

## 附件 II-C: 评估成果和产出绩效的方法

### 在成果层面评价世界粮食计划署的方案绩效

#### 按战略成果分列的成果绩效

1. 应用经修订的 2017-2021 年全组织结果框架（CRF）中方案结果链方式，对世界粮食计划署的方案绩效和其对联国可持续发展目标 2 和 17 的贡献进行了评估。该评估是根据报告年度内活跃的行动的总体成果绩效而得出。本分析以国家办事处 2021 年监测的全组织成果指标为基础。
2. 用于评估成果绩效的方法于 2021 年进行了修订，以解决与基线数据缺乏可比性以及对所有指标应用单一权重相关的限制，无论其使用频率如何。

#### 第 1 步 - 确定在分析中考虑的指标

3. 世界粮食计划署根据每个方案中监测的每个成果指标，针对每个相关的不同目标群体、地点和方式，对绩效进行评级。指标和目标群体、地点和/或模式的每个组合称为“测量”。
4. 以下注意事项适用：
  - a) 仅测量修订后的 CRF 中包含的指标（不考虑特定国家的成果指标）。
  - b) 分析只包含报告了一组完整值的测量 - 基线、年度目标和年度跟踪。如果缺少一个值，则从分析中剔除成果指标测量。如果报告年度记录了多个后续值，则分析中仅包括最近的值。

图 1: 成果指标数据点和组合

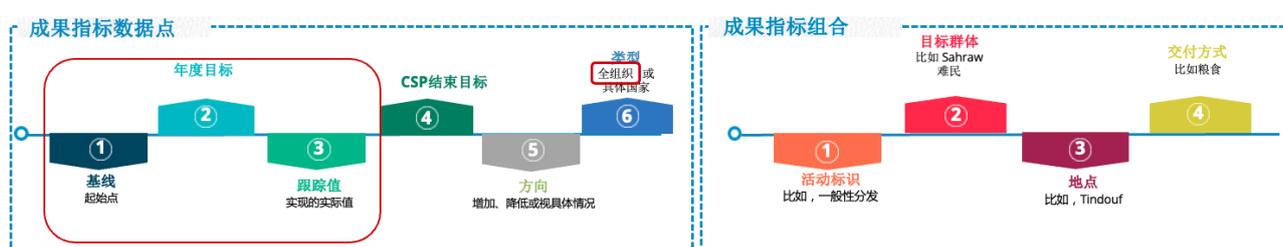


表 1: 在某地某项具体活动援助的某群体的成果指标测量的例子

成果指标	类别	组合	方式	类型	方向	基线	2020 年目标值	2020 年跟踪值
粮食消费得分	具有可接受粮食消费得分的家庭百分比	Adamawa/CAR, 难民粮食/ URT1/URT_GD	食品	全组织	增加	86.76	80	66.4

缩略语: CAR = 中非共和国; GD = 普通分发; URT = 无条件资源交付

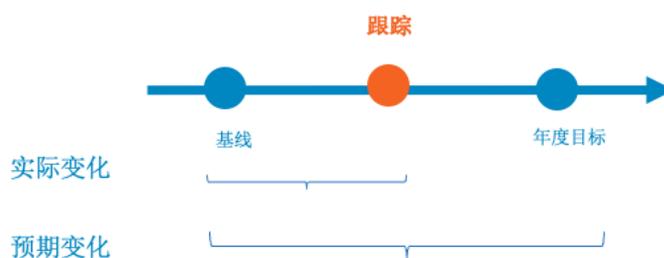
- c) 为避免重复计算，当一个指标有多个类别时，只考虑其中一个。
5. 例如，粮食消费分数 (FCS) 是一个包含三个类别的成果指标，它需要监测三个测量 — 粮食消费分数很差、边缘和可接受的家庭百分比 — 实施的每项活动和使用每种交付方式援助的每个目标群体。
  6. 对于具有多个类别的指标，分析中考虑了以下子指标：
    - 粮食消费分数：
      - 具有可接受粮食消费分数的家庭百分比。
    - 粮食消费分数 - 营养：
      - 从未消费富含血铁的食物家庭百分比（过去 7 天）。

- 从未消费富含蛋白质的食物的家庭百分比（过去 7 天）。
- 从未消费富含维生素 A 食物的家庭百分比（过去 7 天）。
- 基于生计的应对策略指数：
  - 不使用基于生计的应对策略的家庭百分比。
- 就学稳定率/辍学率：
  - 就学稳定率。
- 通过世界粮食计划署支持的汇总系统小农实现的销售价值和数量：
  - 单位价值（美元）。
  - 重量（公吨）。

## 第 2 步 - 在每项成果指标测量中评定绩效

7. 通过将成果指标的年度值与年度目标和在国家战略计划逻辑框架中建立的最新基线进行比较来评估成果绩效。每次测量的评估不仅涉及其最终目标，还涉及其起点。从 2021 年开始，成果分析中应用的计算方法不再包括战略成果类别之外的 CRF 成果链上的汇总。
8. 经修订的 CRF 中有三类指标：
  - **增加指标 (>=)**: 干预后预期值增加的指标。
  - **降低指标 (<=)**: 干预后预期值降低的指标
  - 针对具体情况的指标（方向无法预先确定）：指标值的方向由国家办事处根据预期的方案结果确定。
9. **检查每个指标的年度目标的一致性**：世界粮食计划署基于结果的报告，在全组织层面为选定的成果指标设定了目标，而其它目标则根据具体的行动背景在国家层面设定。
10. 与基线相比，目标不一定代表改进。设定一个反映与基线相比情况恶化的目标可能有行动上的原因。根据 2021 年方法，此类测量被标记以供审查，如果发现有效，则在分析中予以考虑。
11. **绩效评估——成就公式**：用于计算成就经修订的公式将观察到的实际变化 - 部分归因于世界粮食计划署干预 - 与预期变化 - 代表可以观察到的最佳情况 - 进行比较（增加指标示例见附录图 1）。

图 2: 应用成果公式的例子



12. 计算绩效的公式如下所述：

a) 如果目标与基线一致，公式则为：
$$\frac{\text{实际变化}}{\text{预期变化}} = \frac{(\text{最近跟踪} - \text{基线})}{(\text{年度目标} - \text{基线})}$$

这个公式对增加和降低指标都有效。

b) 如果目标与基线不一致，公式则为：
$$\frac{\text{实际变化}}{\text{预期变化}} = \frac{(\text{最近跟踪} - \text{年度目标})}{(\text{基线} - \text{年度目标})}$$

这个公式对增加和降低指标都有效。

13. **绩效评估 - 颜色编码**：每个成果指标测量的绩效评级根据下图所示的标度进行颜色编码<sup>1</sup>。

<sup>1</sup> 2018 年至 2019 年期间对该标度的门槛进行了小幅修订，以使其与全组织其它报告活动保持一致。2019 年之前，门槛是：强：>= 90%，中等：< 90 - >= 50%，弱：< 50%。

3	显著进展，实现或者超过	· 指标的实际变化达到预期变化的 80%或以上。(80% – 80+∞%)
2	有些进展	· 指标的实际变化在预期变化的 50%到 80%之间。(50%-80%)
1	进展不足或无进展	· 指标的实际变化不到预期的 50%。(50-∞% – 50%)

### 第 3 步 - 在每项战略成果类别下，评估全组织成果指标绩效

14. 一旦为每个测量分配了颜色代码，测量就会沿着 CRF 结果链聚合。指标按战略成果、战略结果和战略目标分组。
15. 在任何给定的战略成果类别下，对于每个成果指标，都会计算所有绩效测量的平均值（根据颜色代码），以确定世界粮食计划署范围内的绩效。
16. 例如，在战略成果类别 1.1（维持/增强个人和家庭获得充足食物的机会）下，应用用 180 次测量，计算了 58 个国家办事处的粮食消费分数的国家绩效评定。世界粮食计划署在 FCS 方面的绩效是 180 次测量的平均百分比，根据颜色代码进行评级。
17. 数据可用性：对于每个指标，只有那些报告了足够数据的国家办事处才被纳入绩效分析。“充足的数据”被定义为针对特定成果指标的至少 50% 测量的完整值集（年度跟踪、年度目标和基线）。

**表 2: 在战略成果类别 1.1 下，报告绩效的例子**

成果指标	报告的国家数	报告充足数据的国家数	国家报告率	测量数	完整测量数	报告率	绩效定级	与基线比，稳定改善的测定数	实现目标的测定数
1.1.1 粮食消费得分 – 具有可接受粮食消费家庭的百分比	66	65	98%	203	188	93%	●	115	86
1.1.10 入学率	29	27	93%	35	33	94%	●	20	13
1.1.11 就读率(新)	18	14	78%	25	21	84%	●	15	13
1.1.18 紧急状况准备能力指数	1	1	100%	1	1	100%	●	1	0
1.1.19 参与方案的合格人口比例(覆盖)	21	21	100%	56	55	98%	●	44	32

18. 汇总结果列于附件 II-D 的表格中，表 2 给出了一个例子。以下定义适用：
  - a. **报告的国家数量**：针对一个或多个目标群体、地点、活动和/或模式的指标进行报告的国家数量。
  - b. **报告足够数据的国家数量**：报告足够数据的国家办事处的数量（见上文关于数据可用性的段落）。例如，如果喀麦隆国家办事处测量了 FCS - 可接受 FCS 的家庭比例 - 30 次（针对目标群体、活动、位置和方式的 30 种不同组合），但只有十次测量是完整的，喀麦隆没有包括在粮食计划署对 FCS 的广泛分析中，并被归类为报告数据不足的国家。
  - c. **国家报告率**：报告足够数据的国家数量除以报告国家数量。
  - d. **测量次数**：针对某个指标的测量总数；例如，可以针对三个地点的一项活动和每个地点的三个目标群体测量 FCS，总共进行九次测量。
  - e. **完整测量数**：报告所有三个数据点（年度跟踪、年度目标和基线）的测量数。
  - f. **报告率**：完整测量的数量除以测量的数量。
  - g. **与基线相比显示改善/或稳定的测量次数**：年度跟踪值显示与基线相比有所改善的测量次数。
  - h. **达到或超过目标的测量数**：年度跟踪值达到或超过年度目标值的测量次数。

### 按方案领域分列的成果绩效

19. 按方案领域评估成果绩效的方法与按战略目标评估成果绩效的方法相同，但有两点不同：

- a. 国家办事处不需要报告至少 50% 的测量的完整数据集。
- b. 根据以下指标评估方案领域的绩效。

#### 营养专项活动和营养敏感型方案编制

20. 针对特定营养活动和营养敏感方案编制评估营养方面的成果绩效。
21. 针对营养的活动包括治疗急性营养不良方案，包括艾滋病毒/结核病护理和治疗方案，以及预防营养不良方案，包括预防急性营养不良、发育迟缓和微量营养素缺乏症。
22. 营养敏感型方案编制有助于改善营养成果，并补充其它类型的干预措施，例如一般性分配、基于学校的方案划、粮食系统和小农支持，以及资产创造和生计支持。特定指标用于评估这些干预措施的营养成果绩效。
23. 用于计算绩效指标的指标值基于修订后的 CRF 中列出的营养成果指标，并针对特定的子活动进行监测。

#### 营养专项活动

营养治疗指标	分活动
中度急性营养不良治疗恢复率	治疗中度营养不良
中度急性营养不良治疗死亡率	治疗中度营养不良
中度急性营养不良治疗不反应率	治疗中度营养不良
中度急性营养不良治疗违约率	治疗中度营养不良
参与方案的合规人口比了（覆盖面）	治疗中度营养不良 治疗性喂养（治疗严重急性营养不良）
ART 治疗违约率	HIV/结核病保健和治疗
结核病治疗违约率	HIV/结核病保健和治疗
ART 治疗营养回复率	HIV/结核病保健和治疗
结核病营养回复率	HIV/结核病保健和治疗

营养预防指标	分活动
最低饮食多样性 - 妇女	预防发育迟缓
接受最低可接受饮食的 6-23 月龄儿童百分比	预防发育迟缓
参与方案的合规人口比了（覆盖面）	预防微量营养素缺乏 预防急性营养不良 预防发育迟缓 单独微量营养补充
目标群体中参与足够数量分发的比例（遵从率）	预防微量营养素缺乏 预防急性营养不良 预防发育迟缓

### 营养敏感型干预

营养敏感型指标	分活动
最低饮食多样性 - 妇女	除预防发育迟缓之外所有
接受最低可接受饮食的 6-23 月龄儿童百分比	除预防发育迟缓之外所有
粮食消费分数 - 营养*	全部
优质和营养密集型食品生产增加的百分比	全部
报告营养作物生产增加的目标小农百分比	全部

\* 这是一个具有多个类别的指标，评估中仅包含与“从未消费过的家庭百分比”相关的值。请参阅上文第 3 步的文本，包括汇总的一般规则。

### 基于学校的方案

- 以学校为基础的方案的成果绩效测量考虑了现场学校供餐、替代的带回家的日粮和带回家的日粮。
- 用于计算绩效指标的成果指标值基于修订后的 CRF 中列出的营养成果指标。

学校供膳指标
就读率
入学率
毕业率
在小学两个年级结束时表现出阅读和理解年级水平文本能力的学生百分比
就学稳定率/辍学率
SABER 学校供膳国家能力

\* 这是一个具有多个类别的指标，评估中仅包含与“就学稳定率”相关的值。请参阅上文第 3 步的文本，包括汇总的一般规则。

SABER: 获得更好教育的系统方式。

### 小农支持、生计支持、粮食系统投资和风险管理

- 计算成果绩效以评估资产创造和生计、小农农业市场支持以及气候变化适应和风险管理活动的集体成就。
- 用于计算绩效指标的成果指标值基于修订后的 CRF 中的以下结果指标。

指标
目标社区中报告从增强的资产基础中受益的人口比例
目标社区中报告从增强的生计资产基础中受益的人口比例
目标社区中报告环境效益的人口比例
粮食消费分数 - 具有可接受的粮食消费得分的家庭百分比（活动标记小农农业市场支持/气候变化适应和风险管理/资产创造和生计）
平均减少应对策略指数 — （活动标记为小农农业市场支持/气候变化适应和风险管理/资产创造和生计）
生计应对策略指标 — （小农农业市场支持/气候变化适应和风险管理/资产创造和生计）
世界粮食计划署支持小农采购合同的违约率（百分比）
通过世界粮食计划署支持的农民聚集系统进行销售的目标小农的百分比
世界粮食计划署从小农聚集系统采购的粮食的百分比
小农收获后损失率
通过世界粮食计划署支持的汇总系统实现小农销售的价值和数量*
有证据表明管理气候冲击和风险管理的能力得到提高的目标社区的比例

\* 这是一个具有多个类别的指标，其中与“销售价值”和“销售量”相关的值分别汇总，其绩效指标分别用颜色编码。

### 无条件资源交付

- 计算成果绩效以评估无条件资源交付的结果。

29. 用于计算绩效指标的成果指标值基于修订后的 CRF 中的以下成果指标，并针对一般性分发进行监测。

粮食安全指标	活动标记
粮食消费分数	一般性分发
基于生计的应对战略指数（没有实施应对战略的家庭百分比）	
基于消费的应对战略指数（平均）	
满足基本需要的经济能力	

### 跨领域结果的结果绩效

30. 用于通过跨领域结果评估绩效的方法遵循与通过战略结果评估成果绩效相同的步骤。在步骤 3 中，对于每个跨领域结果指标，计算国家办事处提供的报告足够数据的所有绩效测量的平均值（根据颜色代码），以确定实际粮食计划署的整体绩效。

### 成果分析方法的局限性

#### 衡量结果

31. 尽管在 2021 年进行了修订，但该方法仍然存在两个主要的局限性。首先，当测量值在指标级别进行平均时，它们都被赋予了相同的权重。这意味着根据一项指标报告的绩效并不能反映有关业务的规模，即援助的受益人数量或资金水平。第二个限制与数据可用性有关：只有那些报告足够数据的国家办事处才被纳入绩效分析。虽然这有助于世界粮食计划署避免从不完整的数据中得出结论，但这意味着绩效评级可能并不代表该署的全部业务，而可能只反映具有相关业务的部分国家的业绩。为了缓解这个问题，在每个战略目标下报告的绩效评级都包含了对分析中包含的国家数量的参考。评级应与随附的叙述性解释一起阅读，叙述性解释对绩效进行了更好地概述。

### 在产出层面评估世界粮食计划署的绩效

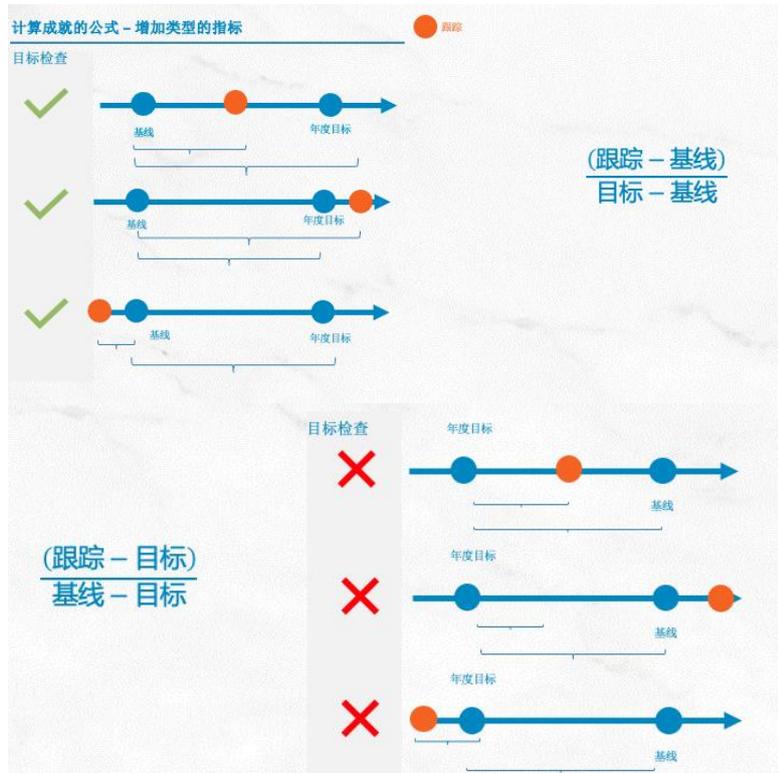
32. 世界粮食计划署通过其产出指标获取短期方案结果。除了援助的受益人数量和分配的交付之外，产出指标还可以衡量援助的社区数量，以及创建、恢复或改善的保健中心、卫生设施、学校花园等的数量。
33. 目前有近 400 个产出指标和 60 个用于监测产出的计量单位。为了评估全组织层面的绩效，将使用相同计量单位的产出指标组合在一起。只有包含在修订后的 CRF 中的产出指标才被汇总，并且在分析中只考虑那些报告了计划值和实际值的指标。

**表 3: 产出指标测定的例子**

产出指标 <sup>2</sup>	测定的数量	完整测定的数量	报告率	国家数	相关的国家	国家报告率	2021 基于需求的目标	2021 年实际值	绩效率
<b>H.1 按类别分, 提供的共享服务数</b>									
为人道主义社区提供的可用仓储能力（立方米）	2	2	100%	2	2	100%	88 287	89 539	101%
与分组成员共享的技术报告数	3	2	67%	3	2	67%	5	5	100%
服务的地点数	8	5	63%	5	4	80%	170	221	130%

## 附录 - 计算实现成就的公式

### 呈现增加的指标



### 呈现降低类型的指标

