

# Recommandations pour la collecte de données et l'échantillonnage :

## Programme International 2025-2028

---

### 1. Introduction

Le Programme dans le Sud Global du Programme International 2025-2028 d'Interaction comprend un total de 35 indicateurs différents. Des fiches spécifiques ont été élaborées pour chacun de ces indicateurs. Cependant, ces fiches n'abordent pas les problématiques générales sur la collecte de données et l'échantillonnage, qui sont donc abordées dans le présent document.

L'objectif principal de ces recommandations pour la collecte de données et l'échantillonnage est d'assurer la qualité des données nécessaires à nos rapports, mais surtout de favoriser l'amélioration de nos projets et organisations. La qualité des données peut être définie selon cinq standards <sup>1</sup>:

- **Validité** : les données mesurent efficacement les indicateurs prévus.
- **Intégrité** : les données sont exemptes d'erreurs, de doublons ou de valeurs manquantes.
- **Précision** : les données présentent un niveau de détail adéquat, incluant une désagrégation appropriée (par âge, genre, etc.).
- **Fiabilité** : les méthodes de collecte de données sont cohérentes afin d'assurer leur stabilité et leur reproductibilité.
- **Actualité** : les données sont disponibles en temps voulu et mises à jour.

*“Les données que vous collectez ne seront jamais exemptes de biais. Par conséquent, vous devez déterminer, avec l'aide de vos partenaires, quelle qualité et quelle quantité de données sont « suffisamment bonnes » pour vos besoins en matière de prise de décision, d'apprentissage et de responsabilité.” (Traduit de MEAL DPro (2019), p. 74)*

Bien que ces recommandations soient principalement destinées aux partenaires du Programme International, nous pensons qu'elles pourraient également être utiles à d'autres membres d'Interaction confrontés à des défis similaires en matière de collecte de données.

Si vous avez des questions ou si vous relevez des incohérences dans ce document, n'hésitez pas à contacter Interaction. Nous espérons que ces recommandations nous aideront à améliorer la qualité de nos données.

---

<sup>1</sup> Adapté de: A Guide to the MEAL DPro (2019) by Catholic Relief Services, the Humanitarian Leadership Academy, and Humentum

## 2. Théorie du Changement et Cadre Logique

Une Théorie du Changement approfondie a été élaborée avec la participation de tous les membres d'Interaction impliqués dans le Programme International 2025-2028. Cette Théorie du Changement a ensuite été utilisée pour construire le Cadre Logique commun du programme. Le tableau suivant présente un aperçu de la structure du cadre logique, illustrant les relations entre les Résultats, les Objectifs spécifiques et la Finalité, avec des exemples d'indicateurs respectifs :

### Cadre Logique du Programme International 2025-2028

		Type d'indicateur	Exemple(s)
<b>Finalité/Impact</b>	<b>Sphère d'intérêt:</b> La finalité représente les changements à long terme et à grande échelle qui dépendent de plusieurs objectifs spécifiques et sont influencés par des facteurs et des acteurs externes.	Indicateur qualitatif	Récits collectés avec la technique du Changement le Plus Significatif (CPS)
↑			
<b>Objectif spécifique</b>	<b>Sphère d'influence:</b> Bien que nous n'ayons pas de contrôle direct sur l'objectif spécifique, nous pouvons l'influencer à travers nos résultats.	Principalement des indicateurs sous forme de pourcentage (%)	% de paysannes et paysans augmentant leur production / % de la population éprouvant une insécurité alimentaire modérée ou grave (FIES)
↑			
<b>Résultat attendu</b>	<b>Sphère de contrôle:</b> Nous avons un contrôle direct sur le résultat attendu par nos activités.	Principalement des indicateurs sous forme numérique (#)	# de paysannes et paysans formés
↑			
↑	Activités & Inputs		Planifier, financer et organiser des formations pour les paysannes et paysans

### 3. Indicateurs numériques et en pourcentage

20 indicateurs sont formulés comme des indicateurs numériques (principalement des indicateurs de résultats), et 11 comme des indicateurs en pourcentage (principalement des indicateurs d'objectif spécifiques). Les indicateurs numériques disposent de fiches simplifiées, car ils ne nécessitent ni calculs complexes ni enquêtes. À l'inverse, les indicateurs en pourcentage requièrent des fiches plus détaillées pour expliciter les calculs et les questions d'enquête. Étant majoritairement collectés via des enquêtes, ils nécessitent également des échantillonnages aléatoires rigoureux, sauf si la population du projet est suffisamment restreinte pour interroger tous les participants/ménages (par ex. 50 participants/ménages). Les fiches d'indicateurs sont disponibles ici sur notre site internet : <https://interaction-suisse.ch/notre-impact/>.

#### ➤ Suivi du nombre de personnes atteintes – Indicateurs OS 1-3

En complément des indicateurs de Résultat et d'Objectif spécifique, nous utilisons un indicateur de suivi du nombre de personnes atteintes. Pour suivre efficacement le nombre de personnes uniques atteintes chaque année par Objectif spécifique, nous recommandons que les partenaires maintiennent une base de données centrale. Cette base de données devrait permettre :

- De répartir les participants du projet par âge et par genre ;
- De renseigner les indicateurs numériques (par ex. : # d'agriculteurs formés en agroécologie) ;
- D'éviter les doubles comptages.

Pour chaque Objectif spécifique, la base de données pourrait être organisée sous forme de tableau comme suit :

OS 1	Homme	Femme	>18 ans	<18 ans	Consultation, dépistage ou soin prodigué	Participant à un groupe de soutien	Etc.
Personne 1	0	1	0	1	1	1	
Personne 2	1	0	0	1	0	1	
Personne 3	0	1	0	1	1	1	
Etc.							


Note: 1 = Oui; 0 = Non

Cette base de données pourrait être construite dans une feuille Excel ou dans un logiciel de suivi spécifique. Elle devrait faciliter la définition d'une base d'échantillonnage (voir ci-dessous). Cependant, dans les projets de grande envergure et complexes, le maintien d'une telle base de données—et par conséquent d'une base d'échantillonnage—peut devenir quasiment impossible.

#### ➤ Échantillonnage aléatoire :

L'échantillonnage aléatoire est une méthode statistique permettant de sélectionner un sous-ensemble de participants/ménages du projet (c'est-à-dire un échantillon) à partir de la population du projet, de

manière que chaque individu ait une chance égale d'être choisi. Cette méthode garantit que l'échantillon est représentatif de la population du projet, réduisant ainsi les biais et permettant de tirer des conclusions généralisables sur la population du projet à partir de l'échantillon.


 Les généralisations faites sur la population du projet ne peuvent cependant pas être considérées comme représentatives de la population entière <sup>2</sup>.

Quelques étapes clés sont suivies pour déterminer un échantillon aléatoire. Premièrement, il faut définir la population du projet, l'unité d'échantillonnage et la base d'échantillonnage :

- **Population du projet** : l'ensemble du groupe sur lequel nous voulons faire une généralisation.
- **Unité d'échantillonnage** : l'élément individuel sélectionné dans la population du projet pour faire partie de l'échantillon et pour lequel nous collecterons des données (généralement des individus ou des ménages).
- **Base d'échantillonnage** : la liste numérotée de toutes les unités uniques composant la population du projet.

Exemple :

- **Population du projet** : 1000 femmes sensibilisées à l'hygiène.
- **Unité d'échantillonnage** : une femme sensibilisée à l'hygiène.
- **Base d'échantillonnage** : une liste comprenant le nom de chaque femme ayant participé à la session de sensibilisation, chacune étant associée à un numéro.

 Dans de nombreuses situations de projet dans le Sud global, il n'est pas faisable de disposer d'une base d'échantillonnage adéquate, car les données démographiques précises n'existent pas. Cependant, l'ensemble du processus de randomisation dépend fortement de la disponibilité d'une base d'échantillonnage. C'est pourquoi nous couvrirons en détail les deux situations dans les pages suivantes.

--- Avec une base d'échantillonnage ---

### 1) Choisissez votre méthode d'échantillonnage aléatoire :

Il existe différentes méthodes pour déterminer un échantillon aléatoire :

---

<sup>2</sup> La population du projet présente souvent des caractéristiques spécifiques (par ex., agricultrices rurales), ce qui la distingue clairement de la population générale (par ex., l'ensemble des habitants d'une région ou d'un pays).



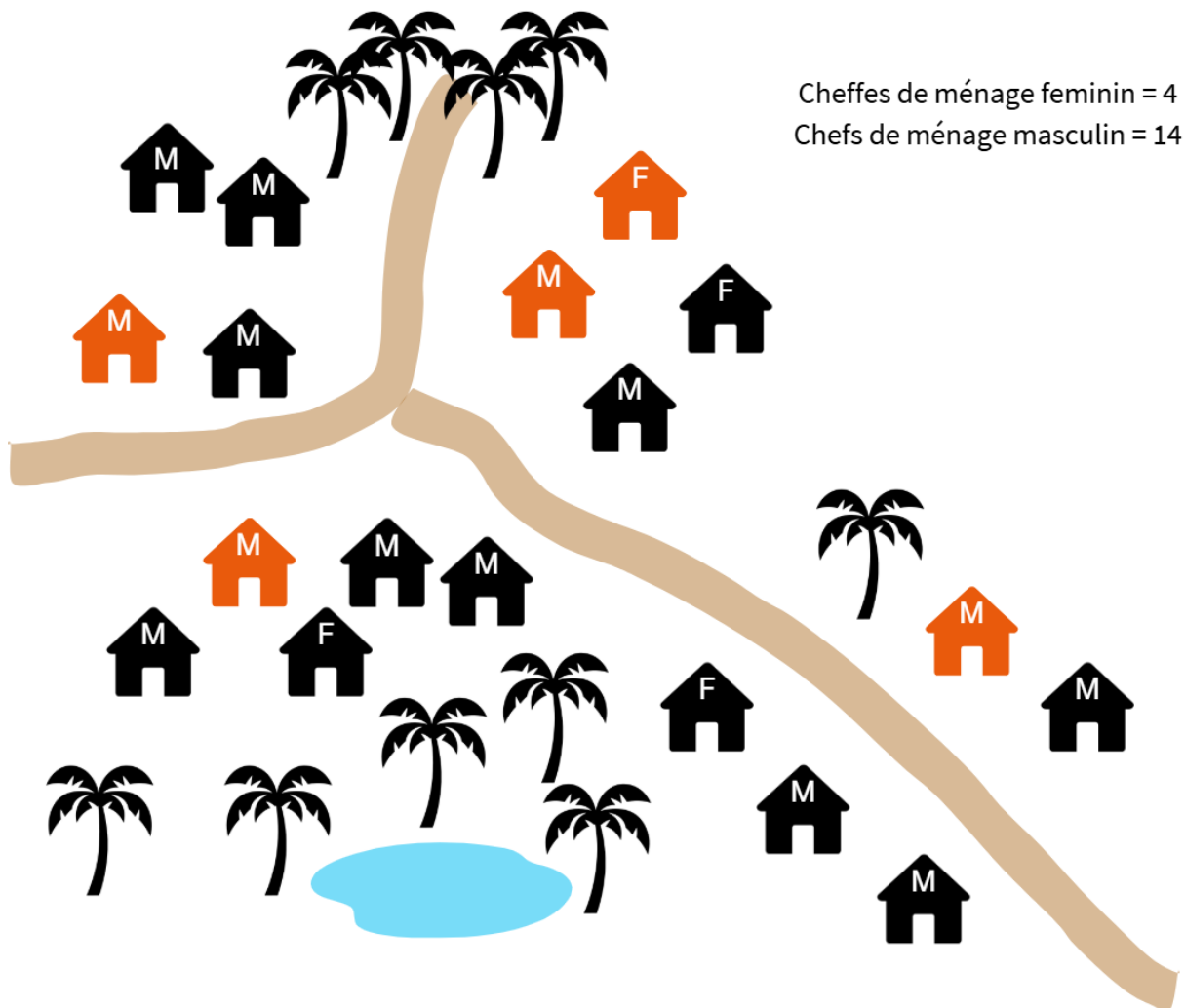
- **Échantillonnage aléatoire systématique :**

Un point de départ est sélectionné de manière aléatoire, puis chaque n-ième membre de la population est choisi.

**Exemple :** sélectionner chaque 5<sup>e</sup> participant sur la liste des participants du projet après avoir choisi un point de départ au hasard.

- **Échantillonnage aléatoire en clusters :**

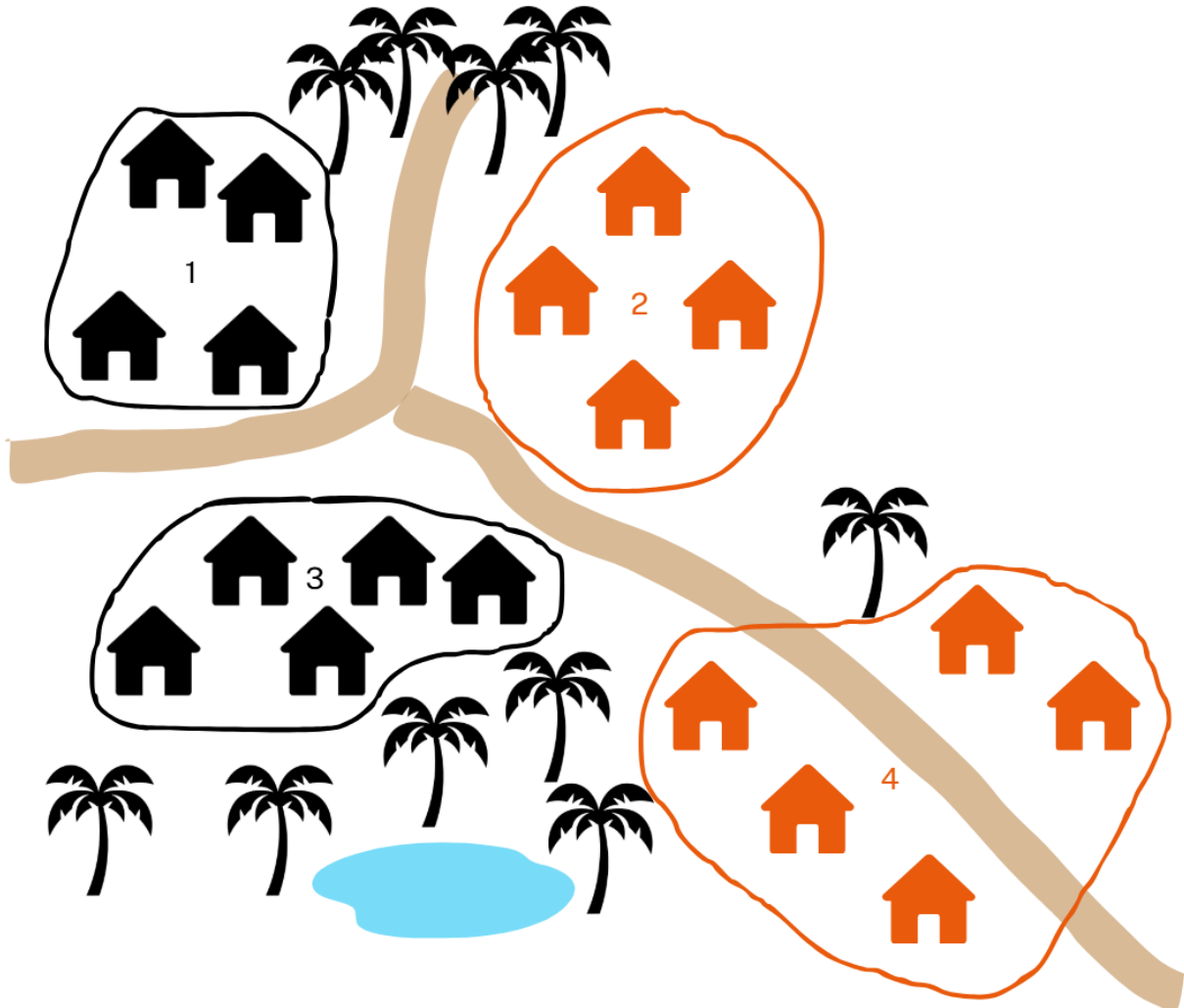
La population est divisée en clusters (généralement sur la base de regroupements géographiques ou naturels). Des clusters entiers sont sélectionnés de manière aléatoire, et tous les membres des clusters sélectionnés sont inclus dans l'échantillon. Si chaque cluster contient trop de personnes, un échantillon représentatif peut être sélectionné au sein de chaque cluster.



Dans cet exemple, 4 chefs de ménage masculin et 1 cheffe de ménage féminin ont été sélectionnés, en respectant leur proportion respective dans la population

- **Échantillonnage aléatoire stratifié :**

La population du projet est divisée en sous-groupes homogènes (strates) selon certaines caractéristiques (par ex., âge, genre). Un échantillon aléatoire est ensuite prélevé dans chaque sous-groupe de manière proportionnelle à sa taille dans la population.



Dans cet exemple, les quartiers (c.-à-d. clusters) N° 2 & 4 ont été sélectionnés aléatoirement. Tous les ménages de ces quartiers seront enquêtés.

## 2) Déterminer la taille de l'échantillon :

Pour qu'un échantillon soit représentatif de la population du projet, il doit être suffisamment grand. Cependant, si la taille de l'échantillon est trop grande, la collecte de données mobilisera du temps et des ressources inutilement. Dans les exemples ci-dessus avec un petit village, les tailles d'échantillon sont sans doute trop faibles pour représenter précisément l'ensemble de la population du projet.

Plusieurs formules existent pour calculer la taille minimale d'un échantillon. Pour nos besoins, nous proposons d'utiliser la formule simplifiée de Yamane :

$$n = \frac{N}{1 + N \cdot e^2}$$

Où :

- n = taille de l'échantillon
- N = taille de la population du projet
- e = niveau de précision/marge d'erreur (généralement 5 %, soit 0,05)

Exemple pour une population de 1 000 femmes sensibilisées à l'hygiène :

$$n = \frac{1000}{1 + 1000 \cdot 0.05^2} = 285.71$$

Ainsi, selon la formule de Yamane, nous devrions interroger **285** femmes pour obtenir un échantillon représentatif. Comme alternative, l'outil de calcul d'IndiKit permet également de déterminer la taille de l'échantillon de manière simplifiée : [calculateur IndiKit](#).

Exemple : Avec un interval de confiance de 95 % et une marge d'erreur de 5 %, la taille d'échantillon requise pour une population de 1 000 est de **278**.

Les deux méthodes donnent des résultats très similaires. Toutefois, les formules de Yamane et d'IndiKit présentent des limites. Par exemple, elles ne tiennent pas compte de la variabilité de la population : plus la population du projet est diversifiée, plus la taille de l'échantillon devra être grande pour assurer le même niveau de précision.

### 3) Sélection aléatoire des unités d'échantillonnage

Exemple : Nous pouvons utiliser la fonction Excel mentionnée ci-dessus pour générer 278 nombres aléatoires. Ensuite, à partir de la liste des 1 000 femmes ayant participé à la séance de sensibilisation, nous identifions les 278 femmes à interroger.

---- Sans base d'échantillonnage ----

#### 1) Choisissez votre méthode d'échantillonnage aléatoire :

Sans base d'échantillonnage, d'autres stratégies doivent être mises en place pour collecter des données pertinentes :

- **Utiliser une liste proxy :**  
Si une liste complète de la population n'est pas disponible, une liste approximative peut être utilisée. Cela peut inclure des registres communautaires, d'institutions publiques ou de prestataires de services qui se rapprochent de la population prioritaire.
- **Méthode du parcours aléatoire :**  
Sélectionner aléatoirement des rues, directions ou ménages (par ex., en lançant un dé ou en tirant au sort pour déterminer quelle direction prendre ou combien de portes sauter).

- **Échantillonnage en boule de neige :**  
Démarrer avec quelques participants sélectionnés aléatoirement (par tout moyen disponible) et leur demander de référer d'autres personnes de la même population. Bien que cela introduise un biais, cette méthode peut être utilisée lorsque les alternatives sont limitées.
- **Échantillonnage aléatoire systématique :**  
Approcher chaque n-ième personne ou utiliser un intervalle de temps aléatoire (par ex., toutes les 10 minutes, interroger la prochaine personne qui passe).
- **Échantillonnage en clusters ou spatial :**  
Diviser la zone géographique en clusters (villages, quartiers, blocs) et en sélectionner certaines de manière aléatoire. Ensuite, interroger tous les membres de ces clusters ou en sélectionner un échantillon aléatoire. En l'absence d'une liste formelle, on peut procéder à un échantillonnage aléatoire basé sur des coordonnées géographiques ou un système de grille. Des outils comme les SIG (Systèmes d'Information Géographique) peuvent aider à identifier des points aléatoires dans une zone spécifique.

Bien que l'échantillonnage aléatoire sans base d'échantillonnage puisse introduire des biais, l'application d'une ou de plusieurs de ces méthodes peut permettre d'approcher une certaine forme d'aléatoire et de réduire les biais.

## 2) Déterminer la taille de l'échantillon :

Pour qu'un échantillon soit représentatif, il doit être suffisamment grand. Toutefois, une taille trop importante engendre des coûts et un temps de collecte excessifs. La taille minimale de l'échantillon peut être calculée avec les mêmes méthodes que ci-dessus (p. 7-8). Cependant, comme la population totale est inconnue, une approximation est utilisée. Cette approximation est particulièrement pertinente pour les grandes populations prioritaires, car les tailles d'échantillon requises restent relativement stables (**ex. : 370 pour une population de 10 000 et 377 pour une population de 20 000**).

## 3) Sélection aléatoire des unités d'échantillonnage :

En l'absence d'une base d'échantillonnage, nous devons simplement commencer la collecte des données en utilisant l'une des méthodes d'échantillonnage suggérées ci-dessus et arrêter une fois que nous avons atteint un nombre suffisant de répondants.

**Exemple :** Nous parcourons aléatoirement un village et interrogeons les ménages jusqu'à atteindre **278 personnes enquêtées**. Ou bien, nous sélectionnons au hasard un nombre suffisant de grappes pour atteindre la taille de l'échantillon.

### ➤ Concernant la désagrégation des indicateurs

---- Avec une base d'échantillonnage ----

La plupart des indicateurs nécessitent idéalement un échantillonnage stratifié pour une répartition selon l'âge (>18, <18) et le genre (Femme/Homme). Cela signifie que nous aurions 4 sous-groupes : Femmes > 18; Femmes < 18; Hommes > 18; Hommes < 18.

Cependant, collecter des indicateurs en pourcentage avec ces 4 sous-groupes peut s'avérer trop complexe pour plusieurs raisons:

1. **Taille d'échantillon élevée :** Pour obtenir des résultats représentatifs avec une marge d'erreur raisonnable (<5 %) pour chaque sous-groupe, des échantillons de grande taille seraient nécessaire, ce qui entraîne une augmentation du temps et des coûts. Voir l'exemple ci-dessous :

	Taille de la population	Taille de l'échantillon *	Marge d'erreur
Hommes > 18	350	183	5%
Femmes > 18	300	169	5%
Hommes < 18	100	80	5%
Femmes < 18	250	152	5%
Somme	1000	584	3%
Population prioritaire	1000	278	5%

\* Calcul effectué avec le calculateur Indikit (IC : 95 % ; ME : 5 %)

Si nous souhaitons une marge d'erreur de 5 % pour chaque sous-groupe, nous devrions interroger 584 personnes, soit plus du double du nombre requis pour obtenir un échantillon représentatif de l'ensemble de la population prioritaire (278).

2. **Complexité des enquêtes ménages :** La plupart des indicateurs en pourcentage sont collectés via des enquêtes ménages, où c'est généralement le chef de ménage qui répond. Cela complique la collecte de données désagrégées selon l'âge et le genre.
3. **Données démographiques limitées :** Si les partenaires ne disposent pas de données démographiques claires de leur population prioritaire, il est impossible de calculer des tailles d'échantillons représentatives pour chaque sous-groupe.

Nous suggérons donc de se concentrer uniquement sur la désagrégation par genre pour les indicateurs en pourcentage. Même dans ce cas, obtenir des données représentatives reste un défi. Par exemple, avec une population de 1 000 personnes, un échantillon représentatif (Intervalle de confiance : 95% et marge d'erreur <5 %) est de 278. Mais si la population est composée à 50 % d'hommes et 50 % de femmes, il faudrait interroger 217 hommes et 217 femmes (434) pour obtenir des résultats représentatifs pour chaque groupe. Cela augmente la taille de l'échantillon de 56 %.

Dans notre cas, où les ressources sont limitées pour le suivi et l'évaluation, la solution la plus faisable est l'échantillonnage stratifié proportionné. Cela signifie que la taille de l'échantillon de chaque sous-groupe est proportionnelle à sa part dans la population :

Stratégie d'échantillonnage proportionné					
1) Déterminer la taille de l'échantillon pour la population prioritaire (IC: 95%; ME: 5%) 2) Déterminer chaque sous-échantillon en fonction de leur proportion dans la population prioritaire : (Taille de l'échantillon pour la population prioritaire / taille de la population prioritaire) × taille du sous-groupe					
	Taille	Proportion	Taille de l'échantillon	Proportion	Marge d'Erreur
Population prioritaire	1000	100%	278	100%	5%
Hommes	750	75%	194,6	75%	6%
Femmes	250	25%	55,6	25%	12%

Cette méthode permet d'obtenir des tailles de sous-échantillon raisonnables qui reflètent la répartition des hommes et femmes dans la population prioritaire. L'échantillon global est ainsi plus représentatif, mais il faut être prudent lors de l'interprétation individuelle des sous-groupes, en particulier lorsque leur taille est réduite, car alors la marge d'erreur est plus élevée.

---- Sans base d'échantillonnage ----

L'absence d'un cadre d'échantillonnage complique grandement la situation, car il est impossible de déterminer la taille de la population prioritaire ou de ses sous-groupes. Comme dans le cas avec une base d'échantillonnage, il faudrait limiter la désagrégation au genre. Une estimation de la répartition des sous-groupes peut être faite en utilisant les données démographiques nationales, par exemple. Toutefois, une prudence accrue est nécessaire lors de l'analyse, car plusieurs biais peuvent être introduits.

➤ **Collecte des données**

La collecte de données relatives aux indicateurs numériques est relativement simple, à condition que leur définition soit claire. Les fiches d'indicateur doivent garantir une collecte de données précise et cohérente. La plupart des indicateurs en pourcentage quant à eux nécessitent des enquêtes quantitatives, et la formulation des questions est essentielle pour collecter des données pertinentes.

Étapes clés pour une collecte de données quantitative de qualité :

**1. Préparer le questionnaire**

- Concevoir le questionnaire en suivant les fiches d'indicateurs pour assurer la clarté et la cohérence.

## 2. Définir la méthode d'échantillonnage

- Sélectionner la méthode d'échantillonnage appropriée et déterminer un échantillon représentatif.

## 3. Sélectionner et former les enquêteurs

- Choisir des enquêteurs qualifiés et leur fournir une formation approfondie.
- S'assurer qu'ils comprennent et maîtrisent chaque question.
- Utiliser des jeux de rôle pour se préparer efficacement.

## 4. Tester le questionnaire

- Réaliser une enquête pilote avec un petit échantillon pour identifier et corriger d'éventuels problèmes.

## 5. Mener les entretiens

- Obtenir le consentement éclairé des répondants en leur expliquant la confidentialité des données.
- Choisir le meilleur moment pour réaliser l'entretien, en tenant compte des disponibilités des répondants.
- Utiliser un langage accessible et instaurer un climat de confiance.
- Insister sur l'importance de leurs réponses et leur garantir que les réponses négatives ou les refus sont acceptables.

## ➤ Nettoyage des données

Le nettoyage des données est le processus d'identification, de correction ou de suppression des erreurs et incohérences dans les données afin d'améliorer leur qualité. Ce processus garantit que les données sont précises, complètes, cohérentes et exploitables pour l'analyse.

Étapes courantes pour le nettoyage des données :

- **Suppression des doublons** : Identifier et supprimer les enregistrements en double afin de garantir l'unicité de chaque point de donnée.
- **Gestion des valeurs manquantes** : Vérifier si les données brutes contiennent des valeurs manquantes. Si ces valeurs posent un problème pour une analyse spécifique, il est possible d'ignorer les lignes ou colonnes concernées.
- **Correction des erreurs** : Identifier et corriger les erreurs dans les entrées de données, telles que les fautes d'orthographe, les valeurs incorrectes ou les incohérences de formatage.
- **Standardisation des données** : Assurer l'harmonisation des formats (formats de date, unités de mesure, etc.) à travers l'ensemble du jeu de données.
- **Filtrage des valeurs aberrantes** : Identifier et décider de conserver ou de supprimer les valeurs extrêmes irréalistes qui pourraient fausser les résultats de l'analyse.

- **Vérification de la cohérence logique :** S'assurer que les réponses suivent une logique réaliste. Par exemple, un garçon de 7 ans ne peut pas avoir répondu à des questions sur sa grossesse.

## 4. Indicateur d'Impact Qualitatif : Changement le Plus Significatif (CPS)

Le programme international prévoit d'utiliser la technique du **Changement le Plus Significatif (CPS)**<sup>4</sup> pour fournir des informations supplémentaires sur l'impact du programme. Un guide spécifique sur le CPS a été créé et est disponible sur notre site internet : <https://interaction-suisse.ch/notre-impact/>. Ici, nous nous concentrerons uniquement sur les questions d'échantillonnage liées à cette technique.

### ➤ Échantillonnage non-aléatoire ou ciblé<sup>5</sup>

L'échantillonnage non-aléatoire ou ciblé est une technique de collecte de données qualitatives dans laquelle le spécialiste Suivi-Evaluation sélectionne délibérément des participants en fonction de caractéristiques spécifiques pertinentes pour le projet ou ses objectifs, plutôt que par sélection aléatoire. Étant donné que l'échantillon est sélectionné de manière non aléatoire, les résultats **ne peuvent pas être généralisés** à l'ensemble de la population. Au lieu de chercher des données quantitatives applicables à une large population, cette approche se concentre sur un petit nombre de **cas riches en informations** qui peuvent fournir des **perspectives approfondies sur des questions clés, y compris des découvertes inattendues**.

*"Dans de nombreuses situations, plus peut être appris en étudiant intensément des cas extrêmes ou inhabituels, qu'avec la représentation statistique du cas moyen"*  
(Patton, 1990, p. 170)

Voici quelques formes courantes d'échantillonnage non-aléatoire ou ciblé :

#### Échantillonnage homogène

- **Objectif** : Se concentrer sur un sous-groupe ayant des caractéristiques, expériences ou expertises similaires.
- **Exemple** : Sélectionner uniquement des femmes occupant des fonctions de leadership local pour explorer les défis spécifiques au leadership féminin.

#### Échantillonnage hétérogène

- **Objectif** : Capturer une diversité de points de vue en incluant des participants aux caractéristiques variées.
- **Exemple** : Sélectionner des participants de différentes tranches d'âge, genres et milieux socio-économiques pour explorer un large éventail d'expériences liées aux pratiques d'hygiène.

#### Échantillonnage de cas typiques

- **Objectif** : Sélectionner des participants représentatifs d'un phénomène donné.
- **Exemple** : Choisir quelques agriculteurs représentant la moyenne des performances dans une région pour analyser les effets d'une formation en agroécologie.

<sup>4</sup> Davies, R. & Dart, J. (2005), The 'Most Significant Change' (MSC) Technique: A Guide to Its Use, version 1.0

<sup>5</sup> Patton, M. Q. (1990), *Qualitative Evaluation and Research Methods*. Sage Publications, Newbury Park, CA.

### Échantillonnage de cas extrêmes ou déviants

- **Objectif** : Étudier des cas atypiques ou exceptionnels susceptibles de fournir des perspectives uniques.
- **Exemple** : Sélectionner le 1% des étudiants les plus performants et ceux en échec total pour comprendre les facteurs contribuant à leur réussite ou leur échec.

### Échantillonnage de cas critiques

- **Objectif** : Sélectionner un petit nombre de cas clés qui peuvent fournir des informations déterminantes sur un phénomène spécifique.
- **Exemple** : Identifier une communauté où une intervention WASH a été particulièrement réussie ou a échoué pour en comprendre les facteurs critiques.

### Échantillonnage en boule de neige (ou en chaîne)

- **Objectif** : Identifier des participants par le biais de recommandations des premiers répondants.
- **Exemple** : Demander aux premiers participants de recommander d'autres personnes susceptibles d'apporter des perspectives importantes (positives ou négatives) sur le projet.

Au début du programme, nous ne recommandons pas une forme spécifique d'échantillonnage raisonné. Nous encourageons plutôt les spécialistes Suivi-Evaluation locaux à **tester différentes options adaptées au contexte** pour la collecte des récits. Dans certains cas, des histoires de changement significatives émergent **naturellement** au fil des activités quotidiennes des agents de terrain, plutôt que par une recherche active<sup>6</sup>. L'hypothèse ici est que les **agents de terrain**, ayant un contact régulier avec les participants du projet, identifieront spontanément des histoires de changement pertinentes.

#### ➤ **Groupe de discussion ou entretiens semi-structurés ?**

Les deux méthodes peuvent être utilisées pour la collecte de données qualitatives :

- Les groupes de discussion permettent des interactions entre les participants et peuvent faire émerger des points de vue divergents. Cependant, ils nécessitent de bonnes compétences en facilitation, notamment pour éviter qu'une personne ne monopolise la discussion.
- Les entretiens semi-structurés sont souvent plus adaptés à la technique CPS, car les récits collectés sont personnels.

#### ➤ **Combien de personnes devons-nous interroger ?**

Les enquêtes qualitatives sont souvent critiquées pour leurs petits échantillons. Toutefois, compte tenu des contraintes de temps et de budget, il est irréaliste d'avoir des échantillons aussi larges que ceux des enquêtes quantitatives.

Cela dit, des échantillons larges ne sont pas nécessaires pour obtenir des données qualitatives pertinentes. Des recherches récentes montrent que :

---

<sup>6</sup> Davies, R. & Dart, J. (2005), The 'Most Significant Change' (MSC) Technique: A Guide to Its Use, version 1.0, p. 24

- Trois groupes de discussion (composés de 8 à 12 personnes) suffisent à capturer environ 80% des thèmes pertinents. Ajouter davantage de groupes n'améliore que marginalement les résultats <sup>7</sup>.
- Six entretiens semi-structurés peuvent également fournir des aperçus sur 80% des thèmes clés <sup>8</sup>.

**La Qualité prime sur la Quantité :**

**- Min 3 groupes de discussion de 8-12 personnes**

**- Min 6 interviews individuels**

Toutefois, ces chiffres s'appliquent principalement à des populations prioritaires relativement homogènes. Il est souvent plus efficace d'investir dans la qualité de la facilitation des entretiens et des groupes de discussion que d'en augmenter le nombre.

---

<sup>7</sup> Guest, G., Namey, E., & McKenna, K. (2017). How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field methods*, 29(1), 3-22.

<sup>8</sup> Gandy, K. (2024). How many interviews or focus groups are enough?. *Evaluation Journal of Australasia*, 1035719X241266964.

## 5. Remarques Finales et Ressources Complémentaires

Ce document fournit des recommandations générales pour améliorer la collecte de données et l'échantillonnage. Cependant, ces processus doivent être adaptés au contexte local. Étant donné nos ressources limitées pour le suivi-évaluation, ainsi que la variabilité et l'instabilité des conditions locales, des hypothèses et compromis seront inévitables dans la collecte des données. Ainsi, les données qualitatives et quantitatives collectées dans le cadre du Programme International doivent être interprétées avec prudence et ne doivent pas être considérées comme aussi rigoureuses que les résultats d'une étude scientifique formelle.

Ressources
INTRAC (2020). <a href="#">Qualitative and Quantitative Methods</a> .
INTRAC (2017). <a href="#">Sampling</a> .
Davies, R. & Dart, J. (2005). <a href="#">Guide sur le Changement le Plus Significatif (CPS)</a> .
Humentum (2019). <a href="#">Guide du MEAL DPro suivi, évaluation, redevabilité et apprentissage pour les professionnels du développement</a> .
Gandy, K. (2024). <a href="#">How many Interviews or Focus Groups Are Enough?</a>
Patton, M. (1990). <a href="#">Qualitative Evaluation and Research Methods</a>

### Remerciements

Toute notre gratitude aux relecteurs et relectrices qui ont apporté des contributions précieuses, améliorant considérablement la qualité de ce document, en particulier, Nicola Malacarne, Mikaël Amsing, ainsi que les membres d'Interaction.

---

Envie d'en savoir davantage sur le programme international d'Interaction 2025-2028  
Resilience 360 :

- # Aperçu du Programme – [Allemand](#) & [Français](#)
- # Notre Impact – [Allemand](#) & [Français](#)



---

© Interaction, Parkterrasse 10, 3012 Berne, Switzerland.  
[info@interaction-schweiz.ch](mailto:info@interaction-schweiz.ch); [www.interaction-schweiz.ch](http://www.interaction-schweiz.ch).

IP 25-28 Recommandations SEA, version de mars 2025.

Auteur : Thibaud ROSSEL, Coordonnateur de Programme Suivi, Evaluation et Apprentissage (SEA), Interaction.  
[Thibaud.rossel@interaction-schweiz.ch](mailto:Thibaud.rossel@interaction-schweiz.ch)