



Photo Credit: Augustus Addo, IWMI

SPIA Feedback: Portfolio Impact Assessment Plan Draft Concept Note

June 2026

This feedback is submitted by the Standing Panel on Impact Assessment (SPIA) in its advisory capacity to the CGIAR system. The observations below are also informed by conversations with the Impact Assessment Community of Practice (IA-COP).

SPIA welcomes the initiative to develop a portfolio-level approach to impact assessment (henceforth IA). The need to demonstrate CGIAR's achievements at scale and coordination around shared questions is important. Our feedback focuses on whether the draft IA plan will produce evidence that is credible, methodologically sound, and useful for decision-making and meets the identified evidence needs of funders.

1. Overarching Concerns

1.1 Mislabeling

The concept note proposes a portfolio-level evidence plan that bundles several fundamentally different types of evidence under a single umbrella: projected benefits and foresight models, adoption and reach surveys, rigorous impact evaluations, process tracing and case studies, and cost-benefit analysis. Each of these methods answers different questions, operates at different certainty levels, and serves different decision-making purposes. Treating them as interchangeable components of an IA plan conflates their methodological foundations and risks undermining the credibility of the evidence produced.

If the plan is intended as a broader portfolio results or overarching MELIAF strategy, it can be labelled and structured as such, with explicit recognition of what each evidence type can and cannot deliver. If it is intended as an IA plan, it should instead focus on questions answerable by IA methods.

1.2 "Impact" is not defined

The concept note does not define what it means by impact. The term is used to encompass projected benefits, reach and adoption, and rigorously measured causal effects. These are categorically different. Without a clear definition, the plan cannot specify what standard of evidence is required to answer the proposed questions, making it impossible to assess whether the proposed methods are adequate.

1.3 The SPIA approach to IA for CGIAR may be helpful

In 2020, SPIA published its [approach to portfolio-level impact assessment](#) (SPIA Technical Note 8). That framework makes two distinctions:

- Accountability vs. learning: Accountability studies document the returns from “big win” innovations at scale, using reach and benefit per beneficiary (with credible counterfactuals). Learning studies test specific steps in theories of change for innovations being scaled, to inform decisions. These require different methods, serve different purposes, and should not be conflated.
- Skewed returns: Returns to agricultural research are highly skewed, i.e., a few innovations with very high returns can justify the entire investment in the CGIAR system. This means portfolio-level impact evidence does not require comprehensive coverage of all impact areas. It requires rigorous evidence on a strategically selected set of innovations, emerging from reach studies for example, where credible causal estimates are feasible. This is a different approach from the one proposed, which seeks to generate evidence across all 5 impact areas and 23 questions simultaneously.

The concept note could reference, engage with, and build upon this existing framework.

1.4 The current IA model is not necessarily failing

The concept note diagnoses CGIAR's bottom-up approach to IA as failing. This may overstate the case. CGIAR is an AR4D organization operating within a complex innovation system. Only a portion of its research leads to innovations, and very few of those innovations reach scale, often over very long-term horizons. With limited and declining funding for IA, the current approach has been opportunistic, focusing on cases with good probability of success and/or where funders and management have a particular interest. Perhaps the question is not whether to replace this model, but how to complement it with enhanced coordination.

2. The three overarching questions

2.1 Progress on Impact Area indicators

The concept note proposes that institution-level impact targets will be set for 2040 through a projected benefits exercise, and that the impact assessment of these targets will be “a major guiding element” for the portfolio plan. However, using IAs to validate projected benefits is methodologically inappropriate as it reduces IA to a confirmation exercise rather than an independent source of evidence.

Furthermore, the plan's own acknowledgement that most evidence will be produced at “moderate to low levels of certainty” is a serious concern. The purpose of portfolio-level impact evidence is to build stakeholder confidence. Evidence of moderate to low certainty risks doing the opposite, particularly if it is presented under an “impact assessment” label that implies rigor.

2.2 Contribution to transformation of food, land, and water systems

Documenting CGIAR's contribution to systemic transformation is conceptually important. However, the two sub-approaches proposed (demonstrating cases where CGIAR “contributed” to solving major challenges and evaluating the effectiveness of “change pathways”) are ambiguous. These are questions that may be valuable for institutional learning, but they should be clearly distinguished from questions that can be answered using causal evidence.

2.3 Cost-effectiveness

Cost-effectiveness is a critical question for funders. However, the document does not demonstrate what cost-effectiveness analysis requires in practice. Specifically:

- The use of the 2025–2030 investment period as the time frame for cost-effectiveness is problematic. Evidence consistently shows that CGIAR innovations take well over 5 years to reach scale and generate measurable impact. The [recent ROI exercise](#) demonstrated that costs were incurred well over the 5-year mark for the few feasible cases studied.
- To conduct credible cost-effectiveness analysis at any level, data on costs, reach and causal impact are prerequisites. The plan does not address how this data will be generated or made available, nor how in-house evidence will complement the commissioned studies proposed.

3. The 23 specific questions

3.1 Most questions are not formulated as research questions answerable by IA methods

The majority of the 23 questions are formulated as target-verification questions (“How many people’s incomes have increased above national poverty thresholds due in part to CGIAR contribution?”), transformation questions (“What are the big development challenges where CGIAR has made a difference?”) or cost-effectiveness questions (“What is the projected value of the 2025–2030 portfolio versus the total investment made?”) rather than research questions that IA can credibly answer.

Better formulated questions could separate effectiveness from scale. For example, instead of “How many women of child-bearing age consume a nutritionally adequate diet due in part to CGIAR contributions?” the plan could ask: “What is the reach of CGIAR innovations that aim to improve dietary outcomes?” and “Which of these CGIAR innovations are most effective at improving dietary adequacy among women of child-bearing age, and how?” The latter is answerable by impact evaluation; the former by reach studies. Together they can inform an estimate of total contribution, but each requires a different method, and the distinction must be explicit.

3.2 Sequencing logic to the questions and methods

Ideally, understanding whether an innovation is being used, and by whom, typically precedes efforts to measure its effects. Reach evidence, like the kind collected by SPIA, helps establish the foundation by identifying which CGIAR innovations have scaled, in which geographies, and through what channels.

The concept note currently treats all 23 questions in parallel, without attention to which are foundational and which are dependent. For example, Question 15 (“How widely are CGIAR innovations used?”) is not simply one question among 23, but a prerequisite for answering most of the others.

This sequencing applies to methods as well. Reach studies should precede impact evaluations, which should ideally precede portfolio-level cost-effectiveness estimates. Each stage produces the evidence that the next stage requires as an input. A plan that does not make this sequence explicit risks commissioning impact studies before the adoption evidence exists to support them, or attempting cost-effectiveness analysis without the causal estimates needed to populate it.

4. Absent but critical elements

4.1 Synthesizing across methods

The plan proposes that projected benefits will be “corroborated by” ex-post impact assessment. But it never specifies what corroboration means methodologically. When a projected benefits model and an impact evaluation produce conflicting estimates, how will the discrepancy be resolved? For instance, SPIA research on [orange fleshed sweet potato \(OFSP\) in Uganda](#) showed that model-based adoption estimates of the biofortified crop were up to 14 times higher compared to ground-truthed estimates (from nationally representative surveys and DNA fingerprinting). Without a framework for synthesizing evidence across methods, the plan risks producing incoherent or contradictory portfolio-level claims.

4.2 No space for null/negative findings

The plan is entirely framed around documenting targets and successes, leaving limited space for findings of no impact, unintended consequences, or innovations that failed to scale. SPIA’s Technical Note explicitly notes that finding evidence of lack of impact is equally valuable feedback into research and scaling efforts.

5. The role of SPIA

5.1 Synthesizing across methods

The concept note proposes to “work with SPIA to ensure country study scope (Phase 2 and 3) is shaped by the priority questions.” We note that SPIA’s credibility with funders and external stakeholders derives from its independence as SPIA selects what to study based on where rigorous evidence is feasible and where evidence gaps are most consequential. SPIA studies already generate evidence on reach, diffusion, and impact of CGIAR innovations. This evidence can and should inform portfolio-level assessments. But the contribution should flow from SPIA’s independent judgment about where credible evidence is achievable, instead of alignment to a pre-

defined framework.

SPIA is prepared to advise on the further development of this plan and to ensure that its own work contributes to and builds on the system's evidence needs, consistent with its mandate for independence, methodological rigor and strengthening IA capacity across CGIAR. In this direction, SPIA will continue supporting IA researchers in the system to design and implement relevant impact assessments that expand the evidence-base and help inform portfolio-wide decision-making.

6. Summary recommendations

1. **Clarifying what the plan is.** It may be more accurate to frame this as a portfolio results strategy, with clear distinctions between evidence types, rather than an impact assessment plan per se.
2. **Defining "impact" explicitly.** It would help to articulate what standard of evidence counts as impact evidence versus other types (projections, narratives, case studies), and how these different types relate to one another in a well-sequenced synthesis framework.
3. **Building on the SPIA approach.** The accountability/learning distinction from SPIA Technical Note 8 could serve as a useful conceptual backbone, with the "big wins" framework offering a credible basis.
4. **Reformulating the questions.** Many of the questions could be strengthened by separating effectiveness questions (answerable by IA) from scale questions (answerable by reach studies) and target-verification questions (answerable by projected benefits).
5. **Being selective.** A tighter set of questions where credible evidence is achievable may ultimately be more persuasive to funders than 23 questions producing evidence of "moderate to low certainty."
6. **Preserving SPIA's independence.** SPIA's evidence can and should inform portfolio-level decisions, though SPIA study selection works best when driven by independent judgment about where credible evidence is feasible.
7. **Allowing for null/negative findings.** The framework would be more credible with external stakeholders if findings of "limited or no impact" are treated as valuable evidence rather than failures.

Note: SPIA welcomes further engagement as this plan is developed further and would be happy to contribute with technical inputs.