



Department of Performance Monitoring and Evaluation

Report on the Assessment of Government Evaluations

Evaluation Title:	Mid-term Evaluation of SARRAH
Evaluation Number:	409
Evaluation Completion Date:	31 December 2013
Period of Evaluation:	3 years
Submitted:	29 January 2014 by Wilma Wessels-Ziervogel
Approved:	

Evaluation Details

Evaluation Title:	Mid-term Evaluation of SARRAH
Evaluation Number:	409
Evaluation Completion Date:	31 December 2013
Created:	09 December 2013 by Mike Leslie
Submitted:	29 January 2014 by Wilma Wessels-Ziervogel
Approved:	
Period of Evaluation:	3 years
Known Cost:	
Known Cost Type:	No Data
Initiated By:	Department for International Development (DFID)
Initiated By Internal:	No
Undertaken By:	Coffey Development International and Human Sciences Research Council
Undertaken By Internal:	No

Assessors

Wilma Wessels-Ziervogel	wilma@southernhemisphere.co.za
-------------------------	--------------------------------

Assessment Documents

Document Name:	Document Type:	Added By:	Added On:
IMPACT EVALUATION OF SARRAH.PDF	Evaluation report	Mike Leslie	09 December 2013
IDEVREAN11003ZA_Mid-Term Evaluation Report_December 13.pdf	Evaluation report	Wilma Wessels-Ziervogel	29 January 2014
IDEVREAN11003ZA_Mid-Term Evaluation Report_December 13_Annexes.pdf	An inception report	Wilma Wessels-Ziervogel	29 January 2014
IDEVREAN11003ZA_NDOH_V1_5_Data collection tool.pdf	Assessment tools	Wilma Wessels-Ziervogel	29 January 2014
IDEVREAN11003ZA_NHI_V3_5_Data collection tool.pdf	Assessment tools	Wilma Wessels-Ziervogel	29 January 2014
IDEVREAN11003ZA_NHIRD_V2_5_Data collection tool.pdf	Assessment tools	Wilma Wessels-Ziervogel	29 January 2014
IDEVREAN11003ZA_QA_V2_5_Data collection tool.pdf	Assessment tools	Wilma Wessels-Ziervogel	29 January 2014
Complete Technical Proposal_240311[1].pdf	The proposal selected to conduct the evaluation	Wilma Wessels-Ziervogel	29 January 2014

Quality Assessment Summary

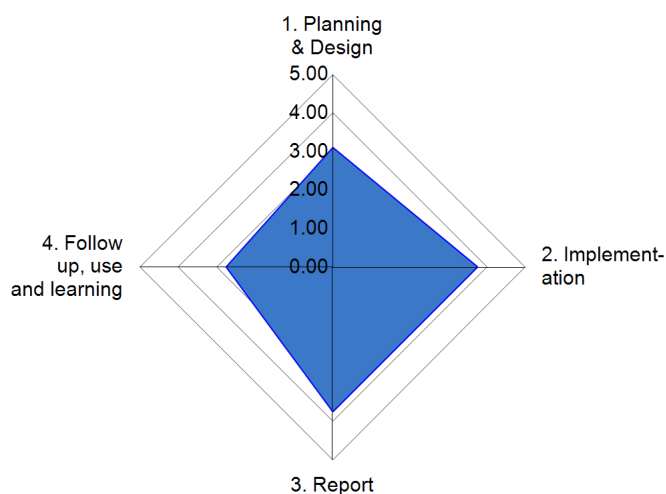
The evaluation received a score of 3.43 on the assessment.

Quality Assessment Scores

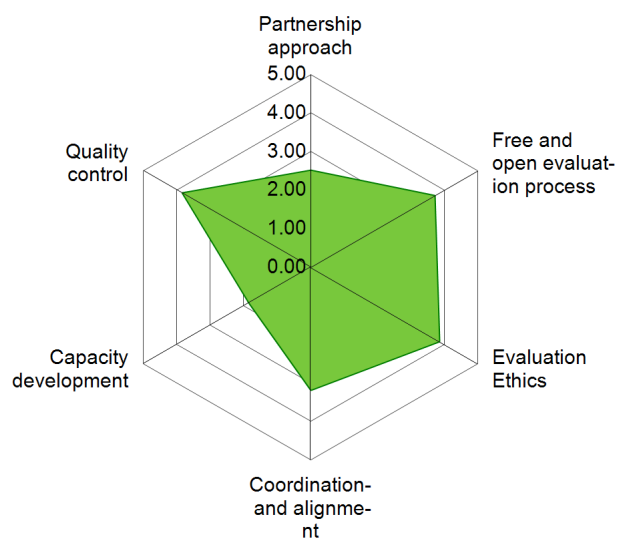
Phase of Evaluation	Score
1. Planning & Design	3.11
2. Implementation	3.77
3. Report	3.75
4. Follow-up, use and learning	2.77
Total	3.43

Overarching Consideration	Score
Partnership approach	2.52
Free and open evaluation process	3.73
Evaluation Ethics	3.87
Coordination and alignment	3.20
Capacity development	1.86
Quality control	3.85
Total	3.43

Scores: Phases of Evaluation



Scores: Overarching Considerations



Phase of Evaluation	Area of Evaluation	Score
1. Planning & Design	1.1. Quality of the TOR	2.25
1. Planning & Design	1.2. Adequacy of resourcing	2.71

Phase of Evaluation	Area of Evaluation	Score
1. Planning & Design	1.3. Alignment to policy context and background literature	5.00
1. Planning & Design	1.4. Appropriateness of the evaluation design and methodology	3.89
1. Planning & Design	1.5. Project management (Planning phase)	3.00
2. Implementation	2.1. Evaluation ethics and independence	4.64
2. Implementation	2.2. Participation and M&E skills development	2.60
2. Implementation	2.3. Methodological integrity	3.90
2. Implementation	2.4. Project management (Implementation phase)	3.00
3. Report	3.1. Completeness of reporting structure	4.42
3. Report	3.2. Accessibility of content	4.00
3. Report	3.3. Robustness of findings	4.42
3. Report	3.4. Strength of conclusions	2.93
3. Report	3.5. Suitability of recommendations	1.57
3. Report	3.6. Consideration of reporting risks and ethical implications	3.14
3. Report	3.7. Project management (Reporting phase)	3.00
4. Follow-up, use and learning	4.1. Resource utilisation	2.00
4. Follow-up, use and learning	4.2. Evaluation use	3.00
Total	Total	3.43

1. Planning & Design

1.1. Quality of the TOR

Standard: 1.1.1. The evaluation was guided by a well-structured and complete TOR or a well-structured and complete internal evaluation proposal

Comment and Analysis: The evaluation ToR was generally well structured although key evaluation questions were not documented.

Rating: 3

Moderation:

Standard: 1.1.2. The purpose of the evaluation stated in the TOR (or an internal evaluation proposal) was clear and explicit

Comment and Analysis: The purpose of the evaluation is stated in the ToR, however the purpose of the mid-term evaluation versus the overall impact evaluation were conflated. As a result mid-term evaluation objectives were later clarified in the inception phase.

Rating: 3

Moderation:

Standard: 1.1.3. The evaluation questions in the TOR (or an internal evaluation proposal) were clearly stated and appropriate for addressing the evaluation purpose

Comment and Analysis: Evaluation questions were not proposed in the ToR or the proposal.

Rating: 1

Moderation:

Standard: 1.1.4. The approach and type of evaluation was suited to the purpose and scope of the evaluation TOR (or an internal evaluation proposal)

Comment and Analysis: The evaluation was incorrectly called an impact evaluation in the ToR. This was later clarified as a mid-term evaluation and the scope adjusted accordingly. The mid-term evaluation as well as using a mixed method approach was well suited to the scope of the evaluation as it allowed for methods to be tailored to the specific structure/objectives and design of specific work streams. This was documented in the Inception Report.

Rating: 3

Moderation:

Standard: 1.1.5. The TOR (or an internal evaluation proposal) identified the intended users of the evaluation and their information needs

Comment and Analysis: The intended users and their information needs were not identified in the ToR.

Rating: 1

Moderation:

Standard:	1.1.6. Key stakeholders were involved in the scoping of the TOR and choosing the purpose of the evaluation
Comment and Analysis:	DFID was responsible in the scoping of the ToR and determining the evaluation purpose and scope. The NDoH (target group) were involved to a lesser degree/not involved as this was perceived to make the evaluation process and findings more objective and credible. Although this was the intention, the limited involvement of NDoH in planning the evaluation meant that there was some disagreement about the scope of what was to be assessed. For example, there was not a shared understanding of the objectives of the SARRAH programme or the criteria against which the SARRAH programme should be assessed. DG!!!! John- was HSLP involved in scoping ToR.
Rating:	2
Moderation:	

1.2. Adequacy of resourcing

Standard:	1.2.1. The evaluation was adequately resourced in terms of time allocated
Comment and Analysis:	Time allocated to the evaluation was not adequate, especially given the challenges in securing interviews with NDoH interviewees (Senior Management) and delays in receiving feedback on the report by DFID. The evaluation time frames were therefore extended by one year.
Rating:	2
Moderation:	

Standard:	1.2.2. The evaluation was adequately resourced in terms of original budget
Comment and Analysis:	TBD
Rating:	N/A
Moderation:	

Standard:	1.2.3. The evaluation was adequately resourced in terms of staffing and skills sets
Comment and Analysis:	The evaluation team consisted of Economists, Health Specialists, Monitoring and Evaluation Experts and Project Managers. This worked well for the evaluation.
Rating:	4
Moderation:	

Standard:	1.2.4. Where appropriate, the evaluation planned to incorporate an element of capacity building of partners/staff responsible for the evaluand
Comment and Analysis:	No capacity building was planned for the evaluand.
Rating:	1
Moderation:	

1.3. Alignment to policy context and background literature

Standard:	1.3.1. There was evidence that a review of the relevant policy and programme environments had been conducted and used in planning the research
Comment and Analysis:	An extensive review was undertaken which included: a review of the health policy context, a historical overview of DFID's work (within the South African health sector), a review of the South African health sector (1970's-2010) and the context of the individual work streams evaluated.
Rating:	5
Moderation:	

Standard:	1.3.2. There was evidence of a review of appropriate literature having been conducted and used in planning the research
Comment and Analysis:	Existing literature was reviewed as part of the evaluation, which included a number of articles in South African and International journals.
Rating:	5
Moderation:	

1.4. Appropriateness of the evaluation design and methodology

Standard:	1.4.1. There was explicit reference to the intervention logic or the theory of change of the evaluand in the planning of the evaluation
Comment and Analysis:	Much time was spent on developing/understanding the theory of change during the inception phase. This was developed through an initial Theory of Change workshop with DFID and the managing agent (HSLP) and subsequent workshops with the managing agent.
Rating:	5
Moderation:	

Standard:	1.4.2. Key stakeholders were consulted on the design and methodology of the evaluation
Comment and Analysis:	DFID and HLSP were mostly involved in determining the evaluation design and methodology. The NDoH were consulted to a lesser degree to ensure the highest level of objectivity. DG!!!!
Rating:	3
Moderation:	

Standard:	1.4.3. The planned methodology was appropriate to the questions being asked
Comment and Analysis:	The methodology employed was appropriate for the evaluation questions which were based on the OECD-DAC (effectiveness, efficiency, impact, relevance and sustainability). A mixed method approach was used for each component/work stream being evaluated. This usually included desktop research, secondary data collation, key informant interviews, focus groups, site visits and surveys. After each work stream was evaluated, a quality review was conducted on each of the individual work stream evaluations and findings/scores adjusted accordingly. A synthesis of the evaluation findings was then conducted and analysed against key evaluation criteria.
Rating:	5
Moderation:	

Standard:	1.4.4. Sampling was appropriate and adequate given focus and purpose of evaluation
Comment and Analysis:	The sample of work streams to be evaluated were carefully selected/prioritised based on their "utility, proportionality and evaluability". This guided the level of effort in evaluating the work stream and is a good tool that could be used for systematically sampling complex programmes. The sample of individuals was adequate as it included NDoH representatives, regional DoH, CSOs, SARRAH appointed consultancies, HLSP and DFID.
Rating:	5
Moderation:	

Standard:	1.4.5. There was a planned process for using the findings of the evaluation
Comment and Analysis:	There was no clear plan for how the evaluation was to be utilised. DG!!!!
Rating:	1
Moderation:	

1.5. Project management (Planning phase)

Standard:	1.5.1. The inception phase was used to develop a common agreement on how the evaluation would be implemented
Comment and Analysis:	During the inception phase there was much consultation between the evaluator, DFID and HLSP to arrive at a common understanding about how to pursue this evaluation. DFID left much space for negotiating this upfront. While a common agreement was reached on this level, the limited involvement of NDoH meant that stakeholders did not necessary agree on the evaluation plan.
Rating:	3
Moderation:	

2. Implementation

2.1. Evaluation ethics and independence

Standard:	2.1.1. Where data was gathered in contexts where ethical sensitivity is high, appropriate clearance was achieved through an ethics review board; e.g. in evaluation involving minors, institutions where access usually requires ethical or bureaucratic clearance, and situations where assurances of confidentiality was offered to participants
Comment and Analysis:	Evaluation ethics were adhered to. The proposal was also reviewed and approved by the HSRC Ethics committee.
Rating:	5
Moderation:	

Standard:	2.1.2. Where external, the evaluation team was able to work freely without significant interference
Comment and Analysis:	The evaluation team worked freely without significant interference from the client, managing agent or the Department.
Rating:	5
Moderation:	

Standard:	2.1.3. The evaluation team was impartial and there was no evidence of conflict of interest
Comment and Analysis:	The evaluation team seemed impartial, although one evaluator was perceived by the Department as having a conflict of interest. DG?!!!!!!
Rating:	4
Moderation:	

2.2. Participation and M&E skills development

Standard:	2.2.1. Key stakeholders were consulted through a formalised mechanism or institutional arrangement
Comment and Analysis:	A steering committee was set up comprising of the evaluation partners (i.e. NDoH, SANAC, and TAC), Monitoring and Evaluation Specialists (DFID and DPME), and the evaluation service providers (HSRC and Coffey International). The purpose of this committee was to provide input on the evaluation strategy, progress and preliminary findings. Although this structure had been set up, meetings have not been regular due to the time constraints of stakeholders. It seems that the NDoH representative was only invited onto the committee at a later stage when problems of access to the Department surfaced. DG!!!!!!
Rating:	3
Moderation:	

Standard:	2.2.2. Where appropriate, an element of capacity building of partners responsible for the evaluand was incorporated into the evaluation process
Comment and Analysis:	No capacity building was implemented for partners.
Rating:	1
Moderation:	
Standard:	2.2.3. Where appropriate, the evaluation team incorporated an element of skills development amongst the evaluators (e.g. students, interns, recent graduates, etc)
Comment and Analysis:	An initial workshop was conducted on brainstorming the methodology options for this evaluation. Junior staff were invited to this discussion and in this way received some capacity development.
Rating:	3
Moderation:	
Standard:	2.2.4. Peer review of the agreed evaluation design and methodology occurred prior to undertaking data collection
Comment and Analysis:	A peer review of the evaluation design of all work streams was undertaken by DFID Monitoring and Evaluation Department prior to the implementation of the evaluations.
Rating:	4
Moderation:	

2.3. Methodological integrity

Standard:	2.3.1. The methods employed in the process of the evaluation were consistent with those planned
Comment and Analysis:	Methods were mostly implemented as planned. There were however challenges in gaining access to documentation required for the evaluation.
Rating:	4
Moderation:	
Standard:	2.3.2. A pilot of data collection instrumentation occurred prior to undertaking data collection
Comment and Analysis:	John ????
Rating:	N/A
Moderation:	

Standard:	2.3.3. Data collection was not compromised by fieldwork-level problems or unplanned diversions from original intentions
Comment and Analysis:	The lack of availability of NDoH interviewees and lack of access to some of the necessary documents compromised the data collection process and resulted in the fieldwork process being delayed and some gaps in the data.
Rating:	3
Moderation:	
Standard:	2.3.4. Forms of data gathering were appropriate given the scope of evaluation
Comment and Analysis:	Multiple methods were used to collect evaluation data and methods were tailored to every work stream being evaluated. This was appropriate for the wide scope of the evaluation.
Rating:	5
Moderation:	
Standard:	2.3.5. The data analysis approach and methods were appropriate and sufficient given the purposes of the evaluation
Comment and Analysis:	The OECD-DAC criteria was used as the framework for data analysis. A thematic analysis was undertaken using Atlas.ti as the data analysis tool.
Rating:	4
Moderation:	
Standard:	2.3.6. Key stakeholders were significantly engaged as part of the methodology
Comment and Analysis:	Stakeholders relevant for each work stream were engaged as part of the data collection process. This included NDoH, regional DoH, CSOs, SARRAH-appointed consultancies, HLSP and DFID.
Rating:	4
Moderation:	
Standard:	2.3.7. The methodology included engaging beneficiaries appropriately as a key source of data and information
Comment and Analysis:	Since the evaluation focussed on impact to a lesser degree, beneficiaries (i.e. those using health care services) were not engaged as part of the process. JOHN!!!!
Rating:	3
Moderation:	

2.4. Project management (Implementation phase)

Standard:	2.4.1. The evaluation was conducted without significant shifts to scheduled project milestones and timeframes
Comment and Analysis:	Time frames were shifted due to the challenges of securing interviews with the NDoH as well as delays in feedback from DFID on the report. This was not however due to challenges within the project management and was successfully negotiated with the donor/client though.
Rating:	3
Moderation:	

3. Report

3.1. Completeness of reporting structure

Standard:	3.1.1. Executive summary captures key components of the report appropriately
Comment and Analysis:	The executive summary provided a good overview of the evaluation report and captured key elements such as the introduction, evaluation approach and method, key findings, lessons learnt and recommendations. It may have been useful to highlight the key strengths and challenges of work streams in terms of the OECD-DAC criteria, since these findings were a key component of the report.
Rating:	4
Moderation:	
Standard:	3.1.2. The context of the development intervention is explicit and presented as relevant to the evaluation
Comment and Analysis:	The context is described within the main body of the report and extensively described in Annex 1 where an overview of DFID and the South African health sector from 1970s to 2010 is provided.
Rating:	5
Moderation:	
Standard:	3.1.3. There is a clear rationale for the evaluation questions
Comment and Analysis:	Evaluation questions are clearly linked to the OECD-DAC criteria used for evaluating development interventions. Since this was a process/mid-term evaluation, the DAC criteria were appropriate to use as it explored all the key elements required (impact, effectiveness, relevance, sustainability and efficiency).
Rating:	4
Moderation:	
Standard:	3.1.4. The scope or focus of the evaluation is apparent in the report
Comment and Analysis:	The evaluation scope in terms of key questions and work streams included and excluded from the evaluation were articulated in the report. A clear rationale was provided for the level of effort exerted in evaluating each work stream, and also the exclusion of certain work streams.
Rating:	5
Moderation:	

Standard:	3.1.5. A detailed methodology is outlined in the relevant section of the report to the point that a reader can understand the data collection, analysis and interpretation approaches used
Comment and Analysis:	Data collection, analysis and interpretation were well documented. A mixed method approach was used in most of the individual evaluations. The standard of quality of findings was then assessed based on the availability/quality of evidence. As part of the synthesis process, a numerical tool was used to summarise the the qualitative data for each work stream against each of the DAC criteria. The weighting used for this tool was clearly described. Atlas.ti data analysis software package was used for analysing qualitative data. Although thematic analysis was used, this was not made explicit in the report.
Rating:	5
Moderation:	
Standard:	3.1.6. Acknowledgement of limitations of all aspects of the methodology and findings are clearly and succinctly articulated
Comment and Analysis:	The limitations of the evaluation were considered under the "research challenges" section. Including the limitations of the performance scoring system would have been useful. A particular "limitation" could perhaps be that scores are negatively skewed when measurement on a particular criterion is premature (for example when it is too soon to measure impact). The scoring system does not make the criterion irrelevant (or not applicable) in this instance, but instead considers the work stream as "performing poorly" or "not performing well".
Rating:	3
Moderation:	
Standard:	3.1.7. Key findings are presented in a clear way; they are made distinct from uncertain or speculative findings; and unused data is not presented in the body of the report
Comment and Analysis:	Key findings were well structured and clearly presented using the DAC criteria as a framework. The findings for each work stream was contextualised by providing an overview of the problem it was trying to address, the intervention of SARRAH to address the problem and then the key evaluation findings according to DAC criteria. Uncertain findings were made explicit, especially through the "Evidence Rating System" tool but also in the narrative report.
Rating:	5
Moderation:	
Standard:	3.1.8. Conclusions and recommendations are clear and succinctly articulated
Comment and Analysis:	Conclusions and recommendations are clear and succinct. Conclusions were structured based on the OECD-DAC criteria which therefore followed the framework used in the findings session, and was thus clear. Key recommendations were presented succinctly and then expanded on in a more narrative section.
Rating:	4
Moderation:	

3.2. Accessibility of content

Standard:	3.2.1. The final evaluation report is user-friendly, written in accessible language and its content follows a clear logic
Comment and Analysis:	The evaluation report was very user-friendly. It utilised clear headings, explained the process (which was complex) in enough detail to understand how the evaluation unfolded, and findings were clear and concise.
Rating:	4
Moderation:	
Standard:	3.2.2. Quality of writing and presentation is adequate for publication including: adequate layout and consistent formatting; complete sentences and no widespread grammatical or typographical errors; consistency of style and writing conventions (e.g. tense, perspective (first person, third person); levels of formality; references complete and consistent with cited references in reference list and vice versa; etc.
Comment and Analysis:	The evaluation report is definitely of publishing standard in terms of layout, formatting, grammar, writing and references.
Rating:	4
Moderation:	
Standard:	3.2.3. Appropriate conventions are used in presentation of data (e.g. use of appropriate statistical language; reporting of p-values where appropriate; not reporting statistically insignificant findings as significant; clarifying disaggregation categories in constructing percentages; not using quantitative language in reporting qualitative data, etc.)
Comment and Analysis:	Appropriate conventions were used for this evaluation. Since primary quantitative data was not collected, statistical language was not used.
Rating:	4
Moderation:	
Standard:	3.2.4. The use of figures and tables is such that it supports communication and comprehension of results; and data reported in figures and tables are readily discernible and useful to a reader familiar with data presentation conventions
Comment and Analysis:	Tables and diagrammes were used throughout the report and this helped delineate findings, challenges, scores, etc. for each of the work streams evaluated. This helped simplify the complex nature of the programme. All tables/diagrammes had clear headings and references to sources that they are drawn from.
Rating:	4
Moderation:	

3.3. Robustness of findings

Standard:	3.3.1. Data analysis appears to have been well executed
Comment and Analysis:	Data analysis seemed to have been done very systematically. Appropriate methods (thematic analysis) and tools (Atlas.ti) were used to analyse data. The "Performance Scoring System" tool helped simply analyse and present data for each of the work streams against each of the indicators. This tool could be useful for future evaluations of complex programmes.
Rating:	5
Moderation:	
Standard:	3.3.2. Findings are supported by available evidence
Comment and Analysis:	Findings were definitely supported by evidence, and where this was lacking the strength of the finding was adjusted. The tool used to test the quality/robustness of evidence (Evidence Rating System) of each criterion provides evaluation practitioners with a good framework for assessing the credibility and evidence-base of evaluation findings.
Rating:	5
Moderation:	
Standard:	3.3.3. The evidence gathered is sufficiently and appropriately analysed to support the argument
Comment and Analysis:	Data was presented through providing an overview of how work streams performed in terms of the DAC criteria. Examples of evidence and quotes were provided where necessary to support the argument.
Rating:	4
Moderation:	
Standard:	3.3.4. There is appropriate recognition of the possibility of alternative interpretations
Comment and Analysis:	This was not observed in the report.
Rating:	N/A
Moderation:	
Standard:	3.3.5. The report appears free of significant methodological and analytic flaws
Comment and Analysis:	No methodological and analytic flaws are evident in this evaluation.
Rating:	4
Moderation:	

Standard:	3.3.6. Relevant limitations of the evaluation are noted
Comment and Analysis:	The limitations of the evaluation were considered under the "research challenges" section. Including the limitations of the performance scoring system would have been useful. A particular "limitation" could perhaps be that scores are negatively skewed when measurement on a particular criterion is premature (for example when it is too soon to measure impact). The scoring system does not make the criterion irrelevant (or not applicable) in this instance, but instead considers the work stream as "performing poorly" or "not performing well". It did however help that where measurement was not possible or premature, this was explained in the narrative section.
Rating:	4
Moderation:	

3.4. Strength of conclusions

Standard:	3.4.1. Conclusions are derived from evidence
Comment and Analysis:	Conclusions were clearly based on the evidence in the findings section.
Rating:	4
Moderation:	

Standard:	3.4.2. Conclusions take into account relevant empirical and/or analytic work from related research studies and evaluations
Comment and Analysis:	The conclusions mainly focussed on the empirical/analytic work of the SARRAH programme.
Rating:	2
Moderation:	

Standard:	3.4.3. Conclusions address the original evaluation purpose and questions
Comment and Analysis:	Conclusions were based on the they key evaluation questions linked to the OECD-DAC criteria, which were in turn linked to the objectives of the evaluation. The link to the objectives of the evaluation could have been brought out more explicitly though.
Rating:	4
Moderation:	

Standard:	3.4.4. Conclusions are drawn with explicit reference to the intervention logic or theory of change
Comment and Analysis:	The theory of change was not articulated or reflected on in the conclusions.
Rating:	1
Moderation:	

3.5. Suitability of recommendations

Standard: 3.5.1. Recommendations are made in consultation with appropriate sectoral partners or experts

Comment and Analysis: John?

Rating: N/A

Moderation:

Standard: 3.5.2. Recommendations are shaped following discussions with relevant government officials and other relevant stakeholders

Comment and Analysis: Discussion of proposed recommendations has not taken place with government or DFID/HLSP. A presentation of the findings and recommendations will happen in the upcoming months, however the report has already been finalised. Stakeholder consultation happened via circulating the individual work stream reports for comment. While the service provider reports limited response, a stakeholder from the NDoH reports that limited information was shared regarding the findings. DG!!! John

Rating: 1

Moderation:

Standard: 3.5.3. Recommendations are relevant to the current policy context

Comment and Analysis: Recommendations formulated in the evaluation synthesis did not seem to be on a policy level. An interviewee articulated that recommendations were not strategic enough.

Rating: 2

Moderation:

Standard: 3.5.4. Recommendations are targetted at a specific audience sufficiently - are specific, feasible, affordable and acceptable

Comment and Analysis: DG !!!!!

Rating: N/A

Moderation:

3.6. Consideration of reporting risks and ethical implications

Standard: 3.6.1. Peer review of the draft evaluation report occurred prior to finalisation of the evaluation report

Comment and Analysis: John!!! The report was extensively reviewed by DFID - M&E department?

Rating: N/A

Moderation:

Standard:	3.6.2. The full report documents procedures intended to ensure confidentiality and to secure informed consent where necessary (in some cases this is not needed - e.g. evaluation synthesis - in which case N/A should be recorded)
Comment and Analysis:	Although the evaluation adhered to high ethical standards, the evaluation report does not report on ethical considerations.
Rating:	2
Moderation:	

Standard:	3.6.3. There are no risks to participants in disseminating the original report on a public website
Comment and Analysis:	There are no risks to participants and the report is already available publicly.
Rating:	4
Moderation:	

Standard:	3.6.4. There are no unfair risks to institutions in disseminating the original report on a public website
Comment and Analysis:	There are no unfair risks to institutions and the report is already available publicly.
Rating:	N/A
Moderation:	

3.7. Project management (Reporting phase)

Standard:	3.7.1. A project closure meeting that reflected on the challenges and strengths of the evaluation process occurred
Comment and Analysis:	A project closure meeting has not been held as yet, but is planned to take place at the end of the project. Ongoing reflections among Senior Evaluators have however taken place.
Rating:	3
Moderation:	

4. Follow-up, use and learning

4.1. Resource utilisation

Standard:	4.1.1. The evaluation was completed within the planned timeframes
Comment and Analysis:	The evaluation period was extended by one year due to delays in setting up fieldwork and obtaining feedback on the report from the funder (DFID).
Rating:	2
Moderation:	

Standard:	4.1.2. The evaluation was completed within the agreed budget
Comment and Analysis:	TBD (Samy)
Rating:	N/A
Moderation:	

4.2. Evaluation use

Standard:	4.2.1. Results of the evaluation have been presented to all relevant stakeholders
Comment and Analysis:	Results have been presented to one Cluster Manager and her unit in response to her request. Presentations are planned for the first quarter of 2014.
Rating:	3
Moderation:	

Standard:	4.2.2. A reflective process has been undertaken by the steering committee (if no steering committee exists then by the evaluation management team or the involved department officials) to reflect on what could be done to strengthen future evaluations
Comment and Analysis:	This has not yet been done. A reflection session is however planned for the evaluation team.
Rating:	3
Moderation:	

Standard:	4.2.3. The evaluation study is seen by interviewed stakeholders as having added significant symbolic value to the policy or programme (e.g. raised its profile)
Comment and Analysis:	One NDoH stakeholder indicated that recommendation were not strategic enough, and did not sufficiently address policy issues. There was however a couple that were valid, although the relevance of the findings (due to the timing of the evaluation), was sometime questionable in terms of the value that it added to the programme as very often recommendation made were already under-way.
Rating:	3
Moderation:	

Standard:	4.2.4. The evaluation study is of conceptual value in understanding what has happened and possibly in shaping policy and practice
Comment and Analysis:	The evaluation was used to improved the programme to some degree. One example is that the evaluation pointed to the inefficiencies at facility level to use the standards developed by the Quality Assurance Department. DG!!!
Rating:	3
Moderation:	
Standard:	4.2.5. Development of a draft improvement plan has been started, but not completed, based on the findings and recommendations set out in the evaluation
Comment and Analysis:	DG!!!
Rating:	N/A
Moderation:	
Standard:	4.2.6. The report is publicly available (website or otherwise published document), except where there were legitimate security concerns *Note: only apply if sufficient time has elapsed since completion of the evaluation
Comment and Analysis:	The report is not yet available. John - why?
Rating:	N/A
Moderation:	
Standard:	4.2.7. There is clear evidence of instrumental use - that the recommendations of the evaluation were implemented to a significant extent *Note: only apply if sufficient time has elapsed since completion of the evaluation
Comment and Analysis:	Since the evaluation report has not yet been formally presented to the NDoH, it is not possible to make this assessment. DG!!!!
Rating:	N/A
Moderation:	
Standard:	4.2.8. There is clear evidence that the evaluation has had a positive influence on the evaluand, its stakeholders and beneficiaries over the medium to long term *Note: only apply if sufficient time has elapsed since completion of the evaluation
Comment and Analysis:	Since evaluation results have not been properly disseminated it is difficult to make this assessment. DG
Rating:	N/A
Moderation:	

References

Coffey International Development. 2013. Mid-term Evaluation SARRAH: Final Report Annexes.

Coffey International Development. 2011. Inception Report.

Coffey International Development. 2011. General and Technical Proposal.

Coffey International Development. Data Collection Tool: Restructuring Review Checklist.

Coffey International Development. Data Collection Tool: Quality Assurance Information Checklist.

Coffey International Development. Data Collection Tool: NHI Information Checklist.

Coffey International Development. Data Collection Tool: NHIRD Document Review Checklist.

Coffey International Development. 2013. Mid-term Evaluation SARRAH: Final Report.

List of Interviewees

Prof. John Seager, Lead Evaluator: Coffey Development International. Telephonic Interview. 22 January 2014.

Dr. Carol Marshall, Cluster Manager: Office of Standards Compliance, National Department of Health. Telephonic Interview. 21 January 2014.