ACADÉMIE HUMANITAIRE · HUMANITAIRES JOBS

# Mini MBA Programmes et MEAL : concevoir, piloter et prouver l'impact

## Manuel complet

Programme de 3 mois exactement (12 semaines, 144 heures) pour concevoir, piloter et évaluer des programmes humanitaires : évaluation des besoins et sévérité, chaîne de résultats et indicateurs, ciblage et transferts monétaires, planification et budget, plan MEAL et qualité des données, échantillonnage, analyse, redevabilité, évaluation, mesure d'impact, optimisation des ressources et gestion adaptative. Exercices chiffrés corrigés, trois examens mensuels avec études de cas, et un diplôme vérifiable de l'Académie Humanitaire.

**https://humanitairejobs.org/fr/academie/mini-mba-programmes-meal/**

Document gratuit de l'Académie Humanitaire. Il peut être imprimé et partagé librement à des fins de formation, sans modification.

# Sommaire


1. Semaine 1 : Évaluer les besoins : données, sévérité et priorités
2. Semaine 2 : Concevoir un programme : chaîne de résultats, indicateurs et cibles
3. Semaine 3 : Ciblage et modalités : qui aider, comment, avec combien
4. Semaine 4 : Planifier et budgéter un programme : du plan de travail au rythme de dépense
5. Semaine 5 : Le système de suivi : plan MEAL, qualité des données et suivi à distance
6. Semaine 6 : Enquêtes et échantillonnage : obtenir des chiffres justes
7. Semaine 7 : Analyser les données : statistiques utiles, pièges et visualisation
8. Semaine 8 : Redevabilité envers les populations : participation, plaintes et retours
9. Semaine 9 : Évaluer : critères, termes de référence, matrice et utilisation
10. Semaine 10 : Mesurer l'impact : contrefactuel, méthodes et contribution
11. Semaine 11 : Optimisation des ressources : coût-efficience, coût-efficacité et arbitrages
12. Semaine 12 : Gestion adaptative, apprentissage et rapports : diriger par les preuves

# Semaine 1 : Évaluer les besoins : données, sévérité et priorités

**Objectifs**

- Situer l'évaluation des besoins dans le cycle de programme humanitaire et choisir le bon type d'évaluation selon la phase de la crise.
- Combiner données secondaires et primaires, repérer les biais, et classer la sévérité d'une zone avec une règle explicite.
- Passer du constat à la décision : analyse de la réponse, faisabilité, valeur ajoutée et choix argumenté des priorités.

## L'évaluation des besoins dans le cycle de programme humanitaire

Le **cycle de programme humanitaire** (HPC) enchaîne cinq étapes : évaluation et analyse des besoins, planification stratégique, mobilisation des ressources, mise en œuvre et suivi, revue et évaluation. Le premier produit collectif est l'**aperçu des besoins humanitaires** (HNO), qui alimente le plan de réponse (depuis 2024, la plupart des pays publient un document unique, le HNRP). Une ONG qui conçoit un programme sans lire le HNO de son pays se prive de la meilleure synthèse disponible et risque de proposer ce que d'autres couvrent déjà.

Le type d'évaluation dépend de la phase de la crise :

| Phase | Outil typique | Délai | Question de décision |
|---|---|---|---|
| Choc soudain | Évaluation rapide initiale multisectorielle (MIRA ou équivalent) | 72 heures à 2 semaines | Où sont les besoins vitaux, et qui intervient où ? |
| Crise installée | Évaluation multisectorielle des besoins (MSNA), enquête ménage représentative | Annuelle | Quelle sévérité par zone et par groupe, combien de personnes dans le besoin ? |
| Programme en cours | Évaluation sectorielle approfondie (nutrition SMART, marché, protection) | Selon besoin | Faut-il ajuster la modalité, le ciblage, la valeur des transferts ? |

Trois notions sont à distinguer. Le **besoin** est l'écart entre la situation observée et un standard (par exemple 15 litres d'eau par personne et par jour selon Sphère). La **vulnérabilité** est la probabilité de subir un préjudice face au choc, qui dépend de l'exposition et des capacités. Les **capacités** sont les ressources des personnes, des communautés et des institutions. Un bon diagnostic décrit les trois ; une liste de manques sans capacités conduit à des réponses qui remplacent ce qui fonctionnait déjà.

> Règle d'or : chaque question d'une évaluation doit servir une décision identifiée à l'avance. Une question « intéressante » qui ne change aucune décision coûte du temps

aux personnes interrogées et dilue l'analyse.

## Les données secondaires : commencer par ce qui existe

La **revue des données secondaires** (SDR) rassemble ce qui est déjà connu avant de collecter quoi que ce soit : HNO, rapports de cluster, enquêtes démographiques et de santé (EDS), analyses IPC, évaluations d'autres ONG, données des services techniques de l'État, suivi des déplacements (DTM de l'OIM), prix sur les marchés, rapports de protection. Une SDR bien faite réduit de moitié la taille d'une collecte primaire, car elle permet de ne poser que les questions sans réponse.

Chaque source est notée sur deux dimensions. La **fiabilité de la source** : méthode connue, échantillon décrit, auteur sans conflit d'intérêts. La **crédibilité de l'information** : confirmée par d'autres sources, cohérente avec ce que l'on sait, récente. Une information d'une source fiable mais vieille de deux ans dans une zone de déplacements massifs est faible ; une information d'une source moyenne confirmée par trois autres est solide.

La SDR se termine par une **analyse des lacunes** : pour chaque question de décision, que sait-on, avec quel niveau de confiance, et que reste-t-il à collecter ? C'est ce tableau, et non une liste de thèmes, qui justifie le questionnaire.

- **Triangulation des sources** : la même information confirmée par des sources indépendantes.
- **Triangulation des méthodes** : une enquête ménage et des groupes de discussion qui convergent.
- **Triangulation des enquêteurs** : des équipes différentes qui obtiennent les mêmes résultats.

## Les données primaires : choisir les méthodes et maîtriser les biais

Quatre méthodes se complètent. L'**entretien avec des informateurs clés** (KII) interroge des personnes qui ont une vue d'ensemble : chef de village, infirmier, enseignante, commerçant, responsable d'une association de femmes. Rapide et peu coûteux, il donne des tendances mais pas des proportions. La **discussion de groupe** (FGD), 6 à 10 personnes homogènes (mêmes sexe et tranche d'âge), explore les perceptions, les normes et les préférences. L'**observation directe** (latrines, marchés, files d'attente) vérifie ce qui est déclaré. L'**enquête ménage** sur un échantillon représentatif est la seule qui produit des proportions généralisables, à condition que l'échantillonnage soit rigoureux (semaine 6).

Les biais les plus fréquents :

- **Biais de sélection** : n'interroger que les villages accessibles par la route, ou les ménages présents en journée (on manque les travailleurs journaliers).
- **Biais de complaisance et d'attente d'aide** : les personnes exagèrent ou minimisent selon ce qu'elles pensent que l'enquêteur veut entendre ou selon l'aide espérée. On le

réduit en expliquant clairement qu'aucune aide n'est promise et en préférant des indicateurs observables.

- **Biais d'enquêteur** : formulation suggestive, présence d'un homme pour interroger des femmes sur la violence. On forme, on teste le questionnaire et on constitue des équipes mixtes.
- **Biais de survie** : dans une zone de forte mortalité ou de départ, les ménages les plus touchés ne sont plus là pour répondre.

> Éthique minimale : consentement éclairé, possibilité de refuser sans conséquence, aucune question sensible sans protocole de référencement (violences basées sur le genre, protection de l'enfance), et données personnelles réduites au strict nécessaire.

## Mesurer la sévérité : règles explicites et personnes dans le besoin

Comparer des zones exige une **échelle de sévérité** commune. Le cadre intersectoriel commun d'analyse (JIAF) et le Cadre intégré de classification de la sécurité alimentaire (IPC) utilisent tous deux cinq phases, de 1 (minimale) à 5 (catastrophique). L'IPC classe une zone selon une **règle des 20 %** : la zone est placée dans la phase la plus sévère qui concerne, avec les phases supérieures, au moins 20 % de sa population.

**Exemple.** Population de 100 000 personnes : phase 1 : 35 % ; phase 2 : 30 % ; phase 3 : 22 % ; phase 4 : 10 % ; phase 5 : 3 %. Phases 4 et 5 réunies : 13 %, sous le seuil. Phases 3, 4 et 5 réunies : 35 %, au-dessus de 20 %. La zone est donc classée en **phase 3**. Le nombre de personnes dans le besoin d'une action urgente (phase 3 ou plus) est 35 000.

Le nombre de **personnes dans le besoin** (PIN) d'un secteur n'est pas la population totale de la zone : c'est la somme des personnes dont la situation atteint le seuil de sévérité défini (souvent phase 3 ou plus). Additionner les PIN de plusieurs secteurs compte plusieurs fois la même personne. On retient donc, pour le chiffre intersectoriel, le maximum par zone et par groupe, ou une analyse au niveau du ménage.

La sévérité doit être **désagrégée** par sexe, âge et handicap. Pour le handicap, le jeu court du Groupe de Washington pose six questions (voir, entendre, marcher, se souvenir, prendre soin de soi, communiquer) ; une personne est comptée comme vivant avec un handicap si elle répond « beaucoup de difficulté » ou « ne peut pas du tout » à au moins une question.

## Analyse de protection, de genre et de marché

Une évaluation qui ne regarde que les manques matériels passe à côté des risques. L'**analyse de protection** identifie les menaces (qui fait quoi à qui), les vulnérabilités des personnes face à ces menaces, et leurs capacités : risque = menace × vulnérabilité / capacité. Elle repère les groupes exclus de l'aide (personnes sans documents, minorités, personnes âgées isolées) et les risques que l'aide elle-même peut créer : distributions qui exposent à des vols, trajets dangereux, tensions entre déplacés et population hôte.

L'**analyse de genre** décrit qui décide, qui contrôle les ressources, qui a accès aux services, et comment la crise modifie ces rôles. Elle évite par exemple de verser l'aide monétaire à un membre du ménage qui n'en contrôle pas l'usage, ou de fixer des horaires de distribution incompatibles avec les tâches des femmes.

L'**analyse de marché** détermine si les transferts monétaires sont faisables : les marchés fonctionnent-ils, les commerçants peuvent-ils répondre à une hausse de la demande sans inflation, l'accès physique et financier est-il sûr ? Les outils de référence (évaluation rapide de marché, cartographie des systèmes de marché) répondent à trois questions : disponibilité des biens, prix et leur évolution, capacité de réapprovisionnement.

> Ne pas nuire : pour chaque option de réponse, on écrit au moins un risque qu'elle crée et la mesure qui le réduit.

## De l'analyse à la décision : l'analyse de la réponse

L'**analyse de la réponse** fait le lien entre ce que l'on sait des besoins et ce que l'organisation va faire. Elle compare des options (nature, cash, coupons, services, plaidoyer) sur des critères explicites : pertinence par rapport aux besoins prioritaires, préférences des personnes, faisabilité (accès, marchés, sécurité, capacités), coût-efficacité, risques et ce que font déjà les autres acteurs (matrice 3W ou 4W : qui fait quoi, où et quand).

Une méthode simple est la **matrice pondérée** : chaque critère reçoit un poids fixé avant de noter les options, pour éviter de choisir d'abord et de justifier ensuite. Par exemple, pertinence 30 %, préférences 20 %, faisabilité 20 %, coût-efficacité 20 %, risques 10 %. Une option qui obtient 4/5 en pertinence, 5 en préférences, 3 en faisabilité, 4 en coût-efficacité et 3 en risques totalise $0{,}3 \times 4 + 0{,}2 \times 5 + 0{,}2 \times 3 + 0{,}2 \times 4 + 0{,}1 \times 3 = 3{,}9$.

La décision finale doit aussi tenir compte de la **valeur ajoutée** de l'organisation : présence, acceptation, expertise, partenaires locaux. Si un acteur couvre déjà correctement un besoin dans une zone, la meilleure décision est souvent d'aller ailleurs ou de compléter, et non de dupliquer.

Le produit final est une **note d'analyse** de quelques pages : besoins prioritaires par zone et par groupe, sévérité, lacunes de couverture, options comparées, recommandation et niveau de confiance. C'est elle que lira la direction, pas la base de données.

## Exercice d'application : classer deux zones et choisir où intervenir

**Situation.** Deux districts touchés par la sécheresse. District A (60 000 habitants) : phase 1 : 30 % ; phase 2 : 32 % ; phase 3 : 25 % ; phase 4 : 11 % ; phase 5 : 2 %. District B (80 000 habitants) : phase 1 : 45 % ; phase 2 : 37 % ; phase 3 : 12 % ; phase 4 : 5 % ; phase 5 : 1 %. Une autre ONG distribue déjà des vivres à 15 000 personnes en phase 3 ou plus dans le district A. Les marchés du district B fonctionnent, ceux du district A sont très perturbés.

**Travail demandé.** 1) Classez chaque district selon la règle des 20 %. 2) Calculez le nombre de personnes en phase 3 ou plus dans chaque district et le nombre non couvert. 3) Recommandez une zone et une modalité, et justifiez en trois phrases.

**Corrigé.** District A : phases 4 et 5 = 13 %, phases 3 à 5 = 38 % : phase 3. Personnes en phase 3 ou plus : 0,38 × 60 000 = 22 800 ; non couvertes : 22 800 − 15 000 = 7 800. District B : phases 3 à 5 = 18 %, sous le seuil ; phases 2 à 5 = 55 % : phase 2. Personnes en phase 3 ou plus : 0,18 × 80 000 = 14 400, toutes non couvertes. Le district B compte plus de personnes non couvertes (14 400 contre 7 800) malgré une classification moins sévère, et ses marchés permettent des transferts monétaires. Une réponse défendable : transferts monétaires ciblés sur les ménages en phase 3 ou plus du district B, et complément en nature dans le district A en coordination avec l'autre ONG. La classification d'une zone ne suffit pas à décider : il faut regarder les nombres absolus et la couverture.

## À retenir

- Chaque question d'évaluation doit servir une décision connue à l'avance.
- Revue des données secondaires d'abord, avec une notation de fiabilité et de crédibilité et une analyse des lacunes.
- Seule une enquête représentative produit des proportions généralisables ; les KII et les FGD donnent tendances et explications.
- Règle des 20 % : une zone prend la phase la plus sévère atteinte, avec les phases supérieures, par au moins 20 % de sa population.
- L'analyse de la réponse compare des options sur des critères pondérés fixés à l'avance, en tenant compte de la couverture existante.

## Quiz

1. Répartition d'une zone : phase 1 : 40 % ; phase 2 : 35 % ; phase 3 : 15 % ; phase 4 : 8 % ; phase 5 : 2 %. Selon la règle des 20 %, la zone est en :
   A. Phase 2
   B. Phase 3
   C. Phase 4
   D. Phase 1

2. Pourquoi ne peut-on pas additionner les personnes dans le besoin de chaque secteur pour obtenir un total intersectoriel ?
   A. Parce que les secteurs utilisent des monnaies différentes
   B. Parce que la même personne peut être dans le besoin dans plusieurs secteurs et serait comptée plusieurs fois
   C. Parce que le total serait trop faible
   D. Parce que les bailleurs l'interdisent

3. Quelle méthode permet d'estimer une proportion représentative de ménages sans accès à l'eau potable ?
   A. Trois entretiens avec des chefs de village
   B. Deux discussions de groupe
   C. Une enquête ménage sur un échantillon aléatoire correctement dimensionné
   D. L'observation d'un point d'eau

4. Selon le jeu court du Groupe de Washington, une personne est comptée comme vivant avec un handicap si :
   A. Elle a un certificat médical
   B. Elle répond « quelques difficultés » à une question
   C. Elle répond « beaucoup de difficulté » ou « ne peut pas du tout » à au moins une des six questions
   D. Elle a plus de 60 ans

5. Dans une matrice pondérée, pourquoi fixer les poids avant de noter les options ?
   A. Pour aller plus vite
   B. Pour éviter de choisir d'abord une option et d'ajuster ensuite les critères pour la justifier
   C. Parce que le bailleur l'exige
   D. Pour obtenir des scores plus élevés

**Corrigé du quiz :** 1. B · 2. B · 3. C · 4. C · 5. B

# Semaine 2 : Concevoir un programme : chaîne de résultats, indicateurs et cibles

**Objectifs**

- Construire une chaîne de résultats cohérente, de l'activité à l'impact, et une théorie du changement dont les hypothèses sont testables.
- Choisir des indicateurs valides et mesurables, avec définition, base de référence, cible et source, et utiliser les indicateurs standard du secteur.
- Fixer des cibles réalistes à partir des données de référence et de la capacité réelle, et intégrer genre, âge et handicap dès la conception.

## La chaîne de résultats : ce qui relève de vous et ce qui n'en relève pas

Une **chaîne de résultats** relie cinq niveaux : **intrants** (argent, personnel, matériel), **activités** (former, distribuer, construire), **extrants** (produits directs : 400 latrines construites, 1 200 ménages formés), **effets** (changements de comportement ou de situation : 80 % des ménages utilisent les latrines), **impact** (changement durable de long terme : baisse des diarrhées chez les enfants de moins de cinq ans).

La frontière décisive est entre extrant et effet. L'organisation **contrôle** ses extrants : si les latrines ne sont pas construites, c'est sa responsabilité. Elle **influence** les effets, qui dépendent aussi des personnes et du contexte. Elle **contribue** à l'impact, avec beaucoup d'autres acteurs. Un rapport qui ne présente que des extrants dit ce qui a été fait, pas ce qui a changé ; un projet qui promet un impact sans effet mesurable promet ce qu'il ne pourra pas démontrer.

> Test du « et alors ? » : pour chaque extrant, demandez « et alors ? » jusqu'à atteindre un changement qui compte pour les personnes. Si la réponse s'arrête à « les gens ont reçu un kit », le programme n'a pas d'effet défini.

Chaque passage d'un niveau à l'autre repose sur des **hypothèses**. Construire des latrines conduit à leur utilisation seulement si elles sont acceptables culturellement, sûres la nuit pour les femmes, entretenues et accompagnées de sensibilisation. Ces hypothèses sont exactement ce que le suivi devra vérifier.

## De la théorie du changement au cadre logique

La **théorie du changement** explique pourquoi et comment le changement attendu se produira, avec ses hypothèses causales, les preuves qui les soutiennent et le rôle des autres acteurs. Le **cadre logique** (logframe) en est la traduction opérationnelle : une matrice de quatre colonnes (logique d'intervention, indicateurs, sources de vérification, hypothèses) et de quatre lignes (objectif global, objectif spécifique, résultats, activités).

La **logique verticale** se lit de bas en haut : si les activités sont réalisées et que les hypothèses de ce niveau se vérifient, alors les résultats sont atteints ; si les résultats sont atteints et que les hypothèses se vérifient, alors l'objectif spécifique est atteint. La **logique horizontale** se lit par ligne : chaque niveau a des indicateurs mesurables et une source qui permet effectivement de les mesurer.

Erreurs fréquentes en revue de propositions :

- Un objectif spécifique qui reformule les résultats (« les ménages ont reçu des transferts ») au lieu d'un changement (« les ménages ciblés améliorent leur consommation alimentaire »).
- Plusieurs objectifs spécifiques dans un petit projet, qui dispersent la mesure.
- Des hypothèses qui sont en réalité des risques que l'organisation devrait gérer (« le personnel est recruté à temps »). Une hypothèse est hors du contrôle du projet ; si elle est critique et très improbable, c'est une hypothèse fatale et le projet doit être reconçu.
- Des sources de vérification que personne ne produira (« statistiques nationales » qui n'existent pas au niveau du district).

## Des indicateurs qui mesurent vraiment

Un bon indicateur est **valide** (il mesure ce qu'il prétend mesurer), **fiable** (deux enquêteurs obtiennent le même résultat), **sensible** (il bouge si la situation change dans la durée du projet), **simple** à collecter et **utile** pour une décision. Le critère SMART (spécifique, mesurable, atteignable, pertinent, temporel) s'applique surtout à la cible associée.

Chaque indicateur clé doit avoir une **fiche de référence** (PIRS) : définition précise, numérateur et dénominateur, unité, désagrégation, méthode et outil de collecte, fréquence, responsable, base de référence et cible. Sans fiche, deux bureaux calculent le même indicateur de deux façons et les chiffres ne s'additionnent pas.

**Exemple de définition rigoureuse.** « Pourcentage de ménages ciblés ayant un score de consommation alimentaire acceptable (supérieur à 35) ». Numérateur : ménages de l'échantillon avec un score supérieur à 35. Dénominateur : ménages ciblés de l'échantillon. Désagrégation : sexe du chef de ménage, présence d'une personne handicapée, statut (déplacé, hôte).

Le secteur dispose d'**indicateurs standard** qu'il vaut mieux utiliser que réinventer, car ils permettent de comparer et d'agréger : score de consommation alimentaire (FCS : seuils 21 et 35 dans la plupart des contextes, 28 et 42 là où la consommation d'huile et de sucre est élevée), indice réduit des stratégies d'adaptation (rCSI), échelle de la faim des ménages (HHS), périmètre brachial (MUAC : moins de 115 mm pour la malnutrition aiguë sévère chez l'enfant de 6 à 59 mois), indicateurs de l'enquête conjointe WASH (JMP), et les registres d'indicateurs des clusters.

## Base de référence et cibles réalistes

Une cible sans **base de référence** n'a pas de sens : « 70 % des ménages avec une consommation acceptable » est un énorme progrès si l'on part de 30 %, et un recul si l'on part

de 80 %. La base de référence se mesure avant le début des activités, avec la même méthode que la mesure finale. Quand elle est impossible (urgence soudaine), on utilise des données secondaires récentes et l'on mesure dès les premières semaines, en le documentant.

Une cible réaliste se calcule, elle ne se devine pas. Elle combine : la base de référence, la **couverture** réelle du projet (part de la population concernée atteinte), l'**efficacité** attendue de l'intervention d'après les preuves disponibles, et les facteurs externes (saison, prix).

**Exemple.** Base : 40 % de ménages avec une consommation acceptable. Le projet atteint 80 % des ménages ciblés. Des projets comparables ont fait passer 50 % des ménages atteints d'une consommation insuffisante à acceptable. Parmi les 60 % de ménages en dessous du seuil, 0,8 × 0,5 = 40 % passent au-dessus : 0,6 × 0,4 = 24 points. Cible : 40 + 24 = 64 %, que l'on arrondira prudemment à 60 % si la soudure est plus dure que prévu.

> Une cible trop ambitieuse ne motive pas : elle oblige à expliquer un échec qui n'en est pas un, et pousse à embellir les données.

## Genre, âge, handicap et inclusion dès la conception

L'inclusion ne s'ajoute pas à la fin : elle se conçoit. Le **marqueur genre avec l'âge** de l'IASC (GAM) évalue les projets sur des éléments concrets : analyse des besoins différenciés, activités adaptées, participation des différents groupes à la conception et à la décision, et mécanismes de retour d'information accessibles. Un projet qui ne collecte pas de données désagrégées ne peut pas montrer que les femmes, les personnes âgées ou les personnes handicapées en bénéficient.

Concrètement : désagréger les indicateurs par sexe, âge (au moins par tranches 0-5, 6-17, 18-59, 60 et plus) et handicap ; adapter les sites (rampes, horaires, distance, files prioritaires) ; prévoir le coût de l'accessibilité dans le budget (souvent 3 à 5 % des coûts des activités concernées) ; consulter séparément femmes et hommes, jeunes et âgés.

Les normes humanitaires d'inclusion des personnes âgées et des personnes handicapées et les lignes directrices de l'IASC sur l'inclusion des personnes handicapées (2019) donnent des actions minimales : **identifier** (données désagrégées), **supprimer les obstacles** (physiques, de communication, d'attitude), **autonomiser** (participation aux décisions) et **renforcer les capacités** des organisations de personnes handicapées.

## Vérifier la conception avant de la soumettre

Avant de soumettre une proposition, une revue interne courte évite la plupart des refus et des mauvaises surprises en cours de projet. Cinq questions suffisent :

- **Cohérence** : chaque activité contribue-t-elle à un résultat, et chaque résultat à l'objectif spécifique ? Une activité sans résultat est à supprimer ; un résultat sans activité est irréaliste.
- **Mesurabilité** : chaque indicateur a-t-il une définition, une base, une cible, une source, un budget et un responsable ?

- **Proportionnalité** : le dispositif de suivi coûte-t-il entre 3 et 10 % du budget (fourchette usuelle), en rapport avec ce que le bailleur et la gestion demandent ?
- **Hypothèses** : les hypothèses critiques sont-elles suivies par un indicateur de contexte ?
- **Personnes** : les communautés ont-elles été consultées sur la modalité, le ciblage et les mécanismes de plainte ?

Le cadre logique n'est pas figé. Il se révise lorsque le contexte change, à condition de documenter la raison, de consulter le bailleur pour toute modification d'objectif ou de cible, et de garder la trace des versions. Un cadre logique qui n'a jamais bougé dans un contexte volatile est souvent un cadre logique que personne n'utilise.

## Exercice d'application : réparer un cadre logique et fixer une cible

**Situation.** Un projet WASH de 12 mois propose : objectif spécifique « 400 latrines construites » ; résultat 1 « les ménages adoptent de bonnes pratiques d'hygiène » ; indicateur de l'objectif spécifique « nombre de kits d'hygiène distribués » ; hypothèse « le personnel sera recruté à temps ». Base de référence : 35 % des ménages se lavent les mains au savon aux moments clés. Le projet atteindra 70 % des ménages de la zone ; dans des projets comparables, 40 % des ménages atteints qui n'avaient pas la pratique l'adoptent.

**Travail demandé.** 1) Identifiez les quatre erreurs. 2) Proposez un objectif spécifique et un indicateur corrects. 3) Calculez une cible réaliste pour le lavage des mains.

**Corrigé.** Erreurs : l'objectif spécifique est un extrant ; le résultat 1 est un effet placé au niveau des résultats ; l'indicateur mesure une distribution et non un changement ; l'hypothèse est un risque interne que l'organisation doit gérer. Objectif spécifique possible : « réduire les risques de maladies liées à l'eau et à l'assainissement dans les ménages ciblés », indicateur : « pourcentage de ménages ciblés qui se lavent les mains au savon aux moments clés (observation et déclaration) ». Cible : 65 % des ménages n'ont pas la pratique ; 0,7 × 0,4 = 28 % d'entre eux l'adoptent ; 0,65 × 0,28 = 18,2 points. Cible : 35 + 18 = environ 53 %.

## À retenir

- L'organisation contrôle ses extrants, influence les effets et contribue à l'impact.
- Le cadre logique se lit de bas en haut (si… alors…) et par ligne (indicateur, source, hypothèse).
- Un indicateur clé a une fiche de référence : définition, numérateur, dénominateur, désagrégation, méthode, fréquence, responsable.
- Cible = base de référence + couverture × efficacité attendue appliquée à la part de la population concernée.
- Désagréger par sexe, âge et handicap dès la conception, et budgéter l'accessibilité.

## Quiz

1. « 1 500 ménages ont reçu un kit d'hygiène » est :
   A. Un impact
   B. Un effet
   C. Un extrant
   D. Une hypothèse

2. Une hypothèse critique et très improbable dans un cadre logique est :
   A. Sans importance
   B. Une hypothèse fatale : le projet doit être reconçu
   C. Un indicateur
   D. Un risque à ignorer

3. Base : 50 % de ménages avec une consommation acceptable. Couverture : 60 % ; efficacité : 50 % des ménages atteints sous le seuil passent au-dessus. Cible calculée :
   A. 65 %
   B. 80 %
   C. 55 %
   D. 60 %

4. Selon les seuils les plus courants, un score de consommation alimentaire de 30 est :
   A. Acceptable
   B. Limite
   C. Pauvre
   D. Impossible

5. Quel élément manque le plus souvent dans un indicateur mal défini ?
   A. Un titre
   B. Un numérateur et un dénominateur précis
   C. Une couleur
   D. Une traduction

**Corrigé du quiz :** 1. C · 2. B · 3. A · 4. B · 5. B

# Semaine 3 : Ciblage et modalités : qui aider, comment, avec combien

**Objectifs**

- Choisir une méthode de ciblage adaptée au contexte et mesurer ses erreurs d'inclusion et d'exclusion.
- Comparer l'aide en nature, les coupons et les transferts monétaires sur des critères explicites, y compris les risques.
- Calculer la valeur d'un transfert à partir d'un panier de dépenses minimum et l'ajuster à l'inflation.

## Pourquoi cibler, et ce que coûte un mauvais ciblage

Avec des ressources inférieures aux besoins, le **ciblage** décide qui reçoit l'aide. Il produit deux erreurs. L'**erreur d'inclusion** : des ménages qui ne remplissent pas les critères reçoivent l'aide (gaspillage, perte de confiance). L'**erreur d'exclusion** : des ménages éligibles ne la reçoivent pas (préjudice direct, souvent pour les plus marginalisés). Réduire l'une augmente généralement l'autre : des critères très stricts excluent des ménages vulnérables ; des critères larges incluent des ménages qui n'en ont pas besoin.

**Calcul.** Sur 1 000 ménages évalués, 400 sont réellement éligibles. Le projet en sélectionne 450, dont 330 éligibles. Erreur d'inclusion : 120 / 450 = 26,7 % des ménages aidés ne sont pas éligibles. Erreur d'exclusion : (400 − 330) / 400 = 17,5 % des ménages éligibles sont laissés de côté.

> En humanitaire, l'erreur d'exclusion est la plus grave quand elle touche la survie. Une politique de ciblage doit dire explicitement laquelle des deux erreurs on accepte le plus, et pourquoi.

Le ciblage a aussi un **coût** : enquêtes, vérification, plaintes, tensions sociales. Quand plus de 70 à 80 % de la population d'une zone est dans le besoin, une distribution générale (ciblage géographique seul) est souvent plus juste et moins coûteuse qu'un ciblage ménage par ménage.

## Les méthodes de ciblage

Les méthodes se combinent souvent en plusieurs étapes :

- **Géographique** : choisir les zones les plus sévères (semaine 1). Simple et peu contestable, mais imprécis à l'intérieur d'une zone.
- **Catégoriel** : critères observables (ménage dirigé par une femme seule, personne handicapée, enfant de moins de deux ans, déplacé récent). Transparent, mais ces catégories ne recoupent qu'imparfaitement la vulnérabilité réelle.

- **Communautaire** : un comité représentatif identifie les ménages selon des critères expliqués publiquement. Utilise la connaissance locale, mais risque la capture par les élites et l'exclusion des minorités. On le contrôle par une vérification aléatoire et un mécanisme de plainte indépendant.
- **Score de vulnérabilité** (proxy means test) : un questionnaire ménage donne un score calculé à partir de variables corrélées à la pauvreté (biens, logement, composition). Plus objectif, mais coûteux, opaque pour les personnes et sensible aux fausses déclarations.
- **Auto-ciblage** : l'aide n'attire que ceux qui en ont besoin (travail contre rémunération à un salaire inférieur au marché, produits peu prisés). Bon marché, mais il peut exclure les personnes qui ne peuvent pas travailler.

Le choix dépend de quatre facteurs : la proportion de personnes dans le besoin, la disponibilité de données fiables, la cohésion sociale (risque de tensions si le ciblage paraît injuste), et le temps disponible. Dans tous les cas, les **critères doivent être publics** et compris par la communauté, et un mécanisme de plainte doit permettre de contester une exclusion.

## Nature, coupons ou transferts monétaires

Les **transferts monétaires** (cash) représentent désormais environ un cinquième de l'aide humanitaire internationale. Les données comparatives montrent qu'à résultats équivalents ils coûtent souvent moins cher à livrer que l'aide en nature, laissent les personnes choisir selon leurs priorités et soutiennent les marchés locaux. Ils ne sont pas toujours la bonne réponse.

| Critère | Nature | Coupons | Cash sans restriction |
|---|---|---|---|
| Marchés non fonctionnels | Adapté | Peu adapté | Inadapté |
| Besoin d'un produit précis (aliment thérapeutique, semences certifiées) | Adapté | Adapté | Peu adapté |
| Besoins variés selon les ménages | Peu adapté | Moyen | Adapté |
| Coût logistique | Élevé | Moyen | Faible |
| Dignité et choix | Faible | Moyen | Élevé |

Les **risques** du cash se gèrent : inflation locale (suivi des prix, plafonds de transferts), insécurité (paiement électronique, distributions fractionnées), détournement par un membre du ménage (choix du destinataire avec analyse de genre), sanctions et lutte antiterroriste (contrôle des prestataires financiers). L'idée que le cash serait détourné vers l'alcool ou le tabac est contredite par de nombreuses études : les dépenses en « biens de tentation » n'augmentent pas significativement.

Le **cash polyvalent** (MPCA) couvre plusieurs besoins de base à la fois ; il se coordonne au sein du groupe de travail cash, qui fixe une valeur de transfert commune pour éviter que les organisations se fassent concurrence.

## Calculer la valeur d'un transfert

La valeur d'un transfert part du **panier de dépenses minimum** (MEB) : le coût mensuel des biens et services dont un ménage a besoin pour couvrir ses besoins de base (alimentation, eau, hygiène, loyer, énergie, transport, communication). Il est calculé pour une taille de ménage type, à partir des prix relevés sur les marchés.

La valeur du transfert n'est pas forcément le MEB entier : on retire ce que le ménage couvre déjà lui-même (revenus, autres aides) et l'on tient compte de l'objectif du programme. **Formule de base : écart = MEB − ressources moyennes du ménage cible.**

**Exemple.** MEB mensuel pour un ménage de 6 personnes : 210 USD, dont 130 pour l'alimentation. Revenu mensuel moyen des ménages ciblés : 75 USD. Écart : 135 USD. Le groupe de travail cash a fixé une valeur commune de 80 % de l'écart pour ne pas créer de dépendance et préserver les revenus propres : 0,8 × 135 = 108 USD par mois.

**Inflation.** Si le prix du panier augmente de 12 % en trois mois, la valeur réelle d'un transfert fixe de 108 USD tombe à 108 / 1,12 = 96 USD en pouvoir d'achat initial. Les programmes prévoient donc une règle d'ajustement : par exemple, révision dès que le MEB varie de plus de 10 %, sur la base du suivi mensuel des prix.

> Taille des ménages : un montant fixe par ménage désavantage les grands ménages. Les programmes utilisent souvent des paliers (1-3, 4-6, 7 personnes et plus) avec un plafond, pour rester simples à expliquer.

## Les mécanismes de paiement et la protection des données

Le **mécanisme de livraison** se choisit après la modalité : remise d'espèces en main propre, agents de paiement, transfert par téléphone mobile, carte prépayée, compte bancaire. Les critères : couverture du réseau, accès des femmes et des personnes âgées aux téléphones, distance aux points de retrait, frais, liquidité des agents, et exigences d'identification (« connaître son client »).

Le paiement électronique réduit les risques de vol et facilite la traçabilité, mais crée des **risques sur les données** : listes de bénéficiaires partagées avec un opérateur, des autorités ou d'autres acteurs, qui peuvent exposer des personnes en danger (réfugiés, minorités). Le Manuel du CICR sur la protection des données dans l'action humanitaire recommande : minimisation des données, finalité définie, accords de partage écrits, évaluation d'impact sur la protection des données avant tout nouveau système, et durée de conservation limitée.

La **réconciliation** compare chaque mois trois listes : ménages enregistrés, paiements envoyés, paiements effectivement retirés. Les écarts (retraits non effectués, doublons, numéros de téléphone partagés par plusieurs bénéficiaires) sont des signaux d'erreur ou de fraude à traiter rapidement.

## Le suivi post-distribution

Le **suivi post-distribution** (PDM) vérifie, deux à six semaines après une distribution ou un transfert, que l'aide est bien arrivée, qu'elle est utilisée et qu'elle produit les effets attendus. Il

porte sur un échantillon aléatoire de bénéficiaires et comporte quatre blocs :

- **Processus** : montant ou quantités reçus, temps d'attente, distance, coût du trajet, sécurité, respect des personnes.
- **Utilisation** : dépenses réalisées avec le transfert, revente de biens en nature, partage.
- **Effets** : score de consommation alimentaire, stratégies d'adaptation, dette.
- **Redevabilité** : connaissance des critères et du mécanisme de plainte, satisfaction, problèmes signalés (demandes de paiement, harcèlement).

Un PDM n'a de valeur que s'il déclenche des décisions. Par exemple : si plus de 10 % des bénéficiaires déclarent avoir payé quelque chose pour être inscrits, on ouvre une enquête et l'on change le processus d'enregistrement ; si 40 % revendent l'huile reçue, on ajuste la composition du panier ou l'on passe au cash.

### Exercice d'application : évaluer un ciblage et fixer un transfert

**Situation.** Sur 2 000 ménages évalués, une vérification indépendante montre que 900 sont réellement éligibles. Le comité communautaire en a sélectionné 1 000, dont 700 éligibles. MEB mensuel : 180 USD ; revenu moyen des ménages ciblés : 60 USD ; règle commune : 75 % de l'écart. Les prix augmentent de 15 % au troisième mois.

**Travail demandé.** 1) Calculez les erreurs d'inclusion et d'exclusion. 2) Calculez la valeur du transfert. 3) Calculez sa valeur réelle après la hausse et proposez une décision.

**Corrigé.** Inclusion : 300 / 1 000 = 30 %. Exclusion : (900 − 700) / 900 = 22,2 %. Les deux erreurs sont élevées : le comité a probablement favorisé certains ménages ; il faut une vérification aléatoire, un mécanisme de plainte indépendant et une révision de la liste. Transfert : écart 120 USD ; 0,75 × 120 = 90 USD. Après la hausse : 90 / 1,15 = 78,3 USD en pouvoir d'achat initial, soit une perte de 13 %. La hausse dépasse le seuil usuel de 10 % : on propose au groupe de travail cash une révision à 90 × 1,15 = 103,5 USD, en vérifiant d'abord que les transferts eux-mêmes ne sont pas la cause de l'inflation locale.

## À retenir

- Erreur d'inclusion = non éligibles aidés / total aidé ; erreur d'exclusion = éligibles non aidés / total des éligibles.
- Le ciblage communautaire doit être contrôlé par une vérification aléatoire et un mécanisme de plainte indépendant.
- Le choix nature, coupons ou cash dépend d'abord du fonctionnement des marchés et du besoin.
- Valeur du transfert = part convenue de (MEB − ressources du ménage), ajustée à l'inflation selon une règle écrite.
- Le suivi post-distribution doit déclencher des décisions, pas seulement produire un rapport.

## Quiz

1. 500 ménages aidés, dont 400 éligibles ; 600 ménages éligibles au total. L'erreur d'exclusion est :
   A. 20 %
   B. 33,3 %
   C. 16,7 %
   D. 25 %

2. Dans une zone où 85 % de la population est dans le besoin, la méthode la plus juste et la moins coûteuse est souvent :
   A. Un score de vulnérabilité ménage par ménage
   B. Une distribution générale avec ciblage géographique
   C. Un tirage au sort
   D. Le ciblage par les autorités

3. Un transfert de 100 USD après une hausse des prix de 25 % vaut en pouvoir d'achat initial :
   A. 75 USD
   B. 80 USD
   C. 125 USD
   D. 100 USD

4. Quelle condition rend les transferts monétaires inadaptés ?
   A. Des ménages aux besoins variés
   B. Des marchés qui ne fonctionnent pas ou ne peuvent pas répondre à la demande
   C. Une forte présence du téléphone mobile
   D. Des personnes qui préfèrent le cash

5. La réconciliation mensuelle des paiements compare :
   A. Le budget et les dépenses du siège
   B. Les ménages enregistrés, les paiements envoyés et les paiements retirés
   C. Les prix de deux marchés
   D. Les rapports de deux bailleurs

**Corrigé du quiz :** 1. B · 2. B · 3. B · 4. B · 5. B

# Semaine 4 : Planifier et budgéter un programme : du plan de travail au rythme de dépense

**Objectifs**

- Construire un plan de travail réaliste : découpage des tâches, dépendances, chemin critique et marges.
- Relier le budget aux activités par des coûts unitaires, et calculer coût par personne et coût par résultat.
- Suivre l'exécution avec le rythme de dépense et détecter tôt les sous-consommations et les dépassements.

## Du cadre logique au plan de travail

Le cadre logique dit quoi ; le **plan de travail** dit qui fait quoi, quand et avec quelles ressources. On part des activités du cadre logique et on les découpe en tâches de 1 à 4 semaines (structure de découpage du travail) : pour « distribuer des transferts à 3 000 ménages », il faut enregistrer, vérifier, contracter un prestataire de paiement, former les équipes, informer les communautés, payer, faire le suivi post-distribution.

Chaque tâche reçoit une **durée**, un **responsable** unique et ses **dépendances** : la plupart sont du type « fin-début » (le paiement ne peut commencer qu'après la signature du contrat). Le diagramme de Gantt montre ces tâches dans le temps, mais il ne suffit pas : il faut identifier le chemin critique.

Le plan de travail intègre aussi les **contraintes du contexte** : saison des pluies qui coupe les routes, période de soudure, fêtes religieuses, élections, calendrier agricole, et délais administratifs (accord des autorités, dédouanement, recrutement). Dans de nombreux contextes, le recrutement d'un poste clé prend 8 à 12 semaines et l'importation de biens 3 à 6 mois ; un plan qui l'ignore est faux dès le premier jour.

> Le premier trimestre d'un projet est presque toujours en retard. Planifier les dépenses de façon linéaire (un douzième par mois) garantit un écart important dès le troisième mois.

## Le chemin critique et les marges

Le **chemin critique** est la plus longue suite de tâches dépendantes ; sa durée est la durée minimale du projet. Toute tâche du chemin critique qui prend du retard retarde le projet entier. Les autres tâches ont une **marge** : elles peuvent glisser sans conséquence, dans la limite de cette marge.

**Exemple.** A : évaluation (3 semaines). B : appel d'offres pour le prestataire de paiement (6 semaines, après A). C : enregistrement des ménages (4 semaines, après A). D : vérification des listes (2 semaines, après C). E : premier paiement (1 semaine, après B et D).

- Chemin A-B-E : 3 + 6 + 1 = 10 semaines.
- Chemin A-C-D-E : 3 + 4 + 2 + 1 = 10 semaines.

Les deux chemins durent 10 semaines : les deux sont critiques et aucune tâche n'a de marge. Un retard d'une semaine sur l'enregistrement ou sur l'appel d'offres repousse le premier paiement. Si l'appel d'offres durait 4 semaines, le chemin A-B-E ferait 8 semaines et la tâche B aurait 2 semaines de marge.

Pour **raccourcir** un projet, on n'agit que sur le chemin critique : accélérer une tâche qui a de la marge ne change rien. Deux techniques : faire en parallèle ce qui était en série (commencer l'enregistrement pendant la fin de l'évaluation, avec un risque de reprise) et ajouter des ressources sur une tâche critique (deux équipes d'enregistrement au lieu d'une).

## Le budget de programme par coûts unitaires

Un budget solide se construit **à partir des activités**, avec des coûts unitaires vérifiables : quantité × coût unitaire × durée ou fréquence. Chaque ligne a une justification (devis, prix de marché, grille salariale, historique). Les bailleurs rejettent surtout les budgets où les montants sont ronds, sans base de calcul, ou sans lien avec le cadre logique.

Le budget distingue habituellement : **coûts directs des activités** (transferts, intrants, formations), **personnel** (avec la part de temps imputée au projet), **coûts de fonctionnement du bureau** partagés entre projets selon une clé de répartition documentée, **suivi et évaluation**, et **coûts indirects** (pourcentage fixé par le bailleur, souvent 7 à 10 %).

**Coûts partagés.** Un bureau coûte 20 000 USD par mois et sert trois projets. La clé la plus défendable est souvent le temps du personnel ou la part du budget : si le projet représente 40 % des dépenses directes du bureau, il prend 8 000 USD par mois. La clé doit être écrite, appliquée de façon constante, et vérifiable par un auditeur.

> Un budget n'est pas un plafond à atteindre : c'est une prévision. Dépenser en fin de projet pour « consommer » le budget est une faute de gestion qui se voit dans les audits.

## Coût par personne et coût par résultat

Deux indicateurs de coût éclairent les décisions, à condition de bien les définir. Le **coût par personne atteinte** divise le coût total par le nombre de personnes aidées. Il est simple mais trompeur : il favorise les programmes qui donnent peu à beaucoup de monde et ne dit rien de l'effet.

Le **coût par résultat** divise le coût par l'effet obtenu : coût par enfant guéri de la malnutrition aiguë sévère, coût par ménage passant d'une consommation insuffisante à acceptable, coût par point d'eau fonctionnel après un an. Il permet de comparer des approches.

**Exemple.** Projet A : 600 000 USD, 4 000 enfants admis, taux de guérison 70 % : coût par enfant guéri = 600 000 / 2 800 = 214 USD. Projet B : 450 000 USD, 3 500 enfants admis, taux de guérison 55 % : 450 000 / 1 925 = 234 USD. Le projet B paraît moins cher par enfant admis (129 contre 150 USD), mais il est plus cher par enfant guéri.

La **part des coûts qui atteint directement les personnes** (ratio de transfert, ou « cost-transfer ratio » pour le cash) mesure l'efficience de la livraison : pour un programme de cash, montant transféré / coût total. Un ratio de 0,80 signifie que 80 cents sur chaque dollar arrivent aux ménages. Il se compare entre programmes similaires, jamais entre modalités ou contextes très différents.

## Suivre l'exécution : le rythme de dépense

Le **rythme de dépense** (burn rate) compare les dépenses réelles cumulées aux dépenses prévues dans le plan de trésorerie, et non à une ligne droite. Il se suit chaque mois, ligne par ligne, avec le responsable financier et le chef de projet ensemble.

**Exemple.** Budget de 12 mois : 900 000 USD. Fin du mois 6 : dépenses prévues cumulées 420 000 USD (le plan tenait compte d'un démarrage lent), dépenses réelles 290 000 USD. Taux d'exécution : 290 000 / 420 000 = 69 %. Au rythme moyen actuel (290 000 / 6 = 48 333 USD par mois), il resterait 610 000 USD à dépenser en 6 mois, soit 101 667 USD par mois : plus du double du rythme actuel. Il faut agir maintenant.

Les causes d'une sous-consommation sont presque toujours opérationnelles : retards d'accès, de recrutement ou d'achats, activités reportées. La réponse suit l'ordre suivant : comprendre les causes ligne par ligne ; accélérer ce qui peut l'être sans perte de qualité ; proposer au bailleur une révision budgétaire ou une extension sans coût **avant** les derniers mois (la plupart des bailleurs exigent une demande au moins 30 à 60 jours avant la fin) ; ne jamais faire d'achats de fin de projet inutiles.

À l'inverse, un dépassement sur une ligne se compense dans la limite de **flexibilité** prévue au contrat (souvent 10 à 20 % entre lignes ou rubriques sans accord préalable). Au-delà, l'accord écrit du bailleur est obligatoire, sinon la dépense risque d'être déclarée inéligible.

## Ressources humaines du programme et organisation

Le plan de dotation découle du plan de travail : combien d'enquêteurs pour enregistrer 3 000 ménages en 4 semaines ? Si un enquêteur réalise 12 entretiens par jour sur 20 jours ouvrés par mois, une équipe de 13 enquêteurs fait 3 120 entretiens en 4 semaines. Le calcul évite à la fois les équipes trop petites (retards) et trop grandes (coûts, qualité).

L'**organigramme du programme** clarifie la responsabilité de chaque résultat et la séparation des fonctions : la personne qui sélectionne les bénéficiaires ne doit pas être celle qui valide les paiements ni celle qui traite les plaintes. La matrice **RACI** (réalise, approuve, consulté, informé) par processus clé (enregistrement, paiement, suivi, plaintes) rend ces séparations visibles et contrôlables.

Enfin, le chef de programme organise un **rythme de pilotage** stable : point hebdomadaire d'équipe sur le plan de travail ; revue mensuelle des indicateurs, du budget et des risques ; revue trimestrielle avec la direction pays, qui peut décider de réallocations. Un programme sans rythme de pilotage découvre ses problèmes dans le rapport final.

### Exercice d'application : chemin critique et rythme de dépense

**Situation.** Tâches : A, évaluation (2 semaines) ; B, recrutement des équipes (5 semaines, après A) ; C, achat des kits (8 semaines, après A) ; D, formation des équipes (1 semaine, après B) ; E, distribution (2 semaines, après C et D). Budget total : 600 000 USD sur 10 mois. À la fin du mois 5, dépenses prévues cumulées : 260 000 USD ; dépenses réelles : 182 000 USD.

**Travail demandé.** 1) Calculez la durée minimale et le chemin critique. 2) Quelle est la marge du recrutement ? 3) Calculez le taux d'exécution et le rythme mensuel nécessaire pour consommer le budget restant.

**Corrigé.** A-C-E : 2 + 8 + 2 = 12 semaines. A-B-D-E : 2 + 5 + 1 + 2 = 10 semaines. Chemin critique : A-C-E, 12 semaines ; le recrutement a 2 semaines de marge. Pour avancer la distribution, il faut agir sur l'achat des kits (lancer l'appel d'offres pendant l'évaluation, ou utiliser un accord-cadre). Exécution : 182 000 / 260 000 = 70 %. Reste : 418 000 USD en 5 mois, soit 83 600 USD par mois contre 36 400 en moyenne jusqu'ici. Il faut identifier les lignes en retard et discuter dès maintenant avec le bailleur d'une révision ou d'une extension sans coût.

### À retenir

- Seul le chemin critique détermine la durée du projet ; accélérer une tâche qui a de la marge ne sert à rien.
- Un budget se construit par activités : quantité × coût unitaire × durée, avec une justification par ligne.
- Le coût par résultat compare mieux les approches que le coût par personne atteinte.
- Le rythme de dépense se compare au plan de trésorerie, pas à une ligne droite, et se suit chaque mois.
- Révision budgétaire ou extension sans coût : à demander tôt, jamais par des achats de fin de projet.

## Quiz

1. A (2 sem.) puis B (4 sem.) et C (3 sem.) en parallèle après A, puis D (1 sem.) après B et C. Durée minimale du projet :
   A. 10 semaines
   B. 7 semaines
   C. 6 semaines
   D. 9 semaines

2. Projet de 300 000 USD, 1 500 enfants admis, taux de guérison 80 %. Coût par enfant guéri :
   A. 200 USD
   B. 250 USD
   C. 240 USD
   D. 160 USD

3. Pour raccourcir un projet, on agit d'abord sur :
   A. Les tâches qui ont le plus de marge
   B. Les tâches du chemin critique
   C. Le budget des coûts indirects
   D. Le rapport final

4. Un programme cash transfère 720 000 USD pour un coût total de 900 000 USD. Le ratio de transfert est :
   A. 0,80
   B. 1,25
   C. 0,72
   D. 0,20

5. À quatre mois de la fin, 40 % du budget n'est pas dépensé à cause de retards d'accès. La meilleure réponse est :
   A. Acheter du matériel pour consommer le budget
   B. Analyser les causes ligne par ligne et demander tôt au bailleur une révision ou une extension sans coût
   C. Ne rien dire au bailleur
   D. Rendre l'argent sans explication

**Corrigé du quiz :** 1. B · 2. B · 3. B · 4. A · 5. B

# Semaine 5 : Le système de suivi : plan MEAL, qualité des données et suivi à distance

**Objectifs**

- Construire un plan MEAL complet et proportionné, du formulaire de collecte jusqu'à la décision de gestion.
- Évaluer et améliorer la qualité des données selon cinq dimensions, et organiser des contrôles qualité réguliers.
- Organiser le suivi dans les zones inaccessibles : suivi à distance, tiers de suivi et triangulation.

## Le plan MEAL : la colonne vertébrale du suivi

Le **plan MEAL** décrit comment chaque indicateur sera mesuré, par qui, quand, et pour quelle décision. Il se présente sous forme de tableau : indicateur, définition, base, cible, source et méthode, fréquence, désagrégation, responsable de la collecte, responsable de l'analyse, utilisation (revue mensuelle, rapport bailleur, ajustement du ciblage).

Il couvre quatre types de suivi :

- **Suivi des activités et des extrants** : ce qui a été réalisé, par rapport au plan (registres de distribution, listes de présence, fiches de chantier).
- **Suivi des effets** : changements chez les personnes (enquêtes de référence, intermédiaires et finales ; suivi post-distribution).
- **Suivi du contexte** : prix, accès, sécurité, mouvements de population, et indicateurs des hypothèses critiques du cadre logique.
- **Suivi de la redevabilité** : plaintes et retours reçus, délai de réponse, satisfaction, participation (semaine 8).

Le coût du plan MEAL doit être dans le budget : personnel, outils, transport, enquêtes, évaluations. Une fourchette de 3 à 10 % du budget du projet est habituelle ; un projet pilote dont on veut démontrer l'effet peut justifier davantage.

> Le meilleur test d'un plan MEAL : pour chaque indicateur, pouvez-vous nommer la réunion où il sera discuté et la décision qu'il peut déclencher ? Sinon, demandez-vous s'il faut le collecter.

## Des outils de collecte au tableau de bord

La chaîne de données a cinq maillons : **collecte**, **saisie et stockage**, **nettoyage**, **analyse**, **diffusion**. La collecte mobile (KoboToolbox, ODK, CommCare) a remplacé le papier dans la plupart des contextes : contrôles automatiques à la saisie (plages de valeurs, champs obligatoires, sauts logiques), géolocalisation et horodatage, envoi direct au serveur.

Un bon formulaire se conçoit à partir du plan d'analyse : on écrit d'abord les tableaux que l'on veut produire, puis les questions nécessaires et seulement elles. Il se teste sur le terrain avec de vrais répondants avant d'être déployé, puis on fige une version ; toute modification ultérieure est documentée, car elle rend les données avant et après incomparables.

Le **nettoyage** suit des règles écrites : doublons, valeurs aberrantes (un ménage de 45 personnes, une dépense alimentaire de 2 000 USD par semaine), incohérences (un enfant de 3 ans marié), entretiens trop courts (moins de 10 minutes pour un questionnaire de 40 minutes). On ne corrige jamais une valeur sans source ; on la signale, on vérifie avec l'enquêteur ou on la retire, et l'on garde un journal de nettoyage.

Le **tableau de bord** montre peu d'indicateurs, avec la cible, la tendance et un code couleur selon des seuils fixés à l'avance. Il est mis à jour au rythme des décisions : hebdomadaire pour les distributions, mensuel pour les indicateurs d'effet intermédiaire.

## La qualité des données : cinq dimensions

Une **évaluation de la qualité des données** (DQA) vérifie cinq dimensions :

| Dimension | Question | Exemple de problème |
|---|---|---|
| Validité | Les données mesurent-elles ce que l'indicateur prévoit ? | On compte les personnes formées au lieu de celles qui appliquent la pratique. |
| Fiabilité | Les méthodes sont-elles stables entre lieux, personnes et périodes ? | Deux bureaux calculent le taux de guérison différemment. |
| Précision | La marge d'erreur est-elle assez faible pour la décision ? | Un échantillon de 50 ménages pour suivre une variation de 5 points. |
| Ponctualité | Les données arrivent-elles à temps pour décider ? | Le rapport mensuel arrive six semaines après la fin du mois. |
| Intégrité | Les données sont-elles protégées contre les erreurs et les manipulations ? | Un même agent saisit, valide et corrige les listes de paiement. |

La méthode la plus concrète est la **vérification par recomptage** : on choisit un échantillon de sites, on recompte à partir des sources primaires (registres, fiches) et l'on compare avec les chiffres rapportés. Un **ratio de vérification** de 1,00 est parfait ; on fixe une tolérance, par exemple entre 0,95 et 1,05. Au-delà de 1,05, les chiffres rapportés sont gonflés ; en dessous de 0,95, des activités ne sont pas rapportées.

**Exemple.** Un centre de santé rapporte 1 240 consultations ; le recomptage des registres en trouve 1 085. Ratio : 1 240 / 1 085 = 1,14. Le centre surdéclare de 14 % : il faut comprendre la cause (double comptage des visites de suivi, pression sur les cibles) avant toute sanction.

## Les contrôles qualité de la collecte

La qualité se joue pendant la collecte, pas après. Les contrôles standard :

- **Formation et test** des enquêteurs, avec un test écrit et des entretiens simulés ; les enquêteurs qui échouent ne collectent pas.
- **Supervision directe** d'une partie des entretiens (par exemple 5 %).
- **Rappels de contrôle** (back-checks) : 10 % des ménages sont réinterrogés sur quelques questions stables (taille du ménage, type d'abri). Un taux de concordance faible pour un enquêteur signale des erreurs ou des entretiens inventés.
- **Contrôles à distance quotidiens** : durée des entretiens, coordonnées GPS (entretiens tous au même endroit), heure (entretiens à 3 heures du matin), distribution des réponses par enquêteur (un enquêteur dont tous les ménages ont exactement 5 membres).

Les données inventées (« curbstoning ») sont plus fréquentes qu'on ne le croit quand les enquêteurs sont payés à l'entretien, sous pression de quotas, ou envoyés dans des zones dangereuses. Les prévenir passe par des objectifs réalistes, une paie journalière plutôt qu'à l'unité, et des contrôles annoncés à l'avance.

> Un responsable MEAL qui ne trouve jamais d'erreur ne cherche pas assez. Publier en interne les erreurs trouvées et corrigées renforce la confiance dans les chiffres.

## Le suivi à distance et les tiers de suivi

Dans les zones inaccessibles au personnel (insécurité, restrictions), le **suivi à distance** combine plusieurs sources pour compenser l'absence de visites :

- **Partenaires locaux** ou personnel local qui ont accès, avec des outils standardisés et des preuves (photos géolocalisées, listes signées).
- **Tiers de suivi** (TPM) : une société ou une organisation indépendante, souvent exigée par les bailleurs, qui vérifie un échantillon d'activités.
- **Appels téléphoniques** à un échantillon aléatoire de bénéficiaires (suivi post-distribution par téléphone).
- **Données à distance** : images satellites (camps, cultures, destructions), données de transactions des opérateurs mobiles, données de prix.
- **Retours de la communauté** par ligne gratuite et messagerie.

Le suivi à distance a des **limites** : échantillons par téléphone biaisés vers les ménages équipés et souvent masculins, risques de pression sur les partenaires, transfert des risques de sécurité vers le personnel local. Il faut le dire dans les rapports, payer correctement les partenaires pour ce travail, et ne pas leur demander de prendre des risques que l'organisation refuserait pour son propre personnel.

La **triangulation** est la règle : une distribution est considérée vérifiée quand au moins deux sources indépendantes concordent (par exemple liste signée et appels à un échantillon de bénéficiaires).

## Rendre le suivi utile : de la donnée à la décision

Le suivi échoue rarement par manque de données ; il échoue parce que personne ne les utilise. Trois mécanismes rendent le suivi utile :

- **Des seuils d'alerte** fixés à l'avance : si le temps d'attente aux distributions dépasse 3 heures, si le taux de plaintes non résolues dépasse 20 %, si le prix du panier augmente de plus de 10 %, une action prédéfinie est déclenchée.
- **Des revues régulières** où l'équipe programme, le MEAL, la logistique et la finance regardent ensemble les mêmes chiffres et décident. Le compte rendu note les décisions, leur responsable et leur échéance.
- **Un retour aux personnes et aux équipes** : présenter aux communautés ce qui a été appris et changé ; montrer aux enquêteurs à quoi leurs données ont servi.

L'indépendance du MEAL compte. Si le responsable MEAL rend compte uniquement au chef de projet dont il mesure les résultats, la tentation de lisser les mauvais chiffres est forte. Une double ligne hiérarchique (opérationnelle au projet, technique à un responsable MEAL pays ou siège) protège la qualité de l'information.

### Exercice d'application : vérifier la qualité des données de trois sites

**Situation.** Un projet nutrition rapporte, pour un mois, les admissions de trois sites. Site 1 : rapporté 320, recompté 312. Site 2 : rapporté 410, recompté 342. Site 3 : rapporté 198, recompté 205. Tolérance : ratio entre 0,95 et 1,05. Les rappels de contrôle sur l'enquête de couverture montrent un taux de concordance de 96 % pour tous les enquêteurs sauf un (61 %), dont les entretiens durent en moyenne 9 minutes contre 35 pour les autres.

**Travail demandé.** 1) Calculez les ratios de vérification et interprétez-les. 2) Que faites-vous des données de l'enquêteur suspect ? 3) Quelles mesures de prévention proposez-vous ?

**Corrigé.** Site 1 : 320 / 312 = 1,03, dans la tolérance. Site 2 : 410 / 342 = 1,20 : surdéclaration de 20 %, à investiguer (double comptage, pression sur les cibles, fraude). Site 3 : 198 / 205 = 0,97, dans la tolérance (légère sous-déclaration). Total rapporté 928 contre 859 recomptés : il faut corriger le rapport mensuel et en informer le bailleur. Les données de l'enquêteur suspect sont mises à part, une partie de ses ménages est réinterrogée ; si la falsification est confirmée, ses entretiens sont refaits ou retirés, et la cause est traitée (quotas, paie à l'entretien). Prévention : paie journalière, contrôles quotidiens de durée et de GPS, rappels de contrôle annoncés, objectifs réalistes.

## À retenir

- Le plan MEAL relie chaque indicateur à une méthode, un responsable, une fréquence et une décision.
- Cinq dimensions de la qualité des données : validité, fiabilité, précision, ponctualité, intégrité.
- Ratio de vérification = chiffre rapporté / chiffre recompté ; au-delà de la tolérance, on investigue avant de sanctionner.
- Rappels de contrôle, durée des entretiens et GPS détectent les données inventées.
- En suivi à distance, une activité est vérifiée quand deux sources indépendantes concordent.

## Quiz

1. Un site rapporte 500 distributions, le recomptage en trouve 400. Le ratio de vérification est :
   A. 0,80
   B. 1,25
   C. 1,20
   D. 0,90

2. Deux bureaux calculent différemment le même indicateur. Quelle dimension de la qualité est en cause ?
   A. Ponctualité
   B. Fiabilité
   C. Intégrité
   D. Précision

3. Un enquêteur réalise des entretiens de 8 minutes pour un questionnaire de 35 minutes, tous au même point GPS. Le plus probable :
   A. Il est très efficace
   B. Des entretiens inventés ou bâclés
   C. Une panne de serveur
   D. Un échantillon trop grand

4. Pourquoi éviter de payer les enquêteurs à l'entretien ?
   A. C'est plus cher
   B. Cela incite à bâcler ou inventer des entretiens
   C. Les bailleurs l'interdisent toujours
   D. Cela réduit la taille de l'échantillon

5. Quelle est la principale limite du suivi post-distribution par téléphone ?
   A. Il est trop cher
   B. L'échantillon est biaisé vers les ménages qui ont un téléphone, souvent représentés par des hommes
   C. Il est interdit
   D. Il ne permet aucune question

**Corrigé du quiz :** 1. B · 2. B · 3. B · 4. B · 5. B

# Semaine 6 : Enquêtes et échantillonnage : obtenir des chiffres justes

**Objectifs**

- Choisir un plan d'échantillonnage (aléatoire simple, systématique, stratifié, en grappes) adapté à la question et au terrain.
- Calculer une taille d'échantillon avec marge d'erreur, correction pour population finie, effet de plan et non-réponses.
- Concevoir un questionnaire fiable et interpréter correctement une proportion et son intervalle de confiance.

## Population, échantillon et représentativité

Une enquête cherche à connaître une **population** (tous les ménages déplacés d'un district) à partir d'un **échantillon**. Les résultats ne sont généralisables que si chaque unité de la population a une **probabilité connue et non nulle** d'être tirée : c'est l'échantillonnage probabiliste. Un échantillon « de convenance » (les ménages proches de la route, ceux présents à la distribution) peut être grand, il reste biaisé.

Il faut d'abord une **base de sondage** : liste des ménages, des sites, des villages, avec leur population. Sans liste fiable, on travaille par étapes : on tire des villages ou des blocs d'un camp, puis des ménages à l'intérieur, en les dénombrant sur place.

Les principaux plans :

- **Aléatoire simple** : tirage au sort dans la liste complète. Le plus simple statistiquement, souvent impraticable sur le terrain.
- **Systématique** : on calcule un pas (population / taille d'échantillon), on tire un point de départ au hasard, puis on prend un ménage tous les « pas ». Pour 2 400 abris et 240 ménages à interroger, pas = 10 ; départ tiré entre 1 et 10, par exemple 7 : abris 7, 17, 27…
- **Stratifié** : la population est divisée en groupes (déplacés et hôtes, zones urbaines et rurales) et l'on tire un échantillon dans chaque groupe. Indispensable si l'on veut des résultats fiables pour chaque groupe.
- **En grappes** : on tire d'abord des grappes (villages), puis des ménages dans chaque grappe. Moins coûteux en déplacements, mais moins précis : il faut un échantillon plus grand.

## Calculer la taille d'échantillon

Pour estimer une proportion, la formule de base est : $\mathbf{n = z^2 \times p \times (1 - p) / e^2}$, où z vaut 1,96 pour un niveau de confiance de 95 %, p est la proportion attendue (0,5 quand on ne sait pas, car c'est le cas le plus exigeant) et e la marge d'erreur.

- Marge 5 % : $1{,}96^2 \times 0{,}25 / 0{,}0025 = 384$ ménages.

- Marge 7 % : 196 ménages.
- Marge 10 % : 96 ménages.

Diviser la marge par deux multiplie l'échantillon par quatre. Le choix de la marge dépend de la décision : pour suivre une variation de 10 points entre deux enquêtes, une marge de 10 % ne suffit pas.

**Correction pour population finie.** Quand la population est petite, l'échantillon nécessaire diminue : $n = n_0 / (1 + (n_0 - 1) / N)$. Pour N = 2 000 ménages : 384 / (1 + 383 / 2 000) = 384 / 1,19 = 323 ménages. Pour N = 100 000, la correction est négligeable.

**Effet de plan.** En échantillonnage en grappes, les ménages d'un même village se ressemblent, ce qui réduit l'information apportée par chacun. On multiplie l'échantillon par l'effet de plan : **DEFF = 1 + (m − 1) × ICC**, où m est le nombre de ménages par grappe et ICC le coefficient de corrélation intra-grappe. Avec 10 ménages par grappe et un ICC de 0,05 : DEFF = 1 + 9 × 0,05 = 1,45. Faute de données, on utilise souvent 1,5 à 2.

**Non-réponses.** On divise par le taux de réponse attendu : avec 10 % de non-réponses, n / 0,9.

> Calcul complet : marge 5 %, p = 0,5, DEFF 1,5, non-réponses 10 % : 384 × 1,5 = 576 ; 576 / 0,9 = 640 ménages, par exemple 64 grappes de 10 ménages.

## Stratification et comparaisons

Si l'on veut des résultats pour chaque strate avec la même précision, il faut la taille d'échantillon **dans chaque strate**. Pour trois districts avec une marge de 5 % chacun, l'échantillon est trois fois plus grand qu'une estimation globale. C'est souvent la principale cause des enquêtes trop chères : on promet des résultats par district, par sexe du chef de ménage et par statut, sans calculer l'échantillon nécessaire.

Quand les strates ont des tailles différentes et que l'on veut un résultat global, on **pondère** : chaque strate compte selon sa part dans la population et non selon sa part dans l'échantillon. **Exemple.** Déplacés : 20 000 ménages, échantillon 400, 60 % en insécurité alimentaire. Hôtes : 80 000 ménages, échantillon 400, 30 % en insécurité alimentaire. Moyenne non pondérée : 45 %. Moyenne pondérée : 0,2 × 60 + 0,8 × 30 = 36 %. Présenter 45 % surestimerait fortement la situation.

Pour **comparer deux enquêtes** (référence et finale) ou deux groupes, il faut un échantillon capable de détecter la différence attendue : détecter un passage de 40 % à 50 % avec 95 % de confiance et 80 % de puissance demande environ 385 ménages par groupe en échantillonnage aléatoire simple, davantage en grappes. Un petit changement exige un grand échantillon.

## Concevoir un questionnaire fiable

Un questionnaire se construit à partir des indicateurs et du plan d'analyse. Règles essentielles :

- **Une idée par question** : « Avez-vous accès à l'eau et au savon ? » est deux questions.
- **Période de rappel définie** : « Au cours des 7 derniers jours » pour l'alimentation, « au cours des 30 derniers jours » pour les dépenses. Une période trop longue produit des

erreurs de mémoire.

- **Modules standard** pour les indicateurs standard (score de consommation alimentaire, rCSI, groupe de Washington), sans les modifier, pour rester comparable.
- **Formulation neutre** : « Que pensez-vous de la distribution ? » plutôt que « Êtes-vous satisfait de la distribution ? ».
- **Ordre** : questions simples au début, sensibles à la fin, et jamais de question sensible sans formation des enquêteurs et sans possibilité de référencement.
- **Traduction et rétrotraduction** dans les langues locales, puis test sur le terrain.
- **Durée** : au-delà de 45 minutes, la qualité des réponses diminue ; supprimer ce qui n'est pas indispensable.

> Le consentement éclairé est lu à chaque répondant : objet de l'enquête, durée, caractère volontaire, absence de lien avec l'aide, confidentialité. Un refus est respecté et remplacé selon la règle prévue, jamais par le voisin le plus pratique.

## Interpréter une proportion et son intervalle de confiance

Un résultat d'enquête est une **estimation**, entourée d'une incertitude. L'intervalle de confiance à 95 % d'une proportion est approximativement : $p \pm 1{,}96 \times \sqrt{p \times (1 - p) / n}$, multiplié par √DEFF en grappes.

**Exemple.** 384 ménages en échantillonnage aléatoire simple, 45 % sans accès suffisant à l'eau. Erreur type : $\sqrt{0{,}45 \times 0{,}55 / 384} = 0{,}0254$. Marge : $1{,}96 \times 0{,}0254 = 5{,}0$ points. Intervalle : 40 % à 50 %.

Conséquences pratiques :

- Si la référence était 45 % (intervalle 40-50 %) et la finale 49 % (intervalle 44-54 %), on ne peut pas affirmer que la situation s'est dégradée.
- Les résultats par sous-groupe ont des intervalles plus larges, car les sous-groupes sont plus petits : 45 % sur 80 ménages dirigés par une femme a une marge d'environ 11 points.
- On présente toujours la taille d'échantillon, et l'intervalle pour les indicateurs clés.

La **signification statistique** ne veut pas dire importance pratique : avec un très grand échantillon, une différence de 1 point peut être significative sans changer aucune décision. Inversement, une différence importante peut ne pas être significative avec un petit échantillon : on dit alors que les données ne permettent pas de conclure, pas qu'il n'y a pas d'effet.

## Organiser une enquête sur le terrain

Une enquête ménage se prépare en semaines, pas en jours. Le calendrier type :

- Semaine 1 : objectifs, indicateurs, plan d'analyse, plan d'échantillonnage, accord des autorités.
- Semaine 2 : questionnaire, traduction, programmation sur tablette, tests.

- Semaine 3 : recrutement et formation des enquêteurs (3 à 5 jours, avec test et pratique), test pilote, corrections.
- Semaines 4 et 5 : collecte, avec contrôles qualité quotidiens et réunions de débriefing.
- Semaine 6 : nettoyage, pondération, analyse, validation avec l'équipe et les partenaires.
- Semaine 7 : rapport et restitution, y compris aux communautés.

**Logistique.** Nombre d'équipes = taille d'échantillon / (entretiens par équipe et par jour × jours de collecte). Pour 640 ménages, 4 entretiens par enquêteur et par jour, 10 jours : 640 / 40 = 16 enquêteurs, soit 4 équipes de 4 avec un superviseur chacune. On prévoit 10 à 20 % de jours de réserve pour les imprévus (pluie, sécurité, accès).

Le **remplacement** des ménages absents ou qui refusent suit une règle écrite (revenir une fois, puis remplacer par le ménage suivant dans la liste tirée), sans quoi les enquêteurs remplacent les ménages éloignés par des ménages proches, ce qui biaise l'échantillon.

## Exercice d'application : dimensionner une enquête de référence

**Situation.** Un projet veut connaître, avec une marge de 5 % et 95 % de confiance, la proportion de ménages en insécurité alimentaire, séparément pour les déplacés (3 000 ménages) et pour les hôtes (40 000 ménages). Échantillonnage en grappes de 12 ménages, ICC estimé à 0,06. Non-réponses attendues : 8 %.

**Travail demandé.** 1) Calculez l'échantillon de base avec la correction pour population finie pour chaque strate. 2) Calculez l'effet de plan et l'échantillon final par strate. 3) Combien de grappes faut-il au total ?

**Corrigé.** Base : 384. Déplacés : 384 / (1 + 383 / 3 000) = 384 / 1,128 = 340. Hôtes : 384 / (1 + 383 / 40 000) = 384 / 1,0096 = 380. Effet de plan : 1 + 11 × 0,06 = 1,66. Déplacés : 340 × 1,66 = 565 ; 565 / 0,92 = 614. Hôtes : 380 × 1,66 = 631 ; 631 / 0,92 = 686. Total : 1 300 ménages. Grappes : 614 / 12 = 51,2 soit 52 grappes ; 686 / 12 = 57,2 soit 58 grappes ; total 110 grappes. Pour un résultat global, les deux strates seront pondérées selon leur poids dans la population (7 % et 93 %).

## À retenir

- $n = z^2 \times p \times (1 - p) / e^2$ : 384 pour une marge de 5 % à 95 % de confiance avec p = 0,5.
- Correction pour petite population : $n = n_0 / (1 + (n_0 - 1) / N)$.
- Effet de plan en grappes : $1 + (m - 1) \times ICC$ ; puis division par le taux de réponse.
- Un résultat par strate exige l'échantillon dans chaque strate ; un résultat global exige une pondération.
- Toujours présenter la taille d'échantillon et l'intervalle de confiance ; non significatif ne veut pas dire absence d'effet.

## Quiz

1. Avec p = 0,5, 95 % de confiance et une marge de 10 %, la taille d'échantillon aléatoire simple est d'environ :
   A. 38
   B. 96
   C. 196
   D. 384

2. Grappes de 20 ménages, ICC de 0,04. L'effet de plan est :
   A. 1,04
   B. 1,76
   C. 1,80
   D. 0,80

3. Strate A : 10 000 ménages, 50 % touchés. Strate B : 30 000 ménages, 20 % touchés. La proportion globale pondérée est :
   A. 35 %
   B. 27,5 %
   C. 30 %
   D. 25 %

4. Référence : 52 % (intervalle 47-57 %). Finale : 55 % (intervalle 50-60 %). Conclusion la plus juste :
   A. La situation s'est nettement améliorée
   B. Les données ne permettent pas de conclure à un changement
   C. La situation s'est dégradée
   D. L'enquête est fausse

5. Pour 2 000 abris et 200 ménages à interroger en échantillonnage systématique, le pas est :
   A. 5
   B. 10
   C. 20
   D. 200

**Corrigé du quiz :** 1. B · 2. B · 3. B · 4. B · 5. B

# Semaine 7 : Analyser les données : statistiques utiles, pièges et visualisation

**Objectifs**

- Choisir les bonnes statistiques descriptives et désagréger pour révéler les inégalités cachées par les moyennes.
- Éviter les pièges classiques : corrélation et causalité, paradoxe de Simpson, dénominateurs trompeurs.
- Analyser des données qualitatives avec rigueur et présenter les résultats pour qu'ils soient compris et utilisés.

## Statistiques descriptives : moyenne, médiane, dispersion

La **moyenne** est sensible aux valeurs extrêmes ; la **médiane** (valeur qui partage l'échantillon en deux moitiés) ne l'est pas. Pour les revenus, les dépenses, les dettes ou les temps d'attente, distributions très asymétriques, la médiane décrit mieux le ménage typique.

**Exemple.** Dettes de dix ménages (USD) : 0, 0, 20, 30, 40, 50, 60, 80, 100, 1 620. Moyenne : 2 000 / 10 = 200 USD. Médiane : (40 + 50) / 2 = 45 USD. Dire « les ménages ont en moyenne 200 USD de dettes » décrit un seul ménage, pas la situation.

La **dispersion** compte autant que le centre : écart type, ou écart interquartile (entre le 25e et le 75e centile) pour les distributions asymétriques. Deux programmes avec un temps d'attente moyen de 2 heures sont très différents si l'un a des attentes entre 1 h 30 et 2 h 30, et l'autre entre 10 minutes et 7 heures.

Pour les indicateurs en proportion, on présente le numérateur et le dénominateur : « 62 % (248 sur 400) ». Les pourcentages calculés sur de très petits effectifs (« 50 % des personnes âgées », soit 2 sur 4) sont à éviter ou à signaler.

> Variation en points et en pourcentage : passer de 20 % à 30 % est une hausse de 10 points de pourcentage, soit une hausse relative de 50 %. Confondre les deux est l'erreur la plus fréquente des rapports.

## Désagréger pour voir les inégalités

Une moyenne globale satisfaisante peut cacher un groupe laissé pour compte. La **désagrégation** (par sexe et âge, handicap, statut, zone, type de ménage) est le moyen de vérifier que le programme atteint tout le monde.

**Exemple.** Taux de satisfaction global : 82 %. Par groupe : ménages sans personne handicapée 86 % ; ménages avec une personne handicapée 54 %. Le programme ne fonctionne pas pour un groupe entier, et le chiffre global le masquait.

Le **tableau croisé** est l'outil de base : un indicateur en ligne, un groupe en colonne. Pour savoir si la différence observée entre deux groupes dépasse ce que le hasard de l'échantillonnage peut produire, on utilise un test (test du khi-deux pour des proportions, test t ou test non paramétrique pour des moyennes) ou l'on compare les intervalles de confiance. Un logiciel le calcule ; le rôle du responsable est de poser la bonne question et d'interpréter.

Attention aux **comparaisons multiples** : si l'on teste 20 différences au seuil de 5 %, on s'attend à en trouver une « significative » par pur hasard. On définit donc à l'avance les comparaisons principales du plan d'analyse, et l'on présente les autres comme exploratoires.

## Corrélation, causalité et paradoxe de Simpson

Deux variables qui évoluent ensemble ne prouvent pas que l'une cause l'autre. Trois explications alternatives existent toujours : le **hasard**, un **facteur commun** (confondant) et la **causalité inverse**.

**Exemple.** Les ménages qui ont participé aux formations agricoles ont de meilleurs rendements. Mais ce sont aussi ceux qui ont plus de terres et d'outils, et qui étaient plus motivés à s'inscrire : la formation a peut-être un effet, mais la comparaison simple ne permet pas de le séparer de ces différences initiales. Mesurer un effet exige un contrefactuel (semaine 10).

Le **paradoxe de Simpson** survient quand une tendance présente dans chaque groupe s'inverse quand on regroupe les données, à cause d'une répartition inégale entre groupes.

| | **Centre A** | **Centre B** |
|---|---|---|
| Cas sévères | 60 guéris sur 100 (60 %) | 110 guéris sur 200 (55 %) |
| Cas modérés | 270 guéris sur 300 (90 %) | 170 guéris sur 200 (85 %) |
| Total | 330 sur 400 (82,5 %) | 280 sur 400 (70 %) |

Ici, A est meilleur dans chaque groupe et au total : pas de paradoxe. Mais si B recevait surtout des cas modérés et A surtout des cas sévères, B pourrait paraître meilleur au total tout en étant moins bon dans chaque groupe. Règle : comparer des performances seulement à **composition comparable**, ou en désagrégeant par le facteur qui change la difficulté.

## Analyser les données qualitatives

Les entretiens et les discussions de groupe expliquent le « pourquoi » derrière les chiffres. Leur analyse suit une méthode, sans laquelle on ne retient que les citations qui confirment ce que l'on pensait déjà.

- **Transcription ou notes détaillées** le jour même, avec les mots des personnes.
- **Codage** : on attribue des codes (thèmes) aux passages. On part de codes liés aux questions d'évaluation, et l'on ajoute ceux qui émergent.
- **Matrice** : thèmes en ligne, groupes (femmes, hommes, jeunes, déplacés, hôtes) en colonne. On voit ainsi quels thèmes reviennent dans quels groupes.

- **Recherche active des contre-exemples** : les cas qui contredisent la tendance générale sont souvent les plus instructifs.
- **Saturation** : on arrête la collecte quand de nouveaux entretiens n'apportent plus de thèmes nouveaux.

On ne présente pas les résultats qualitatifs en pourcentages, mais l'on peut indiquer la fréquence : « ce thème est apparu dans 9 des 12 groupes de discussion, dont tous les groupes de femmes ». Une citation illustre un résultat, elle ne le prouve pas.

> Méthodes mixtes : le chiffre dit « 30 % des ménages ne retirent pas leur transfert », le qualitatif dit « le point de retrait est à 3 heures de marche et les agents manquent souvent de liquidités ». Les deux ensemble mènent à une décision.

## Visualiser pour faire comprendre

Un graphique doit répondre à une question en quelques secondes. Règles pratiques :

- **Barres horizontales** pour comparer des catégories, triées par valeur ; **courbes** pour les évolutions dans le temps ; **cartes** pour les différences entre zones ; éviter les camemberts au-delà de trois catégories et les effets 3D.
- L'axe des barres commence à **zéro** ; sinon une différence de 2 points paraît énorme.
- Un **titre qui dit le message** (« La couverture a baissé dans les zones nord depuis mars ») plutôt qu'un titre neutre (« Couverture par zone »).
- Peu de couleurs, une couleur d'accent pour ce qui compte, et un repère pour la cible.
- La source, la date et la taille d'échantillon en bas du graphique.

Pour un tableau de bord de direction, on limite à 6 à 10 indicateurs clés, avec pour chacun la valeur actuelle, la cible, la tendance et un code couleur dont les seuils sont définis. Un tableau de bord de 40 indicateurs n'est lu par personne.

Sur une carte, on représente des **taux** et non des nombres absolus quand les zones ont des populations différentes : une grande ville aura toujours plus de cas qu'un village.

## Du résultat à la recommandation

Une analyse utile suit une progression en trois temps : **constat** (ce que disent les données), **interprétation** (pourquoi, avec les données qualitatives et le contexte) et **recommandation** (ce qu'il faut faire, par qui, quand). Les recommandations sont peu nombreuses, précises et réalistes, et chacune découle d'un constat.

Il faut aussi indiquer le **niveau de confiance** : fort (plusieurs sources concordantes, échantillon suffisant), moyen, faible (une seule source, petit échantillon). Une direction peut décider sur une preuve moyenne en urgence, mais elle doit savoir qu'elle le fait.

Enfin, la **restitution** des résultats se fait avant le rapport définitif : une séance avec l'équipe et les partenaires permet de corriger les erreurs d'interprétation et de construire l'adhésion aux

recommandations. Les résultats sont aussi partagés avec les communautés, sous une forme adaptée (réunion, affiche, message vocal).

## Exercice d'application : analyser un suivi post-distribution

**Situation.** Suivi post-distribution de 400 ménages. Temps d'attente : moyenne 2 h 40, médiane 1 h 20. Satisfaction globale : 78 %. Ménages dirigés par une femme (120) : 60 % satisfaits ; dirigés par un homme (280) : 85,7 % satisfaits. Le score moyen de consommation alimentaire est passé de 31 à 38 ; la proportion de ménages avec une consommation acceptable est passée de 30 % à 45 %. Dans les discussions de groupe, 8 groupes de femmes sur 10 citent l'heure de distribution (après 15 h) et le trajet du retour.

**Travail demandé.** 1) Que dit l'écart entre la moyenne et la médiane du temps d'attente ? 2) Exprimez la variation de la consommation acceptable en points et en pourcentage. 3) Quelle conclusion et quelle recommandation tirez-vous sur la satisfaction ?

**Corrigé.** La moyenne très supérieure à la médiane indique que la plupart attendent peu, mais qu'un groupe attend très longtemps : il faut identifier quels sites ou quels jours. Consommation acceptable : +15 points de pourcentage, soit +50 % en relatif (on ne peut pas attribuer ce changement au seul programme sans contrefactuel). Satisfaction : vérification, 0,3 × 60 + 0,7 × 85,7 = 78 %. L'écart de 26 points entre ménages dirigés par une femme et par un homme est important et expliqué par les groupes de discussion : horaires tardifs et trajet de retour dangereux. Recommandation : distributions le matin, file prioritaire, points de distribution plus proches, puis vérification au prochain suivi.

## À retenir

- Médiane plutôt que moyenne pour les distributions asymétriques (revenus, dettes, temps d'attente).
- Une variation de 20 % à 30 % est +10 points, soit +50 % en relatif.
- Désagréger révèle les groupes laissés pour compte que la moyenne cache.
- Corrélation n'est pas causalité ; comparer des performances à composition comparable.
- Analyse qualitative : codage, matrice par groupe, recherche des contre-exemples ; une citation illustre, elle ne prouve pas.

## Quiz

1. Dépenses de 5 ménages : 10, 15, 20, 25, 430 USD. Moyenne et médiane :
   A. Moyenne 100, médiane 20
   B. Moyenne 20, médiane 100
   C. Moyenne 25, médiane 20
   D. Moyenne 100, médiane 25

2. Une couverture passe de 40 % à 50 %. C'est :
   A. +10 % en relatif
   B. +10 points de pourcentage, soit +25 % en relatif
   C. +50 points
   D. +25 points

3. On teste 40 différences entre groupes au seuil de 5 %. Combien de résultats « significatifs » peut-on attendre par hasard seul ?
   A. Aucun
   B. Environ 2
   C. Environ 20
   D. Exactement 5

4. Sur une carte comparant des districts de populations très différentes, il faut représenter :
   A. Les nombres absolus de cas
   B. Des taux (par exemple pour 1 000 habitants)
   C. Seulement les noms des districts
   D. Les budgets des ONG

5. Un thème apparaît dans 9 des 12 groupes de discussion. La présentation correcte est :
   A. 75 % de la population pense cela
   B. Ce thème est apparu dans 9 des 12 groupes, en précisant lesquels
   C. Tout le monde pense cela
   D. Ce résultat est statistiquement significatif

**Corrigé du quiz :** 1. A · 2. B · 3. B · 4. B · 5. B

# Semaine 8 : Redevabilité envers les populations : participation, plaintes et retours

**Objectifs**

- Traduire la Norme humanitaire fondamentale (CHS) en pratiques concrètes et mesurables dans un programme.
- Concevoir et piloter un mécanisme de plaintes et de retours accessible, sûr et qui ferme la boucle.
- Organiser la participation réelle des personnes et l'information transparente, y compris sur les décisions difficiles.

## Ce que signifie être redevable envers les personnes

La **redevabilité envers les populations affectées** (AAP) est l'engagement à utiliser son pouvoir de façon responsable : prendre en compte l'avis des personnes, leur rendre compte et répondre de ses actes devant elles. Elle ne remplace pas la redevabilité envers les bailleurs ; elle la complète, et les deux entrent souvent en tension (un bailleur veut des chiffres, les personnes veulent être écoutées).

La **Norme humanitaire fondamentale de qualité et de redevabilité** (CHS), révisée en 2024, fixe neuf engagements envers les personnes et les communautés. En substance : une aide qui respecte leurs droits et leur participation, accessible à temps, pertinente, qui renforce leurs capacités et ne nuit pas, avec des plaintes traitées de façon sûre, une aide coordonnée, une organisation qui apprend, un personnel compétent et bien géré, et des ressources utilisées de façon responsable.

Pour chaque engagement, une organisation peut se poser trois questions : que faisons-nous concrètement ? comment le savons-nous ? que disent les personnes elles-mêmes ? Les indicateurs les plus parlants viennent des personnes : « Pensez-vous que l'aide répond à vos besoins les plus importants ? », « Savez-vous comment faire une plainte ? », « Vous sentez-vous traité avec respect ? ».

> La redevabilité se juge à ce qui change. Un programme qui consulte sans jamais rien modifier enseigne aux personnes que leur avis ne compte pas, et la participation suivante sera plus faible.

## Concevoir un mécanisme de plaintes et de retours

Un **mécanisme de plaintes et de retours** (CFM ou FCRM) reçoit, enregistre, traite et répond à toutes les communications des personnes : questions, suggestions, plaintes simples, plaintes sensibles. Sa conception part de la consultation : quels canaux les personnes utilisent-elles et en qui ont-elles confiance ? Les femmes, les enfants, les personnes âgées et les personnes handicapées n'ont pas les mêmes réponses.

Plusieurs canaux se complètent : ligne téléphonique gratuite, messagerie, boîtes à suggestions (pour l'anonymat, si elles sont relevées par une personne indépendante de l'équipe de distribution), permanences et bureaux d'accueil lors des distributions, visites à domicile pour les personnes qui ne peuvent pas se déplacer, comités communautaires, et retours recueillis lors du suivi.

Le mécanisme définit des **catégories** et des **délais** écrits. Par exemple : demande d'information, réponse sous 3 jours ; plainte de programme (non inscrit, montant erroné), réponse sous 7 jours ; plainte sensible (fraude, abus, exploitation sexuelle, sécurité), transmission sous 24 heures à la personne désignée, selon une filière confidentielle séparée.

Les principes non négociables : **confidentialité**, **sécurité** des personnes qui se plaignent (aucune représaille), **accessibilité** (langues, analphabétisme, handicap), **gratuité**, et **séparation** entre ceux qui traitent les plaintes et ceux qui en font l'objet.

## Fermer la boucle et mesurer la performance

« Fermer la boucle » signifie que chaque personne reçoit une réponse, même quand la réponse est non, et que les tendances des retours alimentent les décisions. La plupart des mécanismes collectent, mais beaucoup ne répondent pas.

Indicateurs de performance utiles :

- Nombre de retours reçus par mois et pour 1 000 personnes aidées, par canal, par sexe et âge.
- Part des retours clôturés dans le délai prévu.
- Part des personnes satisfaites de la réponse (vérifiée par un rappel sur un échantillon).
- Part des personnes qui connaissent au moins un canal (enquête de suivi).
- Nombre de changements de programme issus des retours au cours du trimestre.

**Exemple.** 340 retours en un mois pour 17 000 personnes : 20 pour 1 000. 289 clôturés dans le délai : 85 %. Comparé à un mois précédent de 2 retours pour 1 000, ce n'est pas une dégradation : c'est souvent le signe que le mécanisme est enfin connu et que les personnes lui font confiance.

> Un mécanisme qui reçoit très peu de plaintes est rarement le signe d'un programme parfait : c'est presque toujours le signe d'un mécanisme inconnu, inaccessible ou sans confiance.

## Plaintes sensibles et protection contre l'exploitation et les abus sexuels

Les plaintes d'**exploitation et d'abus sexuels** (EAS) par du personnel humanitaire suivent une filière stricte, fondée sur les six principes fondamentaux de l'IASC : l'EAS constitue une faute grave justifiant le licenciement ; toute activité sexuelle avec un enfant (moins de 18 ans) est interdite quel que soit l'âge de consentement local ; l'échange d'argent, d'emploi, de biens ou de services contre des relations sexuelles est interdit ; les relations sexuelles avec des

bénéficiaires sont fortement découragées car fondées sur une dynamique de pouvoir inégale ; le personnel doit signaler toute préoccupation ; et les responsables doivent créer un environnement qui prévient l'EAS.

Dans le mécanisme, cela signifie : toute personne qui reçoit une allégation la transmet **immédiatement** au point focal désigné, sans enquêter elle-même ; la **survivante ou le survivant** est orienté vers une assistance (soins médicaux, soutien psychosocial, protection, aide juridique) selon ses souhaits ; la personne mise en cause est écartée des contacts avec les bénéficiaires pendant l'enquête ; les informations sont partagées strictement selon le besoin d'en connaître ; le bailleur est informé selon le contrat.

Une approche **centrée sur la survivante ou le survivant** respecte quatre principes : sécurité, confidentialité, respect des choix, non-discrimination. On ne confronte jamais la personne à l'auteur présumé, et l'on ne demande pas de témoignage répété.

## Participation et information

La **participation** va de l'information à la décision partagée. Dans l'échelle classique, on distingue : informer, consulter, impliquer, collaborer, et transférer la décision. La plupart des programmes humanitaires restent au niveau de la consultation ; viser plus haut sur certaines décisions (choix des sites, critères de ciblage, composition des kits, calendrier) améliore la pertinence et l'acceptation.

Les comités communautaires sont utiles mais fragiles : ils reproduisent souvent les rapports de pouvoir existants. On veille à leur composition (femmes, jeunes, personnes handicapées, minorités), on fixe un mandat écrit, et on ne leur confie pas seuls la sélection des bénéficiaires sans contrôle.

L'**information** est une forme d'aide : les personnes ont le droit de savoir qui est l'organisation, ce qu'elle fait, pour qui, selon quels critères, pour combien de temps, comment se plaindre, et que l'aide est gratuite. On utilise les langues et les formats adaptés (radio, messages vocaux, affiches illustrées, réunions séparées pour les femmes).

Les **décisions difficiles** (réduction des rations, fin d'un programme, changement de ciblage) doivent être annoncées tôt, expliquées honnêtement et accompagnées : orientation vers d'autres services, période de transition. Annoncer une coupure le jour de la distribution est une faute de redevabilité qui crée aussi des risques de sécurité pour le personnel.

## Protection des données et intégration dans le cycle de programme

Les retours et les plaintes contiennent des **données personnelles**, parfois très sensibles. Règles minimales : accès limité aux personnes qui traitent les cas ; séparation des données d'identification et du contenu pour les analyses ; anonymisation dans les rapports ; durée de conservation définie ; aucun partage avec des autorités ou des tiers sans base claire et sans évaluation des risques pour les personnes.

La redevabilité doit être intégrée à chaque étape du cycle de programme, et budgétée :

- **Évaluation** : demander aux personnes leurs priorités et leurs canaux préférés.

- **Conception** : critères de ciblage expliqués, mécanisme de plaintes prévu, ligne budgétaire (personnel, ligne téléphonique, matériel d'information).
- **Mise en œuvre** : information, retours traités, ajustements.
- **Suivi** : indicateurs de redevabilité dans le tableau de bord.
- **Évaluation et apprentissage** : les personnes jugent elles-mêmes la pertinence et la qualité de l'aide.

Au niveau collectif, les mécanismes communs (lignes d'information inter-agences) réduisent la confusion pour les personnes, qui ne connaissent souvent pas le nom des organisations. Ils exigent des accords clairs de partage et de transfert des cas.

## Exercice d'application : redresser un mécanisme de plaintes

**Situation.** Programme de 24 000 personnes. En six mois, 31 retours reçus, tous par la boîte à suggestions installée dans le bureau de l'organisation et relevée par le chef de site ; 12 ont reçu une réponse. Une enquête de suivi montre que 22 % des personnes connaissent un moyen de se plaindre (13 % des femmes, 31 % des hommes). Un enquêteur rapporte une rumeur selon laquelle un distributeur demanderait des faveurs sexuelles à des femmes en échange d'une inscription.

**Travail demandé.** 1) Calculez le taux de retours pour 1 000 personnes et le taux de réponse. 2) Identifiez les quatre problèmes principaux. 3) Que faites-vous dans les 24 heures au sujet de la rumeur ? 4) Proposez trois indicateurs de suivi.

**Corrigé.** 31 retours pour 24 000 personnes : 1,3 pour 1 000 en six mois, extrêmement faible. Taux de réponse : 12 / 31 = 39 %. Problèmes : canal unique, peu accessible et situé dans le bureau ; boîte relevée par une personne qui peut être mise en cause (pas d'indépendance) ; faible connaissance, surtout chez les femmes ; boucle non fermée. Dans les 24 heures : transmettre la rumeur au point focal PSEA sans enquêter soi-même, organiser l'écartement préventif du distributeur des activités d'inscription, prévoir l'assistance pour toute survivante identifiée, informer la direction et le bailleur selon les procédures. Indicateurs : retours pour 1 000 personnes par sexe et canal ; part clôturée dans le délai ; part des personnes, par sexe, qui connaissent au moins un canal.

## À retenir

- La CHS fixe neuf engagements ; les meilleurs indicateurs viennent de ce que disent les personnes.
- Un mécanisme de retours a plusieurs canaux choisis avec les personnes, des délais écrits et une filière sensible séparée.
- Fermer la boucle : chaque retour reçoit une réponse et les tendances changent le programme.
- Peu de plaintes signale souvent un mécanisme inconnu ou sans confiance, pas un programme parfait.
- Allégation d'EAS : transmission immédiate au point focal, assistance centrée sur la survivante ou le survivant, aucune enquête improvisée.

## Quiz

1. Un mécanisme reçoit 3 plaintes en six mois pour 20 000 personnes aidées. L'interprétation la plus probable :
   A. Le programme est excellent
   B. Le mécanisme est inconnu, inaccessible ou sans confiance
   C. Les personnes n'ont rien à dire
   D. Il faut supprimer le mécanisme

2. Un agent reçoit une allégation d'exploitation sexuelle contre un collègue. Il doit :
   A. Interroger lui-même le collègue
   B. La transmettre immédiatement au point focal désigné, sans enquêter lui-même
   C. Attendre d'avoir des preuves
   D. En parler en réunion d'équipe

3. Selon les principes de l'IASC, toute activité sexuelle avec une personne de moins de 18 ans est interdite :
   A. Seulement si l'âge de consentement local est de 18 ans
   B. Quel que soit l'âge de consentement local
   C. Seulement pour le personnel international
   D. Seulement pendant les heures de travail

4. Qui ne doit pas relever une boîte à plaintes ?
   A. Un responsable redevabilité indépendant
   B. Le chef de l'équipe de distribution dont les activités font l'objet des plaintes
   C. Un binôme incluant une femme
   D. Un membre de l'équipe MEAL

5. Une réduction des rations est décidée. La pratique conforme à la redevabilité est :
   A. L'annoncer le jour de la distribution
   B. L'annoncer tôt, l'expliquer honnêtement, et orienter vers d'autres services pendant une transition
   C. Ne pas l'annoncer
   D. Laisser les autorités l'annoncer

**Corrigé du quiz :** 1. B · 2. B · 3. B · 4. B · 5. B

-

# Semaine 9 : Évaluer : critères, termes de référence, matrice et utilisation

**Objectifs**

- Choisir le type d'évaluation et les critères adaptés à la décision qu'elle doit éclairer.
- Rédiger des termes de référence précis et construire une matrice d'évaluation rigoureuse.
- Gérer une évaluation de façon indépendante et transformer ses recommandations en décisions suivies.

## Pourquoi et quand évaluer

Une **évaluation** est une appréciation systématique et objective d'une intervention, de sa conception, de sa mise en œuvre et de ses résultats. Elle diffère du suivi : le suivi est continu et interne, il demande « faisons-nous ce que nous avions prévu ? » ; l'évaluation est ponctuelle, souvent externe, et elle demande « faisons-nous la bonne chose, de la bonne manière, et qu'est-ce que cela change ? ».

Elle a deux fonctions, parfois en tension : la **redevabilité** (rendre compte aux personnes, aux bailleurs, au conseil d'administration) et l'**apprentissage** (améliorer ce programme et les suivants). Une évaluation conçue uniquement pour la redevabilité produit des notes ; une évaluation conçue pour l'apprentissage produit des décisions.

| Type | Moment | Question principale |
|---|---|---|
| Évaluation en temps réel | Premières semaines d'une urgence | Faut-il corriger la réponse maintenant ? |
| Évaluation à mi-parcours | Milieu du projet | Que faut-il ajuster pour la seconde moitié ? |
| Évaluation finale | Fin du projet | Quels résultats, quelles leçons pour la suite ? |
| Évaluation d'impact | Pendant et après | Quel changement est attribuable à l'intervention ? |
| Évaluation ex post | 1 à 3 ans après la fin | Les effets ont-ils duré ? |

Avant de commander une évaluation, écrivez qui l'utilisera et pour quelle décision. Si personne ne peut répondre, l'évaluation sera un rapport de plus sur une étagère.

## Les critères du CAD de l'OCDE

Les critères d'évaluation du Comité d'aide au développement de l'OCDE, révisés en 2019, sont au nombre de six :

- **Pertinence** : l'intervention répond-elle aux besoins et priorités des personnes, et reste-t-elle pertinente si le contexte change ?
- **Cohérence** (ajoutée en 2019) : est-elle compatible avec les autres interventions dans le même contexte, internes et externes à l'organisation ?
- **Efficacité** : atteint-elle ses objectifs, y compris pour les différents groupes ?
- **Efficience** : produit-elle ses résultats de façon économique et à temps ?
- **Impact** : quels effets de plus long terme, positifs ou négatifs, voulus ou non ?
- **Durabilité** : les bénéfices continueront-ils après la fin du financement ?

En humanitaire, on utilise aussi les critères adaptés par l'ALNAP : **couverture** (a-t-on atteint toutes les personnes dans le besoin, sans exclusion ?), **connectivité** (lien entre l'urgence et le plus long terme) et **coordination**.

Les critères ne s'appliquent pas tous avec le même poids. Une évaluation en temps réel d'une réponse de trois mois se concentre sur la pertinence, la couverture et l'efficacité ; parler de durabilité à ce stade a peu de sens. Choisir trois ou quatre critères et les approfondir vaut mieux qu'en survoler six.

## Des termes de référence qui obtiennent une bonne évaluation

Les **termes de référence** (TdR) sont le contrat intellectuel de l'évaluation. Des TdR vagues produisent des évaluations vagues. Ils contiennent :

- **Contexte** et description de l'intervention (avec la théorie du changement et le cadre logique en annexe).
- **Objet et utilisateurs** : pourquoi maintenant, pour qui, pour quelles décisions.
- **Portée** : période, zones, composantes, ce qui est exclu.
- **Questions d'évaluation** : cinq à huit questions principales, pas trente. Chaque question est évaluable avec les moyens et le temps disponibles.
- **Approche et méthodes** attendues, en laissant l'équipe proposer, et les données existantes disponibles.
- **Éthique** : consentement, protection des données, protocoles pour les sujets sensibles, participation des personnes.
- **Livrables et calendrier** : rapport initial (inception), restitution, rapport provisoire, rapport final, résumé.
- **Profil de l'équipe**, budget, gouvernance (groupe de référence) et règles d'indépendance.

**Ordre de grandeur.** Une évaluation finale externe de qualité d'un projet multisectoriel coûte souvent entre 25 000 et 80 000 USD selon le terrain et les méthodes. Un budget trop faible conduit à quelques jours de terrain et à un rapport sans preuves ; il vaut mieux moins d'évaluations, mieux faites.

## La matrice d'évaluation et le rapport initial

La **matrice d'évaluation** est l'outil central du rapport initial : pour chaque question, elle précise les sous-questions, les indicateurs ou critères de jugement, les sources de données, les méthodes de collecte et d'analyse, et les limites.

**Exemple.** Question : « Dans quelle mesure les transferts monétaires ont-ils amélioré la sécurité alimentaire des ménages les plus vulnérables ? » Critères de jugement : évolution du score de consommation et du rCSI entre la référence et la fin, comparée à une zone de comparaison ; différences selon le sexe du chef de ménage et la présence d'une personne handicapée. Sources : enquêtes de référence et finale, suivi post-distribution, discussions de groupe. Méthodes : différence de différences, analyse qualitative par groupe. Limites : zone de comparaison partiellement aidée par d'autres acteurs à partir du mois 6.

Le **rapport initial** est le moment de corriger le tir : l'équipe confirme ce qui est faisable, signale les questions impossibles à traiter, propose l'échantillonnage. Le gestionnaire le valide avec le groupe de référence avant le terrain. C'est le meilleur investissement de temps de toute l'évaluation.

La **qualité** d'une évaluation se juge sur la clarté du lien entre **constats** (appuyés par des données), **conclusions** (jugements fondés sur les constats) et **recommandations** (actions qui découlent des conclusions). Une recommandation sans constat correspondant est une opinion de l'évaluateur.

## Indépendance, éthique et gestion de l'évaluation

L'**indépendance** protège la crédibilité : l'équipe d'évaluation n'a pas participé à la conception ni à la mise en œuvre ; le gestionnaire de l'évaluation n'est pas le chef du projet évalué ; l'équipe a accès librement aux données et aux personnes ; le programme peut commenter les faits du rapport provisoire, mais pas imposer de changer les conclusions.

Les normes et règles d'évaluation du groupe des Nations unies pour l'évaluation (UNEG) et les principes de l'ALNAP insistent sur l'**éthique** : ne pas nuire aux personnes interrogées, consentement éclairé, confidentialité, attention aux personnes en détresse et aux risques de représailles, restitution aux communautés.

Le **gestionnaire de l'évaluation** a un rôle actif : il prépare les données, organise l'accès au terrain, réunit le groupe de référence, contrôle la qualité aux étapes clés, protège l'indépendance de l'équipe, et organise la restitution. Les principaux risques sont connus : retards de démarrage qui réduisent le terrain, données de référence inexistantes, accès limité, pression pour adoucir les constats.

> Une évaluation qui ne trouve que des points positifs n'est pas crédible. Un programme qui accepte des constats négatifs bien étayés montre qu'il est sérieux.

## La réponse de la direction et l'utilisation

La **réponse de la direction** (management response) est le document qui transforme une évaluation en décisions. Pour chaque recommandation : acceptée, partiellement acceptée ou

rejetée, avec la justification ; actions prévues ; responsable ; échéance. Elle est rédigée dans le mois qui suit le rapport final et publiée avec lui.

Le **suivi** des actions se fait lors des revues trimestrielles de direction : où en est chaque action ? Les organisations matures publient chaque année un taux de mise en œuvre des recommandations acceptées.

Une recommandation peut être légitimement rejetée : trop coûteuse, hors du mandat, déjà en cours. Mais la justification doit être écrite. Ce qui n'est pas acceptable est l'absence de réponse.

Enfin, la **diffusion** adapte le format au public : résumé de deux pages pour la direction et les bailleurs, présentation pour les équipes, version simple et orale pour les communautés, et publication du rapport complet (par exemple dans la bibliothèque de l'ALNAP) pour le secteur. La publication des évaluations, y compris critiques, est un engagement du Grand Bargain en matière de transparence.

## Exercice d'application : écrire la question et la matrice d'une évaluation

**Situation.** Un projet de 18 mois (eau, hygiène et protection) dans trois camps de déplacés arrive à mi-parcours. La direction pays doit décider dans quatre mois s'il faut étendre le projet à un quatrième camp et comment. Le bailleur a demandé une évaluation mais n'a pas fixé de questions. Budget : 30 000 USD. Les données de référence existent pour l'eau, pas pour la protection.

**Travail demandé.** 1) Quel type d'évaluation et quels critères principaux ? 2) Écrivez trois questions d'évaluation. 3) Pour l'une d'elles, remplissez une ligne de matrice (critères de jugement, sources, méthodes, limites).

**Corrigé.** Évaluation à mi-parcours, formative, avec pour utilisateurs la direction pays et le bailleur ; critères principaux : pertinence, efficacité, couverture, efficience (la durabilité est secondaire à ce stade). Questions possibles : 1) Dans quelle mesure les services d'eau et d'hygiène atteignent-ils les standards prévus pour tous les groupes, y compris les personnes handicapées ? 2) Les activités de protection répondent-elles aux risques prioritaires identifiés par les femmes et les enfants ? 3) Quelles conditions faut-il réunir pour étendre le projet à un quatrième camp avec la même qualité et un coût par personne comparable ? Ligne de matrice pour 2) : critères : correspondance entre risques prioritaires déclarés et activités, utilisation des services, perception de sécurité ; sources : discussions de groupe séparées par sexe et âge, entretiens avec les acteurs de protection, données des services (anonymisées) ; méthodes : analyse qualitative codée, cartographie participative des risques ; limites : absence de référence en protection, d'où une analyse de contribution plutôt qu'une mesure d'évolution.

## À retenir

- Le suivi demande « faisons-nous ce que nous avions prévu ? » ; l'évaluation demande « est-ce la bonne chose, bien faite, et que change-t-elle ? ».
- Six critères du CAD (2019) : pertinence, cohérence, efficacité, efficience, impact, durabilité ; plus couverture, connectivité et coordination en humanitaire.
- Des TdR précis : utilisateurs, décisions, cinq à huit questions évaluables, éthique, livrables.
- La matrice d'évaluation relie chaque question à des critères, des sources, des méthodes et des limites.
- Réponse de la direction : chaque recommandation acceptée ou non, avec action, responsable et échéance.

## Quiz

1. Quel critère a été ajouté par le CAD de l'OCDE en 2019 ?
   A. Pertinence
   B. Cohérence
   C. Efficience
   D. Impact

2. Une évaluation qui demande si toutes les personnes dans le besoin ont été atteintes, sans exclusion, porte sur le critère :
   A. Durabilité
   B. Couverture
   C. Efficience
   D. Connectivité

3. Le moment le plus utile pour corriger une évaluation mal orientée est :
   A. La lecture du rapport final
   B. La validation du rapport initial et de la matrice d'évaluation
   C. La publication
   D. La réponse de la direction

4. Une recommandation qui ne s'appuie sur aucun constat du rapport est :
   A. Une bonne pratique
   B. Une opinion de l'évaluateur, à remettre en question
   C. Obligatoire
   D. Une conclusion

5. Qui ne devrait pas gérer l'évaluation d'un projet ?
   A. Un responsable MEAL pays
   B. Le chef du projet évalué
   C. Un conseiller du siège
   D. Un comité de pilotage

**Corrigé du quiz :** 1. B · 2. B · 3. B · 4. B · 5. B

# Semaine 10 : Mesurer l'impact : contrefactuel, méthodes et contribution

**Objectifs**

- Comprendre le problème du contrefactuel et pourquoi une comparaison avant-après ne mesure pas un effet.
- Choisir et appliquer une méthode quasi expérimentale ou expérimentale adaptée au contexte humanitaire, avec ses conditions de validité.
- Utiliser les approches de contribution quand l'attribution est impossible, et décider ce qu'il est raisonnable de prouver.

## Le problème fondamental : que se serait-il passé sans le programme ?

L'**impact** d'un programme est la différence entre ce qui s'est passé avec le programme et ce qui se serait passé sans lui : le **contrefactuel**. On ne peut jamais observer directement le même ménage avec et sans programme ; toutes les méthodes d'évaluation d'impact sont des façons de construire une estimation crédible de ce contrefactuel.

Deux comparaisons naïves sont trompeuses :

- **Avant-après** chez les bénéficiaires : le changement observé mélange l'effet du programme et tout ce qui a changé en même temps (récolte, prix, sécurité, autres aides). Si la récolte a été bonne, la sécurité alimentaire s'améliore même sans programme ; si elle a été mauvaise, un programme efficace peut paraître sans effet.
- **Bénéficiaires contre non-bénéficiaires** après le programme : les deux groupes étaient différents au départ, justement parce que le ciblage a choisi les plus vulnérables. La comparaison sous-estime alors l'effet, ou le surestime si ce sont les plus motivés qui se sont inscrits (biais de sélection).

> Question à poser devant toute affirmation d'impact : « par rapport à quoi ? ». Si la réponse est « par rapport à avant », ce n'est pas une mesure d'impact.

## L'essai randomisé contrôlé et son usage en humanitaire

Dans un **essai randomisé contrôlé** (ECR), on tire au sort, parmi les personnes éligibles, celles qui reçoivent le programme (groupe de traitement) et celles qui ne le reçoivent pas ou plus tard (groupe de contrôle). Le tirage au sort rend les deux groupes comparables en moyenne sur toutes les caractéristiques, observées ou non. L'effet est simplement la différence des résultats moyens à la fin.

En humanitaire, refuser une aide vitale à des personnes éligibles pour constituer un groupe de contrôle est **contraire à l'éthique**. Mais des ECR restent possibles et ont produit des

connaissances majeures (sur le cash, la nutrition, la santé mentale) dans des situations précises :

- **Ressources insuffisantes** de toute façon pour tout le monde : le tirage au sort est alors une façon équitable de répartir, et le groupe non servi sert de contrôle.
- **Mise en place progressive** : tout le monde recevra le programme, mais l'ordre est tiré au sort (conception en escalier, « stepped wedge »).
- **Comparaison de deux variantes** d'un programme, toutes deux jugées acceptables : cash mensuel ou en une fois, avec ou sans accompagnement.

Les ECR coûtent cher, prennent du temps et répondent à des questions précises ; ils se justifient quand la réponse servira à de nombreux programmes. Pour la plupart des projets, des méthodes quasi expérimentales suffisent.

## La différence de différences

La **différence de différences** compare l'évolution d'un groupe aidé à l'évolution d'un groupe de comparaison non aidé : effet = (après − avant) dans le groupe aidé − (après − avant) dans le groupe de comparaison. Elle élimine les différences initiales stables entre les groupes et les tendances communes (récolte, prix).

**Exemple.** Proportion de ménages avec une consommation acceptable. Groupe aidé : 28 % avant, 52 % après (+24 points). Groupe de comparaison : 35 % avant, 44 % après (+9 points). Effet estimé : 24 − 9 = 15 points. Une comparaison avant-après aurait attribué 24 points au programme ; une comparaison des seuls résultats finaux en aurait attribué 8.

Condition de validité : l'hypothèse des **tendances parallèles**. Sans le programme, les deux groupes auraient évolué de la même façon. On la vérifie si possible avec plusieurs mesures avant le programme, et l'on s'assure que le groupe de comparaison ne reçoit pas une autre aide importante pendant la période (contamination) et ne subit pas un choc spécifique.

En humanitaire, le groupe de comparaison est souvent une zone voisine non couverte faute de financement, ou des ménages éligibles inscrits sur liste d'attente. On ne crée jamais un groupe de comparaison en refusant l'aide pour les besoins de l'étude.

## Appariement et régression sur discontinuité

L'**appariement** construit un groupe de comparaison en associant à chaque ménage aidé un ou plusieurs ménages non aidés qui lui ressemblent sur les caractéristiques mesurées avant le programme (taille, biens, revenus, statut, zone). La méthode du **score de propension** résume ces caractéristiques en une probabilité d'être aidé, et l'on compare des ménages au score proche. Sa limite : elle ne corrige que les différences observées. Si la motivation ou les réseaux sociaux ont influencé l'inscription et ne sont pas mesurés, le biais demeure.

La **régression sur discontinuité** s'applique quand l'éligibilité dépend d'un seuil sur un score : par exemple, ménages avec un score de vulnérabilité de 60 ou plus aidés, en dessous non. Les ménages juste au-dessus et juste en dessous du seuil (57-59 et 60-62) sont presque identiques ; la différence de résultats entre eux estime l'effet du programme autour du seuil. Conditions :

le seuil est appliqué strictement, et les ménages ne peuvent pas manipuler leur score pour passer au-dessus.

**Exemple.** Ménages au score 57-59 : 38 % avec une consommation acceptable. Ménages au score 60-62 (aidés) : 51 %. Effet estimé autour du seuil : environ 13 points. Ce résultat vaut pour les ménages proches du seuil, pas forcément pour les plus vulnérables.

La **série chronologique interrompue** compare la tendance d'un indicateur avant et après le début d'un programme, sur de nombreux points de mesure (par exemple les admissions hebdomadaires en nutrition). Elle est utile quand il n'existe aucun groupe de comparaison, à condition d'avoir assez de points avant l'intervention.

## Approches de contribution : quand l'attribution est impossible

Beaucoup d'effets importants ne se prêtent pas aux méthodes statistiques : un plaidoyer qui change une loi, un renforcement de capacités d'organisations locales, une coordination améliorée. On ne peut pas les comparer à un contrefactuel chiffré. On cherche alors à établir de façon crédible que le programme a **contribué** au changement.

- **Analyse de contribution** (Mayne) : on expose la théorie du changement, on rassemble les preuves pour chaque lien, on examine les explications alternatives, et l'on construit un « récit de contribution » que l'on fait critiquer par des parties prenantes.
- **Traçage de processus** (process tracing) : on teste des hypothèses causales en cherchant des preuves dont la présence ou l'absence permet de confirmer ou d'écarter une explication (par exemple, la reprise mot pour mot d'une note de plaidoyer dans un décret).
- **Récolte des effets** (outcome harvesting) : on identifie d'abord les changements observés chez les acteurs, puis on remonte pour déterminer si et comment le programme y a contribué.
- **Changement le plus significatif** : on collecte des récits de changement auprès des personnes, et des panels sélectionnent les plus significatifs en expliquant pourquoi. Utile pour découvrir des effets non prévus.

> Rigueur ne veut pas dire quantitatif. Une analyse de contribution bien menée, qui examine sérieusement les explications alternatives, est plus crédible qu'une comparaison avant-après chiffrée.

## Choisir la méthode et décider ce qu'il faut prouver

Le choix d'une méthode dépend de quatre questions :

- **Quelle décision ?** Étendre un modèle pilote à l'échelle nationale justifie une preuve forte ; ajuster un programme en cours se contente d'une preuve suffisante.
- **Les preuves existent-elles déjà ?** L'effet du cash sur la consommation alimentaire est très documenté : inutile de le redémontrer avec un ECR dans chaque projet. Il suffit souvent de vérifier que le programme est fidèle au modèle et que les effets intermédiaires sont présents.

- **Qu'est-ce qui est faisable et éthique ?** Existe-t-il un groupe de comparaison sans refuser l'aide ? Un seuil d'éligibilité net ? Une mise en place progressive ?
- **Quel budget et quel délai ?** Une évaluation d'impact sérieuse exige une référence avant le démarrage : elle se conçoit avec le programme, pas à la fin.

Pour la **taille d'échantillon**, une évaluation d'impact doit avoir une puissance suffisante pour détecter l'effet attendu. Détecter un effet de 15 points demande beaucoup moins de ménages qu'un effet de 5 points. Si l'effet attendu est faible et l'échantillon petit, l'étude conclura presque sûrement « pas d'effet significatif », ce qui sera mal interprété comme « pas d'effet ».

Dernier point : présenter honnêtement le **niveau de preuve**. « Les ménages aidés ont amélioré leur consommation de 24 points ; par rapport à une zone de comparaison, l'effet attribuable est estimé à 15 points, sous l'hypothèse de tendances parallèles que les données de deux campagnes précédentes soutiennent » est une phrase qui inspire confiance.

## Exercice d'application : estimer un effet et critiquer la méthode

**Situation.** Un programme de moyens d'existence a soutenu 1 200 ménages. Revenu mensuel médian : ménages aidés 42 USD avant, 71 USD après. Zone voisine non couverte (faute de financement) : 47 USD avant, 58 USD après. Pendant le programme, une autre ONG a distribué des semences dans la moitié des villages de la zone voisine. Le rapport final annonce : « le programme a augmenté les revenus de 69 % ».

**Travail demandé.** 1) D'où vient le chiffre de 69 % et pourquoi est-il trompeur ? 2) Calculez l'effet par différence de différences. 3) Quel est l'effet probable de la distribution de semences sur votre estimation, et que proposez-vous ?

**Corrigé.** 69 % : (71 − 42) / 42, comparaison avant-après qui attribue au programme toute la hausse, y compris celle due au contexte. Différence de différences : (71 − 42) − (58 − 47) = 29 − 11 = 18 USD par mois, soit environ 43 % du revenu initial des ménages aidés. Les semences ont probablement augmenté les revenus de la zone de comparaison, ce qui gonfle sa tendance et sous-estime l'effet du programme (la contamination viole l'hypothèse des tendances parallèles). Proposition : analyser séparément les villages de comparaison sans semences, ou contrôler cette variable ; présenter l'estimation comme une borne basse et indiquer clairement la limite.

## À retenir

- L'impact se mesure par rapport à un contrefactuel ; avant-après n'est pas une mesure d'impact.
- Différence de différences = évolution du groupe aidé − évolution du groupe de comparaison, sous l'hypothèse de tendances parallèles.
- Un ECR en humanitaire est éthique seulement s'il ne prive personne d'une aide disponible : ressources limitées, mise en place progressive, ou comparaison de variantes.
- Régression sur discontinuité : comparer les ménages juste au-dessus et juste en dessous d'un seuil strict d'éligibilité.
- Quand l'attribution est impossible : analyse de contribution, traçage de processus, récolte des effets, changement le plus significatif.

## Quiz

1. Groupe aidé : 30 % avant, 50 % après. Comparaison : 32 % avant, 40 % après. Effet par différence de différences :
   A. 20 points
   B. 12 points
   C. 10 points
   D. 8 points

2. Quelle hypothèse est essentielle à la différence de différences ?
   A. Les groupes sont identiques au départ
   B. Sans le programme, les deux groupes auraient évolué de la même façon
   C. L'échantillon dépasse 1 000 ménages
   D. Le programme a duré plus d'un an

3. Dans quelle situation un tirage au sort est-il éthiquement acceptable en humanitaire ?
   A. Pour refuser une aide vitale disponible à la moitié des éligibles
   B. Quand les ressources ne suffisent de toute façon pas pour tous les éligibles, ou pour l'ordre d'une mise en place progressive
   C. Jamais
   D. Seulement si le bailleur l'impose

4. La principale limite de l'appariement par score de propension est :
   A. Il ne fonctionne qu'avec des ECR
   B. Il ne corrige que les différences mesurées, pas les différences non observées comme la motivation
   C. Il exige un seuil d'éligibilité
   D. Il est interdit par l'OCDE

5. Pour démontrer qu'un plaidoyer a contribué à l'adoption d'une loi, la méthode la plus adaptée est :
   A. Un essai randomisé
   B. Une différence de différences
   C. Une analyse de contribution ou un traçage de processus
   D. Une enquête ménage

**Corrigé du quiz :** 1. B · 2. B · 3. B · 4. B · 5. C

# Semaine 11 : Optimisation des ressources : coût-efficience, coût-efficacité et arbitrages

**Objectifs**

- Appliquer le cadre des 4E (économie, efficience, efficacité, équité) pour juger l'optimisation des ressources d'un programme.
- Réaliser une analyse coût-efficience et coût-efficacité rigoureuse, avec allocation des coûts partagés et analyse de sensibilité.
- Utiliser ces analyses pour arbitrer entre options, sans sacrifier l'équité ni la qualité.

## L'optimisation des ressources : le cadre des 4E

L'**optimisation des ressources** (value for money, VfM) ne signifie pas « le moins cher ». C'est le meilleur usage des ressources pour obtenir les résultats voulus. Le cadre le plus utilisé, issu de la coopération britannique, compte quatre dimensions :

- **Économie** : acheter les intrants de la qualité requise au meilleur prix (prix des kits, salaires, loyers comparés au marché).
- **Efficience** : transformer les intrants en extrants au moindre coût (coût par kit distribué, par ménage enregistré, ratio de transfert).
- **Efficacité** : transformer les extrants en effets (part des ménages qui améliorent leur consommation, taux de guérison).
- **Équité** : atteindre les plus vulnérables et les plus difficiles à atteindre, ce qui coûte souvent plus cher par personne.

On ajoute parfois le **coût-efficacité** comme synthèse : coût par effet obtenu. Les quatre dimensions se tiennent : un programme très économique qui achète des kits de mauvaise qualité, ou très efficient qui évite les zones reculées, n'optimise pas les ressources.

> Atteindre les personnes les plus isolées ou les personnes handicapées coûte plus cher par personne. Un rapport VfM honnête l'explique au lieu de le cacher, et montre que ce surcoût est un choix d'équité.

## Analyse coût-efficience : la méthode

Une **analyse coût-efficience** calcule le coût de production d'un extrant (par exemple le coût par ménage recevant un transfert pendant six mois). Elle exige d'attribuer à l'activité **tous** les coûts qu'elle consomme, pas seulement les coûts directs : personnel du programme, part du personnel support (logistique, finance, RH), bureaux, véhicules, et part des coûts du siège.

Les étapes, inspirées de l'approche d'analyse systématique des coûts développée par des ONG et reprise par l'outil Dioptra :

- **Définir** l'extrant et la période.

- **Rassembler** les dépenses réelles (pas le budget) de toutes les lignes concernées.
- **Allouer** chaque ligne à l'activité selon une clé : 100 % pour un coût exclusivement lié ; part du temps pour le personnel partagé ; part des dépenses directes pour les coûts de support.
- **Diviser** le coût total alloué par le nombre d'extrants.
- **Comparer** avec d'autres programmes ou d'autres périodes, en tenant compte du contexte.

**Exemple.** Programme cash de 6 mois pour 2 000 ménages : transferts 1 080 000 USD ; personnel programme 96 000 ; frais du prestataire de paiement (2 %) 21 600 ; part du support pays 54 000 ; part du siège 42 400. Coût total : 1 294 000 USD. Coût par ménage : 647 USD. Ratio de transfert : 1 080 000 / 1 294 000 = 0,83.

## Analyse coût-efficacité

L'**analyse coût-efficacité** rapporte le coût à un effet : coût par enfant guéri, par ménage sorti de l'insécurité alimentaire, par point de score gagné, par cas de choléra évité. Elle permet de comparer des approches qui visent le même effet.

**Exemple de comparaison.** Deux options pour améliorer la consommation alimentaire de 1 000 ménages sur six mois. Option A, vivres : coût total 520 000 USD, 380 ménages passent à une consommation acceptable (par rapport au contrefactuel). Option B, cash : coût total 450 000 USD, 410 ménages. Coût par ménage ayant progressé : A = 1 368 USD ; B = 1 098 USD. B est plus coût-efficace dans ce contexte.

Trois règles d'interprétation :

- L'effet doit être **attribuable** (semaine 10) : diviser le coût par une amélioration avant-après surestime ou sous-estime grossièrement le coût-efficacité.
- Les comparaisons ne valent qu'entre options visant le **même effet**, dans des contextes comparables.
- Les effets **multiples** compliquent la comparaison : le cash peut aussi réduire la dette et soutenir les marchés. Dans la santé, on utilise des mesures synthétiques (années de vie corrigées de l'incapacité, DALY), rarement disponibles dans les autres secteurs.

Le **retour social sur investissement** (SROI), qui convertit tous les effets en argent, est séduisant mais fragile : les valeurs monétaires attribuées aux effets sociaux sont très discutables. Il faut lire ces ratios avec prudence et toujours examiner les hypothèses.

## Économies d'échelle, coûts fixes et analyse de sensibilité

Le coût unitaire dépend fortement de l'**échelle**. Les coûts fixes (équipe de coordination, système de paiement, bureau) se répartissent sur plus de bénéficiaires quand le programme grandit.

**Exemple.** Coûts fixes : 150 000 USD ; coût variable par ménage : 400 USD. Pour 500 ménages : (150 000 + 200 000) / 500 = 700 USD par ménage. Pour 2 000 ménages : (150 000 + 800 000) / 2 000 = 475 USD. Comparer le coût unitaire d'un projet pilote à celui d'un programme à grande échelle est donc injuste ; on compare à échelle comparable, ou l'on modélise le coût à l'échelle visée.

L'**analyse de sensibilité** teste la robustesse des conclusions quand les hypothèses changent : que devient la comparaison si l'effet de l'option B est 20 % plus faible que prévu ? si les frais de paiement doublent ? Si la conclusion change pour une variation réaliste, il faut le dire : la décision repose alors sur une hypothèse à vérifier.

Le **seuil de rentabilité** d'un investissement est un autre outil utile : un système de paiement mobile qui coûte 60 000 USD à installer et économise 12 USD par ménage et par mois est amorti à partir de 5 000 ménages-mois (60 000 / 12), par exemple 1 000 ménages pendant 5 mois.

## Présenter l'optimisation des ressources aux bailleurs

Les bailleurs demandent de plus en plus une section VfM dans les propositions et les rapports. Une section solide :

- Présente des **indicateurs pour chacun des 4E**, avec leur valeur et une comparaison (période précédente, autre zone, référence externe).
- Montre les **mesures concrètes** prises : achats groupés, accords-cadres, partage de bureaux et de véhicules avec d'autres projets, paiement électronique, réduction des pertes.
- Explique les **surcoûts justifiés** : zones difficiles, accessibilité, sécurité, qualité, localisation (soutien aux partenaires).
- Reconnaît les **limites** : effets non attribuables, données de coûts incomplètes.

Les pièges à éviter : présenter des économies de court terme qui dégradent la qualité (moins de suivi, matériaux moins durables) ; comparer des coûts unitaires de contextes très différents ; afficher un ratio de transfert élevé obtenu en supprimant le suivi et la redevabilité.

> L'optimisation des ressources est aussi un argument pour les coûts de support : un programme sans logistique, sans finance et sans MEAL solides n'est pas efficient, il est risqué.

## Arbitrer : efficience, équité et qualité

Les analyses de coût éclairent les décisions, elles ne les prennent pas. Les arbitrages courants :

- **Couverture contre profondeur** : aider 10 000 ménages avec un transfert qui couvre 50 % du panier, ou 6 000 avec un transfert qui couvre 85 % ? Si un transfert partiel ne permet pas d'atteindre l'effet visé (par exemple éviter les stratégies d'adaptation négatives), aider plus de monde moins bien n'est pas plus efficace.
- **Zones faciles contre zones difficiles** : le coût par personne peut doubler dans les zones reculées, où les besoins sont souvent les plus sévères.
- **Court terme contre durabilité** : réparer un point d'eau coûte moins cher que former un comité de gestion et mettre en place un système de recouvrement des coûts, mais le point d'eau risque de retomber en panne.

Une méthode pour décider est de combiner l'analyse coût-efficacité et des **critères explicites** : seuil minimal de qualité (standards Sphère), priorité aux plus sévères, plafond de coût par personne au-delà duquel on cherche une autre approche. La décision et ses raisons sont écrites, pour pouvoir être expliquées aux personnes, aux équipes et aux bailleurs.

## Exercice d'application : comparer deux options de distribution d'eau

**Situation.** Pour fournir 15 litres d'eau par personne et par jour à 12 000 personnes pendant 12 mois, deux options. Option A, camions-citernes : coûts fixes 30 000 USD, coût variable 0,012 USD par litre. Option B, réhabilitation de deux forages et réseau : investissement 210 000 USD, fonctionnement 2 500 USD par mois. La durée de vie des forages réhabilités est d'au moins cinq ans.

**Travail demandé.** 1) Calculez le coût de chaque option sur 12 mois. 2) Calculez le coût par personne et par an. 3) À partir de combien de mois l'option B devient-elle moins chère ? 4) Quelle recommandation, et quelles conditions ?

**Corrigé.** Volume : 12 000 × 15 × 365 = 65 700 000 litres par an. Option A : 30 000 + 0,012 × 65 700 000 = 30 000 + 788 400 = 818 400 USD, soit 68,2 USD par personne et par an. Option B : 210 000 + 12 × 2 500 = 240 000 USD, soit 20 USD par personne la première année. Seuil : coût mensuel du camionnage ≈ 788 400 / 12 = 65 700 USD ; B est moins chère dès que 30 000 + 65 700 × m > 210 000 + 2 500 × m, soit m > 180 000 / 63 200 = 2,85 mois. Recommandation : B, avec du camionnage pendant les travaux (2 à 3 mois), à condition que l'hydrogéologie confirme le débit et la qualité de l'eau, et qu'un comité de gestion et un financement du fonctionnement soient prévus au-delà du projet.

## À retenir

- 4E : économie, efficience, efficacité, équité ; le moins cher n'est pas l'optimal.
- Coût-efficience = coût total alloué (y compris support et siège) / extrants.
- Coût-efficacité = coût / effet attribuable ; comparer seulement des options visant le même effet.
- Les coûts fixes rendent les petits programmes plus chers par personne : comparer à échelle comparable.
- Analyse de sensibilité : si la conclusion change pour une variation réaliste, la décision repose sur une hypothèse à vérifier.

## Quiz

1. Atteindre les personnes handicapées et les zones reculées augmente le coût par personne. Dans le cadre des 4E, ce surcoût relève de :
   A. L'économie
   B. L'équité
   C. L'efficience
   D. Une erreur de gestion

2. Coûts fixes 100 000 USD, coût variable 300 USD par ménage, 1 000 ménages. Coût par ménage :
   A. 300 USD
   B. 400 USD
   C. 1 300 USD
   D. 100 USD

3. Option A : 600 000 USD pour 400 enfants guéris. Option B : 500 000 USD pour 250 enfants guéris. Laquelle est la plus coût-efficace ?
   A. B, car elle coûte moins cher au total
   B. A : 1 500 USD par enfant guéri contre 2 000 pour B
   C. Elles sont équivalentes
   D. On ne peut pas comparer

4. Pourquoi faut-il inclure une part des coûts de support et du siège dans une analyse coût-efficience ?
   A. Pour gonfler les chiffres
   B. Parce que l'activité consomme réellement ces ressources ; les ignorer sous-estime son coût
   C. Parce que le bailleur paie tout
   D. Ce n'est pas nécessaire

5. Un investissement de 40 000 USD économise 8 USD par ménage et par mois. Il est amorti à partir de :
   A. 500 ménages-mois
   B. 5 000 ménages-mois
   C. 320 000 ménages-mois
   D. 50 000 ménages-mois

**Corrigé du quiz :** 1. B · 2. B · 3. B · 4. B · 5. B

# Semaine 12 : Gestion adaptative, apprentissage et rapports : diriger par les preuves

**Objectifs**

- Installer une gestion adaptative : revues régulières, règles de décision et révision documentée du programme.
- Construire un système d'apprentissage organisationnel qui capitalise et réutilise les leçons, au-delà des rapports.
- Rédiger des rapports aux bailleurs honnêtes et convaincants, et présenter un dossier MEAL complet comme projet final.

## La gestion adaptative

Les contextes humanitaires changent plus vite que les cadres logiques. La **gestion adaptative** consiste à ajuster de façon délibérée et documentée les activités, les méthodes ou les cibles en fonction des données et du contexte, tout en gardant l'objectif. Elle n'est pas de l'improvisation : elle repose sur des hypothèses explicites, un suivi qui les teste, et des moments de décision prévus.

Ses éléments :

- **Des hypothèses de travail** clairement énoncées dans la théorie du changement (« les femmes pourront se rendre aux points de retrait »).
- **Des indicateurs de contexte et d'alerte** qui les testent (temps de trajet déclaré, incidents de sécurité sur les routes).
- **Des règles de décision** écrites à l'avance : « si plus de 20 % des transferts ne sont pas retirés dans les 10 jours, on ouvre des points de retrait mobiles ».
- **Des revues régulières** (trimestrielles, ou mensuelles en crise aiguë) avec l'autorité de décider.
- **Un accord avec le bailleur** sur ce qui peut changer sans demande formelle (activités, méthodes) et ce qui exige un avenant (objectifs, budget global, zones).

Les bailleurs sont souvent plus ouverts qu'on ne le croit à une adaptation bien argumentée. Ce qu'ils refusent, ce sont les changements non annoncés découverts au rapport final.

> Un projet qui n'a rien changé en 18 mois dans un contexte volatile n'a probablement pas regardé ses données, ou n'a pas osé en parler.

## Les revues de programme et l'après-action

La **revue trimestrielle** d'un programme réunit l'équipe programme, le MEAL, la logistique, la finance et, si possible, des représentants des partenaires et des communautés. L'ordre du jour type :

- Progrès des indicateurs clés par rapport aux cibles et aux tendances.
- Exécution budgétaire et rythme de dépense.
- Retours des personnes et des partenaires : thèmes principaux, réponses apportées.
- Contexte et hypothèses : ce qui a changé.
- Décisions : quoi, qui, quand. Et suivi des décisions de la revue précédente.

L'**examen après action** (after action review), issu des armées puis largement adopté, est un outil court (une à deux heures) à utiliser après un événement important : une distribution, une évacuation, une première vague de réponse. Quatre questions : qu'avions-nous prévu ? que s'est-il réellement passé ? pourquoi y a-t-il un écart ? que ferons-nous différemment la prochaine fois ? Il fonctionne si l'on parle de faits, sans recherche de coupables, et si la dernière question produit des actions avec un responsable.

Les leçons doivent être **spécifiques** et **actionnables** : « la communication doit être améliorée » n'est pas une leçon ; « annoncer les dates de distribution par radio et par message vocal au moins 5 jours avant, car 40 % des ménages n'avaient pas reçu l'information » en est une.

## L'apprentissage organisationnel

Les organisations humanitaires produisent beaucoup de rapports et apprennent peu : les mêmes erreurs se répètent d'une crise à l'autre, souvent parce que le personnel tourne vite et que les leçons restent dans des documents que personne ne lit. Un système d'apprentissage efficace :

- **Définit des questions d'apprentissage** prioritaires pour l'organisation (par exemple : « comment maintenir la qualité du suivi quand l'accès est intermittent ? ») et oriente le MEAL de plusieurs projets pour y répondre.
- **Capitalise** sous des formes utilisables : fiches de deux pages, listes de contrôle, modèles de documents, plutôt que des rapports longs.
- **Fait circuler** les leçons au bon moment : au démarrage d'un nouveau projet, on lit les leçons des projets similaires (revue des leçons en phase de conception).
- **Valorise** le signalement des échecs, par exemple par des sessions régulières où les équipes présentent ce qui n'a pas marché et ce qu'elles en ont appris.
- **Partage** avec le secteur : évaluations publiées, contributions aux groupes de travail des clusters, études de cas.

Le rôle du responsable MEAL change alors : d'un producteur de données pour les bailleurs, il devient un **animateur de la réflexion** qui aide les équipes à poser de bonnes questions, à lire les données et à décider.

## Rédiger des rapports aux bailleurs

Un **rapport narratif** de qualité suit la structure du bailleur et répond à trois exigences : **exactitude** (chiffres cohérents avec le rapport financier et les rapports précédents), **analyse** (pas seulement ce qui a été fait, mais ce qui a changé et pourquoi) et **honnêteté** (écarts et difficultés expliqués).

Pour chaque résultat, une structure simple fonctionne : **cible, réalisé, écart, explication, conséquence, mesure**. Par exemple : « Cible : 3 000 ménages. Réalisé : 2 450 (82 %). Écart dû à la fermeture de la route de X pendant 7 semaines. Conséquence : retard de deux cycles de transferts dans 4 villages. Mesure : paiement mobile mis en place en mai ; rattrapage prévu d'ici août, sans coût supplémentaire. »

Erreurs à éviter :

- Des chiffres d'extrants présentés comme des effets (« 12 000 personnes sensibilisées ont amélioré leurs pratiques »).
- Des doubles comptages entre activités (la même personne comptée dans chaque activité) ; il faut distinguer bénéficiaires uniques et participations.
- Des réussites individuelles présentées comme représentatives ; une histoire illustre, elle ne prouve pas.
- L'absence de mention des plaintes reçues et des réponses apportées.

> Le bailleur apprend tôt ou tard les problèmes. Il vaut mieux qu'il les apprenne par vous, avec l'analyse et les mesures prises, que par une évaluation ou un audit.

## La clôture du programme et la durabilité

La **clôture** se prépare dès la conception (stratégie de sortie) et se met en œuvre au moins trois mois avant la fin. Elle comprend :

- **Information** des communautés et des autorités sur la date et les modalités de fin, et orientation vers d'autres services.
- **Transfert** des actifs (points d'eau, équipements) selon les règles du bailleur, avec des accords écrits avec les bénéficiaires de ces actifs.
- **Passation** aux partenaires, aux services techniques ou aux comités, avec un renforcement de leurs capacités réellement planifié et budgété.
- **Enquête finale** et évaluation, puis restitution aux communautés.
- **Archivage** des données et documents, et destruction sécurisée des données personnelles selon la politique de conservation.

La **durabilité** se juge aux conditions réunies : une institution ou une communauté qui assume la responsabilité, des ressources pour le fonctionnement (budget public, cotisations, recouvrement des coûts), des compétences, et une demande réelle des usagers. Sans ces conditions, il vaut mieux le reconnaître et chercher un relais (autre financement, autre acteur) plutôt que d'annoncer une durabilité qui n'existera pas.

## Projet final : le dossier MEAL complet

Le projet final de ce Mini MBA consiste à présenter un **dossier MEAL complet** pour un programme réel ou réaliste de votre choix. Il rassemble tout ce que vous avez appris en douze semaines :

- Une **note d'analyse des besoins** de deux pages : sources, sévérité, groupes prioritaires, lacunes de couverture, options de réponse comparées.
- Une **théorie du changement** et un **cadre logique** avec indicateurs, bases, cibles calculées et hypothèses.
- Une **stratégie de ciblage** et, le cas échéant, le calcul de la valeur des transferts.
- Un **plan MEAL** avec deux fiches de référence d'indicateurs, le plan d'échantillonnage et le calcul de taille d'échantillon de l'enquête de référence.
- Un **mécanisme de plaintes et de retours** : canaux, délais, filière sensible, indicateurs.
- Un **plan d'évaluation** : type, questions, méthode d'estimation de l'effet et ses conditions de validité.
- Une **analyse VfM** : indicateurs des 4E et une comparaison d'options.
- Un **calendrier de revues** et trois règles de décision adaptatives.

Critères de qualité : cohérence entre les parties, calculs justes et vérifiables, réalisme (budget MEAL de 3 à 10 %), place réelle donnée aux personnes, et honnêteté sur les limites. Les études de cas de l'examen du diplôme reprennent exactement ces compétences.

## Exercice d'application : écrire une règle de décision et un paragraphe de rapport

**Situation.** Programme de cash pour 4 000 ménages, paiement mobile. Au troisième mois : 18 % des transferts ne sont pas retirés après 10 jours (6 % au premier mois) ; dans 3 villages sur 20, ce taux dépasse 50 %. Les retours signalent des agents sans liquidités et une route coupée par des groupes armés. Cible du trimestre : 12 000 transferts retirés ; réalisé : 9 840.

**Travail demandé.** 1) Écrivez une règle de décision adaptative pour ce problème. 2) Rédigez le paragraphe du rapport trimestriel sur ce résultat, selon la structure cible, réalisé, écart, explication, conséquence, mesure.

**Corrigé.** Règle : « Si plus de 15 % des transferts d'un village ne sont pas retirés 10 jours après l'envoi, l'équipe vérifie sous 72 heures la liquidité des agents et l'accès ; si la cause est confirmée, un point de retrait mobile ou un agent supplémentaire est organisé avant le cycle suivant, et le bailleur est informé dans le rapport mensuel. » Paragraphe : « Cible : 12 000 transferts retirés. Réalisé : 9 840 (82 %). L'écart provient de 3 villages où plus de la moitié des transferts n'ont pas été retirés, à cause du manque de liquidités des agents et de la coupure de la route de Y. Conséquence : environ 2 160 transferts en attente, et un risque de recours à l'endettement pour les ménages concernés. Mesures : deux points de retrait mobiles depuis la semaine 12, accord avec l'opérateur pour approvisionner les agents, information des ménages par message vocal ; les transferts restent disponibles et seront suivis jusqu'au retrait complet. »

## À retenir

- Gestion adaptative : hypothèses explicites, indicateurs d'alerte, règles de décision écrites, revues régulières, accord avec le bailleur.
- Examen après action : prévu, réalisé, pourquoi l'écart, que faire autrement ; sans recherche de coupables.
- Une leçon est spécifique et actionnable, sinon ce n'est qu'une impression.
- Rapport : cible, réalisé, écart, explication, conséquence, mesure ; bénéficiaires uniques distincts des participations.
- La durabilité se juge à des conditions réunies (responsable, ressources, compétences, demande), pas à une déclaration.

## Quiz

1. Quelle formulation est une leçon apprise utile ?
   A. La coordination doit être renforcée
   B. Annoncer les distributions par radio et message vocal 5 jours avant, car 40 % des ménages n'étaient pas informés
   C. Il faut mieux communiquer
   D. Le projet a été un succès

2. Lors d'un examen après action, la règle principale est de :
   A. Identifier le responsable de l'erreur
   B. Parler des faits sans chercher de coupables, et finir par des actions avec un responsable
   C. Rédiger un long rapport
   D. Inviter le bailleur

3. Un rapport compte 5 000 participants à la formation et 5 000 à la distribution, soit « 10 000 bénéficiaires ». L'erreur est :
   A. Un double comptage possible : il faut distinguer bénéficiaires uniques et participations
   B. Aucune erreur
   C. Une erreur de calcul de pourcentage
   D. Une erreur de taux de change

4. Quel changement exige généralement un avenant formel avec le bailleur ?
   A. Changer de méthode de sensibilisation
   B. Modifier l'objectif spécifique ou les zones d'intervention
   C. Ajuster le calendrier d'une activité de deux semaines
   D. Changer l'ordre du jour d'une réunion

5. La clôture d'un programme se prépare :
   A. Le dernier mois
   B. Dès la conception avec une stratégie de sortie, et se met en œuvre au moins trois mois avant la fin
   C. Après le rapport final
   D. Seulement si le bailleur le demande

**Corrigé du quiz :** 1. B · 2. B · 3. A · 4. B · 5. B