedness and Proportionality were not separately included, these topics are partially touched upon throughout the text.
3. The data in the report seem reliable and cross-checked. The quantitative data are mainly derived from documents of international organisations and internal IRC reports, although some slightly differ (*e.g. 31,922/31,914 of patients triaged*). Quotations are clearly marked in the text and refer consequently to sources. It is, however, not clear who participated in the (Focus Group) Discussions and who provided and discussed the information and

findings under the headings 'DISCUSSION' in the report. The list of interviewees and participants of focus groups is not included in the report (annexes were not available for the evaluator).

4. The data in the report seem reliable and cross-checked. The quantitative data are mainly derived from documents of international organisations and internal IRC reports, although some slightly differ (*e.g. 31,922/31,914 of patients triaged*). Quotations are clearly marked in the text and refer consequently to sources. It is, however, not clear who participated in the (Focus Group) Discussions and who provided and discussed the information and findings under the headings 'DISCUSSION' in the report. The list of interviewees and participants of focus groups is not included in the report (annexes were not available for the evaluator).
5. Findings seem valid and underlying - interrelated processes are explained and validated by different sources (stakeholders, beneficiaries, national and international staff). The evaluation has a strong internal focus, but some external views have been included from persons directly involved in the project as well. The report includes balanced and also critical views of the different internal stakeholders (*e.g. Supervisory positions in the EUI could also have been covered by senior – and well trained staff*).
6. The conclusions are well formulated in a general and descriptive way and seem impartial. Though it is stated that '*the difficult pattern of this epidemic with the weak systems....., has made this intervention an especially difficult challenge*', the specific challenges for this interventions could have been more elaborated and specified for different levels and actors, connecting them with the recommendations in the next chapter.
7. Key recommendations are formulated in a general way and not always clear to who(m) they are addressed. They are mostly forward looking and focussed on the operational level of the project. Given the fact that the Ebola epidemic and nature of response differs considerably from other disasters, it would have been good to include recommendations and lessons learned at strategic level. The report is mentioning several challenges on sustainability, (*e.g. ownership, drop of motivation after handover of IU back to the BO hospital*) and recommendations on this issue could have been included.
8. The report is well written, accessible to read and informative on results, process findings and lessons learned. It contains a rather complete executive summary with most of the necessary information. A final quality check (*e.g. page numbers missing, list of interviewees not included, numbers checked*) could have been beneficial.

Overall assessment: this report is rather useful and of sufficient quality according to SHO standards. With just a few adaptations the quality could have been even more improved.

h. Terre des Hommes

Terre des Hommes received a sum of € 251,374 or 2,49% of the total EBOLA funds. The interventions were mainly geared at Health Care and WatSan in Guinea. An external evaluation of the intervention was carried out in March 2016, and this evaluation was subject of the assessment within the current meta-evaluation as shown in the Quality assessment grid. The evaluation concerns a Joint Humanitarian Response (JHR) for Ebola of seven Dutch humanitarian organisations aggregated in the Dutch Relief Alliance (DRA) and funded by the Dutch Ministry of Foreign Affairs. It is understandable that Terre des Hommes did not perform a separate evaluation of the SHO funded activities, seen the relatively small amount concerned (the total budget of the JHR added up to € 9,9 million), though it is hard to distil from the current evaluation the proportion (some € 145.000 according to an accompanying evaluator letter) that can be attributed to the SHO funding. This SHO co-funding is also not mentioned in the evaluation report.

TABLE 9: TERRE DES HOMMES STOP EBOLA ACTIVITIES IN GUINEA EVALUATION; MARCH 2016

	unacceptable	weak	good	excellent
1. Meeting needs				
2. Appropriate design				
3. Reliable data				
4. Sound analysis				
5. Valid findings				
6. Impartial conclusions				
7. Useful recommendations				
8. Clear report				
Overall assessment			3,0	

1. The ToR exhibits quite clearly the information needs in two separate objectives, 1: to assess the results and the quality of the response and 2: to assess the cooperation, added value of the structure and lessons learned for future collaboration. The elaboration of objective 1. Includes 5 of the DAC criteria and adherence to humanitarian standards such as CHS and Sphere. The report sufficiently meets these information needs, with some omissions such as the lack of discussing humanitarian standards in the quality assessment. Regarding objective 2. The information is somewhat less structured presented, not following the scope of the ToR, though in the end most information needs are satisfactory met.
2. The methodology for this evaluation is very shortly described in the report, but further elaborated in annexes, including an evaluation matrix and considered sufficiently adequate to address all defined questions. A rational for e.g. the number of key informant interviews and of the surveys of beneficiaries is, however, missing as well as attention to gender aspects.
3. Data obtained seem sufficiently reliable. Sources are mostly quoted and can be further traced in the evaluation matrix. Potential biases are unfortunately not described nor the proper application of data collection tools.
4. From the evaluation matrix the various methods of analysis could be derived and they seem sound and consistent. Assumptions are, however, again not described.
5. Findings are elaborately formulated, mostly well-substantiated, enforcing their validity. Limitations to the validity such as the difficulty to attribute certain results to the DRA/SHO funds as so many organisations in the consortium mixed the funds with other funds and as there were so many other actors in the Ebola response are mentioned.
6. Conclusions are separately and clearly formulated. Some of them are, however, more or less repetitions of the facts presented in the findings session and for readability some kind of aggregation or structuring could have been beneficial as well. Points of disagreement are less visible though lessons of wider interest visibly mentioned.
7. Recommendations are addressed at two levels: Ministry of Foreign Affairs (the donor) and the JRE partners. They stem from the conclusions, though without reference and are mostly valid and useful. Certain recommendations e.g. on the continuation or expansion of 'the Stories of Change' as a widely recognised proof of joint learning and collaboration are unfortunately missing.
8. The report is well-written and clearly structured. It could have benefitted from more consistent use of graphs instead of tables and certain information such as the beneficiary responses to knowledge of Ebola and the budget and actual expenditures would preferably be annexed.

All in all this is a well-designed and appropriate evaluation. Unfortunately the contribution of the SHO funds could not be traced and is also not mentioned, but given the circumstances and the total funds involved this is understandable.

i. Unicef Netherlands

Unicef Netherlands received a sum of € 1.657.029 or 16,4% of the total SHO EBOLA funds. The interventions were mainly geared at Health Care, WatSan, education and Child protection provision in Liberia. An external evaluation of the intervention was carried out at corporate level and in Guinea, Sierra Leone and Liberia (covering estimative the 2015 period). This evaluation and the management response to this Evaluation was subject of the assessment within the current meta-evaluation as shown in the below Quality assessment grid.

TABLE 10: STOP DE EBOLA DISASTER, FINAL EVALUATION OF UNICEF, MARCH 2017

	unacceptable	weak	good	excellent
1. Meeting needs				
2. Appropriate design				
3. Reliable data				
4. Sound analysis				
5. Valid findings				
6. Impartial conclusions				
7. Useful recommendations				
8. Clear report				
Overall assessment				3.8

The evaluation focussed on the UNICEF corporate response to the Ebola outbreak emergency coordinated and implemented in West Africa. The total funding for UNICEF's Ebola Response was over US\$ 500 million. The SHO funding is a very small part of the total funding and no reference is made to specific SHO funded activities.

The evaluation is drawing out import lessons to prepare and strengthen UNICEF's approach to addressing and strengthening future health emergencies.

The evaluation is prepared by an evaluation team of 8 persons, supported by UNICEF's evaluation office and an Evaluation Reference Group and complying with UNICEF's evaluation standards. An evaluation Management Response of UNICEF has been prepared separately. Considering the comprehensive nature of the evaluation (multiple countries, regional level and UNICEF coordination level), it is tolerable that the report has been published only recently.

1. The report is comprehensive and covers in a very structured way the different areas and levels to be evaluated. Key evaluation questions are formulated and primary and secondary users identified. The TOR is not included; the report contains a clear description of the focus, methodology, approach used and the restrictions of the focus on strategic challenge. As also stated in the Management Response, the evaluation serves to deepen understanding of key strategic and organisational issues relevant to UNICEF's performance of the Ebola response.
2. The evaluation purpose, objectives and scope are clearly defined in the description of the methodology. Three specific objectives are defined; Accountability, Learning and Strategy. An evaluation framework was developed and used the criteria of effectiveness, efficiency, coordination, relevance and coherence. The prioritization of two programme elements (Health Care Management and Child Protection) is well explained, the report does not attempt to provide detailed information or a technical assessment of the implementation. A participative evaluation process was developed, internal and external stakeholders and beneficiaries were involved and able to provide their feedback. The limitations of the approach are broadly elaborated, just like the consequences on the findings, analysis and conclusions.

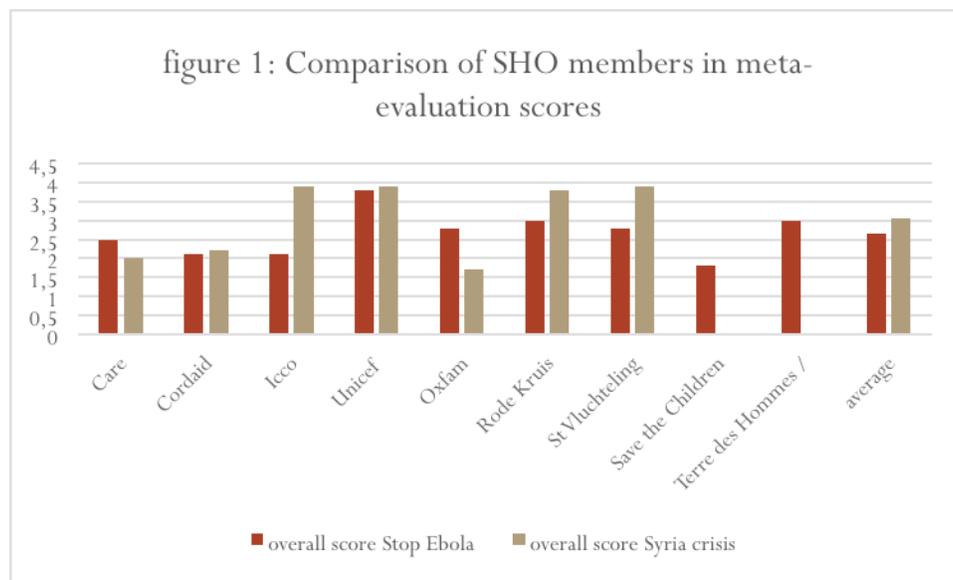
3. The presented data seem to be reliable; data collection was conducted according a mixed methods approach clearly described in the report (lessons review, document review, data analyses, stakeholder consultations, stakeholder polling and case studies of affected communities). The different methods are responding to the complexity of the evaluation. Data on accountability are sufficiently included. Sources are referred to and different opinions are included in a balanced way.
4. The analysis looks sound and is focussing on the strategic challenge of coordinating UNICEF's levels, programmes and operational functions and how these elements combined deliver an effective response. The analysis of key elements is consistently structured in the five chapters of the report. Internal and external factors are broadly taken into account. The clear chronological sequence of developments in the context, management decisions and their implications is well elaborated. The report might have been strengthened with inclusion of the evaluation framework. The analysis is on certain topics somewhat repetitive.
5. The findings look valid, are cross-checked and sufficiently combining the different levels (operational/ strategic level) of the evaluation. Triangulation was conducted using the evidence matrix and findings. Preliminary evidence-based findings were discussed at country and regional level, and with the Evaluation Reference Group and a wider group of UNICEF stakeholders. The report contains many references to internal documents and external studies / reviews on the Ebola Response.
6. The conclusions are clearly and critically formulated, highlighting obtained results and determining in detail where deficiencies have been identified. They are consistently based on the background information, the analysis and the findings presented and clearly addressing different levels (managerial, operational, strategic and accountability). The conclusions include lessons learned and are intended to improve UNICEF's actions and strategy in future situations.
7. The five recommendations look impartial, balanced and well-considered, following the reasoning of conclusions and lessons learned. The recommendations are aimed at different levels within UNICEF (regional office, global management team, the Office of Emergency Programmes and the Programme Division). The recommendations seem useful and practical. The UNICEF's Evaluation Management response includes an elaborated action plan on the 5 key recommendations. Most recommendations are accepted; in some case the level of appropriateness of the level of responsibility (General Management Team) is contested. The actions are included in a time schedule, some of them are actually implemented.
8. The report is well written, informative and well structured. The report follows the UNICEF evaluation quality standards. A balanced and well elaborated executive summery is included in the report, but can easily be read separately. The structure facilitates the readability of the comprehensive topics. Findings, analysis and conclusions and recommendations are linked following the structure of the key evaluation questions.

Overall assessment: the quality of the evaluation is very good and certainly meeting the SOH standards. The report is consistent and informative despite the complexity of the topics evaluated. The conclusions and recommendations are very relevant, critical in a constructive way and forward looking.

5. Discussion and comparison with previous meta-evaluations

In March 2016 SHO commissioned the first meta-evaluation of a complete SHO action: the Syria action of 2013. Not all the members that at that time were evaluated participated in the Ebola response and new members popped in. A comparison is made of the differences in scoring between the SHO members (see figure 1). Although the scoring is a subjective matter and the two meta-evaluations were assessed by different persons, the following can be observed:

- The average score declined somewhat, from 3.1 to 2.6, but this could be caused by the abovementioned evaluator bias.
- Again there are three organisations that score rather poor, below 2.1, one of the three being the same as in 2016, but there is now only one organisation (Unicef) that scores excellent.
- Oxfam and Care improved their score significantly, ICCO and St Vluchteling declined, though St Vluchteling is still above average.



- Evaluations still differ enormously in set-up, budget available (when assessable), ToR's and overall quality. It is evaluator's impression that not many lessons have been learned from the previous meta-evaluation.
- Interesting to see is that again the organisations with international networks (Unicef and Rode Kruis) consistently score above average to excellent.
- Most evaluations mention the difficulty to obtain reliable data especially after the crisis. A suggestion could be to look at alternative evaluation approaches such as real time monitoring and adaptive management techniques.

Similarly to the Syria crisis meta-evaluation The larger evaluations were often entirely or mostly focused at strategic level, whereas the smaller, more project-oriented evaluations looked more at activity and output level, though the better ones also assessed the achievement of outcomes and objectives.

6. General Conclusions and recommendations

1. Evaluations differed in many ways from each other e.g. in size, set-up, execution, scope, timing and period-covered. Some were good to excellent (e.g. Unicef, Rode Kruis, Terre des Homme/JHR) others weak or even unsatisfactory. It is difficult to compare the evaluations as they were actually executed, reducing the effectiveness and usefulness of the current meta-evaluation.

Recommendation:

- 1a. SHO should try to harmonize the evaluations of their member interventions: General Terms of reference, agreed upon before actions are undertaken, preferred timing and execution of evaluations etc. could help.
 - 1b. Participating member organisations should be aware of, have access to and consult each other's ToR's for and actual evaluations. SHO to facilitate such a platform (see also 2.b).
 - 1c. Participating member organisations could be asked to produce some kind of 'management response letter' to enhance credibility and usefulness of the executed evaluations and demonstrate the willingness to learn from them.
2. Participating member organisations were generally not aware of the present meta-evaluation. Representatives hardly knew that the exercise was upcoming and were not always informed. Timing of the meta-evaluation and availability of the evaluation results⁴ were as well an issue as many organisations moved already forward towards other disasters such as the 2017 Famine action.

Recommendation:

- 2a. SHO should hold member organisations accountable for the agreements as laid down in the SHO Organisation Protocol e.g. in timely submitting evaluation reports and
 - 2b. SHO should encourage mutual learning from individual actions by sharing evaluations and drawing attention to actions such as this meta-evaluation. A suggestion would be to have a common cloud-based facility such as a Dropbox folder or "dashboards" where participating members could real-time post and consult respective evaluations and ToR's.
3. In addition to conclusion 1. it is noticed that the quality of evaluations is directly influenced by the budget and time allocated. The weakest evaluations tend to have the smallest budget available and/or time allocated.

Recommendation:

 - 3a. To further improve the quality of reports SOH could suggest a budget for humanitarian action evaluations. A widely accepted range is between 1-5%⁵ of the total budget purely for the external evaluation as such, with smaller programs needing relatively higher budgets than the larger ones. For this Ebola response programme, this would have meant to set aside some € 100-500k for evaluations. It is the impression of the evaluator that this amount is by far not used.
 4. As it is now organised the organisations that cooperate in international networks (e.g. Unicef and Rode Kruis) are allowed to produce overarching evaluations, without specification of the SOH funds involved. This is understandable as these contributions are often in size marginal, yet it is difficult to explain to the Dutch public that some 35% of the total funds are not specifically evaluated. Moreover, also some other SHO members

⁴ One of the evaluations was only available in June 2017, 20 months after the intervention stopped, well beyond the prescribed period of 12 months (Organisatie reglement SHO, 1.16)

⁵ E.g. <http://wiki.interact-eu.net/pages/viewpage.action?pageId=23756932>

(ICCO, Cordaid) do not consistently mention SHO as the (co-)funder of their action, or make the evaluation part of a larger evaluation without specifying the SHO contribution or activities (Terre des Hommes). This all affects the accountability of SHO towards the Dutch public and the recognition of its contributions.

Recommendation:

4a. SHO to be more vigilant and demanding on its members to specify SHO as (co-) funder and to consistently mention the amount and preferably also the actions in the evaluations that contributed to the respective humanitarian interventions.

7. Annexes

1. Terms of reference Meta-Evaluation STOP the EBOLA DISASTER (*attached*)
2. Documents evaluated and consulted (*see: giro555.nl/de-hulp*)
3. Semi-structured interview template used to interview participating member-organisation representatives:
 - a. Do you know about the STOP the EBOLA DISASTER Meta-Evaluation. If so what is your opinion about it;
 - b. Were you consulted in providing material? If so did you provided other material than the evaluation
 - c. Are you interested / aware of the other member organisations evaluations? Did you share your evaluation report with others?
 - d. Do you have any suggestions to increase the usefulness and effectiveness of this meta-evaluation?



Terms of Reference

Evaluation STOP THE EBOLA DISASTER

Background

The 23rd of March 2014 the WHO declared an Ebola outbreak in Guinea. It was the biggest and most complex outbreak of the deadly virus since Ebola was discovered in 1976. For weeks the disease was not recognized and thereby had the change to spread to neighboring countries. Guinea, Sierra Leone and Liberia were hit hardest and the SHO decided to spend the funds raised in those three countries. In total 28639 people were contaminated with the Ebola virus, 11.316 died, most of them in the mentioned three countries.

Guinea, Liberia and Sierra Leone had not experienced the disease before and did not know how to respond to the epidemic. Furthermore, due to conflict in these countries the health systems were not maintained well and they were faced with a shortage of educated health staff. The countries intervened late and slow. Cultural practices, like washing the bodies of people that died, fear for the virus, resistance with the population and strikes of health staff are factors that contributed to a spread of the disease. International interventions were slow as well and also international organizations faced a lack of knowledge and experts.

THE SHO CAMPAIGN

The Samenwerkende Hulp Organisaties (SHO) is an alliance of eleven Dutch aid organisations. In response to major humanitarian disasters (both natural disasters and conflicts) the SHO raises funds in order to provide humanitarian assistance to the victims of the disaster. In addition, the SHO cooperates with media partners such as television, radio, newspapers and various online media to inform the general public about the particular disaster. The aid agencies of the SHO are jointly accountable for the expenditure of the funds, through their own respective mechanisms. This report is an example of joint accountability.

All SHO participants, their local partners and/or their umbrella organisations already worked for years in Sierra Leone, Liberia and Guinea and next to their regular programs responded to the Ebola epidemic. On 5 November 2014 the SHO launched a national campaign in order to scale-up the individual organisations' relief operations. Total funds received on Giro555, the SHO's bank account, were 10,5 million Euro. This was the first time SHO responded to an epidemic.

ORGANISATIONS THAT PARTICIPATED IN THE SHO ACTION

CARE Nederland, Cordaid Mensen in Nood, ICCO en Kerk in Actie, Nederlandse Rode Kruis, Oxfam Novib, Plan Nederland, Save the Children, Stichting Vluchteling, Terre des Hommes, UNICEF Nederland and World Vision

MAIN REASONS AND PURPOSES OF THE EVALUATION

The SHO member organisations jointly raise funds for emergency operations. They spend the raised funds separately, through their respective partners, networks and alliances, with coordination on the ground through the cluster system. The results and expenditure of the emergency operations, funded by SHO, are then jointly reported back to the public.

Since the SHO members do not operate as one organisation during the emergency operation, they need a mechanism to ensure that the individual members perform up to the high standards that the SHO has set for itself. Each member that spends more than EURO 250.000 per emergency operation, has the obligation to conduct an external audit for its own operation. The evaluation reports are used to verify the impact, effectiveness and general quality of the members' operations. The SHO conducts this meta-evaluation to assess the quality and validate the conclusions of the individual evaluation reports.

The meta-evaluation is thus focused on the quality of the evaluation reports. The quality of the implemented projects of SHO member organisations is not assessed by the meta-evaluation, as this has already been accounted for in the individual evaluation reports.

SHO follows the ALNAP¹ definition for meta-evaluations, adapted from Lipsey (2000): systematic synthesis of evaluations providing the information resources for a continuous improvement of evaluation practice.'

The meta-evaluation report will be shared with the SHO Board of directors to monitor the internal adherence to quality standards. It will also be used by individual members as input to improve evaluation practices.

The meta-evaluation takes into account the agreements made with regard to evaluations as stated in the Organisation Protocol ("Organisatie reglement SHO, October 2013", chapters 1.14 - 1.19). This refers to:

- > The OECD DAC criteria for disaster response evaluations, in accordance with the ALNAP Quality Proforma;
- > The Sphere minimum standards in Humanitarian Response during the intervention;
- > The Code of Conduct for The International Red Cross and Red Crescent Movement and NGOs in Disaster Response Programmes during the intervention.

METHODOLOGY AND DELIVERABLES

The quality check of the evaluation reports will be carried out through a desk study of the reports of external evaluations and reviews. Annex 1 provides a quality assessment grid that will be used to assess the quality of these evaluations. This quality assessment grid includes elements such as meeting of objectives indicated in the ToR, quality of the analysis and reliability of the data. The consultant will provide the rationale for scores given and indicate strengths and weaknesses.

¹ ALNAP is the Active Learning Network for Accountability and Performance in Humanitarian Action and was established in 1997. ALNAP is a collective response by the humanitarian sector, dedicated to improving humanitarian performance through increased learning and accountability.

The quality assessment grid will be used for all project evaluation reports shared. Based on the outcomes a synthesis is written with specific attention for each component in the assessment grid.

Based on this desk study a draft report will be developed that will be discussed during a meeting with SHO members. This draft report, no longer than 20 pages (including tables and annexes), should consist of:

- I. An analysis on the quality of the evaluation reports based on the assessment grid mentioned under objective 1;
- II. A section with the conclusions, and lessons learned based on the above analysis;

The meta-evaluation report should be written in **English**

During a meeting with the SHO Board of directors, the consultant will present the outcomes of the research.

After being selected, the consultant will develop a detailed work plan at the beginning of the process. All materials collected in the undertaking of the evaluation process should be lodged with the SHO Secretariat prior to termination of the contract.

RESPONSIBILITIES PARTICIPATING SHO ORGANISATIONS

The process of developing and finalising the meta-evaluation is steered by representatives of the SHO quality group. CARE is responsible for the overall coordination of the meta-evaluation, supported by CORDAID through a Reference Group. This Reference Group will:

- › Select the evaluator(s) and facilitate the communication with the evaluator(s);
- › Facilitate the administrative steps around this evaluation in coordination with SHO back office (contract, payment etc.);
- › Support the consultant in assembling the relevant documentation (i.e. evaluation reports and reviews).

The members of the SHO Quality Group will

- › Provide relevant information within their own respective organisations and send it to the evaluator(s);
- › Provide input into draft report;
- › Agree on the final report.

Participating SHO agencies are required to submit the reports of external and internal evaluations and reviews as input for the meta-evaluation (preferably both in hard copy and electronic format to the SHO secretariat)

Expected workload

Activity	Number of days
Preparation and reading (basic documents)	1
Desk study (analysis different evaluations)	10
Writing meta-evaluation report	2
Discussing Report with SHO members and presentation to board, final adjustments	2
Total working days	15 days

Timeline

Activity	Date (planning)
Deadline for application	1 June 2017
First meeting Reference Group / Consultant	Week 12 June 2017
Draft meta-evaluation report	10 July 2017
Adjustments after quality workgroup has read the meta-evaluation	21 July 2017
Finalisation of the process	1 August

Minimum requirements

The evaluator should have relevant skills and a proven background in humanitarian project implementation and/or evaluations. No specific sector expertise is required. The selection criteria are:

- > Experience with humanitarian aid programs
- > Experience with conducting evaluations
- > Knowledge of the context of the disaster/conflict
- > Fluent in both Dutch and English and able to read French
- > Availability in the indicated period

The consultant should not have evaluated one or more of the individual evaluation reports of the participating SHO member organisations for this SHO program

Applications

Applications can be sent to Inge Leuverink at ile@cordaid.org by 1 June **2017 latest**. The application should contain a short motivation (max 1 page), past experience (max 1 page), a Curriculum Vitae and the consultants' daily rate. Only short-listed candidates will be contacted.

Annex 1 - Quality assessment grid

Assessed report:	
Version: final draft	Date of report:
Assessor:	Date of assessment:

Excellent
 Good
 Weak
 Unacceptable

<p>1. Meeting needs</p> <p>The report adequately meets the information needs expressed in the terms of reference in a way that reflects the stated priorities. The demands which were made during the evaluation process are mentioned, and satisfied when possible.</p>				
<p>2. Appropriate design</p> <p>Key concepts and criteria are precisely defined. The method is described clearly. It is adequate for addressing the questions. Methodological limitations are explained, as well as their consequences on the strength of conclusions, and on the substance of recommendations.</p>				
<p>3. Reliable data</p> <p>Data are sufficiently reliable with respect to the conclusions that are derived from them. Data collection tools have been applied in accordance to standards. Sources are quoted and their reliability is assessed. Potential biases are discussed.</p>				
<p>4. Sound analysis</p> <p>Data are cross-checked, interpreted and analysed systematically and appropriately. Underlying assumptions are clarified. The main external factors are identified and their influence taken into account.</p>				
<p>5. Valid findings</p> <p>The findings are based on evidence through a clear chain of reasoning. The limitations to validity are clearly stated.</p>				
<p>6. Impartial conclusions</p> <p>The conclusions are based on explicit criteria and benchmarks. They are free of personal and partisan considerations. Points of disagreement are reported truthfully. Lessons of wider interest are identified.</p>				
<p>7. Useful recommendations</p> <p>Recommendations stem from conclusions. They are applicable and detailed enough to be implemented by the addressees. The level of recommendations (political, strategic, managerial, etc.) reflects that of the questions.</p>				
<p>8. Clear report</p> <p>The style of the report is interesting for and accessible to the intended users. A short summary stresses the main findings, conclusions, lessons and recommendations in a balanced and impartial way.</p>				
<p>Overall assessment</p> <p>Taking into account the contextual constraints on the evaluation, the report satisfies the above criteria</p>				

