

ANNEX II-C: METHODOLOGY FOR ASSESSING OUTCOME AND OUTPUT PERFORMANCE

Assessing WFP's programme performance at the outcome level

Outcome performance by strategic outcome

1. WFP's programme performance and its contribution to United Nations Sustainable Development Goals 2 and 17 in 2021 are assessed at the strategic objective level using the programme results chain approach of the revised corporate results framework (CRF) for 2017–2021. The assessment is drawn from the overall outcome performance of operations that are active during the reporting year. The analysis is based on corporate outcome indicators monitored by country offices in 2021.
2. The methodology applied to assess outcome performance was revised in 2021 to address limitations in relation to a lack of comparability with baseline data and the application of a single weight to all indicators, regardless of the frequency of their use.

Step 1 – Determining which indicators to consider for analysis

3. WFP rates performance against each of the outcome indicators monitored in each programme, for each of the various target groups, locations and modalities concerned. Each combination of indicator and target group, location and/or modality is referred to as a "measurement".
4. The following considerations apply:
 - a) Only indicators that are included in the revised CRF are measured (no country-specific outcome indicators are considered).
 - b) Only measurements for which a complete set of values – baseline, annual target and annual follow-up – is reported are included in the analysis. If one value is missing, the outcome indicator measurement is excluded from the analysis. If more than one follow-up value is recorded for the reporting year, only the most recent value is included in the analysis.

Figure 1: Outcome indicator data points and combinations

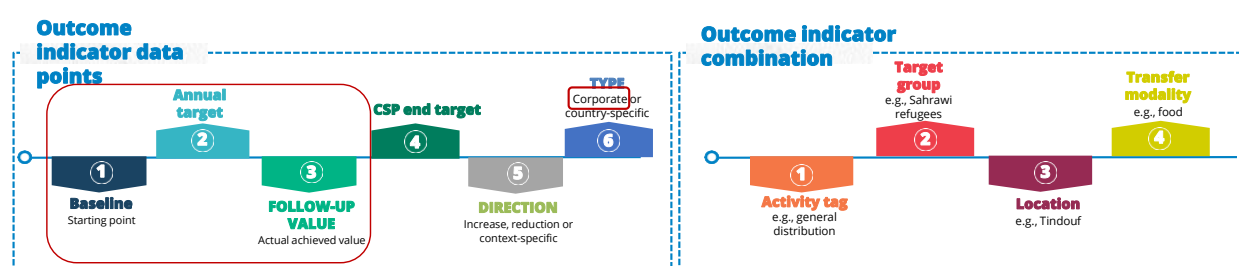


Table 1: Example of outcome indicator measurement for a specific target group assisted under a specific activity in a specific location

Outcome indicator	Category	Combination	Modality	Type	Direction	Baseline	2020 target value	2020 follow-up value
Food consumption score	Percentage of households with acceptable food consumption score	Adamawa/CAR, refugees food/URT1/URT_GD	Food	Corporate	Increase	86.76	80	66.4

Abbreviations: CAR = Central African Republic; GD = general distribution; URT = unconditional resource transfer.

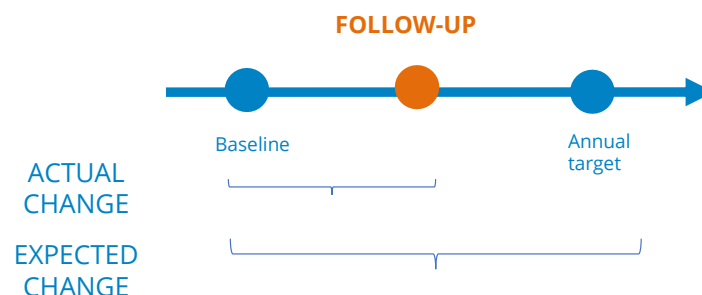
- c) To avoid double-counting, when an indicator has more than one category, only one of them is considered.
5. For example, food consumption score, an outcome indicator with three categories, requires the monitoring of three measurements – percentages of households with poor, borderline and acceptable food consumption score – for each of the target groups assisted under each activity implemented and using each transfer modality.
 6. For indicators with multiple categories, the following sub-indicators are considered in the analysis:
 - Food consumption score:
 - Percentage of households with acceptable food consumption score.
 - Food consumption score – nutrition:
 - Percentage of households that never consumed haem-iron-rich food (in the last seven days).
 - Percentage of households that never consumed protein-rich food (in the last seven days).
 - Percentage of households that never consumed vitamin A-rich food (in the last seven days).
 - Livelihood-based coping strategy index:
 - Percentage of households not using livelihood-based coping strategies.
 - Retention rate/drop-out rate:
 - Retention rate.
 - Value and volume of smallholder sales through WFP-supported aggregation systems:
 - Unit value (USD).
 - Volume (mt).

Step 2 – Rating performance under each outcome indicator measurement

7. Outcome performance is assessed by comparing annual values of outcome indicators against annual targets and the most recent baselines established in the logical frameworks of country strategic plans. Each measurement is assessed in relation to not only its end goal but also its starting point. Beginning in 2021, the calculation methodology applied in outcome analysis no longer includes aggregation along the CRF results chain beyond the strategic outcome categories.
8. There are three types of indicator in the revised CRF:
 - *Increase indicators (>=)*: Indicators for which an increase in value is expected post-intervention.
 - *Decrease indicators (<=)*: Indicators for which a decrease in value is expected post-intervention.
 - *Context-specific indicators* (direction cannot be predetermined): The direction of the indicator value is established by the country office depending on the expected programmatic results.
9. *Checking the consistency of the annual targets for each indicator*: WFP's results-based reporting sets targets for selected outcome indicators at the corporate level, while other targets are set at the country level depending on the specific operational context.
10. The target does not necessarily represent an improvement compared with the baseline. There can be operational reasons for setting a target that reflects a worsening situation compared with the baseline. Under the 2021 methodology, such measurements are flagged for review and, if found to be valid, are considered in the analysis.

11. *Performance assessment – achievement formula:* The revised formula for calculating achievement compares the actual change observed, which is due in part to WFP interventions, to the expected change, which represents the best-case scenario that could be observed (see first figure in the Appendix for an example of an increase indicator).

Figure 2: Example of the application of the achievement formula



12. The formulas for calculating performance are described below:
- a) If the target is set consistently against the baseline, the formula is:

$$\frac{\text{actual change}}{\text{expected change}} = \frac{(\text{latest follow-up} - \text{baseline})}{(\text{annual target} - \text{baseline})}$$

This formula is valid for both increase and decrease indicators.

- b) If the target is not set consistently against the baseline, the formula is:

$$\frac{\text{actual change}}{\text{expected change}} = \frac{(\text{latest follow-up} - \text{annual target})}{(\text{baseline} - \text{annual target})}$$

This formula is valid for both increase and decrease indicators.

13. *Performance assessment – colour coding:* The performance rating for each outcome indicator measurement is colour-coded according to the scale shown below.¹

3	Strong progress, achieved or exceeded	The actual change in the indicator is 80% or more of the expected change (80% – 80+∞%)
2	Some progress	The actual change in the indicator is between 50% and 80% of the expected change (50%–80%)
1	Insufficient or no progress	The actual change in the indicator is less than 50% of the expected change (50-∞% – 50%)

Step 3 – Assessing corporate outcome indicator performance under each strategic outcome category

14. Once a colour code has been assigned to each measurement, the measurements are aggregated along the CRF results chain. Indicators are grouped by strategic outcome, strategic result and strategic objective.
15. Under any given strategic outcome category, for each outcome indicator, an average of all the performance measurements is calculated (in terms of the colour codes) to determine WFP-wide performance.

¹ The thresholds for the scale underwent a minor revision between 2018 and 2019 to bring them into line with other corporate reporting exercises. Before 2019, the thresholds were strong: >= 90 percent, moderate: < 90->= 50 percent, and weak: < 50 percent.

16. For example, under strategic outcome category 1.1, maintained/enhanced individual and household access to adequate food, the country performance rating for food consumption score was calculated for 58 country offices with a total of 180 measurements. WFP-wide performance in terms of the food consumption score is the average of the percentage achievements of the 180 measurements, rated according to the colour codes.
17. *Data availability:* For each indicator, only those country offices that have reported sufficient data are included in the performance analysis. "Sufficient data" is defined as a complete set of values (annual follow-up, annual target and baseline) for at least 50 percent of the measurements against a specific outcome indicator.

Table 2: Example of reporting on performance under strategic outcome category 1.1

Outcome 1.1: Maintained/enhanced individual and household access to adequate food										
Outcome indicators	No. of countries reporting	No. of countries reporting sufficient data	Country reporting rate	No. of measurements	No. of complete measurements	Reporting rate	2021 performance rating	2020 performance rating	No. of measurements showing improvement or stabilization compared with the baseline	No. of measurements achieving or exceeding the target
1.1.1 Food consumption score – percentage of households with acceptable food consumption score	66	65	98%	203	188	93%	●	●	115	86
1.1.10 Enrolment rate	29	27	93%	35	33	94%	●	●	20	13
1.1.11 Attendance rate (new)	18	14	78%	25	21	84%	●	●	15	13
1.1.18 Emergency preparedness capacity index	1	1	100%	1	1	100%	●	●	1	0
1.1.19 Proportion of eligible population that participates in programme (coverage)	21	21	100%	56	55	98%	●	●	44	32

18. The aggregated results are presented in the table in annex II-D and an example is given in table 2. The following definitions apply:
- No. of countries reporting:* The number of countries reporting against an indicator for one or several target groups, locations, activities and/or modalities.
 - No. of countries reporting sufficient data:* The number of country offices reporting sufficient data (see paragraph above on data availability). For instance, if the Cameroon country office measured the food consumption score – the proportion of households with acceptable food consumption score – 30 times (for 30 different combinations of target group, activity, location and modality), but only ten of those measurements are complete, Cameroon is not included in the WFP-wide analysis of the food consumption score and is classified as a country reporting insufficient data.
 - Country reporting rate:* The number of countries reporting sufficient data divided by the number of countries reporting.
 - No. of measurements:* The total number of measurements against an indicator; for example, the food consumption score can be measured for one activity in three locations and for three target groups at each location, giving a total of nine measurements.
 - No. of complete measurements:* The number of measurements for which all three data points (annual follow-up, annual target and baseline) are reported.
 - Reporting rate:* The number of complete measurements divided by the number of measurements.

- g) *No. of measurements showing improvement or stabilization compared with the baseline:* The number of measurements for which the annual follow-up value shows an improvement compared with the baseline.
- h) *No. of measurements achieving or exceeding the target:* The number of measurements for which the annual follow-up value reaches or exceeds the annual target value.

Outcome performance by programme area

19. The assessment of outcome performance by programme area follows the same methodology as the assessment of outcome performance by strategic objective, with two differences:
- a) There is no requirement for country offices to report full data sets for at least 50 percent of their measurements.
- b) Performance in programme areas is assessed based on the following indicators.

Nutrition-specific activities and nutrition-sensitive programming

20. The outcome performance in nutrition is assessed for *nutrition-specific activities* and *nutrition-sensitive programming*.
21. Nutrition-specific activities include treatment of acute malnutrition programmes, including those for HIV/tuberculosis care and treatment, and malnutrition prevention programmes, including the prevention of acute malnutrition, stunting and micronutrient deficiencies.
22. Nutrition-sensitive programming contributes to improved nutrition outcomes and complements other types of intervention such as general distributions, school-based programming, food systems and smallholder support, and asset creation and livelihood support. Specific indicators are used to assess the nutrition outcome performance of these types of intervention.
23. The outcome indicator values used to calculate the performance metrics are based on the nutrition outcome indicators listed in the revised CRF and are monitored for specific sub-activities.

Nutrition-specific activities

Nutrition treatment indicator	Sub-activities
Moderate acute malnutrition treatment recovery rate	Treatment of moderate acute malnutrition
Moderate acute malnutrition treatment mortality rate	Treatment of moderate acute malnutrition
Moderate acute malnutrition treatment non-response rate	Treatment of moderate acute malnutrition
Moderate acute malnutrition treatment default rate	Treatment of moderate acute malnutrition
Proportion of eligible population that participates in programme (coverage)	Treatment of moderate acute malnutrition Therapeutic feeding (treatment of severe acute malnutrition)
Anti-retroviral treatment default rate	HIV/tuberculosis care and treatment
Tuberculosis treatment default rate	HIV/tuberculosis care and treatment
Anti-retroviral treatment nutritional recovery rate	HIV/tuberculosis care and treatment
Tuberculosis nutritional recovery rate	HIV/tuberculosis care and treatment

Nutrition prevention indicator	Sub-activities
Minimum dietary diversity – women	Prevention of stunting
Proportion of children 6–23 months of age who receive a minimum acceptable diet	Prevention of stunting
Proportion of eligible population that participates in programme (coverage)	Prevention of micronutrient deficiencies Prevention of acute malnutrition Prevention of stunting Stand-alone micronutrient supplementation
Proportion of target population that participates in an adequate number of distributions (adherence)	Prevention of micronutrient deficiencies Prevention of acute malnutrition Prevention of stunting

Nutrition-sensitive interventions

Nutrition-sensitive indicator	Sub-activities
Minimum dietary diversity – women	All except prevention of stunting
Proportion of children 6–23 months of age who receive a minimum acceptable diet	All except prevention of stunting
Food consumption score – nutrition*	All
Percentage increase in the production of high-quality and nutrition-dense foods	All
Percentage of targeted smallholder farmers reporting increased production of nutritious crops	All

* This is an indicator with multiple categories for which only the values related to “percentage of households that never consumed” are included in the assessment. Please refer to the text above on step 3, including the general rules on aggregation.

School-based programmes

24. Measurement of the outcome performance in school-based programmes takes into account onsite school feeding, alternative take-home rations and take-home rations.
25. The outcome indicator values used to calculate the performance metrics are based on the nutrition outcome indicators listed in the revised CRF.

School feeding indicators
Attendance rate
Enrolment rate
Graduation rate
Percentage of students who, by the end of two grades of primary schooling, demonstrate ability to read and understand grade-level text
Retention rate/drop-out rate
SABER school feeding national capacity

* This is an indicator with multiple categories, for which only the values related to “retention rate” are included in the assessment. Please refer to the text above on step 3, including the general rules on aggregation.

SABER: Systems Approach for Better Education Results.

Smallholder support, livelihood support, food system investment and risk management

26. The outcome performance is calculated to assess the collective achievement of asset creation and livelihood, smallholder agriculture market support and climate change adaptation and risk management activities.
27. The outcome indicator values used to calculate the performance metrics are based on the following outcome indicators from the revised CRF.

Indicators
Proportion of the population in targeted communities reporting benefits from an enhanced asset base
Proportion of the population in targeted communities reporting benefits from an enhanced livelihood asset base
Proportion of the population in targeted communities reporting environmental benefits
Food consumption score – percentage of households with acceptable food consumption score (activity tagging smallholder agriculture market support/climate change adaptation and risk management/asset creation and livelihood)
Average reduced coping strategy index – (activity tagging smallholder agriculture market support/climate change adaptation and risk management/asset creation and livelihood)
Livelihood coping strategy indicator – (activity tagging smallholder agriculture market support/climate change adaptation and risk management/asset creation and livelihood)
Default rate (as a percentage) of WFP pro-smallholder farmer procurement contracts
Percentage of targeted smallholders selling through WFP-supported farmer aggregation systems
Percentage of WFP food procured from smallholder farmer aggregation systems
Rate of smallholder post-harvest losses
Value and volume of smallholder sales through WFP-supported aggregation systems*
Proportion of targeted communities where there is evidence of improved capacity to manage climate shocks and risks

* This is an indicator with multiple categories for which the values related to “value of sales” and “volume of sales” are aggregated separately and their performance metrics are colour-coded separately.

Unconditional resource transfers

28. The outcome performance is calculated to assess the results of unconditional resource transfers.
29. The outcome indicator values used to calculate the performance metrics are based on the following outcome indicators from the revised CRF and monitored for general distribution.

Food security indicators	Activity tags
Food consumption score	General distribution
Livelihood-based coping strategy index (percentage of households not using coping strategies)	
Consumption-based coping strategy index (average)	
Economic capacity to meet essential needs	

Outcome performance by cross-cutting result

30. The methodology applied to assess performance by cross-cutting result follows the same steps as that for assessing outcome performance by strategic result. In step 3, for each cross-cutting result indicator, an average of all the performance measurements provided by country offices reporting sufficient data is calculated (in terms of the colour codes) to determine WFP-wide performance.

Methodology limitations for outcome analysis

Measuring results

31. Despite the revision undertaken in 2021, two major limitations of the methodology persist. First, when measurements are averaged at the indicator level, they are all assigned equal weight. This means that the reported performance against an indicator does not reflect the size of the operations concerned in terms of the number of beneficiaries assisted or the level of funding. The second limitation is related to data availability: only those country offices that report sufficient data are included in the performance analysis. While this helps WFP to avoid drawing conclusions from incomplete data, it means that the performance rating might not be indicative of the entirety of WFP's operations and may instead reflect performance in only a subset of countries with relevant operations. To mitigate this issue, the performance rating reported under each strategic objective includes a reference to the number of countries included in the analysis. Ratings should be read in conjunction with the accompanying narrative explanations, which provide a better overview of performance.

Assessing WFP's performance at the output level

32. WFP captures short-term programmatic results through its output indicators. In addition to numbers of beneficiaries assisted and transfers distributed, output indicators can measure the numbers of communities assisted, and health centres, sanitary facilities, school gardens, etc., created, restored or improved.
33. There are currently almost 400 output indicators and 60 units of measurement for monitoring outputs. To assess performance at the corporate level, some output indicators that use the same unit of measurement are grouped together. Only output indicators included in the revised CRF are aggregated and only those for which planned and actual values are reported are considered in the analysis.

Table 3: Example of output indicator measurement

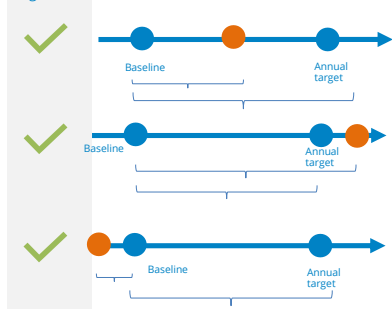
Output indicator	No. of measurements	No. of complete measurements	Reporting rate	No. of countries	Relevant countries	Country reporting rate	2021 needs-based target	2021 actual value	Performance rate
H.1 Number of shared services provided, by type									
Warehousing capacity (m ²) made available to the humanitarian community	2	2	100%	2	2	100%	88 287	89 539	101%
Number of technical reports shared with cluster partners	3	2	67%	3	2	67%	5	5	100%
Number of locations served	8	5	63%	5	4	80%	170	221	130%

Appendix – Outcome achievement formula

Increase indicator

Achievement formula – INCREASE TYPE INDICATORS ● FOLLOW-UP

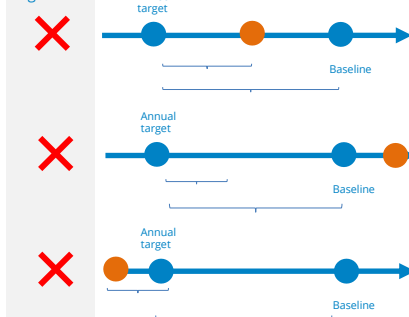
Target check



$$\frac{(follow - up - baseline)}{target - baseline}$$

$$\frac{(follow - up - target)}{baseline - target}$$

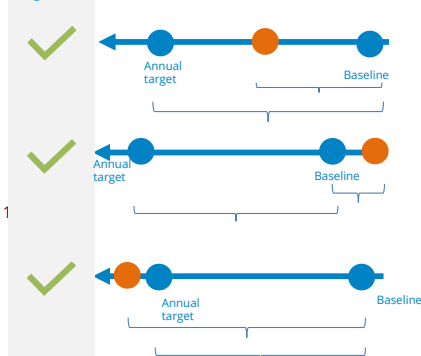
Target check



Decrease-type indicator

Achievement formula – DECREASE TYPE INDICATORS ● FOLLOW-UP

Target check



$$\frac{(follow - up - baseline)}{target - baseline}$$

$$\frac{(followup - Target)}{Baseline - Target}$$

Target check

