

1. Overview

Growth and Employment in States		
GEMS1 – Meat & Leather	2010-2015	
Audit dates	17 th – 21 st November 2014	
Overall final ratings ¹	MUST	399/470=85%
	RECOMMENDED	129/160=80%
Coverage	Full Audit	
	All control points checked	
DCED Standard	Version VI, January 2013	

Signed:



29th January 2015.

GEMS 1 Kevin Billing, Team Leader

Date, Place

Kano, Nigeria.



9 February 2015
Auckland New Zealand

Auditor Alexandra Miehlbradt

Date, Place

Table of Contents

1. Overview	1
2. Key Audit Findings	3
3. Summary of the Programme and Results Measurement System	6
4. Summary of the Audit Process	9
5. Detailed Scoring of the Control Points	11
6. Summary of Areas that Require Improvement	18
Annexes	19

¹ An overall rating of 100% implies that the project meets the compliance criteria and has a strong measurement system of acceptable quality within the boundaries of what the programme has set itself to measure, not that it is has a perfect measurement system. The totals and percentages are based on not rounded figures from Annex 1 and results have been rounded for inclusion in the main report.

Acronyms

A&B	Advocacy and Business Membership Organisation Development
BDO	Business Development Officer
BMO	Business Membership Organisation
DCED	Donor Committee for Enterprise Development
FF	Feed Finishing
FLG	Finished Leather Goods
GEMS 1	Growth and Employment in States – GEMS 1 Meat & Leather
ICF	Intervention Control Framework
IM	Intervention Manager
M4P	Making Markets Work Better for the Poor
MRM	Monitoring and Results Measurement
RC	Results Chain
SQ	Skin Quality

2. Key Audit Findings

Articulating the Results Chain	
<ul style="list-style-type: none"> • Results chains have been articulated for all interventions. They are mostly logical and thorough. • Most results chains are supported by adequate research and analysis. • Programme staff are familiar with the results chains and can give specific examples of how they use them to guide their activities. • There is a system to review results chains and evidence of some review. • Results chains include systemic change when appropriate. 	<ul style="list-style-type: none"> • There are some discrepancies in the logic and detail of the results chains. • A few results chains are not supported by adequate and documented research and analysis, including analysis of the likelihood of sustainability. • Some partners cannot adequately describe the logic of interventions. • The system for reviewing results chains is not always applied consistently and thoroughly with adequate documentation.
Defining Indicators of Change	
<ul style="list-style-type: none"> • There are indicators defined for almost all changes in results chains. Many indicators are relevant and appropriate. • The universal impact indicators are included. • In practice, GEMS1 gathers information on indicators of sustainability. • Programme staff clearly understand indicators and can give specific examples of how they illustrate programme progress. • Projections are made and updated when sufficient information is available. 	<ul style="list-style-type: none"> • A few changes are missing indicators. Some indicators are inappropriate or insufficiently clear. • Indicators of sustainability are often not documented and sometimes not thorough. • Some projections are not clearly and adequately supported.
Measuring Changes indicators	
<ul style="list-style-type: none"> • Sector information and some baseline information are gathered. • There is a documented plan to collect information for each intervention. Staff are able to provide details on how studies have been or will be conducted. • Studies use mostly good practices, triangulation and quality control in data collection. • Qualitative information is gathered; it is usually thorough and appropriate. 	<ul style="list-style-type: none"> • There is usually no clear and documented plan to gather baseline information. Baseline information gathered is often inadequate. • Documented plans for information collection are often not sufficiently clear and detailed. • There is inadequate documentation of studies. • Some sample sizes are insufficient. There is too little quality control in data analysis and some calculations lack clarity. • Qualitative information is sometimes not documented and occasionally not thorough.

Estimating Attributable Changes	
<ul style="list-style-type: none"> There is a plan to assess attribution for each intervention. The plans have usually been carried out and occasionally augmented. Assessment is sometimes in accordance with good practices. 	<ul style="list-style-type: none"> The methodologies to assess attribution are sometimes not sufficient given the other factors that may affect indicators. Some practices in assessment of attribution were inadequate; calculations are not always clear.
Capturing Wider Changes in the System or Market	
<ul style="list-style-type: none"> There are documented plans for assessing wider systemic changes. The plans and implementation of them are mostly adequate and in accordance with good practice. 	<ul style="list-style-type: none"> Some of the plans for assessing wider systemic change are not sufficiently clear. The studies sometimes suffered from inadequate samples and insufficient quality control and clarity in data analysis.
Tracking Programme Costs	
<ul style="list-style-type: none"> Costs are adequately tracked annually and cumulatively. Costs, including overhead, are logically and clearly allocated per output. 	
Reporting Results	
<ul style="list-style-type: none"> The programme produces quarterly and annual reports with clear estimates of programme wide impacts and context. The programme has a well-designed aggregation system. Overlap is taken into consideration in aggregation. Contributions from other public and private entities are acknowledged. Outreach is disaggregated by gender. The results of systemic change are reported. 	<ul style="list-style-type: none"> At the time of the audit, the aggregation system had not yet been used for reported figures. Aggregation of figures to date was not yet systematic and clear. Some results are not divided by direct and indirect. Annual results are not published.
Managing the System for Results Measurement	
<ul style="list-style-type: none"> There is a system and supportive culture for using information on results in management. Managers can give many clear examples of using information in decision-making. The system is now supported by sufficient human and financial resources. Staff are clear on their results measurement roles and responsibilities. The system has been institutionalised. 	<ul style="list-style-type: none"> The system for using information on results is not adequately documented. The written MRM guidance does not cover all necessary topics. Management roles in quality control are not sufficiently robust.

Final ratings

“Must” control points:

Percentage	Description	Programme Rating
91-100	Strong results measurement system	
81-90	Reasonable results measurement system	√
71-80	Moderate results measurement system	
61-70	Moderate results measurement system with notable weaknesses	
51-60	Weak results measurement system	
41-50	Weak results measurement system	
31-40	Weak results measurement system	
21-30	Weak results measurement system	
11-20	Weak results measurement system	
0-10	Weak results measurement system	

“Recommended” control points:

Percentage	Description	Programme Rating
81-100	Results measurement system with strong additional features	
61-80	Results measurement system with some additional features	√
41-60	Results measurement system with some additional features	
21-40	Results measurement system with few additional features	
0-20	Results measurement system with few additional features	

3. Summary of the Programme and Results Measurement System

Summary of GEMS1

The Growth & Employment in States (GEMS) program in Nigeria is a partnership among the Federal Government of Nigeria, the Department for International Development (DFID) and the World Bank, which aims to reduce poverty in Africa's most populous country. Its goal is increased growth, incomes and jobs in selected states. The purpose of the program is an improved business environment in selected states.²

GEMS1 is one component programme under the overall GEMS programme. It is funded solely by DFID and implemented by GRM. GEMS1 focuses on the meat and leather sectors. The programme uses a Making Markets Work Better for the Poor (M4P) approach in order to bring about sustainable improvements in market outcomes by altering the incentives that the market system offers to participants. GEMS1 is a five-year programme that started in April 2010 and is scheduled to end in September 2015. The programme has had two contract amendments that have increased both the budget and targets; the current budget is GBP8.79M.

Objectives of GEMS1

The GEMS1 objective is "to create sustainable improvements in the Meat and Leather Industries which generate increased growth, incomes and employment."³ GEMS1 has a strong focus on fostering systemic changes that will underpin continued pro-poor growth of the sectors after the program has ended. Over the life of the programme, GEMS1 aims to increase the profits of 33,000 enterprises representing 100,000 working people, increase incomes by GBP24.7M and generate 4,400 additional jobs.

Focus of GEMS1

GEMS1 focuses on three sectors: meat, leather and cross-cutting issues. GEMS1 divides its work in these three sectors into areas called "outputs" as follows:

Sector	Outputs
Meat	Output 1: Feed Finishing
	Output 2: Meat Processing
Leather	Output 3: Finished Leather Goods
	Output 4: Finished Leather
	Output 5: Skin Quality
Cross-Cutting	Output 6: Advocacy and BMO Development
	Output 7: Access to Finance

For each output, GEMS1 has identified opportunities to increase industry competitiveness and promote pro-poor growth by addressing key constraints.

History of Results Measurement at GEMS1

GEMS1 was established under a "design and implement" contract in March 2010 and thus began with a 9-month inception phase. The consultant responsible for results measurement during the inception phase introduced the DCED Standard into the program and initiated the establishment of a

² GEMS1 Mid Term Report 2014, May 2014

³ Ibid.

results measurement system based on the framework of the Standard.

The program was approved to start implementation in March 2011. In June 2011, GEMS1 contracted a full-time results measurement manager/adviser familiar with the Standard, who had already completed two short assignments with GEMS1 on results measurement. She more fully designed and began to operationalize the results measurement system in GEMS1.

In late 2011 and 2012, GEMS1 was involved in two exercises encompassing all of the GEMS programs. The first was the establishment of a GEMS portfolio logframe and the adjustment of the GEMS1 logframe to fit within this overall logframe. The second was the development of an overall GEMS results measurement guide. GEMS1 contributed to this guide and tweaked its system to comply with the guide. These two activities required significant involvement of the GEMS1 team, particularly the results measurement manager/adviser.

Late in 2012, GEMS1 commissioned a pre-audit review of its results measurement system to assess the system and provide recommendations for improvement.

During 2012-2013, GEMS1 was gradually able to increase the amount of time that the results measurement manager/adviser spent operationalizing the results measurement system in the program. A key part of this was the manager/adviser dedicating considerable time to building the capacity of the intervention managers in key aspects of the results measurement system. During this time, GEMS1 also regularly brought in a consultant to provide short-term assistance in results measurement as well as an experienced local researcher to help with specific studies.

However, GEMS1 realized that these arrangements did not provide enough human resources to adequately assess, analyse, report on and use results in decision-making. Therefore, after the second contract amendment was finalized late in 2013, GEMS1 contracted the consultant and experienced researcher as full-time staff, increasing the results measurement team to three full-time positions. This team now supports a growing staff and programme, which is expanding into additional states in Northern Nigeria.

This history has meant that the results measurement system in GEMS1 has developed gradually as resources were added to enable the system to be fully operationalized and embedded in the program. The additional resources and management commitment have enabled improvements in capacity and upgrades to the system over time. The lack of sufficient human resources early in the program adversely affected the quality of results measurement for early interventions. However, the steady improvements in the results measurement system have led to later interventions enjoying better results measurement than earlier ones. It is apparent that the team is continuing to improve the results measurement system with several significant enhancements just started or planned. This on-going commitment to improving results measurement is commendable.

Key features of the GEMS1 results measurement system

The heart of the GEMS1 results measurement system is the intervention control framework (ICF). This document, prepared for each intervention, summarizes a planned intervention and its expected results. The ICF includes a cover page summarizing the background and key features of the intervention, an intervention results chain, intervention measurement plan and projections of expected results with supporting calculations when possible. The ICF is also used to track the progress of the intervention and results. The projections are updated with estimated results as information is gathered on changes at each level of the results chain. In addition, key aspects of progress and decisions are kept in the ICF, as well as a summary of the data gathered on the intervention results.

The results measurement team leads time-bound assessments of intervention results, primarily at the level of target enterprises, but sometimes also at the level of programme partners and/or service providers. The results measurement team designs and manages data collection and analysis; enumerators are contracted for larger studies. These time-bound assessments are complimented by regular information gathering by the intervention manager and business development officers. The results measurement team triangulates information from various sources in order to estimate the results of the interventions at the each level of an intervention results chain. The results measurement team aggregates the estimated results from each of the interventions quarterly for reporting to DFID. Generally, GEMS1 conducts several assessments of each intervention to check both for results and for the sustainability of those results. The programme accrues sustainable results for two years after an intervention, after which no further results for that intervention are reported.

GEMS1 reports results once an initial assessment has been done. However, further assessments and data analysis sometimes lead to an adjustment of estimated results, which are then included in the following quarterly report. GEMS1 has taken advantage of multiple assessments and additional analysis to improve the quality of its estimated results over time.

Interventions in the cross-cutting output on Advocacy and BMO Development, are considered “pre-condition” interventions that support other interventions. As such, these interventions are only monitored at the lower levels of the results chain. When work in these interventions achieves a specific result, such as opening access to a particular source of credit or enabling a BMO to offer a specific service to members, a new intervention results chain is developed focused on that specific activity. Higher-level results are then measured through this new intervention.

Using information on results in management

Information on results is shared as it is gained through meetings among the intervention teams, results measurement team and managers. There is a constant focus on using this information to improve interventions, develop new interventions and discontinue ineffective interventions. This regular interaction is punctuated by the quarterly reporting cycle and more formal “technical review meetings” which occur approximately twice a year.

Management structure

A team leader heads GEMS1, supported by a deputy team leader and a part-time technical director. The deputy team leader and five intervention managers lead implementation in the various outputs. They are supported by eight business development officers. The results measurement team, consisting of a manager/RM adviser, an RM officer and a field research officer, supports the implementation team. GEMS1 brings in technical consultants to assist in addressing specific aspects of intervention design and implementation. In addition, GEMS1 works closely with government and private sector partners. However, all results measurement is managed in-house. While GEMS1 gathers information from partners, the partners are not responsible for any data gathering beyond their normal operations. Generally, information from partners is cross-checked with other sources.

4. Summary of the Audit Process

The audit reviewed a representative sample of all current and past interventions in GEMS1. It was, thus, a full audit. The auditor arrived at the representative sample using a 2-stage combination of purposeful and stratified random sampling.

GEMS1 has 63 interventions across its three sectors – meat, leather and cross-cutting. By calculating the square root of the interventions, the auditor determined that (at least) eight interventions should be reviewed. GEMS1 classifies its interventions into the following categories: idea, proof of concept, planned, foundation, on going, complete, on hold and discontinued. These categories were checked to determine which were appropriate for sampling in the audit. Three categories – idea, proof of concept and on hold – were excluded from the eligible sample of interventions, as they were not developed enough to have sufficient documentation for the audit. Essentially, these categories represented ideas that had not yet been turned into interventions. With this exclusion there were 54 interventions eligible for the audit. The table below shows the sectors, outputs, stage of implementation, number of interventions and number of interventions eligible for the audit.

Sector and Output	Title	Stage of implementation	Number of interventions	Number of eligible interventions
Meat Sector				
Output 1	Feed Finishing	Mature	10	8
Output 2	Meat Processing	Mature	21	18
Leather Sector				
Output 3	Finished Leather Goods	Work in progress	7	7
Output 4	Finished Leather	Infancy	2 (1 discontinued & 1 foundation)	1
Output 5	Skin Quality	Early stage of implementation	2 (1 completed & 1 ongoing)	2
Cross-Cutting Issues				
Output 6	BMO Development	Mature	16	15
Output 7	Access to finance	Early stage of implementation	6	3
Total interventions			63	54

In order to get a representative sample, it was determined that the interventions should be evenly spread among the three sectors with three interventions in each, for a total of nine. It was also determined that there should be sufficient mature interventions in the sample in order to thoroughly assess the control points in the Standard related to measuring changes in indicators (section 3), estimating attributable changes (section 4) and capturing wider changes in the system or market (section 5). An analysis of the outputs and interventions was undertaken according to: 1) the percentage of eligible interventions in each output, 2) the reported outreach, income and jobs per output and intervention according to the latest annual report, and 3) the percentage of the project budget spent on each output. Based on this analysis, Output 1 was chosen for the meat sector, Outputs 3 and 5 for the leather sector and Output 6 for the cross-cutting issues. Then three interventions were randomly selected in Outputs 1 and 6.⁴ One intervention was randomly selected in Output 5 because it only includes 2 interventions. One intervention was purposively selected and one intervention randomly selected in Output 3. This process yielded the following sample for the audit:

⁴ One intervention selected in Output 6 – A&B1 – was later found to be an early version of three other interventions: A&B2, A&B4 and A&B7. Therefore A&B2 was substituted for A&B1 in the sample.

Sector and Outputs	Interventions
Meat Sector: Output 1: Feed Finishing	FF3: Feed Finishing for Sallah 2012
	FF5: Feed Finishing Year Round with Feed Tech (Kaduna)
	FF6: Year Round Feed Finishing with Feed Master (Zamfara)
Leather Sector: Output 3: Finished Leather Goods Output 5: Skin Quality	FLG1: Creating Better Business Linkages
	FLG2: Access to Finance – NERFUND Loan
	SQ1: Improving Skins Quality through Use of Industrial Salt – Pilot
Cross-cutting Issues: Output 6: Advocacy and BMO Development	A&B2: Policy Lobbying Demonstration through LAPAN
	A&B4: Association Development – AFLEM Kaduna
	A&B14: Advocacy for Access to Finance

The in-country portion of the audit was conducted in Kano from the 17th through the 21st of November 2014. Document review preceded the in-country visit. Further assessment and report writing followed the visit.

For each intervention, the audit reviewed the intervention control framework containing the results chain, measurement plan and supporting information; available background information; data, analysis and available reports from measurement studies and on-going monitoring; and consultant terms of reference and reports. For the overall GEMS1 programme, the audit reviewed annual and quarterly reports, the programme organogram, cost allocation and value for money analyses, staff terms of reference and technical review sessions notes. A full list of the documents reviewed is provided as Annex 3.

For each intervention, interviews were held with the Intervention Manager(s), Business Development Officer(s), associated members of the Results Measurement Team, consultants when relevant and selected key partners when relevant. For the overall programme, interviews were held with the Results Measurement Team, the Team Leader, the Deputy Team Leader, the Technical Director and the Finance and Administration Manager. The full list of interviews conducted is provided as Annex 4.

5. Detailed Scoring of the Control Points

The program rates 399 out of 470 possible points for the MUST control points and 129 out of 160 possible points for the RECOMMENDED control points. The maximum scores have been adjusted to exclude the “Not Applicable” compliance criteria. All compliance criteria were verified. For details of the ratings for the overall programme, sectors and interventions, see Annex 1.

The table below shows the rating for each control point and summarizes the audit findings for that control point. More detailed findings per sector are provided in Annex 2.

Control Point	M/R	Max. Score	Rating	Justification
Section 1: Articulating the Results Chain				
1.1 An appropriate, sufficiently detailed and logical results chain(s) is articulated explicitly for each of the interventions.	M	30	26	<ul style="list-style-type: none"> Results chains have been articulated for all interventions. They are mostly logical and thorough. There are some discrepancies in the logic and detail of some results chains.
1.2 Each results chain is supported by adequate research and analysis.	M	30	28	<ul style="list-style-type: none"> Most results chains are supported by adequate and documented research and analysis. A few results chains are not supported by adequate research and analysis, particularly analysis of the likelihood of lasting impact. In some cases, research and analysis underlying the results chains are not sufficiently documented. For almost all interventions, assumptions are clear but for one intervention, significant assumptions are not adequately outlined.
1.3 Mid and senior level programme staff are familiar with the results chain(s) and use them to guide their activities; key partners can explain the logic of interventions.	M	30	29	<ul style="list-style-type: none"> Intervention staff are familiar with the results chains and can give specific examples of how they use them to guide their activities. Many partners can describe the logic of the interventions; however, a few are not clear on intermediate results.
1.4 The results chain(s) are regularly reviewed to reflect changes in the programme strategy, external players and the programme circumstances.	M	20	16	<ul style="list-style-type: none"> A system exists to review results chains but it is not sufficiently elaborated or documented. Most results chains have been reviewed regularly, however a few have not; the reasons for changes are not adequately documented.
1.5 The results chain(s) include the results of broader systemic change at key levels.	REC	10	10	<ul style="list-style-type: none"> Expected systemic change is included in results chains when appropriate.
1.6 The research and analysis underlying the results chain(s) take into account the risk of displacement.	REC	10	8	<ul style="list-style-type: none"> Displacement is usually mentioned in the ICF, however with no or unclear justification. Staff can usually explain why displacement is unlikely.

Control Point	M/R	Max. Score	Rating	Justification
Section 2: Defining Indicators of Change				
2.1 There is at least one relevant indicator associated with each change described in the results chain(s).	M	20	16	<ul style="list-style-type: none"> There are indicators documented for almost all changes in the results chains. Most indicators are relevant. Some documented indicators are not relevant, unclear, not thorough or not well defined, for example, they are missing the unit of analysis or don't adequately describe what will be assessed.
2.2 The universal impact indicators are included in the relevant results chain(s).	M	10	10	<ul style="list-style-type: none"> The universal impact indicators are included in the ICFs.
2.3 There are specific indicators that enable the assessment of sustainability of results.	M	20	17	<ul style="list-style-type: none"> GEMS1 measures results over time to assess sustainability and staff gather specific and usually thorough information on the likelihood of sustainability of results. The indicators of sustainability are often not adequately documented, particularly those that assess sustainability of the business model supporting service and product provision to beneficiaries.
2.4 Mid and senior level programme staff understand the indicators and how they illustrate programme progress.	M	20	20	<ul style="list-style-type: none"> Implementation staff clearly understand the indicators and how they illustrate programme progress.
2.5 Anticipated impacts are realistically projected for key quantitative indicators to appropriate dates. ⁵	REC	30	24	<ul style="list-style-type: none"> Expected results are projected when sufficient information is available; this includes projections for the universal impact indicators. Supporting calculations for projections are documented, but are not always clearly linked to the data from studies.
Section 3: Measuring Changes in Indicators				
3.1 Baseline information on key indicators is collected.	M	20	13	<ul style="list-style-type: none"> There are some plans to gather baseline data, however they are usually not sufficiently clear and thorough. Some baseline information is gathered, however it is often not sufficiently thorough. In particular, although sector information is gathered including on the existing practices of target beneficiaries, often not enough baseline data is gathered on the actual beneficiaries included in impact studies, thus making assessment of results less reliable.

⁵ Only intervention level projections in the intervention control frameworks have been considered for this control point. The log frame targets have not been considered.

Control Point	M/R	Max. Score	Rating	Justification
3.2 Information for each indicator is collected using methods that conform to good research practices.	M	40	30	<ul style="list-style-type: none"> • There is a documented plan to collect information on the indicators for each intervention. The documented plans are sometimes insufficiently clear, detailed and thorough. In particular the planned methods for data collection are not well-described using terms from the MRM guidance. However staff can usually add the necessary depth. • Assessment studies usually use good practices and quality control in data collection, entry and cleaning. There is attention to data quality and usually data from several sources is triangulated. • Assessment studies are inadequately documented; there is little written information on methodology, quality control and data analysis. • Sample sizes for some studies were inadequate, particularly when missing data reduced the effective sample size for data analysis. • There is insufficient quality control in data analysis, resulting in some calculation or analysis mistakes. It is sometimes difficult to trace figures fully from reporting back to respondent data.
3.3 Qualitative information on changes at various levels of the results chain is gathered.	M	20	17	<ul style="list-style-type: none"> • For some, mainly newer, interventions, adequate qualitative information is included in measurement plans and is thoroughly gathered in practice. • For some, mainly older, interventions, qualitative information to be gathered is not adequately documented in the measurement plan and is not always sufficiently thorough in practice.
3.4 Reported changes in indicators that are extrapolated from pilot figure are regularly verified.	REC	20	17	<ul style="list-style-type: none"> • For programme strategies rolled out over a number of years, data quality is regularly improved – studies are conducted several times and in different areas to improve the extent to which data is representative. • The use of later studies to re-estimate results is occasionally not appropriate or not sufficiently clear.

Control Point	M/R	Max. Score	Rating	Justification
Section 4: Estimating Attributable Changes				
4.1 Attributable changes in all key indicators in the results chains are estimated using methods that conform to established good practice.	M	50	37	<ul style="list-style-type: none"> • There are documented plans to assess attribution at the level of beneficiaries and sometimes at the level of partners and service providers. However, the plans are not sufficiently detailed and do not always include adequate methods to assess the strength of expected causal links. • The plans to assess attribution are usually carried out and occasionally augmented. • Assessment of attribution at the level of partners and service providers is mostly adequate but would often benefit from more systematic use of participant opinion or other qualitative tools. • At the level of beneficiaries, the adequacy of methods to assess attribution is variable. For the skin quality and business linkage interventions, the methodology is appropriate and adequate. For the newer feed finishing interventions, the methodology compares beneficiaries to non-beneficiaries through several studies but may not adequately take into account possible differences between the two groups. For the old feed finishing intervention and the finished leather goods finance intervention, the methodology did not adequately take into account the counterfactual. • In some cases, the methodology to assess attribution was degraded in practice by challenges or inadequacies in data collection. • For some interventions, calculations used to estimate attributable results are clear and traceable; for others they are not completely clear and traceable and/or how attribution was taken into account is not explicit.
Section 5: Capturing Wider Changes in the System or Market				
5.1 The results of systemic change at key levels in the results chain(s) are assessed using methods that conform to established good practices.	REC	50	43	<ul style="list-style-type: none"> • There are documented plans for assessing wider systemic change when appropriate. While the documented plans are not always sufficiently clear, staff are able to provide details. When applicable, the plans were carried out effectively. • In feed finishing, the methodology to assess attribution compared an indirect beneficiary group with a non-beneficiary group as described above, which may not account for

Control Point	M/R	Max. Score	Rating	Justification
				<p>possible differences between the groups.</p> <ul style="list-style-type: none"> Estimates of systemic results are supported by clear calculations, but the figures are not always easily traceable to the studies.
Section 6: Tracking Programme Costs				
6.1 Costs are tracked annually and cumulatively.	M	20	20	<ul style="list-style-type: none"> Costs are tracked annually and cumulatively using a clear accounting system. Annual and cumulative totals of all program-related costs spent in country have been produced annually.
6.2 Costs are allocated by major component of the programme.	REC	20	20	<ul style="list-style-type: none"> The accounting system enables the separation of costs by output. It includes a clear and logical system to allocate overhead costs per output. Annual and cumulative estimates of costs per output have been produced annually. Allocation of implementation and overhead costs is clear and documented.
Section 7: Reporting Results				
7.1 The programme produces a report, at least annually, which clearly and thoroughly describes results to date.	M	30	21	<ul style="list-style-type: none"> GEMS1 has a well-designed system for aggregation that transparently takes overlap into account, but the system was only being operationalized at the time of the audit and it had not yet been used for reported figures. The system that had been used for aggregation for reported figures up through the May 2014 annual report takes overlap into account and is documented, but it is not systematic and adequately clear. The programme produces quarterly and annual reports with clear estimates of programme wide impacts and extensive context and qualitative results.
7.2 Contributions of other publicly funded programmes and private contributions are acknowledged.	M	10	10	<ul style="list-style-type: none"> Contributions from other public and private entities are acknowledged in the text of the reports.
7.3 Reported changes in key indicators are disaggregated by gender.	M	10	8	<ul style="list-style-type: none"> Outreach is disaggregated by gender using appropriate methodologies. Income and jobs are sometimes disaggregated by gender in the ICFs but not in reports, as this is not required by the donor's reporting format. The reasons for not disaggregating all key indicators by gender are clear and justified but not documented.

Control Point	M/R	Max. Score	Rating	Justification
7.4 Results of systemic change and/or other indirect effects are reported.	REC	10	8	<ul style="list-style-type: none"> The results of systemic change are reported both quantitatively and qualitatively. When the results are reported quantitatively, they are not always separated by direct and indirect.
7.5 Results are published.	REC	10	0	<ul style="list-style-type: none"> A summary of results is not made publicly available.
Section 8: Managing the System for Results Measurement				
8.1 The programme has a clear system for results measurement that ensures that findings are used in programme management and decision-making.	M	40	33	<ul style="list-style-type: none"> There is a system and supportive culture for using information on results to manage the programme portfolio, sector strategies and individual interventions. The system ensures that results information is regularly and effectively integrated into management decision-making. However, the system is not adequately documented. Managers can provide many clear examples of using information on results in programme decision-making. Both written and verbal guidance on the results measurement system is provided to staff. However, the written guidance does not adequately cover all necessary elements of the system. There is insufficient guidance, for example, on indicators of sustainability, data gathering as it is conducted in practice, how to disaggregate figures by gender and reporting templates.
8.2 The system is supported by sufficient human and financial resources.	M	30	28	<ul style="list-style-type: none"> The system is now supported by sufficient human and financial resources. There are three dedicated results measurement staff and all implementation staff are involved in results measurement. Before this year, there were not sufficient human and financial resources for results measurement and this affected the quality of the MRM system and its implementation, particularly for the earliest interventions. Staff are clear on their roles and responsibilities and most roles and responsibilities in the system are appropriately allocated. However, managers do not play a strong enough role in quality control of analysis and documentation as well as systematic implementation of the formal aspects of the results measurement system. Minimal additional skills in assessing the progress of BMO capacity and the various

Control Point	M/R	Max. Score	Rating	Justification
				factors affecting advocacy are needed.
8.3 The system is integrated with the management of the programme.	M	20	20	<ul style="list-style-type: none"> • There is abundant evidence that the results measurement system has been institutionalised, such as regular discussions on results and inclusion of results measurement in meetings and job descriptions. • Staff state that results measurement is indivisible from other aspects of their jobs. • All staff can provide examples of results measurement activities undertaken in the last month.

6. Summary of Key Areas that Require Improvement

Articulating the Results Chain

- Ensure results chains are logical and sufficiently detailed.
- Document the system for reviewing results chains. Systematically review results chains according to the system and document reasons for changing results chains or leaving them as is.

Defining Indicators of Change

- Ensure all indicators are relevant, clear, thorough and well defined.
- Document indicators of sustainability.
- Ensure that calculations for projections are adequately clear and supported.

Measuring Changes in Indicators

- Develop, document and carry out adequate plans to collect baseline data.
- Ensure plans to collect information on indicators are clearly documented.
- Clearly document each impact study, including methodology, attribution strategy, sampling, quality control measures, data processing and analysis methodology, limitations, findings and the analysis of data that underpins the reported findings.
- Ensure that sample sizes are sufficient, appropriate to the size of the intervention and magnitude of the results and take into account the potential for non-responses.
- Increase quality control in data analysis.
- Ensure that the documentation of calculations and findings clearly traces figures to study findings and explicitly indicates how attribution was taken into account and sources of data triangulated.

Estimating Attributable Changes

- At the level of beneficiaries, ensure that attribution methodologies adequately take into account the counter-factual and potential influence of external factors.
- Ensure that calculations used to estimate attributable results are clear, explicit and well supported.

Capturing Wider Changes in the System or Market

- At the level of beneficiaries, ensure that attribution methodologies for the results of systemic changes adequately take into account the counter-factual and potential influence of external factors.
- Ensure that calculations used to estimate attributable results of systemic change are clear, explicit and well supported.

Reporting Results

- Complete revisions to the new aggregation system and fully put it into use. Ensure that aggregation is systematic and figures are easily traceable.
- Publish a summary of results so that it is publicly available.

Managing the System for Results Measurement

- Revise written guidance on the results measurement system so that it adequately covers all necessary elements of the system and documents the system for using information on results in management.
- Strengthen the role of managers in the systematic implementation of the system and quality control in data analysis and documentation.

Annexes

1. Overall and sector specific ratings
2. Sector specific findings
3. List of documents reviewed
4. List of interviews conducted