

3-Year System Business Plan Companion Document:

Action 4: Strengthen program performance management

Prepared by: CGIAR System Management Office, following consultations with Science Leaders and the Strategy Impact Monitoring and Evaluation Committee of the System Council ('SIMEC').

Document status: **An implementation plan** for input by the System Council at SC7 (November 2018) in advance of final approval by the System Management Board in December 2018.

Action for SMB10: **Endorse the implementation plan** for inclusion as a 2019-2021 Business Plan Companion document to the System Council's SC7 meeting.

A. Implement a 12-Point Framework for action on program performance management: status report and plans

1. A 12-point framework for action is being implemented, and the following actions are planned over the business plan period:

Business Plan Companion: Strengthen program performance management
Action 4

Area of action	Status/Comment
PAST: RESULTS REPORTING	
1. <i>CGIAR Annual Performance Report</i>	New style version piloted for 2017 performance, delivered in September 2018 (presented in this meeting)
2. <i>CRPs/Platform reporting requirements</i>	Revised reporting requirements: 2017 templates and guidance can be seen in this link , and updated versions are being produced for 2018 reporting and the 2019 POWB. Quality assurance procedures for all underlying data to be systematized (pilot in 2019).
3. <i>Common results reporting indicators</i>	A first set of definitions and guidance being used for reporting on 2017 is in this link . Updated versions of definitions and guidance will be posted by December 2018. Results are in the Annual Report, which has links to underlying databases.
4. <i>Outcome and impact case studies</i>	Template and guidance piloted for 2017 reporting (available in this link), updated version being produced for 2018 reporting. Results: some included in the 2018 annual report, others being added to the dashboard.
5. <i>Program Results Dashboard</i>	Under construction. An early demonstration version with partial (real) data will be available to view at the 7 th System Council meeting in November 2018.
6. <i>Learning from past Evaluations</i>	Valuable lessons from past evaluations are being used in the development of the first business cycle. Agreement on a new evaluation workplan for this business cycle will ensure that lessons and recommendations emerging from evaluations can inform, at key decision times, independent verification and learning in program performance management, portfolio design and/or program implementation.
7. <i>Impact Assessments</i>	Important for the 2019-2021 business plan will be sufficient investment and strategic scheduling of impact assessments to allow for findings to provide robust evidence of the long-term impact of research investments.
PRESENT: IMPROVED MANAGEMENT OF CURRENT RESEARCH PROGRAMS	
8. <i>Within-cycle reviews and evaluations</i>	Within a 3-year cycle, appropriate reviews and evaluations will be planned and carried out on relevant topics and at specific junctures to support improved management of program performance.
9. <i>Program Performance Management Standards</i>	Section B below describes how Program Performance Management Standards will be implemented for 2019-2021. Formal adoption by the System Council and System Management Board by early 2019 of the ISPC Quality of Research For Development Framework ('QoR4D') is proposed to ensure that relevant elements of the QoR4D framework can be used as the benchmark to assess selected performance standards, and can be well integrated into the approved CGIAR System Risk Management Framework.
FUTURE: IMPROVED DECISION-MAKING ON FUTURE PROGRAMMING AND FUNDING	
10. <i>Quality at Entry Assessment</i>	Programs will need to demonstrate that they have passed performance standards.
11. <i>Planning landscape</i>	The Planning landscape is described in this business plan and will inform decision-making for the future, including on fund allocation for the next cycle.
12. <i>Allocation criteria and tools</i>	For development by end-2021 as part of new portfolio development process

Why “Program Performance Management”?

3. In AR4D, unlike in many development projects, there is no simple way *at the system (Portfolio) level* to use monitored outputs and outcomes as a basis for judging the quality of research management. The reasons include:
 - a. Outcomes and impacts of CGIAR AR4D result from **complex chains involving multiple organizations**, so that the contribution of CGIAR cannot be simply ‘monitored’ but needs to be estimated via rigorous impact assessments (which are expensive and therefore appropriately employed selectively). This is reflected in the endorsed **CGIAR Performance Management Conceptual Framework**¹, with its three spheres of ‘control, influence and interest’.
 - b. **The long period of time** required for much of AR4D uptake to get to practical outcomes and impacts, in the circumstances of rural low-income agriculture (often 5-20 years or more) – so that ‘currently measured outcomes’ flow from past research outputs and only indirectly (if at all) reflect the current research Portfolio².
 - c. **Comparing apples and oranges:** The research programs produce very different types of outputs and outcomes, over very different time frames. It is not possible to make simple judgments about whether a new variety is ‘better’ than a new water management technique or a change in international policy. Bibliometrics (monitoring statistics on research publications) is the traditional means of addressing this, but is also fraught with difficulty when comparing diverse fields, and can give strong incentives to researchers to prioritize peer-reviewed papers over practical outcomes³.
 - d. By its very nature, research is a step into the unknown, and it is not expected that all lines of research will be equally successful⁴. The art of research management, rather than trying to ensure equal attainment of output targets

¹ See Conceptual framework in SC3-03, 17 Nov 2016 http://www.cgiar.org/wp-content/uploads/2016/11/SC3-03_Towards-PerformanceMgmtSystem_17Nov2016.pdf

² Another big difference from development projects is that the target geographical area for R4D outputs is normally much larger and more diverse than the actual operational area where the research is carried out and may not even be clearly known at the beginning of a research program, which makes ‘baselines’ difficult.

³ Wouters, P. et al. (2015). The Metric Tide: Literature Review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management). UK HEFCE; Rijcke, Sarah de et al (2016), ‘Evaluation Practices and Effects of Indicator Use—a Literature Review’, *Research Evaluation*, 25 (2016), 161–69

⁴ “The business of science is intensely frustrating. Most experiments fail, most great ideas come to nothing, and most genuine discoveries turn out to be of modest importance. Years of effort can easily be wasted on what turns out to be a mirage.” Dr Rupert Beale, University of Cambridge (2018), London Review of Books in an article on CRISPR.

across all research lines, is **to manage a portfolio of research to get the best overall return**, including supporting some ‘high risk, high return’ lines that are potential game changers but may take many years to deliver⁵. In terms of cost-benefit analysis, it has been demonstrated that a few high performing research innovations can yield returns that amply repay the cost of the entire CGIAR Research Portfolio⁶.

- e. The highly technical nature of most research means that **subject matter specialists in touch with the latest thinking and developments are needed to assess its scientific quality**, efficiency of design and (often) other aspects such as the appropriateness of partnerships. This argues against trying to micromanage R4D performance from a distance.
 - f. There’s a very real risk of the issue becoming “what’s measured, matters”: **poorly chosen metrics can set up strong ‘perverse incentives’** (or ‘goal displacement’) for research programs and individual researchers to, for example, over-claim outputs and outcomes (or split reporting to increase numbers), to set R4D targets that are less demanding and include less ‘high-risk, high return’ work, and/or to focus their attention on deliverables such as publications, instead of outcomes. These ‘perverse incentives’ are not a rare phenomenon: they have consistently been recorded from previous ‘results - based management’ attempts by CGIAR and in research establishments elsewhere⁷. Excessive ‘command and control’ also runs the risk of penalizing creativity and demoralizing researchers, making it harder for CGIAR to attract and retain top talent.
 - g. Finally, the **environment for AR4D is changing rapidly**, with new technologies and new types of skills required. In new programs, **key research personnel and programs may also have changed** from those who were responsible for past results. Thus, the historical success of past research outputs and outcomes from a particular AR4D program may provide some general comfort to investors, but not necessarily be predictive of future success.
4. For the reasons above, the approach proposed to performance management is a combination of careful *evaluation* of the past program, combined with assurance that *current research management systems* are designed, positioned and managed to

⁵ Examples of high risk, high return lines are attempts to develop a malaria vaccine, or C4 rice. See also Perrin, B (2002) How to—and how not to – evaluate innovation. *Evaluation*: 8: 13-28

⁶ Raitzer, D; Kelley, T (2008) Benefit–cost meta-analysis of investment in the International Agricultural Research Centers of the CGIAR, *Agricultural Systems* 96:108-123

⁷ For relevant CGIAR experience, see [Immonen S, Cooksy LL \(2014\): Using performance measurement to assess research: Lessons learned from the international agricultural research centres](#) *Evaluation* 20: 96–114. For international experience see Footnote 1.

deliver impacts as effectively and efficiently as possible. This proposal operationalizes the CGIAR Performance Management Framework and also builds on recent experience elsewhere⁸.

5. Does this mean that monitoring results is not important for performance management?
On the contrary: **it is expected that CRP management and governance bodies will monitor the progress of their own research program closely.** They are close to the research and can interpret monitoring data sensibly, for example they are in a position to know whether a missed milestone is reasonable or not. It is expected that CRP management will use monitoring results, along with other information, to manage the research portfolio, including cutting back unpromising lines of research and reallocate resources to more promising lines. Part of the performance management assessment is to check how effectively CRP management does this⁹.

B. Implement Program Performance Management Standards for 2019

6. The main objectives of the proposed CGIAR Program Performance Management Standards are:
- a. To **provide assurance** to Funders and other stakeholders that program management standards are high, and that they can invest with confidence. This means ensuring for example that research design and partnerships are fully focused on delivering impact; that research is of high scientific quality; that research managers are taking tough decisions when necessary e.g. to stop funding some areas and reallocating funding to others; and that other aspects of management systems are in place to promote a variety of agreed system objectives.
 - b. To **improve program performance management** across CGIAR wherever needed. *The effort required here - for programs to manage optimally to meet their programmatic objectives - should not be underestimated in a system where management of financial and human resources, projects, ethics and intellectual property are nearly all the responsibility of Centers rather than program managers.* The requirement to meet the standards should serve as a guarantee that program performance management is consistently good across CGIAR, and not just in the top-performing programs.
 - c. To **focus program efforts** on a limited number of well-defined high-priority areas identified jointly by key stakeholders, in each program cycle, to *complement (not*

⁸ For example IDRC's RQ+ framework and performance management in the non-CGIAR IARCs

⁹ It is important however to recognise the current limitations on the power of CRP management, who report that they normally only are able allocate W1/2 funding (on average about 20% of total funding in 2017, and as low as 7% of total in one CRP) and do not have direct management control over the development and direction of bilateral projects which comprise most of the portfolio.

replace) the more complex analysis carried out in program evaluations and appraisals. The standards shine a light on specific aspects of management (that can get ‘lost’ in an overall appraisal with many elements to consider) and provide a strong incentive for managers to fix any problems within the business cycle.

<p style="text-align: center;">Good management means:</p> <p style="text-align: center;">High science quality</p> <p style="text-align: center;">Decisions are made in an open, transparent and equitable way</p> <p style="text-align: center;">Efficient and effective performance</p> <p style="text-align: center;">Risks are well managed</p> <p style="text-align: center;">Credibility is sustained</p>

7. **The proposed 2019-2021 standards are set out in Table 1 below.** They will:
- a. For each three-year cycle, focus on a few simple but powerful, high priority program management areas.
 - b. Be set to provide immediate assurance to the System Management Board and System Council in cycle 1 (2019-2021) that the fundamentals of good research program management (such as accessible documentation) are in place, and as a first step towards meeting more challenging standards of excellence in later cycles (2022 and beyond).
 - c. Set targets at a level which move the whole system forward -: all programs within the CGIAR portfolio should be able to meet them within the three-year cycle, given appropriate commitment and investment.
 - d. Change and/or evolve over successive 3-year cycles, to reflect agreed priorities, and “ratchet upwards” (e.g. target percentages will increase over cycles).
 - e. Be assessed once per three-year cycle via a ‘desk review’ of online documentation – by the end of year 2 to give programs time to invest (where needed) to meet standards before the next cycle if they do not pass the assessment first time. A draft timeline is in **Table 2** below.
 - f. Programs that do not pass standards in the first assessment will be given specific feedback on improvements required, and they will self-report (by the end of the same business cycle) against measures taken to resolve any problems.

- g. The entity that carries out 'quality at entry' will usually accept this self-report but will have the right to carry out its own checks if needed.
- h. The entity that undertakes the assessment may vary over successive business plan cycles, should the standards change (point d. above)¹⁰. For the period 2019-2021, it is proposed that [entity/entities TBC] provides that objective assessment, based on an assessment outline reviewed and endorsed by the System Council.
- i. Contribute to the data that supports the CGIAR System operating according to a 'combined assurance' model in respect of System-wide opportunities and risks (refer companion document to Action 8 of the CGIAR System 3-year Business Plan).
- j. The [ISPC Quality of Research For Development Framework](#), upon formal approval by the System Council, will be incorporated into the performance standards assessment process (refer appendix 1 with indicative mapping to the [CGIAR System Risk Management Framework](#)).

¹⁰ Responsibility for assessment is being discussed in the context of the System Council's discussions on the appropriate Terms of Reference and workplans for its science, impact assessment and evaluation bodies/functions, and the introduction of the CGIAR System Internal Audit Function and its strengthening role over 2018. Decisions that are anticipated to be taken by virtual means by the System Council immediately prior to SMB10 should inform identification of the entity.

TABLE 1 LIST OF PROGRAM PERFORMANCE MANAGEMENT STANDARDS PROPOSED FOR FIRST CYCLE, 2019-21

Overarching aim	Management standard for all programs to meet in first cycle (2019-2021)	Proposed (<i>draft</i>) metric, assessed by appropriate independent body
Provide assurance to Funders and other stakeholders that all projects in program are appropriate and relevant (by Cycle 1)	1. Program has a transparent and logical process for selection, prioritization and inclusion of new projects and withdrawal of projects from program, based on the theory of change and factors such as comparative advantage, scientific merit, potential value for money.	Qualitative assessment (using agreed rubric) of quality of process and documentation.
CGIAR recognized as a global leader for the science of gender in agriculture, integration of high- quality gender research throughout the CGIAR research portfolio (by Cycle 2-3)	2. Correct reporting of gender within the research portfolio. (Note that the management standard is part of a wider set of agreed actions toward meeting CGIAR gender objectives.)	Agreed target for OECD and other gender markers (defined by gender group) appropriately applied.
Provide assurance to Funders and other stakeholders that CGIAR pooled budgets (W1/2) are effectively and efficiently managed (by Cycle 1)	3. Program has transparent systems for planning and managing budgets to reach program objectives, and clear and efficient division of responsibility between Programs and their implementing partners (including Centers).	a. Annual Plan of Work and Budget makes clear logical links between budgets and activities b. Budget holder responsibilities for key Program staff are clearly assigned and documented for W1/2 funding.
Provide assurance to Funders and other stakeholders that the program is managed effectively to further stated objectives and SRF targets. (Cycle 1)	4. Program progress and priorities are regularly reviewed, and logical and transparent decisions are taken about (re)prioritization of W1/2 funding, including activities to expand or cut back.	Qualitative assessment (using agreed rubric) of the quality of analysis and process.
Provide assurance to Funders and other stakeholders that CGIAR results reporting is high quality and credible and supported throughout by high-quality evidence. (Cycle 2-3)	5. Program reporting to CGIAR (annual reports, common reporting indicators, outcome-impact case studies) is of adequate quality and the evidence presented is properly archived, linked and accessible.	Qualitative assessment (using agreed rubric) of the quality of program reporting, supported by random sampling to look at specific aspects in more detail.
CGIAR programs and projects adequately transparent to international standards, such as IATI https://iatistandard.org/en/about/iati-standard/ (Cycle 2-3)	6. All key program and project documents accessible and findable to be viewed electronically by System Organization and system advisory bodies.	An agreed list of key documents is available in agreed CGIAR repositories, with working links.

TABLE 2 DRAFT TIMELINE FOR PERFORMANCE STANDARDS USE AND FUND ALLOCATION IN FIRST CYCLE

Actual dates	Cycle years	Action	Responsibility
Nov 18	Year 0	Standards for 2019-21 confirmed by SC. Metrics and responsibilities for measurement agreed.	SMO, SMB, SC
Dec 18	Year 0	Preparation of Annual Plans of Work and Budget (POWB): Programs need to include any investment needed to meet standards in their POWB for Year 1. Most actions in cycle 1 will not require significant extra funds. However, to improve evidence of results there may be a need for additional investment in monitoring and impact assessment.	Programs
Sept-Oct 2019	Year 1 Q4	Programs need to include any investment needed in their POWB for Year 2, as above.	Programs
July-Aug 2020	Year 2 Q2	Independent assessment of performance standards for each program. In principle this will be a desk study based on assessment of data available online. (Availability of data online is a key issue which needs to be sorted out in first cycle.) Stable funding in cycle: no cuts	TBC
Oct-Nov 2020 (processes from proposal to approval)	Year 2 Q2	Discussion on performance standards and reports for current cycle. Performance standards proposed for following cycle, and agreed by program leaders, SMB and SC.	SMO, Program leaders, SMB, SC
Jan 2021	Year 3 Q1	Programs draft their revised proposals/implementation plans for the next 3- year cycle. This should include any investments needed to meet the agreed next cycle of standards. For programs that not meet certain current standard(s), this should also include an annex with a written response to the independent assessment, with information on any improvements already made and a performance improvement plan to meet the relevant standard(s) before end of year.	Programs
March 2021	Year 3 Q1	Programs that failed standard(s) in previous cycle: annual report includes a section on improvements made in response.	Programs
May 2021	Year 3 Q2	Quality at Entry (QaE) assessment for new proposals/implementation plans.	TBC

Business Plan Companion: Strengthen program performance management
Action 4

Actual dates	Cycle years	Action	Responsibility
		Passing the standards is a 'necessary but not sufficient' condition for future funding. Passing performance assessment or satisfactory self-report on improvement required as 'entry ticket'. If self-reporting is not convincing, limited checks could be carried out.	
Nov 2021	Year 3 Q4	Approval process for new proposal/implementation plan	

Appendix 1

QoR4D as a source of assurance for the CGIAR System Risk Management Framework

As noted above, it is envisaged that the System Management Board will endorse, and the System Council will formally adopt a CGIAR Quality of Research for Development ('QoR4D framework') by early-2019, thus providing the 'benchmark' to provide assurance for relevant performance standards linked to the November 2017 approved CGIAR System Risk Management Framework.

An indicative linkage of the ISPC's proposed QoR4D framework to CGIAR's agreed System-wide risk families is set out below.

(I) KEY ELEMENTS

QoR4D KEY ELEMENT	SUMMARIZED DEFINITION
RELEVANCE	The importance, significance and usefulness of the research objectives, processes and findings to the problem context and to society, and CGIAR's comparative advantage to address the problems.
CREDIBILITY	The quality of science, implying that the research findings are robust, and sources of knowledge are dependable.
LEGITIMACY	The research process is fair and ethical and perceived as such.
EFFECTIVENESS	That research generates knowledge, products and services that stimulate actions that address the problem and contribute to solutions and innovations.

(ii) INDICATIVE MAPPING - QoR4D AS SOURCE OF ASSURANCE LINKED TO CGIAR SYSTEM RISK MANAGEMENT FRAMEWORK

*** TO NOTE THAT OTHER ASSURANCE PROVIDERS (INTERNAL AUDIT, SPIA, EVALUATION SERVICES, SC INTELLECTUAL PROPERTY GROUP, SMB ETC) PROVIDE ASSURANCE ON VARIOUS AREAS OF CGIAR'S OPPORTUNITIES AND RISKS, TO CONTRIBUTE TO AN OVERALL COMBINED ASSURANCE APPROACH (AS ADOPTED IN THE RISK MANAGEMENT GUIDELINES OF THE CGIAR SYSTEM)*

	CGIAR "Opportunities and Risk " draft indicators	Assurance via ISDC under QoR4D framework			
		Relevance	Effectiveness	Legitimacy	Credibility
CGIAR is no longer a front runner	Compelling shared research agenda reinforces Funders commitments				
	CGIAR science is relevant and cutting edge				
	Seizing the "next thing" gives sustainable competitive advantage				
	CGIAR's research agenda aligns with international community priorities				
CGIAR loses its central role	Intellectual assets produced by CGIAR are managed innovatively and leveraged				
	CGIAR a desired and supportive partner				
	Diversity and predictability of funding maintains CGIAR as a global player				
	Research activities are well deployed and coordinated in targeted countries				
	CGIAR genebanks demonstrate their unique role				
Non adherence to appropriate values	CRPs and Platforms deliver on the objective of the SRF				
	Ethical research practices employed to achieve research results				
	CGIAR values and desired behaviors strengthen its credibility and attractiveness				
	Adequate processes are in place to prevent or detect inappropriate use of funds				
Unsatisfactory evidence and assurance received	Clear and transparent financing of the CGIAR portfolio				
	Impact evidenced by hard data				
	Funds used in accordance with approved annual work programs and budgets				
	Activities are implemented for CRPs, platforms and flagships as agreed with Funders				
	Delivery of the CGIAR Portfolio is adequately evidenced and visible				
Poor execution undermines capability	Project assessment and lifecycle management are effective				
	Adequate use of intellectual property and licensing tools maximizes accessibility and/or impacts, including via the production of International Public Goods				
	Top talent is attracted to and retained by CGIAR and Centers				
	Effective and efficient operations minimize costs and protect key assets (people, system, data) against threats (internal, external, cyber, etc).				
	Financial stability requirements met by all Centers				
	CGIAR System adds significant value to outweigh related costs and constraints				