

# Report:

# CGIAR Standard Indicators

**An Assessment of the Standard Indicators for  
the One CGIAR's Common Results Framework**

Prepared by: Rachel Claus, Rachel Davel, & Brian Belcher



---

Any views expressed in this publication are those of the authors. They do not necessarily represent the views of the One CGIAR, the authors' institutions, or financial sponsors.

---

## Acknowledgements

The Sustainability Research Effectiveness Program is supported by an Ashoka Research Chair, the Canada Research Chairs Program, and the Canadian Social Sciences and Humanities Research Council (SSHRC). This report was supported by the One CGIAR System Management Office. We thank Julien Colomer, Frank Place, Tom Zeppenfeldt, Angela Gimenez Barrera, and Claudia Velasco for their contributions to this work.

## Produced by

Sustainability Research Effectiveness Program  
Royal Roads University  
2005 Sooke Road  
Victoria, British Columbia  
V9B 5Y2  
Canada

Telephone: +1-250-391-2600 ext.4407  
Email: [brian.belcher@royalroads.ca](mailto:brian.belcher@royalroads.ca)  
Website: <https://researcheffectiveness.ca>

© SRE 2023



This research was carried out by the Sustainability Research Effectiveness Program at Royal Roads University in collaboration with the Program Performance division of the One CGIAR System Management Office.

The work of the One CGIAR System Management Office is supported by the CGIAR Trust Fund.

## Table of Contents

Table of Contents .....	ii
List of Tables .....	iii
List of Figures.....	iii
List of Acronyms.....	iii
Executive Summary.....	iv
Introduction .....	1
Methods .....	2
Results .....	3
Output Indicators .....	3
Outcome Indicators.....	5
Impact Indicators .....	7
Application of the Standard Indicators and Reporting in the Results Dashboard .....	11
Conclusion and Recommendations .....	13
References .....	14
Appendix 1. Rulebook.....	15

## List of Tables

<i>Table I. Summary of SMART assessments for the standard indicators.</i> .....	v
<i>Table II. Summary of SMART assessments for the NPIs.</i> .....	v
<i>Table III. Summary of error types.</i> .....	v
<i>Table IV. Thematic areas not covered by the standard indicators, but observed in the NPI set.</i> .....	vi
Table 1. Standard indicators included in the review. ....	1
Table 2. Summary of SMART assessments for the standard indicators. ....	3
Table 3. SMART assessments of the output-level standard indicators. ....	3
Table 4. SMART assessments of the outcome-level standard indicators.....	5
Table 5. SMART assessments of the impact-level standard indicators. ....	7

## List of Figures

<b>Figure 1.</b> Relevant Excerpts from report of result code #187.....	7
<b>Figure 2.</b> Results Dashboard – Summary of key results reported .....	12
<b>Figure 3.</b> Results Dashboard – Capacity change section.....	12

## List of Acronyms

CO2	carbon dioxide
CRF	Common Results Framework
GIS	Geographic Information System
ha	hectare
KASRB	knowledge, attitudes, skills, relationships, and behaviour
km	kilometer
MELIA	Monitoring, Evaluation, Learning, and Impact Assessment
OICR	Outcome Impact Case Report
PRMS	Performance and Results Management System
SDG	Sustainable Development Goals
SIDS	Standard Indicator Description Sheets
SMART	Specific, Measurable, Achievable, Relevant, Time-bound
Tg	teragram
USAID	United States Agency for International Development

## Executive Summary

### Introduction

One CGIAR developed a Common Results Framework (CRF) comprising 32 standard indicators to standardize results management and streamline reporting. The aim is for all Initiatives, Impact Area Platforms, and, ideally, non-pooled projects to use the standard indicators for monitoring and reporting across the entire One CGIAR portfolio. As part of the effort to ensure that the CRF is functional and appropriately structured for standardized results management and reporting, we conducted two reviews: 1) a systematic analysis of the standard indicators, including a review of how they were applied and reported in the 2022 annual reporting by Initiatives and presented in the CGIAR Results Dashboard; and 2) a review of a random sample of 246 non-pooled indicators (NPIs) from a set of 3933 indicators from non-pooled CGIAR projects to assess how well the standard indicators represent non-pooled projects. A good indicator provides a practical way to measure something that cannot easily be measured directly. Changes in the indicator should be closely correlated with changes in the phenomenon of interest. Good indicators are specific, measurable, achievable, relevant, and time-bound (SMART). The overall set of indicators should be coherent, representative, comparable, and aggregable. We analyzed the quality of both sets of indicators to inform good practice in indicator design and use. The two analyses are presented in two separate reports<sup>1</sup>, summarized together in this executive summary.

### Method

The output and outcome standard indicators were collected from the Standard Indicator Description Sheets (SIDS) used for annual reporting, and the impact standard indicators were collected from the online CGIAR MEL ToC tool. The NPIs were collected by an independent consultant for a prior review with a request sent to CGIAR Centers to compile indicators and results reported for past projects and programs. Our team used that dataset for our analysis. The first step was to develop rules for each assessment, including defining the SMART appraisal criteria and specifying the protocol to document observations. The next step was to systematically review each indicator according to the established definitions and rules. For the assessment of the standard indicators, the application of the indicators in the Results Dashboard, the SIDS, and the Initiative annual reports were also reviewed. A stratified random sample of 246 of NPIs was selected and mapped to the CRF to assess their fit and inform any necessary improvements to the CRF based on practical examples of past indicators developed for non-pooled CGIAR projects and programs.

*Limitations:* Assessment of the standard indicators was limited by inconsistent documentation across the SIDS, the online CGIAR MEL ToC Tool, the Results Dashboard, and Initiative annual reports. The NPI dataset was compromised by inconsistent reporting of data from contributing Centers and missing contextual information about respective projects/programs to guide interpretation.

### Key Results

Table I presents a summary of the SMART assessments for the standard indicators, and Table II presents a summary of the SMART assessments for the NPIs. Even assuming time-boundness to be met by annual reporting requirements, few of the indicators fully satisfied the SMART criteria. Only 2 of the 32 standard indicators fully satisfied the SMART criteria, and only 4 of 246 NPIs sampled fully satisfied the SMART criteria.

---

<sup>1</sup> CGIAR Standard Indicators Report, Assessment of One CGIAR's Non-pooled Indicators Report

Table I. Summary of SMART assessments for the standard indicators.

	<i>Specific</i>	<i>Measurable</i>	<i>Achievable</i>	<i>Relevant</i>	<i>Time-bound</i>
Yes	10	15	26	19	2
Partial	-	-	-	1	-
No	22	1	0	3	30
Uncertain	-	16	6	9	-

Table II. Summary of SMART assessments for the NPIs.

	<i>Specific</i>	<i>Measurable</i>	<i>Achievable</i>	<i>Relevant</i>	<i>Time-bound</i>
Yes	74	150	198	90	22
Partial	-	-	-	91	-
No	172	8	7	10	224
Uncertain	-	88	41	55	-

There were several common errors in indicator formulation that contributed to the failure to meet SMART criteria. Some errors had implications for the SMART scores at the individual indicator level, while other errors had implications for comparability and reporting at the aggregate level (Table III).

Table III. Summary of error types.

<b>Error Type</b>	<b>Details of Error</b>	<b>Analysis where Error was Observed</b>
Insufficient specification of variable (Specificity error)	It is unclear what is to be measured; manifests in NPIs that have brief variable descriptions or use vague language	Standard indicator analysis NPI analysis
Overly complicated specification of variable (Specificity error)	Obscures what is to be measured; manifests in NPIs that contain unnecessary/convoluted information or additional detail that limits what is to be measured (i.e., conditionality)	NPI analysis
More than one variable (Specificity error)	Specificity error that manifests in NPIs that contain two or more variables	Standard indicator analysis NPI analysis
No metric (Measurability error)	Does not state what aspect of the variable is to be measured; can infer binary assessments (i.e., did/did not happen)	NPI analysis
Unclear metric (Measurability error)	Does not adequately specify the aspect of the variable that is to be measured; manifests in NPIs with poor selection of the metric	Standard indicator analysis NPI analysis
Weak relationship to result (Relevance error)	Results in NPIs not linked to a result statement and NPIs that are mismatched to their result	Standard indicator analysis NPI analysis
Different indicators used for the same result (Aggregation error)	Many proxies could be used for the same result; no standardization creates issues for comparability at the aggregate level	Standard indicator analysis NPI analysis
Different metrics used for the same/similar indicators (Aggregation error)	Many metrics could be used to measure a given indicator; no standardization creates issues for comparability at the aggregate level	Standard indicator analysis NPI analysis
Multiple units of measurement used for the same metric (Aggregation error)	Many units of measurement could be used to measure a given metric; no standardization requires data be converted for comparability at the aggregate level	NPI analysis
No unit of measurement specified (Aggregation error)	Any given unit of measurement could be selected at random across the system when none is specified; affects consistency and comparability at the aggregate level	Standard indicator analysis NPI analysis
Excessive specification of the variable (Aggregation error)	Too much contextualization turns like indicators into distinct and different ones; creates issues for comparability at the aggregate level	NPI analysis

Only 22.8% of NPIs could be mapped directly to a single standard indicator in the CRF, 46.3% could be mapped to multiple standard indicators, and 30.9% could not be mapped to any standard indicator. The varying metrics and units of measurement used across the sample of NPIs to indicate the same phenomena or result pose issues for aggregation and accurate reporting of portfolio results. Thematic areas not covered by the standard indicators and that we deemed relevant to consider when revising the CRF are presented in Table IV.

*Table IV. Thematic areas not covered by the standard indicators, but observed in the NPI set.*

<b>Result Level</b>	<b>Thematic Area</b>
Activity	Project management
	Engagement/outreach
	Training delivery
	Partner participation (in CG activities)
Output	Crop breeding
	Genebank conservation and access
Outcome	Improved relationships/networks/fora
	Improved market linkages
Impacts	Agricultural production
	Production losses
	Soil quality/fertility
	Food safety
	Income
	Commodities/products sold
	Employment
	Livelihood diversification
	Economic growth (large-scale)

## Conclusion and Recommendations

The analyses reveal variable quality in One CGIAR indicators and demonstrate that there is scope to improve the standard indicators within the CRF, as well as better formulate project and program (Initiative) level indicators. Considering the analyses, we offer the following recommendations:

Recommendations to improve the standard indicators, guidance, and CRF:

- *Revise the standard indicators and guidance to specify the most relevant variables, metrics, and units of measurement of interest to the CGIAR, and ensure they are appropriate proxies for the result level and category.*
- *Expand the existing result categories and standard indicators to be inclusive of key results in the NPIs that could not be mapped.*

Recommendations to improve future NPI development:

- *All indicators should have only one variable, and be framed as clearly as possible to enable compatible data collection, aggregation, and reporting.*
- *Provide guidance and formats for developing contextualized and project-specific indicators based on standard indicators.*
- *CGIAR researchers and staff require resources and/or training to develop and use high-quality indicators.*

## Introduction

One CGIAR developed a Common Results Framework (CRF) to standardize results management and reporting. The CRF includes 32 standard indicators to measure outputs, outcomes, and impacts, further subcategorized according to thematic areas of activity and influence. An indicator is a “quantitative or qualitative variable that represents an approximation of the characteristic, phenomenon or change of interest (for instance, efficiency, quality or outcome). Indicators can be used to monitor research or to help assess, for instance, organizational or research performance” (MARLO, 2015, para.33). The aim is for all Initiatives and Impact Area Platforms to use the standard indicators in the CRF to enable the aggregation of contributions of the entire portfolio. Currently, summary results and supporting evidence of the One CGIAR’s contributions are reported and linked to the Results Dashboard using a new Performance Result Management System (PRMS). This report presents an analysis of the standard indicators (Table 1).

**Table 1.** Standard indicators included in the review.

<i>Result Level</i>	<i>Result Category</i>	<i>Standard Indicators</i>
Output	Knowledge product	Number of peer reviewed publications published in the reporting year
		Number of other information products published in the reporting year
		Number of MELIA study knowledge products (and subtypes) published in the reporting year
	Innovation development	Number of innovations profiled
Capacity sharing for development	Number of people trained	
Outcome	Uptake of information product	Altmetric Attention Score
	Innovation use	#people using the innovation disaggregated by gender
		Number of innovations being used (stage 4)
		Other quantitative measure of CGIAR innovation use, e.g. area
	Number of innovation packages designed	
Capacity change	Change in capacity of key (a) individuals (b) organizations (government, civil society, private sector), (c) networks (e.g. multistakeholder platforms)	
Policy change	Number of new or significantly updated policies/strategies/legal instruments/programs/ budgets/investments	
Impact	Poverty reduction, livelihoods and jobs	#poor people benefiting from relevant CGIAR innovations
		#people assisted to exit poverty
	Nutrition, health and food security	#people benefiting from relevant CGIAR innovations
		#people meeting minimum dietary energy requirements
		#people meeting minimum micronutrient requirements
		#cases of communicable and noncommunicable diseases
	Gender equality, youth and social inclusion	#women’s empowerment and inclusion in the agricultural sector
		#women benefiting from relevant CGIAR innovations
		#youth benefiting from relevant CGIAR innovations
		#women assisted to exit poverty
	Environmental health and biodiversity	#ha under improved management
		#km3 consumptive water use
		#ha deforestation
		#Tg nitrogen application
		#plant genetic accessions available and safely duplicated
	Climate adaptation and mitigation	#tonnes CO2 equivalent emissions
		#plans with evidence of implementation
#\$ climate adaptation investments		
#people benefiting from climate adapted innovations		
#people benefiting from the implementation of adaptation plans		



A good indicator provides a practical way to monitor change in meaningful units over time. Changes in the indicator should be closely correlated with changes in the phenomenon of interest. That is, changes in the value of an indicator from a baseline level indicate changes in a phenomenon that is difficult or impossible to measure directly. In results management literature, good indicators follow the SMART heuristic; they are specific, measurable, achievable, relevant and time-bound (Morrison, 2022). However, there are many versions of the SMART heuristic. We developed the following definitions for evaluating the standard indicators:

- *Specific* means that the indicator statement clearly specifies the variable that will be measured/assessed. Developing specific indicators provides clarity for effective results management.
- *Measurable* means that it is feasible to measure the indicator. Developing indicators that can be feasibly measured is important to ensure that the work necessary to monitor and report progress can be done.
- *Achievable* means that it is feasible for a measurable change in the indicator to be realized within the reporting period.
- *Relevant* means that the indicator directly measures the result of interest or serves as an appropriate proxy for the result. It must be appropriately classified to its result type and category, and be useful for assessing productivity (for outputs) or progress in a process of change (for outcomes) or changes in socio-economic/environmental conditions (for impacts). Developing relevant indicators is important to ensure that the indicators provide an adequate measure of the variable (result) of interest.
- *Time-bound* means that the indicator has an explicit timeframe within which to be measured.

A SMART indicator is not necessarily the most accurate and precise one, but is one that meets learning and reporting needs within resource constraints (Neerman, 2021). Indicators need be sensitive to changes in the result of interest. In other words, changes in the indicator should be incremental and continuous, not discontinuous or sudden, such that there will be a detectable change in the indicator if there is a change in the result of interest. Primary or secondary data should be available at the appropriate scale and at reasonable cost, including data on current conditions at baseline.

We systematically assessed each of the standard indicators to determine the extent to which they fulfilled the SMART criteria, and reviewed their application in the Results Dashboard and annual reports.

The assessment answers the following questions:

1. Do the individual CRF standard indicators satisfy the criteria of good indicators?
2. Does the overall set of standard indicators adequately represent the overall CGIAR portfolio of results?
3. How have the standard indicators been applied in practice?
4. What are the lessons for improved indicator formulation and use?

This report presents the analysis by result level: outputs, outcomes and impacts. We first present a summary of the SMART assessments of the standard indicators, followed by observations on their comprehensiveness and appropriateness. Next, we present observations and suggestions to improve the application of standard indicators in the results dashboard reports. We conclude with recommendations and guidance to improve the formulation and application of the standard indicators.

The standard indicator results are reported in annual Initiative reports (in ‘Section 4. Initiative key results’) and the [CGIAR Results Dashboard](#). The Standard Indicator Description Sheets (SIDS) for outputs and outcomes were used as guidance documents for reporting against standard indicators at the Initiative level.

## Methods

The review was carried out according to the following steps:

**Step 1.** Develop rules (Appendix 1) to guide the assessment. The rules specify how to document observations (Yes, Partial, No, Uncertain; justifications) and identify common errors across the standard indicators. Definitions for each

result type and result category were drawn from the CGIAR glossary and SIDS as a reference to classify and map each indicator to the CRF.

**Step 2.** Systematically review each standard indicator according to the established definitions and rules.

**Step 3.** Assess whether the indicators provide adequate representation of the intended CGIAR results, individually and collectively.

**Step 4.** Review the use of standard indicators in the CGIAR Results Dashboard and Initiative-level reporting to understand how the indicators are applied. We reviewed a sample of three Initiative reports (one from each action area):

- Resilient Aquatic Food Systems Initiative (Resilient Agrifood System Action Area)
- Low-Emission Food Systems Initiative (Systems Transformation Action Area)
- Market Intelligence and Product Profiling Initiative (Genetic innovation Action Area)

## Results

Table 2 summarizes the results of the SMART assessment, which suggests the standard indicators vary in their quality. The CRF standard indicators are intended to be a general all-encompassing set; however, the lack of specificity in some indicators could lead to inconsistent interpretations. For several indicators, it was not clear how they would be measured (assessed as ‘uncertain’); these were typically the same indicators assessed to be ‘not specific’ because they did not clearly define the variable of interest. As a result, several indicators’ measurability assessments were contingent upon the specificity assessments. Most standard indicators were achievable. None were assessed to be not achievable. However, some indicators are not specified sufficiently to know whether they could be observed in the reporting period (i.e., uncertain of what to observe). While all the standard indicators are linked to a result level and result category, some are mismatched. Only two of the standard indicators were explicitly time-bound. However, according to the SIDS, it is intended that each is to be monitored and reported on an annual basis.

**Table 2.** Summary of SMART assessments for the standard indicators.

	<i>Specific</i>	<i>Measurable</i>	<i>Achievable</i>	<i>Relevant</i>	<i>Time-bound</i>
Yes	10	15	26	18	2
Partial	-	-	-	2	-
No	22	1	0	3	30
Uncertain	-	16	6	9	-
Total	32	32	32	32	32

## Output Indicators: Observations and Opportunities for Improvement

### *SMART Assessment Results*

Table 3 provides an overview of the SMART assessments of the output standard indicators.

**Table 3.** SMART assessments of the output-level standard indicators.

<i>Result Category</i>	<i>Standard Indicator</i> <sup>2</sup>	<i>SMART Assessment</i> <sup>3</sup>	<i>Suggestions for Improvement</i>
Knowledge product	Number of peer reviewed publications published in the reporting year	SMART	<ul style="list-style-type: none"> <li>• None</li> </ul>

<sup>2</sup> \* denotes indicators requiring minor revision. \*\* denotes indicators requiring more major revision.

<sup>3</sup> ? indicates uncertainty whether the indicator satisfied the criterion. / indicates partial satisfaction of the criterion. \_ indicates the criterion was not satisfied. The letter indicates that the indicator satisfied the criterion.

	Number of other information products published in the reporting year*	_MART	<ul style="list-style-type: none"> <li>Specify mutually exclusive sub-categories of “other information products” to be counted</li> </ul>
	Number of MELIA study knowledge products (and subtypes)*	SMAR_	<ul style="list-style-type: none"> <li>Make time-boundedness explicit (Number of evaluation studies produced in the reporting year)</li> </ul>
Innovation development	Number of innovations profiled*	_MAR_	<ul style="list-style-type: none"> <li>Clarify what “profiled” means: does it mean registered in CGSpace, or at stage 9 innovation readiness?</li> <li>Make time-boundedness explicit</li> </ul>
Capacity sharing for development	Number of people trained**	SMA/_	<ul style="list-style-type: none"> <li>Consider using this as an indicator of capacity change outcomes</li> <li>Output indicator could be framed as “Number of trainings delivered”</li> <li>Make time-boundedness explicit</li> </ul>

Three of the five output indicators adequately specify the variable in their current formulation: ‘number of people trained’, ‘number of MELIA study knowledge products (and subtypes)’, and ‘number of peer reviewed publications published in the reporting year’. The other standard output indicators need further specification. ‘Number of other information products’ is elaborated in the guidance (by type), but does not specify what is to be included and excluded in the count. As a result, the count includes many different kinds of products: reports, presentations, briefs, blog posts, working papers, book chapters, videos, posters, datasets, internal documents, manuals, brochures, conference papers, other, theses, newsletters, books, abstracts, training materials, extension materials, internal reports, tools, websites, conference proceedings, journal items, opinion pieces, templates, case studies, infographics, photo reports, press items, audio, data papers, maps, news items, newsletter articles, other (toolkit), and source code. Indicators formulated like this encourage a lack of coherence and comparability in reporting. It would therefore be useful to develop a few mutually exclusive sub-categories of other information products.

The standard indicator ‘number of people trained’ was assessed to be partially relevant. While the indicator is linked to a result category (Capacity Sharing for Development), it is mismatched to its result level. The standard indicator, ‘number of people trained’, provides a better proxy of Capacity Change than of Capacity Sharing for Development. Other measures like ‘number of trainings delivered annually’ by type should be considered as an indicator of Capacity Sharing for Development. Such measures would further distinguish the results presented for Capacity Sharing for Development from results for Capacity Change.

The SIDS specify that counts for knowledge products are to be provided for each reporting year. Although we can assume each standard indicator will be reported annually, it should still be made explicit in all indicator statements, as done in the ‘number of peer-reviewed publications produced in the reporting year’ standard indicator.

*Comprehensiveness*

The output indicators represent the main kinds of outputs produced by the CGIAR, but omit an important output of genebanks. The conservation of germplasm indicated by ‘#plant genetic accessions available and safely duplicated’ is misclassified as an indicator of the Environmental Health and Biodiversity impact area (see impact indicators section). Additionally, ‘number of people trained’ is an indicator of an outcome (sphere of influence), not an output. This suggests that it may be useful to include an explicit indicator that counts ‘number of trainings delivered annually’ by type (i.e., short-term, long term) to provide an accurate and comprehensive picture of CGIAR’s delivery of capacity building activities. In the Results Dashboard, it appears that the 511 results specified refers to the number of reports in the details section to reflect training events delivered. It would be useful to also collect information on the types of training and their counts (i.e., how many of what type of training). This could be done by collecting and aggregating the details of the type of training (i.e., either by pre-defined types or observing commonly reported trainings by

keyword terms) specified for each result in the details section. Alternatively, it could follow the short-term and long-term definitions for trainees. If ‘number of people trained’ is restructured as an outcome standard indicator in the Capacity Change result category, it would act as a more appropriate proxy than the current standard indicator ‘change in capacity of key (a) individuals (b) organizations (government, civil society, private sector), (c) networks (e.g. multistakeholder platforms)’. Aside from linking to policy briefs and trainings delivered to policymakers, there is no causal logic in the CRF for what outputs lead to policy change. Although there is a lot of emphasis on partnerships and engagement as important means by which CGIAR aims to contribute to change, there are no indicators of outputs such as policy dialogues or other links to policy processes.

## Outcome Indicators: Observations and Opportunities for Improvement

### SMART Assessment Results

Table 4 provides an overview of the SMART assessments of the outcome standard indicators.

**Table 4.** SMART assessments of the outcome-level standard indicators.

<i>Result Category</i>	<i>Standard Indicator<sup>4</sup></i>	<i>SMART Assessment<sup>5</sup></i>	<i>Suggestions for Improvement</i>
Uptake of information product	Altmetric Attention Score	SMAR_	<ul style="list-style-type: none"> <li>Make time-boundedness explicit (i.e., for reporting period)</li> </ul>
Innovation use	# people using the innovation disaggregated by gender*	SMAR_	<ul style="list-style-type: none"> <li>Spell out ‘#’ as ‘number of’</li> </ul>
	Number of innovations being used (disaggregated by stage)*	SMAR_	<ul style="list-style-type: none"> <li>Make time-boundedness explicit (i.e., in reporting year)</li> <li>Consider tracking the change in use of specific innovations (i.e., moving from one innovation use level to another)</li> </ul>
	Other quantitative measure of CGIAR innovation use, e.g. area**	_ _??_	<ul style="list-style-type: none"> <li>Specify the most important variables and metrics to be counted (i.e., area; number of users) and units of measurement to be reported</li> </ul>
	Number of innovation packages designed**	_MA?_	<ul style="list-style-type: none"> <li>Specify who designs innovation packages. If other actors, it is an outcome indicator; if CGIAR entities, it is an output indicator (if the latter, then the indicator overlaps with ‘number of innovations profiled’)</li> </ul>
Capacity change	Change in capacity of key (a) individuals (b) organizations (government, civil society, private sector), (c) networks (e.g. multistakeholder platforms)**	_MA/_	<ul style="list-style-type: none"> <li>Specify metric(s) to be used to report capacity change</li> </ul>
Policy change	Number of new or significantly updated policies/strategies/legal instruments/programs/budgets/investments	SMAR_	<ul style="list-style-type: none"> <li>Make time-boundedness explicit (i.e., in reporting year)</li> </ul>

The capacity change standard indicator is problematic for two reasons. First, it includes three possible scales of capacity change (i.e., individuals, organizations, networks), which makes aggregation and comparability challenging.

<sup>4</sup> \* denotes indicators requiring minor revision. \*\* denotes indicators requiring more major revision.

<sup>5</sup> ? indicates uncertainty whether the indicator satisfied the criterion. / indicates partial satisfaction of the criterion. \_ indicates the criterion was not satisfied. The letter indicates that the indicator satisfied the criterion.

Second, it is unclear how capacity change is to be measured/assessed. These details are currently ambiguous in the SIDS guidance. The Dashboard does not adequately or clearly distinguish between results reported for Capacity Sharing for Development and Capacity Change either. The standard indicator for capacity change does not specify what should be counted as capacity change and what should be counted as people trained. For example, one PRMS report (result code #4264) under capacity change counts the number of farmers reached through knowledge sharing. This is provided as evidence of one of the 7 results of organizational capacity change for the United States Agency for International Development (USAID). However, USAID contributed funds to deliver the information to farmers; USAID did not experience capacity change. This example showcases there is ambiguity in the indicator's framing as the result has been misattributed (i.e., incorrectly assigned as an outcome). In another example, a reported instance of capacity change for the Government of Assam (result code #3594) duplicates another (result code #3154) that reports farmers being reached by access to information and quality seeds of new varieties. The Government of Assam is listed as a partner, but there is no indication that the Government's capacity changed. As organizations' capacity change is counted in this indicator (as well as individuals' and networks'), it is reasonable to assume that an organization listed under capacity change in the Dashboard would have experienced organizational capacity change. In practice, it appears the Dashboard counts funding organizations or partner organizations that supported the delivery of capacity-building activities. It may be that these kinds of reports of reach would be more appropriately placed in the Capacity Sharing for Development section because they count the number of farmers reached by knowledge sharing activities, which is similar to training delivery. Result code #4264 on fall army worm knowledge sharing does not have a corresponding activity reported in the Capacity Sharing for Development section of the Dashboard. Another PRMS report for capacity change documents adoption of new cropping practices (result code #2769), but its connection to capacity change is unclear and we question why this is not reported under innovation use instead. The capacity change indicator simply repeats the results statement/category rather than providing an indicator of the expected result (e.g., 'capacity change>change in capacity...'). Guidance is needed on what metric to use to measure capacity change. The current Capacity Change indicator at the outcome level ('change in capacity of key individuals [...]') is not mutually exclusive of the output level indicator for Capacity Sharing for Development ('number of people trained').

'Other quantitative measure of CGIAR innovation use, e.g. area' does not adequately specify the variable; it includes any (and all) quantitative measures of use. Its lack of specificity had knock-on effects for assessing other criteria. Such an open-ended indicator encourages a lack of coherence across the portfolio, such as reporting in different units. This example also illustrates the ambiguity in framing that can be broadly interpreted (i.e., 'other quantitative measure'). The Results Dashboard does not report on these other quantitative measures of innovation use, likely because the varied reporting across Initiatives make aggregation impossible. Relevance and achievability could not be assessed as the indicator is not adequately specified in its current formulation.

'Number of innovation packages designed' does not specify what is to be counted or how it is different from innovation profiling. Being an outcome indicator implies that other actors (sphere of influence) are doing the packaging. If the CGIAR is designing innovation packages, this indicator would be more appropriately used as an output indicator, under the Innovation Development result category, because it would be within the CGIAR's sphere of control. If other actors are expected to produce innovation packages, it should be clear in the indicator statement. Results for the 'number of innovation packages designed' are not reported in the annual reports or in the Results Dashboard.

### *Comprehensiveness*

The set of outcome standard indicators refers to the main areas of CGIAR's intended contributions (policy change, innovation, adoption/uptake of knowledge, capacity change), but the reporting of innovation use and capacity change in the Results Dashboard raises questions about their appropriateness. Evidence reported under Innovation Use in the Dashboard seems to reflect adoption of CGIAR knowledge instead of the number of innovation packages that have been designed and who has designed them. For example, one result (result code #187) reports on the use of simulations

and breeding pipeline management tools that were deployed by CGIAR breeders. While this result discusses a suite of tools, it is not clear in the Dashboard whether this is considered an “innovation package”. The number of users is noted in the narrative of the report, but the numbers are different than the innovation use numbers reported at the bottom of the form (Figure 1). It is also not clear how the gender disaggregated and organizationally disaggregated data on use were collected because it is not specified. Reporting on output and outcome capacity indicators appears to be conflated at times. Instead of reporting on qualitative changes in capacity in the Dashboard, it appears that the number of people reached (by various CGIAR activities) is being reported under the Capacity Change standard indicator. The CGIAR should consider how adoption, reach, and training are intended to be conceptualized in the CRF and reported.

**Description:**

Simulations and breeding pipeline management tools were deployed to CGIAR breeders to optimize breeding schemes aligned with the desirable characteristics of farmer- and market demanded crop varieties. The suite of tools has proven to be useful to CGIAR breeders who stated that they optimized breeding schemes in 86% of their pipelines assessed. The tools and simulation results are publicly accessible on [www.excellenceinbreedingplatform.org](http://www.excellenceinbreedingplatform.org). Access requires registration before use. So far, there are 68 CGIAR users, 146 non-CGIAR users that have registered, and 63 non-CGIAR users that have uploaded data from across the world. The tools are useful to breeding programs that would like to optimize their breeding schemes.

[.....]

## Innovation use

**Number of people using the innovation, disaggregated by gender:**

**Female:** 30

**Male:** 100

**Other quantitative measures of innovation use (e.g. area):**

**1. Unit of measure:** Not Provided

**Quantity:** Not Provided

Figure 1. Relevant Excerpts from report of result code #187.

## Impact Indicators: Observations and Opportunities for Improvement

### SMART Assessment Results

Table 5 provides an overview of the SMART assessments of the standard impact indicators.

Table 5. SMART assessments of the impact-level standard indicators.

<i>Result Category</i>	<i>Standard Indicator</i> <sup>6</sup>	<i>SMART Assessment</i> <sup>7</sup>	<i>Suggestions for Improvement</i> <sup>8</sup>
Poverty Reduction, Livelihoods and Jobs	#poor people benefiting from relevant CGIAR innovations**	_ <i>?A?</i> _	<ul style="list-style-type: none"> <li>Reformulate to clarify how people are expected to benefit (i.e., nature and degree of benefit)</li> <li>Specify that this indicator is to count realized benefits, not projections</li> </ul>

<sup>6</sup> \* denotes indicators requiring minor revision. \*\* denotes indicators requiring more major revision.

<sup>7</sup> ? indicates uncertainty whether the indicator satisfied the criterion. / indicates partial satisfaction of the criterion. \_ indicates the criterion was not satisfied. The letter indicates that the indicator satisfied the criterion.

<sup>8</sup> All indicators need to spell out contraction ‘#’ as ‘number of’, and implement minor reformulation to make time-boundedness explicit (e.g., at conclusion of 3 year program cycle).

Report: One CGIAR Standard Indicators

	#people assisted to exit poverty**	_?AR_	<ul style="list-style-type: none"> <li>• Provide standard definition of “exit poverty” (i.e., poverty threshold)</li> <li>• Specify that this indicator is to count realized benefits, not projections</li> </ul>
Nutrition, Health and Food Security	#people benefiting from relevant CGIAR innovations**	_?A?_	<ul style="list-style-type: none"> <li>• Reformulate to clarify how people are expected to benefit (i.e., nature and degree of benefit)</li> <li>• Specify that this indicator is to count realized benefits, not projections</li> </ul>
	#people meeting minimum dietary energy requirements*	S?AR_	<ul style="list-style-type: none"> <li>• Define the threshold and unit of measurement (e.g., kcal)</li> </ul>
	#people meeting minimum micronutrient requirements*	S??R_	<ul style="list-style-type: none"> <li>• Provide definition of “minimum micronutrient requirements”</li> </ul>
	#cases of communicable and noncommunicable diseases**	_??R_	<ul style="list-style-type: none"> <li>• Reformulate to clarify which diseases are the focus of the indicator or remove</li> </ul>
Gender Equality, Youth and Social Inclusion	#women’s empowerment and inclusion in the agricultural sector**	_??R_	<ul style="list-style-type: none"> <li>• Focus the indicator on a single variable (empowerment or inclusion, not both)</li> <li>• Specify what metric is to be used</li> <li>• Ensure that the indicator is mutually exclusive of others in its result category</li> </ul>
	#women benefiting from relevant CGIAR innovations**	_?A?_	<ul style="list-style-type: none"> <li>• Reformulate to clarify how women are expected to benefit (i.e., nature and degree of benefit)</li> <li>• Specify that this indicator is to count realized benefits, not projections</li> </ul>
	#youth benefiting from relevant CGIAR innovations**	_?A?_	<ul style="list-style-type: none"> <li>• Reformulate to clarify how youth are expected to benefit (i.e., nature and degree of benefit)</li> <li>• Specify that this indicator is to count realized benefits, not projections</li> </ul>
	#women assisted to exit poverty**	_??R_	<ul style="list-style-type: none"> <li>• Provide standard definition of “exit poverty” (i.e., poverty threshold)</li> <li>• Specify that this indicator is to count realized benefits, not projections</li> </ul>
Environmental Health and Biodiversity	#ha under improved management*	_?AR_	<ul style="list-style-type: none"> <li>• Provide definition of “improved management”</li> <li>• Specify that this indicator is to count realized benefits, not projections</li> </ul>
	#km3 consumptive water use*	_?AR_	<ul style="list-style-type: none"> <li>• Reformulate to clarify which water uses are the focus of the indicator (i.e., irrigation, drinking, processing, etc.), or remove</li> </ul>
	#ha deforestation*	_?AR_	<ul style="list-style-type: none"> <li>• Clarify variable to be measured (number of hectares saved from deforestation or number of hectares deforested)</li> </ul>
	#Tg nitrogen application**	_???_	<ul style="list-style-type: none"> <li>• Reformulate to clarify, or remove</li> </ul>
	#plant genetic accessions available and safely duplicated**	_MA_ _	<ul style="list-style-type: none"> <li>• Focus on one variable (available or duplicated, not both)</li> <li>• Move to output result level (as this is an indicator of genebank output)</li> </ul>

Climate Adaptation and Mitigation	#tonnes CO2 equivalent emissions*	SMAR_	<ul style="list-style-type: none"> <li>Reformulate to clarify the context of the emissions to be counted (i.e., from agriculture, sequestered emissions, offset emissions, etc.)</li> <li>Specify scale and location of emissions to be measured (i.e., all emissions globally, from target countries, from particular agro-ecosystems, etc.)</li> </ul>
	#plans with evidence of implementation**	_MA_ _	<ul style="list-style-type: none"> <li>Delete or move. This is an outcome-level indicator of policy change; the current policy outcome indicator does not capture implementation, so it may be useful to have this indicator</li> </ul>
	#\$ climate adaptation investments**	_MA_ _	<ul style="list-style-type: none"> <li>Delete or move. This is an outcome-level indicator of policy change; the current policy change outcome indicator does not have a specific climate adaptation focus, so it may be useful to have this indicator, or include a climate tag as done for other results provided in the dashboard</li> <li>Specify currency unit</li> </ul>
	#people benefiting from climate adapted innovations**	_?A?_	<ul style="list-style-type: none"> <li>Reformulate to clarify how people are expected to benefit (i.e., nature and degree of benefit)</li> <li>Specify that this indicator is to count realized benefits, not projections</li> </ul>
	#people benefiting from the implementation of adaptation plans**	_?A?_	<ul style="list-style-type: none"> <li>Reformulate to clarify how people are expected to benefit (i.e., nature and degree of benefit)</li> <li>Specify that this indicator is to count realized benefits, not projections</li> </ul>

The SIDS do not provide any guidance for impacts, and results for impacts have yet to be reported in the Dashboard or annual reports. There is no evidence that they are being used at the current stage of the CGIAR program cycle, even to collect baseline data. Many of the impact indicators represent phenomena that require further clarification to provide good proxies of the result to be measured (for example, ‘number of people assisted to exit poverty’, ‘number of people meeting micronutrient requirements’). Specific thresholds, such as the universal/national poverty line, need to be specified to facilitate clear and consistent reporting and data collection.

The analysis revealed three particularly problematic indicators that either require fundamental revision or should be considered for removal from the CRF:

‘Nutrition, health and food security># cases of communicable and noncommunicable diseases’. Even if it was possible to get a reliable count and show a change from year to year, it would be impossible to attribute the change to the CGIAR. If there are particular diseases likely to be significantly impacted by the CGIAR, the indicator should focus on those diseases, with consideration of where and how data will be collected. Otherwise, this indicator should be removed from the set.

‘Environmental health and biodiversity> #Tg nitrogen application’. It is not clear what result this indicator is meant to proxy or what the intended direction of the desired change is. This indicator either needs fundamental revision considering the result it aims to indicate, or should be removed from the set.

‘Environmental health and biodiversity> #km3 consumptive water use’. It is not clear what result this indicator is meant to proxy. It is also not clear whether water use for irrigation, drinking, or product processing is intended to be measured. The indicator needs to be considered for what exact result or change in state it is aiming to proxy, and be reformulated for clarity. Otherwise, it should be removed from the set.

Six of the impact standard indicators count the number of people benefiting from innovations relevant to the impact area, disaggregated by the type of innovation (i.e., whether it is related to climate, food security, poverty, or social



inclusion). This kind of disaggregation is not useful, as many innovations are likely to contribute to benefits across multiple impact areas. For example, an innovation that leads to improvements in food security may also lead to reductions in poverty. Therefore, the most common error in the set of impact standard indicators is inadequate specification of the type and magnitude of benefits expected to be measured by the indicator. Vague terms leave ambiguity about to what is to be measured (and how) and may be interpreted differently by different users. ‘Benefits’ need to be specified. What benefits does the CGIAR intend for people to receive/gain from the CGIAR’s research and development? There is a risk with the current framing and lack of guidance that Initiative leaders could be tempted to extrapolate innovation use numbers to claim benefits rather than assessing actual realized benefits, which would be misleading. There is potential for some indicators to overlap considerably, particularly those framed like ‘#people benefiting...’. ‘Number of people benefitting...’ in the Nutrition, Health and Food Security impact area would include those who benefit from improved nutrition, health and/or food security. Each could be measured in different ways, and would likely imply double- or triple-counting people. Many of the benefits in the gender/inclusion impact area will be captured by collecting gender- and age-disaggregated data from other impact areas. For example, ‘#women assisted to exit poverty’ overlaps with ‘#people assisted to exit poverty’, which makes having a separate indicator counting women redundant. Additionally, because being assisted to exit poverty is not defined, if the figure will be extrapolated from innovation use, it will also be redundant.

Use of symbols like ‘#’ should be avoided and instead be written out in full for consistency and clarity. Insufficient specificity in indicators caused problems for measurability, which were assessed as uncertain because there were several possible ways to measure the variable. For example, we were uncertain how nitrogen application would be measured for the standard indicator, ‘#Tg nitrogen application’.

Three of the indicators were mismatched to their associated result type, and were also assessed to not be appropriate proxies for their results. For example, ‘#plant genetic accessions available and safely duplicated’ is a measure of a genebank output and therefore is not an indicator of Environmental Health and Biodiversity impact. A better proxy for biodiversity could be defined. Two of the indicators under the Climate Adaptation and Mitigation impact area refer to specific policy change outcomes: “#\$ climate adaptation investments” and “#plans with evidence of implementation”. These indicators are not mutually exclusive from the outcome standard indicator for policy change: “Number of new or significantly updated policies/strategies/legal instruments/programs/ budgets/investments”.

None of the impact standard indicators are time-bound. Baseline data and time-boundedness are of particular importance when considering impact-level results. Changes in state or flow need to be analyzed compared to baseline scenarios. The scale and scope of data collection for the expected impacts need to consider the time lags inherent to the achievement of social benefits from projects and programs.

### *Comprehensiveness*

Many indicators in the current set of impact standard indicators do not adequately specify the benefits that will be realized, which leaves the set prone to repetition because the variables are not mutually exclusive. Much of the repetition across the impact standard indicators could be reduced by collecting age- and gender-disaggregated data and specifying data sources for each indicator. To date, none of the impact standard indicators are reported in the Results Dashboard or the annual reports.

Standard indicators for some impact areas only focus on one of their listed aims. For example, all indicators under the Poverty, Livelihoods and Jobs impact area relate to poverty, with no indicators for livelihoods or employment impacts. In the Climate Mitigation and Adaptation impact area, there is no indicator for climate resilience. In the Environmental Health and Biodiversity impact area, there is no indicator for biodiversity. Representativeness of the kinds of impact contributions the CGIAR aims to influence for each impact area should be considered. If needed, the impact area should be revised to more accurately represent the impacts to which the CGIAR can realistically contribute.

Some CGIAR themes are not represented in the standard indicators. For example, much of WorldFish impacts will be achieved through increased protein consumption. Currently the standard indicators measure contributions to dietary energy and micronutrients, so there is likely scope to include an indicator on macronutrients.

The indicators based on the number of people benefiting do not qualify the range, diversity, or magnitude of benefits. This is key information that is currently overlooked that would provide a clearer and more meaningful picture of the CGIAR's contributions. Given the range of possible measures that could be used to collect data on the standard indicators, revisions to the indicators' formulation must also consider data sources and collection methods. It should be clear in the indicator guidance (to be developed) how they will be measured (primary or secondary data) and at what scale they will be measured (geographic scope).

There are twenty standard impact indicators. Some impact areas have up to five standard indicators. Considering the cost of collecting, analyzing, and storing primary data on realized benefits, some indicators could be removed. The most important results and indicators should be prioritized according to what is realistic and practical for the CGIAR to influence. Currently, the CRF is structured to provide indicators for each impact area, which are broad categories that could reflect multiple kinds of results. The current structure of the impact areas in CRF requires users to guess the result from the standard indicator, which will increase scope for misinterpretation. It would be helpful to first develop impact result statements that correspond to each impact area, and then develop appropriate indicators of the results. For example, an impact statement could be: People in CGIAR target countries are meeting minimum micronutrient requirements, and then develop a corresponding indicator for that, such as: number of people with vitamin A deficiency. However, the implications associated with the selection of the type of number (cardinal or ratio) for the indicator should be considered. Ratios may be more appropriate than cardinal numbers for certain impact results, because they control for fluctuations in population growth and other demographics well beyond the scope of control of any project or program. For example, more people could experience vitamin A deficiency year to year due to population growth, but the relative proportion of the population experiencing the condition could still be lower. However, ratio numbers pose challenges for aggregation across contexts, which can only be done by averaging the numbers collected, which obscures the meaning of the number at global or CGIAR target country scales.

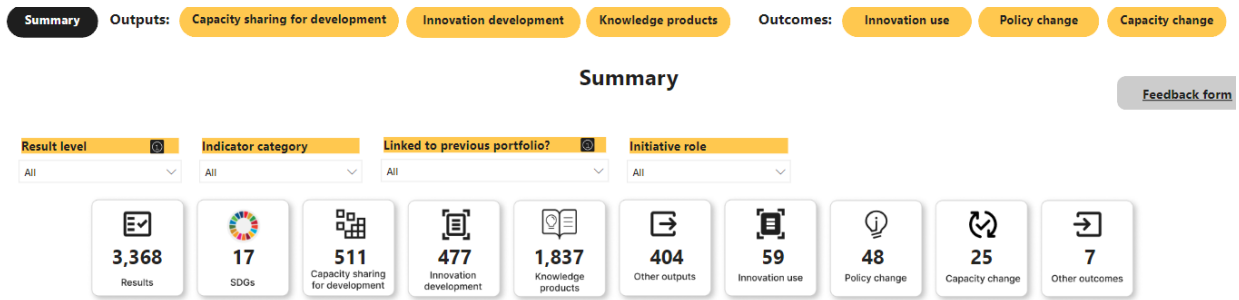
## **Observations and Suggestions to Improve Application of the Standard Indicators and Reporting in the Results Dashboard**

### **Ensure the reported evidence is reliable (PRMS reports and linked evidence)**

Reliability of reported evidence in “Result details” is questionable, particularly for outcomes. Several of the detailed reports of “policy change” do not link evidence of the policy change that has happened. Often the details only link to a PRMS report. Of those that do link to evidence, some provide links to reports (or outputs), not Outcome Impact Case Reports (OICRS) or outcome evaluation studies. For example, the policy change detail for a policy change influenced by the RAqFS Initiative states that the Initiative supported the Ministry of Fisheries and Marine Resources in the design of Solomon Islands Community Based Coastal and Marine Resource Management Strategy 2021-2025 (result code #2522). The PRMS report does not say what the support was or provide evidence that it happened. In another example, the reported policy change on Government of Tanzania gazettes grazing lands secured through joint village land use planning and participatory rangeland management innovations (Livestock and Climate Initiative, result code 4149) links to 3 outputs from 2021 in the PRMS linked pdf. For outputs, we found that scaling readiness (result code 953) itself was reported as a proven innovation, with link to the Sartas et al. article from 2020, which describes the approach, not proof that it works. Moreover, it is unclear why this links evidence from 2020, but is reported as a proven innovation for 2022.

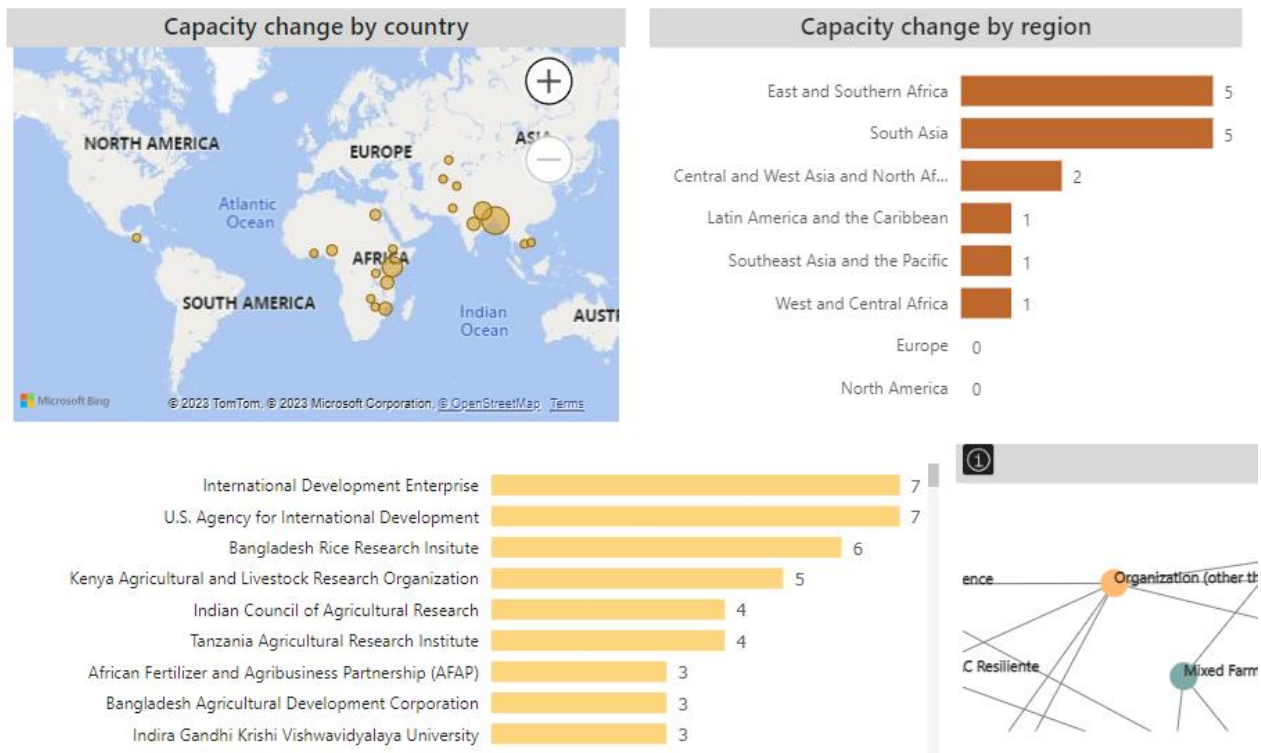
**Language in the Results Dashboard should be precise and specific so that the presentation of results is self-explanatory. All graphs should have informative labels to assist with the interpretation of data presented in the Dashboard.**

On the Summary page, ‘25 capacity change’ needs to be reformulated to clearly state what is being reported (Figure 1). We assume this is aggregated from the organizations and counts listed in the unlabeled bar chart in the capacity change window (Figure 2), which would be more accurately described as the number of organizations supporting capacity change.



**Figure 2.** Results Dashboard – Summary of key results reported.

Three graphs are missing titles or are poorly labeled. For example, in the ‘capacity change’ section, there is a list of organizations with bar charts presenting numbers (Figure 2). Presumably this is meant to convey how many times an organization has experienced capacity change, but it is not clear. It is also not clear what it means that USAID has had 7 capacity changes. When a user clicks on the bars, the results details are presented for what comprises the result. However, as discussed above, the reported 7 results for USAID capacity development support might be more appropriately captured under the Capacity Sharing for Development as part of an indicator for people reached or training delivery.



**Figure 3.** Results Dashboard – Capacity change section.

## Conclusion and Recommendations

Some of the standard indicators at the output and outcome result levels are well formulated and follow the principles of quality result indicators, while others, especially at the impact result level, could be improved. Common errors include vague or improper framing of indicators and the lack of standardized units of measurement, resulting in the inability to aggregate. Other errors in formulation were minor, such as the use of contractions, acronyms, or symbols. The standard indicators are not entirely comprehensive of all results to which the CGIAR will contribute. There is scope to improve the standard indicators, the SIDS, and the CRF to be more precise and useful for annual reporting and strategic decision-making on the CGIAR's productivity and contributions to outcomes and impacts. Based on the analysis, we recommend the following:

### **Recommendation 1. Revise the standard indicators and guidance for clarity and specificity.**

To this end, we suggest:

- Avoid using terms like 'other information product' or 'other quantitative measure'. Organize the most important ones in mutually exclusive sub-categories.
- Indicators statements should be as self-explanatory as possible. Relative framing like 'improved management' and vague framing like 'change in capacity', and 'benefits' require further specification. It should be clear how each indicator is to be measured to ensure that reporting is consistent, comparable, and aggregable. We suggest replacing indicators counting 'number of people benefiting...' to qualify the ways in which people are expected to benefit, supported by corresponding explanation of causal logic between innovation use, policy change, or capacity change to reach those benefits. Ideally, indicators measuring benefits should be mutually exclusive.
- Including time-bound framings such as 'annual', 'annually', 'by [year]', etc. in the set of standard indicators will help address the lack of time-boundedness and improve clarity.
- Avoid the use of acronyms/symbols and shorthand when developing indicators and guidance.
- There needs to be consistent standard indicator formulation across all areas of use in the CGIAR system (e.g., SIDS guidance, Results Dashboard, annual reporting, PRMS, etc.).

### **Recommendation 2. Provide format for precise indicator formulation at the project/work package/Initiative scale.**

The CRF and SIDS provide a general framing of the standard indicators. However, the generalized framing may not provide sufficient guidance for Initiative- or project-specific reporting. The framing/format of the standard indicators and corresponding guidance should achieve the following:

- Standard indicators need to be clear and precise about what they will measure, and they must maintain aggregability and comparability for reporting yet remain generic enough so that contextual details can be added as needed for Initiative-specific and project-level indicator formulation. Indicators should only be measured if the change being assessed results from work done by the Initiative/project as referenced in the Theory of Change.
- The standard indicators need to be applicable across the CGIAR's portfolio, using a standard metric and unit of measurement. They need to be as clear and specific as possible at the aggregate level to avoid ambiguity and multiple (mis)interpretations of what data will be measured and aggregated.
- The CRF needs to specify how data will be collected and aggregated for each indicator and provide corresponding guidance. At the initiative and work package levels, the standard indicators should collect an array of disaggregated information that enables more specific and detailed Initiative/Centre-level reporting (and strategic decision-making). The standard indicators should achieve aggregability. Disaggregate data that Centres can capture include: gender, age, actor type, geographic scope/scale (georeferenced data), landscape type, crop/livestock type, specific innovation or general innovation type, specific output or general output type, output topic/theme, etc. for aggregation into broad terms defined by the standard indicators.

## References

MARLO. (2015). Glossary of terms used in MARLO. Retrieved from <https://marlo.cgiar.org/glossary.do>

Morrison, M. (January, 2022). History of SMART Objectives. Retrieved from <https://rapidbi.com/history-of-smart-objectives/>

Neerman (February, 2021). What are SMART Indicators in Monitoring and Evaluation. Retrieved from <https://neerman.org/blogs/what-are-smart-indicators-in-monitoring-and-evaluation/>

## Appendix 1. Rulebook

### Section 1: Standard Indicator Suitability to Result Type

Key definitions to refer to for suitability assessment:

**Output:** “The products, goods, and services of research and the research process in terms of knowledge, foras and processes”; conceptualized in the sphere of control

**Outcome:** “Change in knowledge, attitudes, skills, and relationships that manifest as changes in behavior (who is doing what differently) resulting in whole or in part from research and its related activities”; conceptualized in the sphere of influence

**Impact:** “A change in state or a change in flow resulting in whole or in part from a chain of events to which the research has contributed. Impacts can be economic, socio-cultural, institutional, environmental, technological or of other types”; conceptualized in the sphere of interest

Column D: Result level (output/outcome/impact) classification

Function of the field: Indicate suitability of SIDS classification.

How to fill it in:

- Select **Yes** if the standard indicator measures something that matches the definition of the original SIDS result level type
- Select **No** if the standard indicator measures something that matches a different definition than the original SIDS result level type
- Select **Uncertain** if it is unclear what the standard indicator measures, and could either fit the definition of multiple result level types
  - \*Rule: if **Uncertain** of result level type, category (section 2, field 1) assessment must also select **Uncertain**

Column E: Justification

Function of the field: Justify the suitability assessment done in previous field.

How to fill it in:

- If **Yes**, specify aspects of the standard indicator that meets the result level definition; consider where in the spheres (control, influence, interest) the indicator fits
- If **No**, explain why the standard indicator doesn't fit, specify aspects of the indicator that correspond to another result level definition; consider where in the spheres the indicator fits OR doesn't match at all
- If **Uncertain**, explain why uncertain; specify aspects of the standard indicator that are inconsistent with other result level definition(s); consider where in the spheres the indicator fits OR doesn't match at all

Column F: Reclassification

Function of the field: Identify the best-fit result level.

How to fill it in:

- If **Yes**, note N/A (not applicable)
- If **No**, note the result level that the standard indicator matches the best (note: only one option can be selected)
- If **Uncertain**, note two or more result levels that the indicator could fit (depending on the range of appropriate possibilities)

Interpretation:

**Yes** means that the standard indicator's result level is suitable

**No** means that the standard indicator's result level is not suitable

**Uncertain** means the standard indicator could fit multiple result levels

## Section 2: Standard Indicator Suitability to Result Category

Key definitions to refer to for suitability assessment:

(Outputs)

**Knowledge Product:** “an intellectual asset produced by the CGIAR, as an output of research and development activities” (e.g., reports, briefs, extension, training/e-learning materials, books/chapters, data/databases, data collection/analysis tools, video, audio, images/graphics, maps, GIS, software/code, models, digital/mobile apps, websites/data portals/online platforms)

**Innovation Development:** “a new, improved, or adapted output of groups of outputs such as technologies, products and services, policies, and other organizational and institutional arrangements with high potential to contribute to positive impacts when used at scale”

**Capacity Sharing for Development:** “people trained by the CGIAR, with the aim of leading to changes in knowledge and behavior”

(Outcomes)

**Uptake of Information Product:** Knowledge products (non-innovations) are being used

**Innovation Use:** “the extent to which an innovation is being used, by which type of users and under which conditions”

**Capacity Change:** “changes in capacity of organizations resulting from CGIAR interventions”

**Policy Change:** “policies, strategies, legal instruments, programs, budgets, or investments at different scales (local to global) that have been modified in design or implementation, with evidence that the change was informed by CGIAR research”

(Impact Area)

**Poverty Reduction, Livelihoods and Jobs**

**Nutrition, Health and Food Security**

**Gender Equality, Youth and Social Inclusion**

**Environmental Health and Biodiversity**

**Climate Adaptation and Mitigation**

Column G: Result category suitability

**Function of the field:** Indicate standard indicator’s suitability with original classification.

How to fill it in:

- Select **Yes** if the standard indicator measures something that matches the definition of the original SIDS result category
- Select **No** if the standard indicator measures something that matches a different definition than the original SIDS result category
- Select **Uncertain** if it is unclear what the standard indicator measures, and could either fit the definition of multiple result category

Column H: Justification

**Function of the field:** Justify the suitability assessment done in previous field.

How to fill it in:

- If **Yes**, specify aspects of the standard indicator that meets the result category definition
- If **No**, explain why the standard indicator doesn’t fit, specify aspects of the standard indicator that correspond to another result category definition; indicate if a new result category needs to be drafted
- If **Uncertain**, explain why uncertain; specify aspects of the standard indicator that are inconsistent with other result category definition(s); indicate if a new result category needs to be drafted

Column I: Reclassification

Function of the field: Identify the best-fit result category.

How to fill it in:

- If **Yes**, note N/A (not applicable)
- If **No**, note the result category that the indicator matches the best (note: only one option can be selected)
- If **Uncertain**, note two or more result categories that the indicator could fit (depending on the range of appropriate possibilities)
  - \*Exception: If a new result category needs to be drafted, fill in the open text field(s) with a suggestion for a better category

Interpretation:

**Yes** means that the standard indicator's result category was suitable

**No** means that the standard indicator's result category was not suitable AND/OR we suggest a 'better' result category

**Uncertain** means the standard indicator could fit multiple results categories

### Section 3: Assessment of Standard Indicator Uniqueness/Mutual Exclusivity

Column J: Standard indicator statement is unique

Function of the field: Indicate whether the standard indicator is unique (i.e., statement not repeated across the CRF).

How to fill it in:

- Select **Yes** if the wording of the standard indicator statement is not repeated
  - *Exceptions:*
    - If there are slight wording differences, then note: Yes (only slight wording differences make it unique)
    - If the word order is different, then note: Yes (though b/c rearranged wording)
- Select **No** if the wording of the standard indicator repeats in another indicator

Column K: Standard indicator statement is mutually exclusive

Function of the field: Indicate whether what the standard indicator is measuring is mutually exclusive (i.e., conceptually involves double-counting same measures).

How to fill it in:

- Select **Yes** if the wording of the standard indicator statement is mutually exclusive
  - *Exceptions:*
    - If uncertain of what a standard indicator is measuring b/c of interpretation (i.e., based on sections 1 or 2), note: Yes (unless it means \_\_\_\_, which then overlaps with [standard indicator])
- Select **No** if the wording of the standard indicator repeats is not mutually exclusive

List of standard indicators to refer to for mutual exclusivity assessment:

(OUTPUTS)

(Knowledge Products)

**No. peer reviewed publications**

**No. of other information products**

**MELIA:** The knowledge product is part of a MELIA study

(Innovation Development)

**Number of innovation profiles**

(Capacity Sharing for Development)



**Number of people trained**

(OUTCOMES)

(Uptake of Information Product)

**Altmetric score**

(Innovation Use)

**Number of innovations being used (stage 4):** Connected next user (rare): “The innovation is used by some organizations connected to partners involved in the initial innovation development”

**# beneficiaries using the innovation, disaggregated by gender**

**Other quantitative measure of CGIAR innovation use, e.g. area**

**Number of people (m/f) reached by CGIAR innovations**

**Number of ha (or other unit of measure) affected by CGIAR innovation**

**Number of innovation packages designed**

(Capacity Change)

**Change in capacity of key (a) individuals (b) organizations (government, civil society, private sector), (c) networks (e.g. multistakeholder platforms)**

**Organizational change**

(Policy Change)

**Number of new or significantly updated policies/strategies/legal instruments/programs/budgets/investments**

(IMPACT AREA)

(Poverty Reduction, Livelihoods and Jobs)

**#poor people benefiting from relevant CGIAR innovations**

**#people assisted to exit poverty**

(Nutrition, Health and Food Security)

**#people benefiting from relevant CGIAR innovations**

**# people meeting minimum dietary energy requirements**

**# people meeting minimum micronutrient requirements**

**# cases of communicable and noncommunicable diseases**

(Gender Equality, Youth and Social Inclusion)

**#women’s empowerment and inclusion in the agricultural sector**

**#women benefiting from relevant CGIAR innovations**

**#youth benefiting from relevant CGIAR innovations**

**#women assisted to exit poverty**

(Environmental Health and Biodiversity)

**#ha under improved management**

**#km3 consumptive water use**

**#ha deforestation**

**#Tg nitrogen application**

**#plant genetic accessions available and safely duplicated**

(Climate Adaptation and Mitigation)

**#tonnes CO2 equivalent emissions**

**#plans with evidence of implementation**

**#\$ climate adaptation investments**

**#people benefiting from climate adapted innovations**

**#people benefiting from the implementation of adaptation plans**

Column L: Justification

Function of the field: List where the standard indicator overlaps/double-counts another standard indicator.

How to fill it in:

Report: One CGIAR Standard Indicators

- If **Yes**, leave blank
- If **No**, list one or more standard indicators where there is overlap/potential for double-counting (can make some other notes if needed, can note how many other standard indicators overlap)

Column M: The Indicator has a Standardized Unit

Function of the field: Assess whether the standard indicator includes a unit.

How to fill it in:

- If the indicator includes a unit, document ‘Specifies the unit’ with the unit in brackets
- If the indicator has a unit, but it could be better qualified, document the unit in brackets and add notes for suggested improvement
- If the indicator does not include a unit, document ‘The unit of measure is missing’
  - For qualitative indicators, add ‘(but it is a qualitative indicator)’

Column N: Data Collection Method for the Indicator

Function of the field: Assess whether the standard indicator includes the way to measure it.

How to fill it in:

- If a measurement tool or data collection method **is mentioned**, document ‘Way to measure is noted’ and can include the method in brackets
- If a measurement tool or data collection method **is implied**, document ‘Way to measure is implied’ and can include implied method (e.g., by counting, from reporting, from attendance lists, etc.)
- If a measurement tool or data collection method is **not mentioned**, document ‘Way to measure is unclear’ and can include side notes

Column O: Requires Primary/Secondary Data Collection

Function of the field: To capture best guesses of who will be responsible for collecting, or who is most likely to collect indicator data.

How to fill it in:

- If the **CGIAR should collect the data/evidence**, then write ‘CGIAR will collect’
- If **Others already collect the data/evidence**, then write ‘Others collect’
- **Where uncertain, but likely to be one or the other**, add ‘[the most likely (CGIAR/others)]?’
- **Where uncertain, and equally likely to be either**, then write ‘Who collects? Uncertain if CGIAR or others’.

Column P: Comments

Function of the field: Document any questions or comments about the standard indicator.

How to fill it in:

- Write questions about the standard indicator.
- Write comments/observations about the standard indicator.

## Section 4: SMART Assessment

Column Q: Is the standard indicator specific?

Function of the field: Assess whether the standard indicator clearly specifies the variable to be measured.

How to fill it in:

- Select **Yes** if the standard indicator fully satisfies the specific criterion
- Select **No** if the standard indicator does not satisfy one or all aspects of the specific criterion

### Column R: Justification (S)

Function of the field: Justify the specific assessment done in previous field.

How to fill it in:

- If **Yes**, document ‘specifies the variable’
- If **No**, document ‘the variable is insufficiently defined’
- Can note what aspects of the criterion were and were not met.

### Column S: Is the standard indicator measurable?

Function of the field: Assess whether the variable is feasible to measure.

How to fill it in:

- Select **Yes** if the standard indicator fully satisfies the measurable criterion
- Select **No** if the standard indicator does not at all satisfy the measurable criterion
- Select **Uncertain** if there is insufficient information to make the assessment
  - \*rule: If ‘yes’ to one aspect, but ‘uncertain’ to another, then select **Uncertain**
  - \*rule: If ‘no’ to one aspect, but ‘uncertain’ to another, then select **Uncertain**

### Column T: Justification (M)

Function of the field: Justify the measurable assessment done in previous field.

How to fill it in:

- If **Yes**, then document ‘The indicator is feasible to measure (easy to collect or affordable to collect or others collect both current and baseline if required)’
- If **No**, then document ‘The indicator is not feasible to measure (intensive to collect or expensive to collect, no baseline exists)’
- If **Uncertain**, then document ‘Uncertain of feasibility to collect (many potential collection methods with varying affordability, collection intensity, or data availability, uncertain if there is a baseline)’

### Column U: Is the standard indicator achievable?

Function of the field: Assess whether is it feasible for the standard indicator to be realized within the reporting period.

How to fill it in:

- Select **Yes** if the standard indicator fully satisfies the achievable criterion
- Select **No** if the standard indicator does not at all satisfy the achievable criterion
- Select **Uncertain** if there is insufficient information to make the assessment
  - \*rule: If ‘yes’ to one aspect, but ‘uncertain’ to another, then select **Uncertain**
  - \*rule: If ‘no’ to one aspect, but ‘uncertain’ to another, then select **Uncertain**

\*Assumption: Can be observed within the reporting period (annually).

### Column V: Justification

Function of the field: Justify the achievable assessment done in previous field.

How to fill it in:

- If **Yes**, then document ‘The indicator can be observed within the expected timeframe’
- If **No**, then document ‘The indicator cannot be observed within the expected timeframe’
- If **Uncertain**, then document ‘Critical information missing to make the assessment’

Column W: Is the standard indicator relevant?

Function of the field: Assess whether the indicator is related/connected to its result type/category (i.e., relevant to the goal or objective being measured) and is useful to measure progress along the ToC or the change process.

How to fill it in:

- Select **Yes** if the standard indicator fully satisfies the relevant criterion
- Select **Partial** if the standard indicator partially satisfies the relevant criterion (e.g., answer ‘yes’ to one aspect, but answer ‘no’ to another)
- Select **No** if the standard indicator does not at all satisfy the relevant criterion
- Select **Uncertain** if there is insufficient information to make the assessment
  - \*rule: If ‘yes’ to one aspect, but ‘uncertain’ to another, then select **Uncertain**
  - \*rule: If ‘no’ to one aspect, but ‘uncertain’ to another, then select **Uncertain**

Column X: Justification

Function of the field: Justify the relevant assessment done in previous field.

How to fill it in:

- Consider which sub-criteria apply:

	A	B
1	Indicator is linked to a result	The indicator is relevant (and appropriate) as a proxy for the result of interest
2	Indicator is linked to a result, but the result is poorly framed/ meaningless/mismatched (link is problematic)	Indicator is not relevant (and appropriate) as a proxy for the result of interest
3	Indicator is not linked to a result, so connection unclear	It is not clear whether the indicator is relevant (or appropriate) as a proxy for the result of interest

- If **Yes**, then document A1 and B1
- If **Partial**, then document either A1 and B2, or A2 and B1, or A3 and B1
- If **No**, then document A3 and B2
- If **Uncertain**, then document B3 with any A1, A2, or A3

Column Y: Is the standard indicator time-bound?

Function of the field: Assess whether the standard indicator has a set/explicit timeframe for achieving the goal or objective.

How to fill it in:

- Select **Yes** if the standard indicator fully satisfies the time-bound criterion
- Select **No** if the standard indicator does not at all satisfy the time-bound criterion

Column Z: Justification

Function of the field: Justify the time-bound assessment done in previous field.

How to fill it in:

- If **Yes**, then document ‘Provides a timeframe/window for measurement (e.g., date to date); frequency of collection noted’
- If **No**, and it is an output or outcome indicator, then document ‘Not explicit (implies EoP achievement)’

Report: One CGIAR Standard Indicators

- If *No*, and it is an impact indicator, then document ‘No baseline or target listed to be achieved by a certain time’

## Section 5: Common Errors

### Column AA: Nature of Error

Function of the field: Document issues identified through the SMART assessment and note any other observations on errors made in the process.

How to fill it in:

- List and describe all errors (can draw from the below list or document new ones)

*List of Potential Error Observations:*

- The indicator measures two or more things
- The indicator repeats the result statement
- The indicator statement is not an indicator (is actually the means/tool to measure, is an assumption, is the result, is an objective, etc.)
- Geographic scope/scale not explicit
- Indicator can fit many standard indicators (i.e., the standard indicators are not mutually exclusive)
- Ambiguity in the indicator makes it difficult to map

### Column AB: Style or Structural Observations

Function of the field: Document how the standard indicator is framed and note any issues with formulation.

How to fill it in:

- Build a chain of components based on the standard indicator's original structure (can draw from the below list or document new ones):
  - e.g., [component] + [component] + [component] + [component]
- Describe any formulation issues (can draw from the below list or document new ones)

*List of Framing Components:*

- Variable [focal part of the indicator; the *what* or the *who* (if 'number of people...')]
- Measure [specifies what aspect of the variable will be measured]
- Unit [specifies measurement system]
- Method [qualifies approach/tool/instrument used to measure; the *how*]
- Target [quantifies goal to reach]
- Actor [actor mentioned in the indicator]
- Geography [qualifies location information and/or scope/scale; the *where*]
- Verb [qualifies the action being tracked/measured]
- Rate [qualifies the speed of occurrence]
- Magnitude [relative adjective; e.g., increased, decreased]
- Frequency [qualifies temporal information of when data is collected/measured; *how often*]
- Timeframe [qualifies *when* to be measured]
- Readiness/state [qualifies the stage/state]
- Qualifier [descriptor of the variable]
- Activity [action undertaken by the project]
- Output [output/innovation produced by the project]
- Outcome [KASRB, nature of behaviour change/influence]
- Impact area [qualifies impact area]
- SDG [number/target]

*List of Formulation Observations:*

- The indicator is a quantitative/qualitative indicator
- The indicator contains sub-indicators with similar/related dimensions
- The indicator is not framed as an indicator
- Result/indicator contains acronyms
- Indicator uses symbols/contractions (e.g., ‘no.’ or ‘#’ to denote ‘number of’)
- The indicator includes information in round/square brackets
- Indicator refers to the activity/output
- No indication if a baseline/endline was planned for this indicator
- A target would be useful for this indicator

## **Section 6: Formulation guidance and suggested application of the standard indicator at WP/initiative level**

Column AC: Formulation guidance and suggested application of the standard indicator at the WP/initiative levels

Function of the field: Document guidance to better formulate standard indicator statements (to increase clarity, address errors, etc.), and specify how the indicator should be formulated at the WP/initiative level (that enables aggregation).

How to fill it in:

- Describe how to improve the indicator framing
- Provide an example of how the indicator could be better framed (at standard indicator level), and/or suggested formulation for WP/initiative level