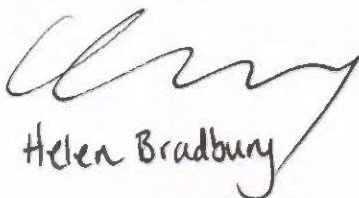


1. Overview

Program	Alliances Lesser Caucasus Programme	
Audit visit dates	January 23-27, 2017	
Overall final ratings ¹	MUST	440/460 = 96%
	RECOMMENDED	85/110 = 78%
Coverage	All control points were checked. All outcomes were checked. The audit is limited to ALCP's work in the Ajara region and excluded its work in the SJ and KK regions, which were the focus of a previous Audit in 2014. It excludes 4 interventions from the total set of ALCP's interventions (28).	
DCED Standard	Version VII, April 2015	

Signed:

Team Leader 
Helen Bradbury

Date, place *Manenti, March 13th 2017*



Auditors Ben Fowler March 6, 2017

Table of Contents

1. Overview	1
2. Summary of the Programme and Results Measurement System	3
3. Summary of the Audit Process.....	5
4. Summary of Findings.....	7
5. Summary of Key Areas for Improvement	13
Annexes.....	14

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.

Acronyms

ALCP	Alliances Lesser Caucasus Programme
AJ	Ajara
BDO	Business Development Officer
DCED	Donor Committee for Enterprise Development
IA	Impact Assessment
IG	Intervention Guide
KK	Kvemo Kartli
MRM	Monitoring and Results Measurement
M4P	Making Markets Work for the Poor
NAIC	Net Attributable Income Change
RC	Results Chain
RM	Results Measurement
SJ	Samtskhe-Javakheti
WEE	Women's Economic Empowerment

2. Summary of the Programme and Results Measurement System

Summary of Programme

- The Alliances Lesser Caucasus Programme (ALCP) is a market development project implemented by Mercy Corps Georgia and funded by the Swiss Agency for Development and Cooperation. It works in dairy, beef, sheep and honey sub-sectors in the Kvemo Kartli (KK), Samtskhe-Javakheti (SJ) and Ajara (AJ) regions in Southern Georgia.
- The audit focused exclusively on ALCP's activities in Ajara and did not examine its operations in SJ and KK, which were the subject of a previous audit.
- However, ALCP applies a similar approach to results measurement across AJ, KK and SJ.
- The Goal of ALCP is to contribute to poverty alleviation and the transition to a durable market economy for the livestock sector in KK, SJ and AJ. The programme is run according to the M4P (Making Markets Work for the Poor) approach. Sustainability is built in through a minimum co-investment of 35% from the market players with whom it invests. ALCP's target is to reach 24,000 households. This is 20% of poor households in the programme area, who will benefit directly and indirectly through improved services, markets and operating environment, with increased income from sales, reduced production & transaction costs, increased net worth and employment.
- The programme's first phase ran from 2008 to 2011 covering only Samtskhe Javakheti (SJ). In 2011 a separate programme was launched that covered KK, while the work in SJ was expanded. This continued until 2014. In 2014, a new programme was launched that combined Mercy Corps' work in SJ and KK, while also expanding to focus on AJ. This is the current phase that is being audited (but only for AJ). Mercy Corps expects a new round to launch in 2017 that will maintain its work in the three regions while also expanding to the neighbouring countries of Azerbaijan and Armenia.
- The project reports its results in terms of three outcomes: supporting functions, core markets and cross-cutting issues. The latter consists of directly-focused gender interventions, information and governance interventions.

Key Features of the Results Measurement System

- A defining feature of ALCP's RM system is that its impact assessments are conducted at a regional level rather than an intervention-specific level. That is done for several reasons. One is that ALCP is implementing various interventions in a small geographic area. That makes it difficult to isolate the impact of specific interventions, given that the beneficiaries of specific interventions are likely also engaged in others and that its veterinary intervention is used by a large majority of the livestock-owning population in its target areas. Another reason is that the income benefits of its interventions in Outcome 3 (e.g., information access) are difficult to capture directly. A third reason is that the combination of ALCP's various interventions likely creates synergies that make the ultimate results greater than the sum of each intervention separately.
- This area-based approach to conducting the impact assessment may not collect the full impact of some of ALCP's smaller interventions. Consequently ALCP's impact assessment methodology is largely conservative. When this occurs, ALCP will conduct follow-up mini-survey to understand

the impact of those interventions specifically, though the results will not be added to the impact assessment results.

- The impact assessment methodology establishes a retrospective baseline for two years prior to the date of the impact assessment. This is to ensure that the impact assessment includes questions on all of the key interventions that the project has launched. Two years was selected because many of the phase's interventions started at that time and recall prior to that period has proved in KK and SJ to be less accurate. ALCP triangulates the retrospective baseline figures against multiple sources, including its initial quantitative and qualitative baseline studies, national statistics, and trends drawn from their monthly monitoring data. ALCP's baseline sources include an area-wide quantitative baseline study, a market research study, focus group discussions, and key informant interviews.
- Because the impact assessment is conducted at the level of the entire region rather than for each intervention, it may not fully capture the impact of the smaller interventions. The programme intends to follow-up with additional surveys where this occurs, but the methodology for doing so has not been fully developed yet. The impact assessment methodology considers any person who has participated in even one of ALCP's interventions to be a beneficiary.
- By using an area-based approach, the IA methodology makes it more challenging to determine whether treatment and non-treatment farmers are similar or different. To mitigate this, ALCP tests how homogenous the treatment and control groups are by examining both groups against a number of demographic factors, and validates that both groups live in the same geographic locations.
- The random selection of interviewees at a regional level in the IA plan means that the results reported may not capture the impacts of the interventions that are omitted from the sample (e.g., farmers selling to a new dairy in a village not selected as part of the sample). ALCP conducts mini-survey to assess the NAIC generated by those interventions, but does not include it in their final reporting to donors.
- Given that waiting until the end of the phase would not provide information on ALCP's progress against its targets, ALCP also develops monthly estimates of its progress against key quantitative indicators including particularly scale and NAIC. This is done only for interventions in outcomes 1 and 2; NAIC is not estimated for interventions in outcome 3. To estimate this, ALCP receives quantitative data from its partners in monthly datasheets. ALCP has identified ratios that it uses to estimate NAIC from those interventions. These ratios are drawn from mix of expert opinion and secondary source research. For all interventions, except its veterinary work (for which the numbers reached are too large to calculate easily), ALCP receives a list of the names of every customer that its partners serve. ALCP reviews this list to assess overlap. These results are reported to the donor, but noted as being estimated prior to the completion of the impact assessment. ALCP's experience in KK and SJ suggests that its monthly estimates are sometimes conservative and sometimes optimistic. When it was discovered that the scale numbers for Roki in KK from the impact assessment were lower than that from the monthly data sheets, ALCP adjusted down its figures for AJ too in order to continue to be conservative.
- ALCP has not yet conducted an impact assessment for its work in Ajara given the challenges of surveying during winter. It plans to do so in April or May 2017. It has already completed its impact assessments for SJ and KK in 2016.
- ALCP has a strong focus on collecting qualitative information. It has a separate monitoring plan specifically to document the qualitative information that it collects. It periodically conducts mini-surveys that allow it to collect additional documentation on an intervention.

- ALCP is committed to achieving women's economic empowerment. Its RM system includes both quantitative and qualitative research to understand its gender results. Outcome 1 and 2 interventions include activity boxes in many results chains that are specifically oriented to ensuring that women engage in and benefit from the intervention (i.e., gender sensitive interventions) whereas certain Outcome 3 interventions take a gender overt approach to specifically target issues that can improve WEE. Qualitative research is targeted to include equal numbers of female and male respondents so that both perspectives are reflected.
- ALCP implements its interventions in stages. Once its outputs have been achieved, it will consider scaling up the intervention with additional activities. When interventions have been expanded, extra boxes are added in the results chain that depict the added activities and expected results.
- ALCP defines systemic change as including a number of different types of changes, including imitation (i.e. crowding in, copying), business expansion and unintended effects. Of these, imitation is depicted visually in the results chain using grey boxes and some others are presented in the results chains as specific boxes. Other types of systemic change (e.g., business expansion) are monitored primarily using observation and are not presented directly in the results chain.
- ALCP holds regular monitoring action plan (MAP) meetings. The MAP meetings are used to review the progress of each intervention relative to expected progress and make adjustments were results are not as expected. They are a monitoring and management feedback loop and a way that the two are integrated.
- For some interventions that have spread beyond the programme area, such as Women's Rooms, ALCP measures them in each of the three regions but captures nationwide impact and systemic change outside the programme area from the KK office where these interventions started.

Evolution of the Results Measurement System

- ALCP Ajara has been working to comply with the DCED Standard since May 2011. Prior to requesting this audit, Alliances KK contracted a pre-audit review in February 2012 that was conducted by a consultant familiar with the DCED Standard, and conducted a full DCED Audit in November 2013 of the ALCP SJ and KK during Phase 1.

3. Summary of the Audit Process

The first step was to determine how many sectors of components ALCP Ajara was operating in. A review of its portfolio determined that ALCP works in just one sector: the livestock sector. It has three separate outcomes, as outlined above, but these were not sufficiently distinctive to count as separate components. Therefore the audit's sampling methodology did not treat these outcomes as separate sectors that would be each sampled separately.

Next the audit team determined the number of interventions to be audited. ALCP chose to exclude three interventions whose monitoring documentation was not yet complete:

- 1.4.1 Elva / ALCP AJ-1

- 2.3.5 Ecotours Georgia / ALCP AJ -1
- 2.3.5/3.1.3/3.3.2 Eco Films Ltd / ALCP AJ -1

Another intervention was omitted from the list of interventions:

- 3.1.1 Advisory Committee Meetings

A total of 24 interventions remained for consideration. By calculating the square root of these interventions, the auditors determined that at least five interventions needed to be reviewed. To obtain a representative sample of the portfolio, the auditors stratified the interventions into those representing more than 10% of the portfolio, and those representing less. Two interventions were selected randomly from among the three interventions whose budgets were at least 10% of the portfolio. Then the remaining three were selected at random from the full set of interventions.

Outcome	Intervention Number and Name
1	1.1.1 Roki/ALCP AJ- 1
2	2.2.1 Star Consulting/ALCP AJ-2
2	2.3.3 Natural Produktsia / ALCP AJ- 1
2	2.3.6 Bakuri/ ALCP AJ -1
3	3.1.10 Business Association of Ajarian Beekeeping /ALCP AJ – 1

Following the initial selection of the audit interventions, ALCP explained that the selection did not fully capture its work on gender. Consequently, the auditor agreed to add a sixth intervention, to make a total of 6 interventions.

Outcome	Intervention Number and Name
3	3.2.5 Women's Rooms

The resulting list of interventions included two that had calculated NAIC, and four that had not (as explained above). It also included at least one intervention from each of ALCP's three outcomes.

For each audited intervention, the monitoring framework was reviewed, including the results chain, monitoring plan, key quantitative changes, key qualitative changes, and supporting documents. Other intervention-specific documents included investment plans, secondary data, data sheets, and their aggregation system.

For ALCP Ajara as a programme, the M&E manual, progress reports, job descriptions, organizational chart, background research (e.g. sector analysis), list of interventions and logframe were reviewed. A list of documents reviewed is included as Annex 3.

For ALCP as a programme overall and its Ajara office, interviews were held with the Team Leader, Deputy Team Leader, M&E Coordinator, M&E Officer, Information Manager, Information Officer, representative of ALCP's research partner, Finance and Compliance Director, M&E Assistant, Ajara Manager, DRR, Gender and Governance Coordinator, and Information Management Officer. For the selected interventions, interviews were held with the Business Development Officers, BDO Assistants, M&E Officer and Information Officer. The full list of interviews conducted is included in Annex 4.

The audit was originally to have been conducted by two auditors, Ben Fowler and Sadia Ahmed, but visa

issues meant that only Ben conducted the audit.

4. Summary of Findings

The program rates 96% (out of 460 possible points) for the MUST control points and 78% (out of 110 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. The table below summarizes the ratings per section. For details of the ratings for the overall programme and each sector, see Annex 1.

Sections	Level	Percent
Section 1	Must	94%
	Rec	85%
Section 2	Must	97%
	Rec	33%
Section 3	Must	96%
	Rec	90%
Section 4	Must	80%
Section 5	Rec	100%
Section 6	Must	100%
	Rec	n/a
Section 7	Must	97%
	Rec	100%
Section 8	Must	98%

The tables below show the rating for each control point and summarize the audit findings per section of the Standard. More detailed findings per sub-sector are provided in Annex 2.

Section 1: Articulating the Results Chain

Control Point	M/R	Max. Score	Rating
1.1 An appropriate, sufficiently detailed and logical results chain(s) is articulated explicitly for each of the interventions.	M	30	26
1.2 Each results chain is supported by adequate research and analysis.	M	30	28
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	30
1.4 The results chain(s) are regularly reviewed to reflect changes in the programme strategy, external players and the programme circumstances.	M	20	19
1.5 The results chain(s) include the results of broader systemic change at key levels.	REC	10	10
1.6 The research and analysis underlying the results chain(s) take into account the risk of displacement.	REC	10	7

Strengths	Weaknesses
<ul style="list-style-type: none"> All interventions have a results chain that includes the main activities and expected results. These results chains usually logically reflect the links between activities and intervention goals. The changes in boxes are usually well articulated. Results chains are supported by a market study for the sector and information from the partner collected in an Investment Plan. Internal assumptions supporting the rationale for the links between boxes are documented as part of the Intervention Rationale. Staff have an in-depth knowledge of the results chains for their interventions. They can provide concrete examples of how they are using the results chains to inform their decision-making. The programme has a system for regularly reviewing results chains. This is described in the manual. Certain manifestations of systemic change (i.e. imitation) are reflected in most results chains. 	<ul style="list-style-type: none"> In some cases, steps are skipped between boxes in the results chains. In others, too many changes are combined in one box. This is sometimes the result of adding the expansion activities into the existing results chain boxes. In certain cases, the links between boxes are not causal. The external assumptions (beyond the control of ALCP) upon which the success of each intervention depends are rarely explicitly documented. Displacement is often not documented at farmer level. Although the programme doesn't expect it to happen, the reasons are not documented and in some cases (such as the Batumi Agro Market), there is a reasonable possibility that farmer-level displacement could happen as higher standards displace those farmers who don't adopt the new system.

Section 2: Defining Indicators of Change

Control Point	M/R	Max. Score	Rating
2.1 There is at least one relevant indicator associated with each key change described in the results chain(s).	M	20	19
2.2 Information to be collected includes qualitative information on changes at various levels of the results chain.	M	20	19
2.3 A small number of indicators at the impact level can be aggregated across the programme	M	10	10
2.4 There are specific Indicators that enable the assessment of sustainability of results.	M	20	19
2.5 Mid and senior level programme staff understand the indicators and how they illustrate programme progress.	M	20	20
2.6 Anticipated impacts are realistically projected for key quantitative indicators to appropriate dates.	REC	30	10

Strengths	Weaknesses
<ul style="list-style-type: none"> Every change described in the results chains has at least one indicator and typically multiple indicators. ALCP includes qualitative information 	<ul style="list-style-type: none"> For many of the reviewed interventions, there are no projections for key quantitative indicators. Projections are used mainly to develop expectations for the project client in

<p>gathering for all of its interventions. The qualitative indicators focus consistently on understanding the reasons for changes.</p> <ul style="list-style-type: none"> • Although many monitoring plans do not include the common impact indicators, they are described in the manual and captured in practice. • Each intervention has indicators to check the assessment of the sustainability of results. • Staff understand the indicators that are important for their work. • Projections are developed for some of the key quantitative indicators. The calculations for these projections are documented. 	<p>the investment plan, and so focus mainly on output indicators, rather than including a focus on what contribution the intervention will make to NAIC and other impact indicators. The projections are not always reviewed annually.</p>
--	--

Section 3: Measuring Change in Indicators

Control Point	M/R	Max. Score	Rating
3.1 Baseline information on all key indicators is collected.	M	10	10
3.2 Information for each indicator is collected using methods that conform to good research practices.	M	40	38
3.3 Programmes have a mechanism for assessing and understanding differentiated results by gender.	M	20	20
3.4 Programmes monitor to identify unintended effects.	REC	20	18

Note: the score for Section 3 assesses ALCP's plan for reconstructing a baseline through the impact assessment, but does not assess the implementation of the baseline since the impact assessment has not yet been conducted.

Strengths	Weaknesses
<ul style="list-style-type: none"> • ALCP has an appropriate plan for reconstructing a baseline through the impact assessment. It will triangulate the results with other qualitative and quantitative information to test its adequacy. • ALCP has detailed monitoring plans outlining what information it will collect and when for all indicators. • ALCP has a system to monitor unintended effects. • ALCP has an appropriate system to capture differentiated results by gender. This includes both gender-disaggregated indicators and qualitative research that identifies gender-relevant issues and differences in results. 	<ul style="list-style-type: none"> • In some cases, qualitative information is not collected according to plan or not collected at the scheduled time. • Unintended effects are defined by ALCP as being a type of systemic change, which creates a bias to look for positive effects and changes within the sector rather than broader effects. Sometimes unintended effects have been described as instances that an intervention did not work as intended.

Section 4: Estimating Attributable Changes

Control Point	M/R	Max. Score	Rating
4.1 Attributable changes in all key indicators in the results chains are estimated using methods that conform to established good practice.	M	20	17

Note: the score for Section 4 assesses ALCP's plan for conducting its impact assessment, since the impact assessment has not yet been conducted.

Strengths	Weaknesses
<ul style="list-style-type: none"> ALCP has a plan for conducting an impact assessment. ALCP's approach to attributing partner-level results is appropriate and uses several methods of triangulation to verify the estimate. The quality assurance procedures that are planned to be implemented by ALCP's research partner in the impact assessment follow good practice and are comprehensive. 	<ul style="list-style-type: none"> ALCP's planned impact assessment only measures revenue, not net income, but ALCP plans to report this as NAIC. This almost certainly overestimates the NAIC for many interventions (though among the studied interventions likely not for dairy), as beneficiaries often need to incur additional costs (e.g., purchasing veterinary drugs) in order to realize the additional revenue. This also likely overstates the number of beneficiaries (income scale), as they are calculated based on the number of beneficiaries who surpass the average revenue of the non-beneficiary group. Again, this would not apply to the interventions that lowered costs for farmers, such as dairy.

Section 5: Capturing Wider Changes in the System or Market

Control Point	M/R	Max. Score	Rating
5.1 The results of systemic change at key levels in the results chain(s) are assessed.	REC	20	20

Note: the score for Section 5 assesses ALCP's plan for assessing systemic change at the beneficiary level. Because no systemic changes have yet been identified at the beneficiary level that have contributed to NAIC (though many have been identified at the partner level), the implementation of the plan has not been scored.

Strengths	Weaknesses
<ul style="list-style-type: none"> ALCP defines what it means by systemic change and depicts certain aspects of it (i.e., imitation by other market actors) in many of its intervention results chains. The plan to assess systemic change is of high quality, appropriate for the context, takes attribution into account, is timely and documented. 	

Section 6: Tracking Programme Costs

Control Point	M/R	Max. Score	Rating
6.1 Costs are tracked annually and cumulatively.	M	20	20
6.2 Costs are allocated by major component of the programme.	REC	n/a	n/a

Strengths	Weaknesses
<ul style="list-style-type: none"> A clear, accounting system is in place to track costs and produce annual and cumulative totals of all programme-related costs spent in country. The programme has annual and cumulative totals of all programme-related costs spent in country. 	

Section 7: Reporting Results

Control Point	M/R	Max. Score	Rating
7.1 The programme clearly and appropriately aggregates programme-wide impact at least annually.	M	30	28
7.2 The programme produces a report at least annually which describes results to date.	M	30	30
7.3 Results of systemic change are reported.	REC	10	10
7.4 Results are published.	REC	10	10

Strengths	Weaknesses
<ul style="list-style-type: none"> ALCP produces biannual reports that report results to date. The report provides estimates of the achievement to date of the common impact indicators. The experience of ALCP in SJ and KK indicates that the methodology used for estimating NAIC is of high standard and conservative relative to the actual results likely to be obtained through the impact assessment. Substantial qualitative information is provided in the report in a separate annex that breaks down qualitative findings per intervention. The report outlines qualitative achievements and discusses progress towards the sustainability of results. The perspectives of stakeholders are also presented. Sustainability of the results is discussed. The report recognizes the financial contributions of ALCP's clients and third parties 	<ul style="list-style-type: none"> There were some instances that the full-time equivalent jobs figures were overstated due to an incorrect conversion of seasonal employment to FTE.

<p>(e.g., other market system participants who do not have a contractual relationship with ALCP) to the projects. The role of the government's agri-credit programme is also recognized.</p> <ul style="list-style-type: none"> • The results of systemic change are reported in the report via a systemic change log. These results are credible and consider ALCP's attribution. • Results are published on ALCP's website. 	
---	--

Section 8: Managing the System for Results Measurement

Control Point	M/R	Max. Score	Rating
8.1 The programme has a clear system for using information from the results measurement system in management and decision-making.	M	30	30
8.2 The system is supported by sufficient human and financial resources.	M	20	20
8.3 The system is well managed and integrated with programme management.	M	40	38

Strengths	Weaknesses
<ul style="list-style-type: none"> • Tasks and responsibilities in relation to results measurement are appropriate and documented. • Staff are able to accurately describe their tasks and responsibilities in results measurement. • Managers can explain to what extent underlying assumptions in the results chains are proving to be valid, and can cite decisions they have made based on the information provided by the results measurement system. • ALCP clearly uses monitoring information to inform their decision-making, including particularly through the monthly Monitoring Action Plan meetings. • The results measurement system has clearly been institutionalized. Responsibilities are specified in the terms of reference for staff. There are regularly meetings in which monitoring findings are reviewed and decisions are taken. • The team appears to have adequate financial and human resources to implement its system. • To guide their MRM activities, technical staff draws on the skills of their MRM colleagues and have access to an M&E manual. • The system for quality assurance is generally 	<ul style="list-style-type: none"> • The MRM manual is in some cases outdated as it does not describe ALCP's current practice in some cases. For example, that outcome results chains are not actually revised annually and that common impact indicators like income scale are not included and measured for all Outcome 1 and Outcome 2 interventions.

appropriate, with multiple team members checking outputs from the results measurement system.	
---	--

5. Summary of Key Areas for Improvement

Articulating the Results Chain

- Ensure that all of the results chains have undergone a quality assurance check.
- Consistently assess the likelihood and document the potential for displacement at the beneficiary level, as noted in the results measurement manual.

Defining Indicators of Change

- Make projections for all key quantitative indicators.

Measuring Changes in Indicators

- Develop a more robust system for calculating the full time equivalent jobs created from new full-time, part-time and seasonal work generated by ALCP's interventions that is less susceptible to calculation errors.
- Consider improving the file management system so that worksheets that monitoring staff are using for their calculations are maintained in a single place.
- The wording of unintended effects in the project documentation and its inclusion as a type of systemic change is such that there could be a bias towards observing instances of positive unintended effects. Therefore make clear that staff should also be aware of negative unintended effects.
- Ensure baseline information is gathered in a timely manner for all interventions.
- Project impact for key quantitative indicators for beneficiary impact level (e.g., NAIC) to at least the end of the project.

Estimating Attributable Changes

- In the impact assessment, assess changes in net incomes (i.e. profit) of beneficiaries rather than revenue.

Reporting Results

- Ensure that the overlap calculations for each intervention (e.g., Women's Rooms) are being calculated correctly and that there is not possibility that the same beneficiaries are recorded multiple times.²
- Ensure that the calculation of full-time equivalent jobs is correct and does not overstate the jobs created through seasonal work. Ensure that seasonal jobs are not reported as part-time jobs.

Managing the System for Results Measurement

- Ensure that the results measurement manual is kept fully up-to-date

² In practice, this has not affected results because ALCP does not report the scale results from Outcome 3, which is why no marks were deducted.

Annexes

1. Overall and sub-sector specific ratings
2. List of documents reviewed
3. List of interviews conducted

Annex 2: Sector Specific Findings

Sector: Livestock

Interventions: 2.3.6 BAM, 3.1.10 / 2.3.5 ABBA, 1.1.1 Roki, 2.2.1 Star Consulting, 3.2.5 Women's Rooms, 2.3.3 Natural Produktsia

Strengths	Weaknesses
Articulating the Result Chains	
<ul style="list-style-type: none"> • All interventions have a results chain that includes the main activities and expected results. These results chains usually logically reflect the links between activities and intervention goals. The changes in boxes are usually well articulated. • Results chains are supported by a market study for the sector and information from the partner collected in an Investment Plan. • Internal assumptions supporting the rationale for the links between boxes are documented as part of the Intervention Rationale. • Staff have an in-depth knowledge of the results chains for their interventions. They can provide concrete examples of how they are using the results chains to inform their decision-making. • Certain manifestations of systemic change (i.e. imitation) are reflected in most results chains. • Displacement is noted at the farmer level for Roki and Women's rooms. 	<ul style="list-style-type: none"> • For Roki, box 12 is not the result of 11, but from 8, 9 and 10. Box 9 should lead to 10. • For Natural Produktsia, Box 10 should not lead to 9 and Box 7 shouldn't lead to Box 8; inclusion of many changes in Boxes 7, 8 and 12 make measurement more confusing. • For ABBA, productivity is defined as increased sales and motivation, which overlaps with an existing box. ABBA gaining new members is an important change which is missing in the results chain between boxes 10 and 15. Box 16 doesn't lead to box 17. Many of the causal links between boxes are unclear. Boxes 13 and 7 don't seem to directly lead to 11. Box 4 should lead to Box 7. • For BAM, Box 10 does not lead to box 14. • For Star Consulting, Box 8 does not cause Box 7. Box 7 should have a box link to it. Box 16 doesn't lead to Box 17. • For Star Consulting, and ABBA, staff could describe having reviewed the results chains in the past year but there wasn't any written documentation of this in the RC Updates log. • For Women's Rooms, Box 13 should lead to Box 14 and 14 to 15. Boxes 9-12 seem to be caused by more than Boxes 7-8 (i.e. intermediary steps). Box 15 includes changes at two different levels. Box 14 is unclear in 'gender sensitive' and how it differs from Box 15. • The external assumptions (beyond the control of ALCP) upon which the success of each intervention depends were rarely explicitly documented. An example would be the effects of social norms on the success of the Women's Rooms intervention. • For Star Consulting, BAM, ABBA and Natural

Strengths	Weaknesses
	<p>Produktsia, displacement is not documented at farmer level. For Star Consulting and BAM, there is a reasonable expectation that the interventions could create displacement for farmers who choose not to improve their quality standards.</p>
Defining Indicators of Change	
<ul style="list-style-type: none"> • Every change described in the results chains has at least one indicator and typically multiple indicators. • ALCP includes qualitative information gathering for all of its interventions. The qualitative indicators focus consistently on understanding the reasons for changes. • Although many monitoring plans do not include the common impact indicators, they are described in the manual and captured in practice. • Each intervention has indicators to check the assessment of the sustainability of results. • Staff understand the indicators that are important for their work. • Projections are developed for some of the key quantitative indicators. The calculations for these projections are documented. 	<ul style="list-style-type: none"> • For Roki, Box 9 states 'provide services' but indicator is '# vets upgraded'. Box 10 states access and informed, but indicator is only access - not informed. Box 11 does not measure consultation services. • For BAM, there are no questions that directly measure whether income has been safeguarded or not. • For Natural Produktsia, no indicators on selling to BAM and HoReCa sectors in Box 12; Box 7 doesn't capture production • For Star Consulting, indicators for Box 6 don't capture all aspects of the change. Indicators for GSI1 are not all appropriate. Box 17 doesn't measure the increased income received. • For ABBA, indicators for Box 4 don't capture all of the relevant changes. Box 17 doesn't actually measure income. Also, several relevant indicators of qualitative sustainability are missing, such as members' satisfaction with ABBA, willingness of ABBA officials to continue, willingness of beekeepers and others to keep paying dues etc. • For Women's Rooms, several of the indicators for Boxes 14-17 have been designed to not require a baseline, yet qualitative information suggests the baseline value was not zero and so the indicator will not provide the full picture of the change. Indicator 15 doesn't capture change in women's involvement being publicized. Some qualitative indicators (e.g., intention to continue using the services, satisfaction of all actors with the changes). • None of the reviewed interventions have projections for all indicators. Projections are used mainly to develop expectations for the project client in the investment plan, and so

Strengths	Weaknesses
	<p>focus mainly on output indicators, rather than including a focus on what contribution the intervention will make to NAIC and other impact indicators. Star Consulting, ABBA and BAM have particularly few projections.</p> <ul style="list-style-type: none"> • Although in some cases new projections have been added when new phases of the intervention were started, none of the initial projections have been updated following their creation for any intervention.
Measuring Changes in Indicators	
<ul style="list-style-type: none"> • ALCP has an appropriate plan for reconstructing a baseline through the impact assessment. It will triangulate the results with other qualitative and quantitative information to test its adequacy. • ALCP has detailed monitoring plans outlining what information it will collect and when for all indicators. • ALCP has a system to monitor unintended effects. • ALCP has an appropriate system to capture differentiated results by gender. This includes both gender-disaggregated indicators and qualitative research that identifies gender-relevant issues and differences in results. 	<ul style="list-style-type: none"> • For Star Consulting, no final qualitative monitoring date was planned after August 2016, and information was not collected on several of the qualitative indicators. • For BAM, the investment plan is noted as the source of all quantitative baseline data, despite not being able to supply it all. A follow-up baseline study was conducted but is not described in the monitoring plan. • For ABBA, no method has been included to verify the reported honey sales, and indicators for Box 7 are only to be collected at the IA, which is late for a key indicator. Part-time jobs are being converted to full-time equivalent jobs incorrectly. Box 15 captures revenue, not sustainability (i.e., profit). • For Women's Rooms, there is no specific plan for final qualitative questions that are to be asked at the end of the intervention (though a mini-survey is planned). The methodology for defining what is considered a women-led initiative (vs. a mixed or men-led initiative) is not fully clear. • In some cases, qualitative information is not collected according to plan or not collected at the scheduled time. For example, Natural Produktsia hasn't fully collected information on Box 16. For Roki, qualitative information on indicator 14.1 was not collected according to plan. • Unintended effects are defined by ALCP as being a type of systemic change, which creates a bias to look for positive effects and changes within the sector rather than broader effects. Sometimes unintended

Strengths	Weaknesses
	effects have been defined as instances that an intervention did not work as intended.
Estimating Attributable Changes	
<ul style="list-style-type: none"> ALCP has a plan for conducting an impact assessment. The impact assessment will collect results for all of the interventions, because it is an area-based IA methodology. ALCP's approach to attributing partner-level results is appropriate and uses several methods of triangulation to verify the estimate. The quality assurance procedures that are planned to be implemented by ALCP's research partner in the impact assessment follow good practice and are comprehensive. 	<ul style="list-style-type: none"> ALCP's planned impact assessment only measures revenue, not net income, but ALCP plans to report this as NAIC. This almost certainly overestimates the NAIC, as beneficiaries often need to incur additional costs (e.g., purchasing veterinary drugs) in order to realize the additional revenue. This also likely overstates the number of beneficiaries (income scale), as they are calculated based on the number of beneficiaries who surpass the average revenue of the non-beneficiary group.
Capturing Wider Changes in the System¹	
<ul style="list-style-type: none"> ALCP defines what it means by systemic change and depicts certain aspects of it (i.e., imitation by other market actors) in nearly all of the assessed results chains. In two interventions (BAM and ABBA) non-imitation types of systemic change are included in the results chains without grey boxes. 	
Managing the System for Result Measurement	
<ul style="list-style-type: none"> Managers can explain to what extent underlying assumptions in the results chains are proving to be valid, and can cite decisions they have made based on the information provided by the results measurement system. Staff are able to accurately describe their tasks and responsibilities in results measurement. The team appears to have adequate financial and human resources to implement these interventions. The system for quality assurance is appropriate. 	

¹ As beneficiary level impacts of systemic change (indirect impacts) are not expected to have happened yet, control points 5.1.3-5 are not applicable for the interventions.

Part I: Must CPS

Code	Control Point	Compliance Criteria broken down for scoring (out of 10)	Livestock	Sector 2	Sector 3	Sector 4	Program	Average Score
1.1	An appropriate, sufficiently detailed and logical results chain(s) is articulated explicitly for each of the interventions.	1.1.1 A results chain is developed and documented for each intervention.	10.0					10
		1.1.2 Each results chain shows all key changes arranged in logical order, demonstrating as far as possible how the selected intervention leads to achievement of development goals.	8.0					8
		1.1.3 Each results chain is sufficiently detailed so that changes at all levels can be assessed quantitatively and/or qualitatively.	8.3					8
1.2	Each results chain is supported by adequate research and analysis	1.2.1 The programme has documented critical external assumptions that affect the logic of the results chain.	8.2					8
		1.2.2 The programme has documented reasons that support the logical links of the results chain	10.0				10	
		1.2.3 The documentation explains how the changes outlined in each results chain are likely to lead to lasting impact.	10.0				10	
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.	1.3.1 Mid and senior level programme staff can describe the respective results chain(s) covering their work.	9.5					10
		1.3.2 Mid and senior level programme staff can give examples of how they will use or how they have used (for programmes more than 1 year old) results chain to guide their work.	10.0				10	
		1.3.3 Key partners can describe the logic of interventions that is reflected in results chains. (W/A)	10.0				10	
1.4	The results chain(s) are regularly reviewed to reflect changes in the programme strategy, external players and the programme circumstances.	1.4.1 The programme has a clear system for reviewing the results chain(s) at least once a year.	8.8				10.0	10
		1.4.2 Use: The programme has evidence to show that the results chain(s) have been reviewed at least once in the last year.	10.0					9
		2.1.1 Quantitative and/or qualitative indicators are defined for each change in the results chain(s).	10.0					10
2.1	There is at least one relevant indicator associated with each change described in the results chain(s).	2.1.2 The indicators are relevant to the associated changes in the results chain(s).	8.7				9	
2.2	Information on changes at various levels of the results chain.	2.2.1 Assessment of changes includes qualitative information.	10.0					10
2.2	Information on changes at various levels of the results chain.	2.2.2 Qualitative information enables an appropriate assessment of why changes are or are not taking place and the character, depth and sustainability of changes at various levels of the results chain.	9.0					9
2.3	A small number of indicators at the impact level can be aggregated across the programme.	2.3.1 Each results chain includes either the common impact indicators, or alternative aggregatable indicators at the closest feasible level in the results chain. If any results chain does not include these indicators, written justification is provided explaining why not.	10.0					10
2.4	There are specific indicators that enable the assessment of sustainability of results.	2.4.1 Specific indicators (qualitative and/or quantitative) are defined that enable assessment of sustainability of results in the results chains.	10.0					10
		2.4.2 The indicators are relevant and appropriate to assessing the sustainability of results at key levels of the results chains.	9.0					9
		2.5.1 Mid and senior level programme staff can describe the indicators covering their work.	10.0					10
2.5	Mid and senior level programme staff understand the indicators and how they illustrate programme progress.	2.5.2 Mid and senior level programme staff can give examples of how they will use or how they have used (for programmes more than 1 year old) information on changes in indicators to inform their strategy and implementation decisions.	10.0					10
		3.1.1 A documented plan is in place to gather baseline information	10.0					10
		3.1.2 Use: The programme has baseline data which show the status of key indicators before activities have led to change	9.7					10
3.1	Baseline information on all key indicators is collected.	3.2.1 A documented plan is in place to collect information for each indicator at appropriate times.	9.7					10
3.2	Information for each indicator is collected using methods that conform to good research practices.	3.2.2 The plan is thorough, realistic and in accordance with good research practice. It shows for each indicator what information will be collected, when and how the information will be collected and how each indicator will be calculated or described.	9.0					9
		3.2.3 Use: The programme can demonstrate that it used the plan to collect information.	10.0					10
		3.2.4 Use: The programme can demonstrate that information collection conformed to established good practices (in terms of research design, timing, sampling, quality control, etc.)	8.7					9

3.3	Programmes have a mechanism for assessing and understanding differentiated results by gender.	3.3.1	Programmes have a system for assessing and understanding differentiated results by gender.	10.0				10.0	10
		3.3.2	Use: Programmes have used this system to assess and understand differentiated results by gender.	10.0					10
4.1	Attributable changes in all key indicators in the results chains are estimated using methods that conform to established good practice.	4.1.1	The programme has a documented plan for assessing and estimating the attribution of observed changes to programme activities for each of the key indicators in the results chain.	10.0					10
		4.1.2	The methods chosen to assess and estimate attribution link back to the results chains, are appropriate to the programme context and conform to established good practices.	7.0					7
		4.1.3	Use: The programme has used the plan to estimate attributable change in indicators.						
		4.1.4	Use: The programme can demonstrate and staff can explain the methods used to assess and estimate attribution and how the methods conform to established good practices.						
		4.1.5	Use: Figures are supported by clear calculations; assumptions are outlined if necessary.						
6.1	Costs are tracked annually and cumulatively.	6.1.1	A clear, accounting system is in place to track costs and produce annual and cumulative totals of all programme-related costs spent in country.	10.0				10.0	10
		6.1.2	Use: The programme has annual and cumulative totals of all programme-related costs spent in country.	10.0					10
		7.1.1	The programme has a documented system for estimating programme-wide impacts for common impact indicators (and/or other high level common indicators) at least annually.	10.0				10.0	10
7.1	The programme clearly and appropriately aggregates programme-wide impact at least annually.	7.1.2	The system for estimating programme-wide impacts is clear, thorough and appropriate.	10.0				10.0	10
		7.1.3	Use: The programme can clearly explain how the estimates were derived and show supporting calculations. These calculations take overlap into account. (W/A)	8.0				8.0	8
		7.2.1	The programme has a report with clear estimates of programme-wide impacts for the common impact indicators (and/or other high level common indicators) produced in the last year. The report outlines the context needed to understand the quantitative results.	10.0				10.0	10
7.2	The programme produces a report at least annually which describes results to date.	7.2.2	The report outlines qualitative achievements and discuss progress towards the sustainability of results.	10.0				10.0	10
		7.2.3	When reported changes are due in part to the work of other publicly-funded programmes and private contributions, they are acknowledged in the report.	10.0				10.0	10
		8.1.1	The programme has a documented system in place to show how information from the results measurement system will inform management decision making.	10.0				10.0	10
8.1	The programme has a clear system for using information from the results measurement system in management and decision-making.	8.1.2	The system is realistic and ensures that results information is regularly and effectively integrated into management decision making.	10.0				10.0	10
		8.1.3	Use: Managers can explain to what extent underlying assumptions in the results chain(s) are proving to be valid, and can cite decisions they have made based on the information provided by the results measurement system.	10.0					10
8.2	The system is supported by sufficient human and financial resources.	8.2.1	The programme can show that sufficient human and financial resources are available and have been allocated to manage and implement the results measurement system.	10.0				10.0	10
		8.2.2	Staff are able to accurately describe their tasks and responsibilities in results measurement.	10.0				10.0	10
		8.3.1	Tasks and responsibilities in relation to results measurement are appropriate and documented.	10.0				10.0	10
8.3	The system is well managed and integrated with programme management.	8.3.2	Evidence exists of the results measurement system having been institutionalized, for example in the form of inclusion in programme management documents, job descriptions, staff performance reviews, regular meetings etc.	8.0				8.0	8
		8.3.3	All programme staff have access to written guidance (e.g. manual or staff guide) on how to implement all elements of the results measurement system.	10.0				10.0	10
		8.3.4	The programme systematically checks the quality of all results measurement activities and outputs in the system.						
Totals									440
Percentage									96%

Part II: Recommended CPs

Compliance Criteria broken down for scoring [out of 10]		Livestock	Sector 2	Sector 3	Sector 4	Program	Average Score
Code	Control Point						
1.5	The results chairs or another tool outline what broader systemic changes are expected, and how the programme expects to influence them.	10					10
1.6	The research and analysis underlying the results chair(s) take into account the risk of displacement.	7					7
2.6	Anticipated impacts are realistically projected for key quantitative indicators to appropriate dates.	2.6.1					5
		2.6.2					5
		2.6.3					-
3.4	Programmes monitor to identify unintended effects.	3.4.1	8				8
		3.4.2	10				10
5.1	Systemic changes are assessed using appropriate methods.	5.1.1					10
		5.1.2	10				10
		5.1.3					
		5.1.4					
		5.1.5					
6.2	Costs are allocated by major component of the programme. (Applicable only to programmes with more than one main intervention)	n/a				n/a	n/a
7.3	Results of systemic change are reported.	10					10
7.4	Results are published.	10					10
7.4.1		A document with the results and costs described in sections 7.1-7.3 is made publicly available.				Totals	
		Percentage				85	
						78%	