

FOOD AND
NUTRITION
TECHNICAL
ASSISTANCE

Guide d'Echantillonnage

Robert Magnani

Food and Nutrition Technical Assistance Project
Academy for Educational Development
1825 Connecticut Ave., NW,
Washington, DC, 20009-5721
Tel: 202-884-8077
Fax: 202-884-8432
Courrier électronique : fanta@aed.org
Site Web: <http://www.fantaproject.org>



Cette publication a bénéficié du soutien fourni par l'Office de la Santé et de la Nutrition, Bureau des Programmes globaux, Agence des Etats-Unis pour le Développement international, aux termes de l'Accord coopératif No. HRN-A-00-98-00046-00, Projet d'Assistance technique pour l'Alimentation et la Nutrition (FANTA), accordé à l'Académie pour le Développement de l'Education. Un soutien complémentaire a été apporté par l'Office de Nourriture pour la Paix, Bureau de Réponse humanitaire. Des versions précédentes du guide ont été élaborées avec un financement du Projet de Suivi de l'Alimentation et de la Nutrition (IMPACT) (Contrat No. DAN-5110-Q-00-0014-00, Commande 16), géré par l'International Science and Technology Institute, Inc. (ISTI). Les opinions exprimées ici incombent aux auteurs et ne reflètent pas forcément les vues de l'Agence des Etats-Unis pour le Développement international.

Publiée en décembre 1997

Publiée en français 2001

Des exemplaires du Guide peuvent être obtenus auprès :

Food and Nutrition Technical Assistance
Project (FANTA)
Academy for Educational Development
1825 Connecticut Avenue, NW,
Washington, DC 20009-5721
Tél : 202-884 8000 Fax : 202-884 8432
Courrier électronique : fanta@aed.org
Site Web : <http://www.fantaproject.org>

Table des matières

- 1. Objectif du Guide**
- 2. Définir les objectifs de mesure**
- 3. Déterminer les tailles nécessaires de l'échantillon**
- 4. Choisir l'échantillon**
- 5. Analyser les données**

Figures

- 3-1: Exemples d'information nécessaire pour déterminer la taille de l'échantillon, indicateurs génériques de l'alimentation infantile et juvénile au Titre II
- 3-2: Valeurs de Z_α et Z_β
- 3-3: Exemples de calculs de la taille de l'échantillon exprimés en tant que proportions
- 3-4: Tailles de l'échantillon nécessaires pour les combinaisons choisies de P_1 et les changements ou différences entre groupes de comparaison devant être détectés (pour $\alpha = 0,95$ et $\beta = 0,80$)
- 3-5: Exemple de calcul de taille d'échantillon pour un indicateur exprimé comme moyenne
- 3-6: Nombres typiques de ménages à contacter pour trouver un individu de référence pour les indicateurs sanitaires génériques du Titre II (en supposant six personnes par ménage)
- 3-7: Exemple de calcul de la taille de l'échantillon pour les enquêtes de suivi pour des indicateurs exprimés en tant que proportions
- 3-8: Exemple de calcul de la taille de l'échantillon pour les enquêtes de suivi pour un indicateur exprimé en tant que moyenne

- 4-1: Etapes de la sélection d'un échantillon systématique aléatoire de grappes avec PPT
- 4-2: Exemple - sélection d'un échantillon systématique aléatoire de grappes avec PPT
- 4-3: Etapes de la sélection d'un échantillon systématique aléatoire de grappes avec probabilité égale
- 4-4: Exemple – sélection d'un échantillon systématique aléatoire de grappes avec probabilité égale
- 4-5: Etapes d'utilisation de la méthode de segmentation pour choisir les ménages de l'échantillon
- 4-6: Exemple d'une grappe hypothétique qui a été divisée en six segments
- 4-7: Carte d'une grappe d'échantillons hypothétiques montrant les points de départ possibles
- 4-8: Exemple de la conception d'un échantillon et de la sélection pour une zone de comparaison
- 5-1: Procédures pour calculer les probabilités d'échantillonnage pour les éléments de l'échantillon (P_i) pour les plans d'échantillonnage par grappe deux stades
- 5-2: Exemple de calcul des probabilités de sélection, poids d'échantillonnage et poids d'échantillonnage standardisés – données hypothétiques

Annexes

- Annexe 1 : Liste des indicateurs génériques du Titre II
- Annexe 2 : Tailles de l'échantillon nécessaires pour les combinaisons choisies de P_1 et changements ou différences entre groupes de comparaison devant être détectés (pour $\alpha = 0,95$ et $\beta = 0,90$)

Remerciements

Le Guide a été rédigé par Robert Magnani, de l'Ecole de Santé publique et de Médecine tropicale de l'Université de Tulane, pour le projet IMPACT. Les auteurs tiennent à faire mention de Janet Rice, Université de Tulane, pour ses commentaires fort utiles sur les avant-projets. Eunyong Chung, de l'Office de la Santé et de la Nutrition du Bureau global de l'USAID, a apporté des connaissances précieuses et un soutien en général pour la mise au point du présent Guide. Le Bureau de Nourriture pour la Paix a joué un rôle très important, encourageant et soutenant l'élaboration du Guide. Bruce Cogill, Anne Swindale et Patrick Diskin, du Projet IMPACT, ont fourni de nombreux commentaires et une assistance importante. Une mention toute spéciale doit être faite de l'éditrice, Dorothy B. Wexler et de la conseillère au collationnement de texte, Stacy Swartwood. Les promoteurs collaborateurs ont joué un rôle essentiel au niveau de la mise au point de ce Guide. Ce Guide leur est consacré.

A propos de cette série

Cette série de Guides d'Indicateurs génériques du Titre II a été mise au point par le Projet d'Assistance technique pour l'Alimentation et la Nutrition (FANTA) et ses projets prédécesseurs (IMPACT, LINKAGES), dans le cadre du soutien USAID des Organismes collaborateurs pour la mise au point de systèmes de suivi et d'évaluation devant être utilisés dans le cadre des programmes du Titre II. Ces guides doivent fournir la base technique pour les indicateurs et la méthode recommandée pour collecter, analyser et notifier les indicateurs génériques mis au point en consultation avec les organisations bénévoles privées (OBP) en 1995/1996.

Voici une liste des guides disponibles:

Agricultural Productivity Indicators Measurement Guide par Patrick Diskin

Anthropometric Indicators Measurement Guide par Bruce Cogill

Food Security Indicators and Framework for use in the Monitoring and Evaluation of Food Aid Programs par Frank Riely, Nancy Mock, Bruce Cogill, Laura Bailey et Eric Kenefick

Household Food Consumption Indicators Measurement Guide par Anne Swindale et Punam Ohri-Vachaspati

Infant and Child Feeding Indicators Measurement Guide par Mary Lung'aho

Sampling Guide (Guide d'Echantillonnage) par Robert Magnani

Water and Sanitation Indicators Measurement Guide par Pat Billig

1. Objectif du Guide

L'échantillonnage a pour but de réduire le coût de la collecte de données sur une population en réunissant une information provenant d'un sous-ensemble au lieu de la population entière. Les enquêtes par sondage sont souvent le moyen le plus faisable de réunir les données nécessaires pour l'évaluation des programmes du Titre II. Ce guide montre comment choisir des échantillons de communautés, de ménages et/ou d'individus pour de telles enquêtes de manière à les utiliser de pair avec des indicateurs adéquats et des plans d'étude d'évaluation pour arriver à des conclusions valides quant à l'efficacité des programmes au Titre II. Le guide souligne l'utilisation des méthodes d'échantillonnage probabiliste, jugées essentielles pour assurer l'objectivité des évaluations du programme. Des estimations des caractéristiques de la population tirées des enquêtes sur sondage faites en suivant les directives suggérées représenteront probablement de manière approximative la « véritable » valeur de la population dans les limites d'une marge d'erreur spécifiée avec une probabilité connue.

Le guide a été rédigé pour des lecteurs disposant d'une connaissance limitée en matière d'échantillonnage. Toutefois, une connaissance fondamentale en statistiques facilitera l'utilisation du guide. Le matériel est présenté étape par étape dans la séquence susceptible d'être adoptée lors de l'évaluation du Titre II. Quatre phases principales sont décrites :

1. Définir les objectifs de mesure de l'enquête

Ici, il s'agit de ce que l'enquête elle-même espère accomplir. Cette phase traite aussi bien la substance de l'étude – à savoir, quels sont les progrès qu'un groupe cible a fait face aux objectifs du projet – que de la question statistique liée au degré de précision des données.

2. Déterminer les conditions liées à la taille de l'échantillon

Cela explique comment calculer les tailles d'échantillon après avoir décidé de *ce qui* est mesuré et du *degré de précision* avec lequel cela doit être mesuré. La méthode consiste en trois grandes étapes. Premièrement, il faut déterminer le nombre total d'éléments de

l'échantillon. Pour cela, des formules sont présentées afin d'identifier le nombre de personnes qu'il faut échantillonner en fonction du critère suivant : est-ce que les progrès doivent être mesurés au niveau des changements de la *proportion* de la population avec telle caractéristique donnée ou des changements de la *moyenne* d'un indicateur donné (par exemple, calories totales consommées par habitant et par jour). Deuxièmement, le nombre total d'éléments doit être converti en nombre de ménages qu'il faut contacter. Troisièmement, le nombre total de ménages doit être traduit en unités pratiques (grappes et sujets de ces grappes) auxquelles rendra visite l'équipe de l'enquête.

3. Choisir l'échantillon

Ici, l'on définit l'échantillonnage probabiliste et l'on explique pourquoi il est recommandé. Ensuite, cette phase explique, étape par étape, les diverses manières dont sont choisis les grappes et les éléments susmentionnés, en fonction des circonstances (surtout si l'on connaît ou non la taille de la grappe). Des suggestions sont également faites pour traiter les problèmes opérationnels.

4. Analyser les données

Ici, l'on traite des questions statistiques liées au calcul des poids et des erreurs

types qui se présentent suite à la combinaison de méthodes utilisées pour choisir les grappes et les éléments. Des formules sont présentées pour le calcul de poids entrant dans diverses combinaisons classiques.

Ce guide est différent d'autres guides d'échantillonnage destinés au personnel de terrain car il apporte de brèves explications sur le bien-fondé des diverses procédures et pratiques d'échantillonnage. En effet, on pense que le personnel de terrain sera mieux équipé pour adapter les procédures aux circonstances locales s'il comprend le bien fondé sous-jacent et la raison d'être de ces procédures. Des exemples à titre d'illustration des calculs et procédures sont donnés tout au long du guide.

Le guide a été rédigé pour traiter des questions d'échantillonnage susceptibles de se présenter lors des évaluations du programme du Titre II, mais aucun document de ce type ne saurait anticiper entièrement toutes les nuances pouvant apparaître lors des applications effectives. Aussi, les utilisateurs du guide devraient-ils savoir à l'avance qu'ils devront, à l'occasion, consulter des personnes ayant une expertise en matière d'échantillonnage.

2. Définir les objectifs de mesure

La première étape de la conception d'une enquête consiste à définir les objectifs de mesure. C'est particulièrement important quand l'enquête doit dégager les données primaires pour les évaluations des programmes. Quand les objectifs sont clairement spécifiés, les questions adéquates peuvent être incluses au protocole d'enquête et un plan d'échantillonnage pertinent pourra être conçu pour ces protocoles. Quand elles sont définies de manière inadéquate, l'enquête et l'évaluation du programme risquent de faire fausse route dès le début.

Pour définir les objectifs de mesure, il faut notamment répondre aux trois questions suivantes:

- Qu'est-ce qui doit être mesuré ?
- Chez qui ?
- Avec quel niveau de précision nécessaire ?

Qu'est-ce qui doit être mesuré ?

La réponse à la question « qu'est-ce qui doit être mesuré ? » est généralement posée en termes de *variables* ou

d'*indicateurs*.¹ Les indicateurs recommandés pour les divers programmes du Titre II sont présentés dans d'autres guides du projet IMPACT et, par conséquent, ils ne sont pas examinés de manière approfondie dans ce rapport. Aux fins d'évaluation, ils se rapportent aux types de résultats ou d'effets que le programme souhaite avoir : par exemple, réduction des niveaux d'enfants souffrant de malnutrition, âgés de six à 59 mois ou augmentations des taux d'enfants allaités exclusivement au sein de moins de six mois.

Il convient de faire attention à deux autres aspects de la question « qu'est-ce qui doit être mesuré ? » : (1) est-ce que les changements dans le temps seront mesurés ou alors les changements au sein du groupe cible comparés à un groupe/région témoin, et (2) quels sont les éventuels facteurs interférents

¹ Dans l'idéal, ce qui est mesuré sera défini par des « tableaux fictifs » pour le rapport final. Ces tableaux indiqueront la manière dont les données mesurées dans le cadre d'une enquête donnée seront utilisées pour l'analyse.

pouvant se présenter lorsqu'on utilise un groupe témoin ou une région témoin.

Le premier aspect (qu'il s'agisse de mesurer les changements dans le temps ou les différences entre les groupes du projet et les groupes témoins) est important parce que les conditions liées à la taille de l'échantillon varieront de manière considérable suivant la formule choisie (voir Chapitre 3 pour une discussion plus approfondie). L'importance du second aspect réside dans le fait que tout facteur interférent pouvant biaiser les résultats de l'analyse de données avec des groupes témoins ou des régions témoins doit être identifié dès le départ. Evidemment, tous les efforts doivent être faits pour choisir une zone témoin qui est aussi analogue que possible à la zone du programme. Mais, vu qu'il existera presque inévitablement certaines différences, toute variable ou tout facteur dont on pense qu'il pourrait influencer les indicateurs de résultat pour l'évaluation doit être spécifié. Ainsi, il pourra être mesuré dans les protocoles d'enquête. Cet aspect est discuté plus dans le détail dans le Guide du Suivi et de l'Évaluation du Projet IMPACT.

Chez qui ?

La question « chez *qui* ? » est la base pour définir (1) la population à laquelle on pourra extrapoler de manière valide les résultats de l'enquête, et (2) la portée des opérations d'échantillonnage et de travail de terrain qui seront entreprises. Ici, plusieurs aspects doivent être pris en compte.

Domaines : Un domaine concerne une population ou un sous-groupe spécifique pour lequel on souhaite obtenir des estimations d'enquête séparées. Pour les évaluations des projets du Titre II, les domaines regroupent généralement, soit

(1) la population générale de la zone ciblée par le projet, soit (2) la sous-population des bénéficiaires du projet. Lorsqu'on utilise des groupes ou des régions témoins, cette sous-population constitue un domaine supplémentaire.

Les domaines doivent être définis dès le départ puisque la taille de l'échantillon doit être déterminée sur la base par domaine (voir Chapitre 3). Par exemple, si l'on souhaite comparer de nouvelles régions du projet à d'anciennes régions du projet, ces deux groupes devront être identifiés comme des domaines séparés. Cela garantira une taille d'échantillon dans chaque groupe suffisante pour faire des comparaisons valides. Le mécanisme visant à diviser la population étudiée en sous-groupes ou domaines séparés est appelé *stratification*.

Désigner une population ou un sous-groupe particulier présentant un intérêt pour l'étude en tant que domaine est la seule manière de vérifier que la taille de l'échantillon est suffisante pour mesurer de manière fiable les changements dans le temps pour les sous-groupes ou les différences entre les groupes de comparaison.

Univers de l'enquête : L'univers concerne la population et/ou la zone géographique pour laquelle des inférences ou déductions peuvent être faites à partir des données de l'enquête. Dans ce guide, l'univers sera normalement la population de la zone géographique couverte par le projet qui est en train d'être évaluée. S'il existe un groupe témoin, son univers sera la population définie géographiquement de non-bénéficiaires à partir desquels

l'échantillon témoin a été choisi. Si les ressources ne suffisent pas pour couvrir l'univers entier de l'enquête, un univers partiel pourra être utilisé – trois des cinq circonscriptions couvertes par un projet, par exemple. Toutefois, cela limite la capacité à généraliser les données de l'enquête (et les résultats de l'évaluation) à un univers plus petit.²

Unités de mesure et répondants : Les unités de mesure sont les personnes *auxquelles* se réfèrent les données de l'enquête et les répondants sont les personnes *à partir desquelles* l'information est obtenue. Les unités de mesure sont généralement définies par les indicateurs. Les indicateurs des programmes du Titre II concernent généralement les ménages, à l'exception des indicateurs sur l'alimentation infantile où ce sont des nourrissons ou enfants de moins de 24 mois qui constituent les unités. Répondants et unités de mesure peuvent être la même chose, mais pas forcément. Ils sont la même chose pour des indicateurs pour lesquels les répondants doivent donner des informations sur eux-mêmes. Ils diffèrent quand l'information est obtenue à partir de ce que l'on appelle des répondants indirects ou *délégues*. Par exemple, l'information sur les indicateurs liés à l'alimentation infantile

² La capacité à généraliser reste limitée, même si les circonscriptions sont choisies aléatoirement. Cela est dû au fait que la randomisation exige un nombre suffisamment important d'unités d'échantillonnage pour dégager des estimations non biaisées. A cet effet, il est utile de choisir un échantillon de circonscriptions qui « représentent » l'ensemble plus grand de circonscriptions si l'on peut démontrer empiriquement que les circonscriptions de l'échantillon sont analogues du point de vue caractéristiques clés et que l'exécution du programme ne favorise pas les circonscriptions choisies pour l'évaluation.

est généralement obtenue auprès de la mère ou de la personne qui s'occupe de l'enfant. Les indicateurs au niveau ménage sont généralement mesurés en obtenant une information de la part d'un informant qui connaît bien la situation du ménage.

Les unités de mesure et les répondants pour chaque indicateur doivent être identifiés dès le début de la conception de l'enquête, car ils influencent la formulation du plan d'échantillonnage ou plan de sondage et la qualité des données de l'enquête collectées.

Quel est le niveau de précision nécessaire ?

Le niveau ou degré de *précision* nécessaire pour une enquête concerne l'ordre de grandeur de l'erreur au niveau des estimations de l'enquête jugé tolérable pour telle ou telle activité. Les enquêtes peuvent être conçues pour fournir des estimations très précises ou uniquement des approximations sommaires. Le degré diffèrera en fonction des ressources disponibles et des utilisations escomptées des données de l'enquête.

Pour les enquêtes conçues pour mesurer le changement dans le temps ou les différences entre des groupes de comparaison, la précision est spécifiée en termes du *plus petit* changement ou de la *plus petite* différence avec le groupe de comparaison qui peut être mesuré de manière fiable. Ce dont il s'agit sera décidé par le concepteur de l'enquête en fonction du niveau de signification statistique et de puissance (voir Chapitre 3, notamment Sections 1 et 3.6). Cela dépendra également des objectifs du programme qui ont été spécifiés pour un indicateur donné.

Pour les enquêtes conçues pour mesurer le changement dans le temps ou les différences entre les groupes de comparaison, la précision est spécifiée en termes du plus petit changement ou de la plus petite différence entre les groupes de comparaison qui peut être mesuré de manière fiable.

La tâche consistant à spécifier la précision de l'enquête pour les évaluations du programme du Titre II est simplifiée quand les objectifs du

programme sont stipulés en termes d'indicateurs objectivement vérifiables, avec des cibles de performance. Par exemple, un programme du Titre II au Mozambique pour la période 1997-2001 stipule une amélioration de l'emmagasinage et de la transformation alimentaire dans les exploitations agricoles d'ici 2001 et spécifie que cette amélioration sera mesurée partiellement au niveau de l'éventuel accroissement de 30% dans l'adoption de meilleures techniques d'emmagasinage d'ici la fin du projet.

3. Déterminer les tailles nécessaires de l'échantillon

1. Facteurs influençant les décisions liées à la taille de l'échantillon

La taille de l'échantillon nécessaire pour une enquête donnée est déterminée par ses objectifs de mesure. Pour les enquêtes conçues pour mesurer soit des changements au niveau des indicateurs dans le temps, soit des différences au niveau des indicateurs entre les régions du projet et les régions témoins, cinq facteurs déterminent la taille de l'échantillon nécessaire pour un indicateur donné pour chaque série d'enquêtes et/ou groupe de comparaison. Les deux premiers facteurs concernent des caractéristiques de population et les trois derniers sont choisis par le concepteur de l'enquête ou le responsable de l'évaluation. Il s'agit des facteurs suivants :

- nombre d'unités de mesures au sein de la population cible
- niveau initial de l'indicateur
- ordre de grandeur du changement ou des différences entre les groupes de comparaison susceptible d'être mesuré de manière fiable

- degré de confiance avec lequel on souhaite pouvoir dire qu'un changement observé ou une différence entre les groupes de comparaison, de l'ordre de grandeur susmentionné, n'aurait pas eu lieu par hasard (niveau de signification statistique), et
- degré de confiance avec lequel on souhaite pouvoir dire qu'un changement effectif ou une différence effective, de l'ordre de grandeur susmentionné, pourra être détecté (puissance statistique).

Aux fins d'exemple, pour une évaluation conçue pour mesurer des changements dans le temps (par exemple, une conception prétest/post-test d'un groupe), probablement que l'on souhaite mesurer une diminution d'un point de pourcentage dans la proportion d'enfants de six à 59 mois qui ont un poids insuffisant avec un degré de confiance de 95% et une puissance de 80%. Si 40% des enfants, d'après les estimations, étaient d'un poids insuffisant au moment de l'enquête initiale, l'objectif serait de

mesurer un changement dans la prévalence des enfants d'un poids insuffisant, passant de 40% à 20% et d'avoir (1) un degré de confiance de 95% montrant que la baisse ne se serait pas produite par hasard et (2) un degré de confiance de 80% sachant que l'on peut détecter un tel changement s'il a effectivement eu lieu (puissance). Les calculs liés à la taille de l'échantillon répondent à deux questions : (1) Combien d'enfants âgés de six à 59 mois (unité de mesure) seront-ils nécessaires pour atteindre les objectifs susmentionnés ?, et (2) Combien de ménages faudra-t-il choisir pour trouver ce nombre d'enfants ?

Pour une évaluation uniquement post-test, comparant les zones du projet et les zones témoins, l'on suppose que la même différence – 20 points de pourcentage – est souhaitée entre les deux pour un indicateur spécifié. La taille de l'échantillon devra être déterminée pour être en mesure de détecter de manière fiable une telle différence entre les deux régions. Le même principe s'applique à la conception prétest et post-test avec zones de traitement et zones témoins, sauf que la taille de l'échantillon devra être fixée pour vérifier une détection fiable de la différence dans le degré de changement.

2. Besoins d'information au départ

Avant de commencer à déterminer la taille de l'échantillon, une information doit être collectée concernant deux aspects :

- la composition des ménages, et
- les niveaux ou taux *escomptés* ou normaux des indicateurs devant être mesurés.

La composition des ménages concerne la proportion du total des ménages pouvant avoir un individu ou des individus dans le sous-groupe en question. (L'utilisation de cette information est expliquée ci-après à la Section 3.1.2 de la page 16 décrivant comment convertir le nombre d'éléments nécessaires en nombre de ménages qu'il faut contacter.) Par exemple, l'indicateur générique du Titre II « Pourcentage de nourrissons recevant des aliments supplémentaires pendant deux semaines après la diarrhée » exigera des estimations de la proportion de ménages susceptibles d'avoir des enfants de moins de 24 mois. La source habituelle d'information pour la composition des ménages est le recensement le plus récent de la population. Dans l'idéal, l'on disposera de données pour la zone cible du programme devant être évaluée. Si tel n'est pas le cas, l'on utilisera des données pour le niveau suivant d'agrégation (par exemple, département, province ou région) ; et si ces données ne sont pas non plus disponibles, il faudra utiliser des données au niveau national.

L'information sur les niveaux ou taux *escomptés* des divers indicateurs qui seront mesurés n'est pas aussi facile à obtenir. Par exemple, pour le pourcentage de nourrissons recevant des aliments supplémentaires après la diarrhée, deux choses doivent être déterminées : (1) la proportion de ces enfants susceptibles d'avoir eu un épisode diarrhéique dans les deux semaines précédant l'enquête, et (2) la proportion susceptible d'avoir reçu des aliments supplémentaires suivant un épisode diarrhéique pendant la période juste avant l'enquête. Les sources possibles d'information sont les enquêtes précédentes qui ont pu être faites dans le

pays ou dans un pays voisin, les données du Ministère de la Santé ou d'autres organismes publics ou encore les meilleures estimations de personnes connaissant bien la situation. La Section 3.3 ci-après fournit des directives sur ce qu'il faudrait faire si aucune source fiable d'information n'est disponible.

La Figure 3-1 fournit des exemples du type de données préliminaires sur la population en question devant être disponibles avant de commencer le travail consistant à fixer la taille de l'échantillon pour une enquête donnée des indicateurs sur l'alimentation du nourrisson et de l'enfant du Titre II.

Figure 3-1 : Exemples d'information nécessaire pour déterminer la taille de l'échantillon, indicateurs génériques de l'alimentation infantile et juvénile au Titre II

- A. Information sur la composition de la population :
1. Nombre moyen de personnes par ménage
 2. Proportion de la population totale qui a :
 - a. Des enfants âgés de moins de 0-59 mois.
 - b. Des enfants âgés de moins de 24 mois.
 - c. Des nourrissons âgés de moins de six mois.
 - d. Des nourrissons âgés de six à 10 mois.
- B. Information sur les niveaux ou taux *escomptés* de la population cible :
1. Proportion d'enfants âgés de six à 59 mois qui sont chroniquement.
 2. Proportion d'enfants âgés de six à 59 mois qui ont un poids insuffisant.
 3. Proportion de nourrissons de moins de six mois qui ont été allaités dans l'heure - huit heures suivant la naissance.
 4. Proportion de nourrissons de moins de six mois allaités uniquement.
 5. Proportion de nourrissons âgés de six à 10 mois qui reçoivent des aliments complémentaires.
 6. Proportion d'enfants de moins de 24 mois qui ont souffert d'un épisode diarrhéique pendant les deux semaines précédant l'enquête.
 7. Proportion d'enfants de moins de 24 mois qui ont souffert d'un épisode diarrhéique pendant les deux semaines précédant l'enquête et qui ont reçu une alimentation continue.
 8. Proportion d'enfants de moins de 24 mois qui ont souffert d'un épisode diarrhéique pendant les deux semaines précédant l'enquête et qui ont reçu des aliments supplémentaires.
-

3. Calculs de la taille de l'échantillon

3.1 Calculer le nombre d'éléments de l'échantillon et de ménages : Deux étapes permettent de déterminer la taille de l'échantillon de l'enquête :

1) calculer le nombre d'éléments de l'échantillon nécessaires pour répondre aux conditions de mesure d'un indicateur donné, et

2) calculer le nombre de ménages qu'il faudra contacter pour trouver le nombre d'éléments nécessaires dans le cadre de la première étape.

Les formules de ces calculs sont présentées aux Sections 3.1.1 et 3.1.2 ci-après.

3.1.1 Calculer le nombre d'éléments de l'échantillon : Les indicateurs peuvent être exprimés comme une proportion, une moyenne ou un total. Dans le premier cas, un indicateur peut être exprimé comme une proportion de la population, à l'exemple du pourcentage de nourrissons jusqu'à l'âge de six mois qui sont allaités exclusivement ou qui sont rabougris (voir la liste des indicateurs génériques de l'alimentation infantile et juvénile du Titre II donnée à la Figure 3-6). Dans le second cas, quand ils sont exprimés comme une moyenne ou un total, les indicateurs peuvent être donnés en termes de la quantité d'un produit donné, d'un bien ou du nombre total de personnes, par exemple, l'apport calorique moyen par jour et par habitant dans une population donnée.

3.1.1.1 Indicateurs exprimés en tant que proportions: La formule suivante (Equation de Base 1) peut être utilisée pour calculer la taille de l'échantillon nécessaire pour les indicateurs exprimés en tant que pourcentage ou proportion. Notons que les tailles de l'échantillon

obtenues concernent chaque série d'enquêtes ou chaque groupe de comparaison.

Equation de Base 1 : Proportions

$$n = D [(Z_{\alpha} + Z_{\beta})^2 * (P_1 (1 - P_1) + P_2 (1 - P_2)) / (P_2 - P_1)^2]$$

CLE :

n = taille minimale de l'échantillon nécessaire par série d'enquêtes ou groupe de comparaison.

D = effet de conception (dans les équations suivantes, cela est supposé être la valeur *implicite* de 2 - voir Section 3.4 ci-après).

P_1 = le niveau estimé d'un indicateur mesuré en tant que proportion au moment de la première enquête ou pour la zone témoin.

P_2 = le niveau *escompté* de l'indicateur soit à une date future, soit pour la zone du projet de sorte à ce que la quantité ($P_2 - P_1$) soit l'ordre de grandeur du changement que l'on souhaite qu'il détecte.

Z_{α} = le score-Z correspondant au degré de confiance que l'on souhaite avoir pour conclure qu'un changement observé de la taille ($P_2 - P_1$) n'aurait pas eu lieu par hasard (α - niveau de signification statistique), et

Z_{β} = le score-Z correspondant au degré de confiance que l'on souhaite avoir pour dépister avec certitude un changement de la taille ($P_2 - P_1$) si un tel

changement a effectivement eu lieu (β – puissance statistique).

Z_α et Z_β ont des valeurs « standard » en fonction de la fiabilité souhaitée. Elles

sont fournies ci-après dans la Figure 3-2. Notons que plus le pourcentage est élevé, plus le programme peut être sûr de mesurer des résultats exacts.

Figure 3-2 : Valeurs de Z_α et Z_β

α (alpha)	Z_α	β (beta)	Z_β
0,90	1,282	0,80	0,840
0,95	1,645	0,90	1,282
0,975	1,960	0,95	1,645
0,99	2,326	0,975	1,960
		0,999	2,320

L'utilisation de la formule pour calculer les indicateurs exprimés en tant que proportions est indiquée ci-

après à la Figure 3-3. Des paramètres standard Z_α et Z_β pris de la Figure 3-2 sont incorporés.

Figure 3-3 : Exemples de calculs de la taille de l'échantillon exprimés en tant que proportions

Exemple 1 : Supposons que l'on veut mesurer un accroissement de 10 points de pourcentage dans la proportion de ménages démontrant un comportement correct de lavage des mains. Supposons qu'au moment de la première enquête, environ 50% des ménages avaient probablement de bonnes pratiques de lavage des mains. Dans ce cas, $P_1 = 0,50$ et $P_2 = 0,60$. En utilisant les paramètres standard avec un niveau de signification de 95% (α) et une puissance de 80% (β), les valeurs de la Figure 3-2 de $Z_\alpha = 1,645$ et $Z_\beta = 0,840$ sont choisies. En insérant ces valeurs dans la formule ci-dessus, l'on obtient le résultat suivant :

$$\begin{aligned}n &= 2 [(1,645 + 0,840)^2 * ((0,5)(0,5) + (0,6)(0,4))] / (0,6 - 0,5)^2 \\ &= 2 [(6,175 * 0,49) / 0,10^2] \\ &= 2 [(3,02575) / 0,01] = 2 (302,575) = 605,15, \\ &\text{soit 606 ménages par série d'enquêtes.}\end{aligned}$$

Exemple 2 : Supposons qu'un programme d'alimentation des enfants cherche à accroître la proportion de nourrissons de moins de six mois recevant uniquement le lait maternel de 20 points de pourcentage sur une période de 5 ans. Supposons qu'au départ, l'on pense qu'environ 60% des nourrissons de la population cible sont allaités exclusivement pendant six mois. Par conséquent, $P_1 = 0,60$ et $P_2 = 0,80$. Vu que le programme souhaite être certain de pouvoir détecter un accroissement de 20 points de pourcentage si un tel accroissement a effectivement lieu, on choisit une puissance de 90%, de pair avec le niveau de signification standard de 95%. Par conséquent, $Z_\alpha = 1,645$ et $Z_\beta = 1,282$. La taille de l'échantillon nécessaire est donc:

$$\begin{aligned}n &= 2 [(1,645 + 1,282)^2 * ((0,6)(0,4) + (0,8)(0,2))] / (0,8 - 0,6)^2 \\ &= 2 [(8,567 * 0,40) / 0,20^2] \\ &= 2 [(3,4268) / 0,04] = 2 (86,67) = 171,34, \\ &\text{soit 172 nourrissons par série d'enquêtes.}\end{aligned}$$

La Figure 3-4 permet au concepteur de l'enquête qui connaît l'ampleur du changement et le niveau de précision souhaité de choisir les tailles de l'échantillon sans devoir faire les calculs montrés ci-dessus. Les tailles de l'échantillon montrées ont été calculées en utilisant la formule de base dans cette

section : les valeurs pour les niveaux initiaux de l'indicateur (P_1) se situent dans une fourchette allant de 0,10 à 0,50 et les changements/différences au niveau d'un indicateur donné d'ampleur spécifiée ($P_2 - P_1$) se situent dans une fourchette allant de 0,05 à 0,30. Ce tableau concerne des valeurs de $\alpha = 0,95$

et $\beta = 0,80$. Un tableau comparable pour $\alpha = 0,95$ et $\beta = 0,90$ se trouve en Annexe

2. La Section 3.5 fournit des directives sur la manière de choisir les paramètres.

Figure 3-4 : Tailles de l'échantillon nécessaires pour les combinaisons choisies de P_1 et les changements ou différences entre groupes de comparaison devant être détectés (pour $\alpha = 0,95$ et $\beta = 0,80$)

Changement/différence devant être détecté ($P_2 - P_1$)												
P_1	Augmentation						Diminution					
	0,05	0,10	0,15	0,20	0,25	0,30	0,05	0,10	0,15	0,20	0,25	0,30
0,10	1 074	309	152	93	63	45	679	111	21	•	•	•
0,15	1 420	389	185	110	73	51	1 074	216	70	23	3	•
0,20	1 717	457	213	124	81	56	1 420	309	114	49	21	7
0,25	1 964	513	235	134	86	60	1 717	389	152	73	37	19
0,30	2 161	556	251	142	90	62	1 964	457	185	93	51	29
0,35	2 310	587	262	147	92	62	2 161	513	213	110	63	38
0,40	2 408	605	268	148	92	62	2 310	556	235	124	73	45
0,45	2 458	611	268	147	90	60	2 408	587	251	134	81	51
0,50	2 458	605	262	142	86	56	2 458	605	262	142	86	56
0,55	2 408	587	251	134	81	51	2 458	611	268	147	90	60
0,60	2 310	556	235	124	73	45	2 408	605	268	148	92	62
0,65	2 161	513	213	110	63	38	2 310	587	262	147	92	62
0,70	1 964	457	185	93	51	29	2 161	556	251	142	90	62
0,75	1 717	389	152	73	37	19	1 964	513	235	134	86	60
0,80	1 420	309	114	49	21	7	1 717	457	213	124	81	56
0,85	1 074	216	70	23	3	•	1 420	389	185	110	73	51
0,90	679	111	21	•	•	•	1 074	309	152	93	63	45

3.1.1.2 Pour les indicateurs exprimés en tant que moyennes ou totaux: La formule suivante peut être utilisée pour calculer la taille de l'échantillon nécessaire pour des indicateurs qui sont exprimés en tant

que moyennes ou totaux. La taille de l'échantillon doit être calculée pour chaque série d'enquêtes ou groupe de comparaison.

Equation de Base 2 : Moyennes ou totaux

$$n = D [(Z_{\alpha} + Z_{\beta})^2 * (et_1^2 + et_2^2) / (X_2 - X_1)^2]$$

CLE :

n = taille minimale de l'échantillon nécessaire par série d'enquêtes ou groupe de comparaison.

D = effet de conception pour les enquêtes à grappes (utiliser la valeur *implicite* de 2, tel que discuté dans la Section 3.4).

X₁ = le niveau estimé d'un indicateur au moment de la première enquête ou pour la zone témoin.

X₂ = le niveau *escompté* de l'indicateur soit à une date future, soit pour la zone du projet pour que la quantité (X₂ - X₁) soit la taille de l'ampleur du changement ou des différences entre les groupes de comparaison que l'on souhaite être en mesure de détecter.

et₁ et et₂ = écart type *escompté* pour les indicateurs pour les séries respectives d'enquête ou les groupes de comparaison étant comparés

Z_α = le score-Z correspondant au degré de confiance que l'on souhaite avoir pour conclure qu'un changement observé de la taille (X₂ - X₁) n'aurait pas eu lieu par hasard (signification statistique), et

Z_β = le score-Z correspondant au degré de confiance que l'on souhaite avoir pour dépister avec certitude un changement de la taille (X₂ - X₁) si un tel changement a effectivement eu lieu (puissance statistique).

La principale difficulté en utilisant la formule susmentionnée, c'est qu'elle demande des informations sur l'écart type de l'indicateur utilisé pour le calcul de la taille de l'échantillon. La solution préférée consiste à utiliser des valeurs d'une enquête précédente réalisée dans le contexte dans lequel est mis en œuvre le programme qui fait l'objet de l'évaluation. Si de telles données ne sont pas disponibles, l'on peut utiliser des données d'une autre partie du pays ou d'un pays voisin avec des caractéristiques analogues. De telles données sont souvent présentées dans les rapports d'enquête.

La formule ci-dessus est appliquée à la Figure 3-5. Les valeurs types pour Z_α et Z_β fournies dans la Figure 3-2 sont utilisées.

Figure 3-5 : Exemple de calcul de taille d'échantillon pour un indicateur exprimé comme moyenne

Supposons que l'on souhaite mesurer les effets d'un programme de Titre II sur la consommation quotidienne de calories par habitant. Au début du programme, l'on suppose qu'une moyenne de 1 700 calories par habitant et par jour sont consommées par la population ciblée par le projet. Le but du programme est d'arriver à un accroissement de 20%, soit 340 calories par habitant et par jour. Par conséquent, $X_1 = 1\,700$ et $X_2 = 2\,040$. Des données d'une enquête récente faite dans une région voisine indiquent une consommation moyenne de calories par jour et par habitant de 1 892 et un écart type de 1 136.

En calculant la taille de l'échantillon nécessaire, l'on suppose un rapport constant entre l'écart type et l'apport moyen de calories dans les deux séries d'enquêtes. En utilisant les données du pays voisin, nous pouvons arriver à un écart type approximatif en utilisant le ratio entre la moyenne du pays voisin et l'écart type et en l'appliquant aux moyennes estimées X_1 et X_2 (1 700 et 2 040). Par conséquent, l'écart type estimé (et_1) est calculé en calculant le ratio de la moyenne de la population de référence ou population voisine par rapport à l'écart type (1 892/1 136), soit 1,6655 et en l'appliquant à la moyenne X_1 estimée : $1\,700/1,6655 = 1\,021$. Le second écart type (et_2) est calculé en appliquant le même ratio de la moyenne du pays voisin à l'écart type (1 892/1 136), soit 1,6655 et en l'appliquant à la moyenne estimée X_2 : $2\,040/1,6655 = 1\,225$.

En fonction des paramètres standard avec un niveau de signification de 95% et une puissance de 80%, les valeurs de la Figure 3-2 de $Z_\alpha = 1,645$ et $Z_\beta = 0,840$ sont choisies. En insérant ces valeurs à la formule ci-dessus, l'on arrive au résultat suivant :

$$\begin{aligned}n &= 2[(1,645 + 0,840)^2 * (1021^2 + 1225^2) / (2\,040 - 1\,700)^2] \\ &= 2[(6,175)(2\,543\,066) / (340)^2] \\ &= 2[15\,705\,432 / 115\,600] = 2(135,843) = 272 \\ &\text{soit 272 ménages par série d'enquêtes.}\end{aligned}$$

3.1.2 Déterminer le nombre de ménages qui doivent être contactés : Les calculs ci-dessus permettent au concepteur de l'enquête de connaître le nombre d'éléments de l'échantillon qui doivent être contactés pour mesurer les changements/différences au niveau des indicateurs clés. Mais tous les ménages

n'auront pas un membre qui entre dans la catégorie indiquée dans l'indicateur (par exemple, des enfants de moins de six mois ou des enfants qui ont eu un épisode diarrhéique les deux semaines précédentes), et il faudra donc contacter un nombre plus important de ménages pendant les travaux de terrain de

l'enquête que le nombre d'éléments indiqués. Ensuite, l'étape suivante consiste à convertir la taille de l'échantillon exprimée en termes *d'éléments* en taille de l'échantillon exprimée en termes de *ménages*. (Notons que la procédure de calcul de la taille de l'échantillon en deux étapes n'est pas nécessaire pour les indicateurs mesurés au niveau des ménages car, par définition, les indicateurs au niveau ménage peuvent être mesurés pour chaque ménage choisi dans un échantillon donné.)

Par exemple, pour l'indicateur « Pourcentage de nourrissons/enfants de moins de 24 mois allaités pendant la première heure de la vie, » normalement 6%-8% seulement de la population ciblée par le projet comprendra des nourrissons/enfants de moins de 24 mois. (La fourchette reflète le niveau de fécondité au sein de la population ; plus la fécondité est élevée, plus on aura une proportion importante de nourrissons et d'enfants au sein de la population.) Par conséquent, au sein d'une population où le nombre moyen de personnes par ménage est de 6, l'on peut s'attendre à ce que 36%-48% des ménages comprennent des nourrissons/enfants de moins de 24 mois.

L'étape suivante consiste à convertir cette information en nombre de ménages qu'il faut contacter pour trouver le nombre d'éléments nécessaires. Supposons que dans ce cas, on a déterminé qu'il faudrait 300 nourrissons/enfants de moins de 24 mois pour obtenir l'indicateur susmentionné sur l'allaitement maternel. Supposons en outre qu'il s'agit d'une population à fécondité élevée, alors le nombre de ménages sera calculé de la manière suivante : $300 / (0,08 * 6)$, soit $300 / 0,48$, égal à un total de $n=625$ ménages.

Notons également que, dans la discussion susmentionnée, l'on suppose que *tous* les sujets présentant les conditions requises pour un indicateur trouvé dans un ménage de l'échantillon doivent être inclus dans l'échantillon (à savoir, tous les enfants âgés de six à 59 mois). Une autre stratégie consisterait à échantillonner uniquement un sujet présentant les conditions requises par ménage (quel que soit le nombre qu'on peut y trouver). Toutefois, cela demanderait que l'on contacte un nombre plus important de ménages, d'où des coûts plus élevés pour le travail de terrain de l'enquête. La stratégie consistant à les compter tous simplifie également le calcul des poids de l'échantillonnage car il ne faudra pas calculer une fraction d'échantillonnage à l'intérieur d'un ménage (voir Figure 5-1, D, pour la formule de ce calcul). Le problème lorsqu'on choisit tous les sujets concernés au sein des ménages de l'échantillon, c'est que les niveaux estimés d'erreur d'échantillonnage pour les estimations de l'enquête peuvent être biaisés vers le bas suite au groupage en grappes au sein des ménages à moins que l'on utilise des méthodes complexes pour estimer les erreurs d'échantillonnage. Du point de vue pratique, l'ampleur de ce biais est généralement petite et, à moins que l'on ne dispose d'un statisticien et d'un logiciel adéquat lors de l'étape de l'analyse, il s'agit d'un risque qui est souvent accepté dans le cadre des enquêtes (voir Chapitre 5, Section 3).

Parfois, il est pourtant nécessaire de ne choisir qu'un sujet par ménage car l'échantillonnage de tous les sujets présentant les conditions requises aboutit à des interviews qui sont trop longues. Seulement, les données détaillées sur la proportion de ménages qui comprennent

des sujets répondant aux critères pour les divers indicateurs sont rarement (voir jamais) disponibles, faisant qu'il est difficile de déterminer à l'avance le nombre exact de ménages qu'il faudrait contacter pour arriver à une taille d'échantillon souhaitée. Peut-être que la meilleure solution à ce niveau est de calculer le nombre de ménages qu'il faudra calculer, tel que décrit ci-dessus, et ensuite d'ajouter une *marge de sécurité* de 20% à 25% de ménages en

plus pour compenser car, dans certains cas, certains sujets concernés au sein d'un ménage ne sont pas choisis. Il s'agit toutefois de reconnaître que c'est une approximation rudimentaire.

La Figure 3-6 présente des indications sur les nombres typiques de ménages qu'il faudra contacter pour trouver une seule unité de mesure pour chaque indicateur sanitaire générique du Titre II.

Figure 3-6 : Nombres typiques de ménages à contacter pour trouver un individu de référence pour les indicateurs sanitaires génériques du Titre II (en supposant six personnes par ménage)

Indicateur	Individu de référence	% de la population ¹	Nb de ménages nécessaires ¹
% rabougris	enfants 6-59 mois	14%-19%	1,2/0,9
% d'un poids insuffisant	enfants 6-59 mois	14%-19%	1,2/0,9
% allaités dans les 8 heures suivant la naissance	nourrissons < 24 mois	6%-8%	2,8/2,1
% allaités uniquement	nourrissons < 6 mois	1,5%-2%	11,1/8,3
% recevant des aliments complémentaires	nourrissons 6-10 mois	1%-1,5%	16,7/11,1
% ayant eu un épisode diarrhéique pendant les 2 dernières semaines	enfants < 24 mois	6%-8%	2,8-2,1
% ayant eu un épisode diarrhéique qui reçoivent une alimentation continue	enfants < 24 mois avec épisode diarrhéique	1,5%-2% ²	11,1/8,3
% ayant eu un épisode diarrhéique qui reçoivent des aliments supplémentaires	enfants < 24 mois avec épisode diarrhéique	1,5%-2% ²	11,1/8,3

1. Les fourchettes indiquées concernent respectivement des populations à fécondité faible/élevée.

2. Suppose une fréquence de 25% de la diarrhée dans les deux semaines précédentes. L'on peut remplacer avec les estimations locales.

3.2 Choisir des indicateurs pour déterminer la taille de l'échantillon :

Un certain nombre d'indicateurs seront mesurés dans le cadre d'une enquête classique d'une évaluation d'un programme du Titre II. Dans l'idéal, les conditions pour chaque indicateur seront prises en compte lorsqu'on détermine la taille de l'échantillon nécessaire pour une enquête donnée. Mais quand le nombre d'indicateurs devant être mesurés est important, cela représente une lourde tâche.

Ce problème est généralement traité de l'une des deux manières suivantes. Une option consiste à déterminer les indicateurs qui risquent d'être les plus exigeants du point de vue taille de l'échantillon et d'utiliser la taille de l'échantillon nécessaire pour cet indicateur. De cette manière, on satisfait aux conditions de tous les autres indicateurs. Dans la plupart des cas, il s'agira de l'indicateur dont l'unité de mesure est celle trouvée le plus rarement au sein de la population cible. Par exemple, dans la Figure 3-6, il s'agira du « Pourcentage de nourrissons âgés de six à 10 mois recevant des aliments complémentaires, » car seul 1% à 1,5% de la population cible sont des nourrissons dans cette fourchette d'âge. Le grand avantage de cette procédure, c'est qu'elle garantit automatiquement une taille d'échantillon adéquate pour tous les indicateurs devant être mesurés. L'inconvénient, c'est qu'une taille d'échantillon plus grande sera choisie que celle nécessaire pour certain ou un grand nombre d'indicateurs.

La seconde approche consiste à identifier un petit nombre d'indicateurs dont on pense qu'ils sont les plus importants aux fins d'évaluation du programme et de limiter les calculs de la

taille de l'échantillon à ces indicateurs. Cela garantira une taille d'échantillon adéquate pour les indicateurs clés. Le problème, c'est que la taille risque de ne pas être adéquate pour d'autres indicateurs qui sont plus exigeants du point de vue taille. Mais puisqu'on a déjà jugé qu'ils étaient d'importance secondaire, cela peut constituer un compromis raisonnable, surtout quand le temps et les ressources sont limités.

Un autre compromis quand il faut tenir compte des aspects financiers et logistiques consiste à calculer la taille de l'échantillon pour les indicateurs clés et pour les indicateurs les plus exigeants du point de vue taille de l'échantillon et ensuite, de choisir entre les deux la taille d'échantillon la plus grande qui est faisable. Cette option garantit des tailles d'échantillon adéquates pour les indicateurs clés et les meilleures estimations possibles pour les indicateurs plus exigeants au vu des ressources disponibles. Cette approche est souvent utilisée.

3.3 Choisir les valeurs initiales des indicateurs quand celles-ci ne sont pas connues :

La valeur initiale d'un indicateur exprimée en tant que proportion – à savoir P_1 – provient dans l'idéal de l'information disponible grâce à d'autres enquêtes qui ont été faites dans un contexte donné (tel qu'indiqué à la Section 2 ci-dessus, des enquêtes précédentes ont peut-être été réalisées dans la région ou dans un pays voisin ; les données du Ministère de la Santé ou d'autres organismes publics ; ou quand de telles informations ne sont pas disponibles, les opinions et estimations provenant de personnes connaissant bien la situation et reposant sur les meilleures sources disponibles). En choisissant une

valeur pour P_1 , il vaut mieux se rapprocher de la valeur de 0,50. La raison en est la suivante : la variance des indicateurs qui sont mesurés en tant que proportions atteignent leur maximum quand ils approchent de 0,50. Le plus sûr, c'est de toujours choisir $P_1 = 0,50$ car cela garantira une taille d'échantillon adéquate qu'elle que soit la valeur effective de P_1 . Toutefois, cela entraînera également des échantillons qui sont plus grands que nécessaire dans le cas où la valeur effective de P_1 est très différente de 0,50. Aussi, l'approche recommandée est de deviner au mieux en fonction de l'information disponible et de choisir la valeur de P_1 avoisinant 0,50. Par exemple, si l'on pense qu'un indicateur se situait dans la fourchette 0,30-0,40 au moment de l'enquête initiale, on devrait choisir 0,40.

Pour les indicateurs exprimés en tant que moyennes ou totaux, les estimations sont nécessaires, non seulement pour les données initiales au moment de la première enquête ou de la zone témoin (X_1), mais également pour l'écart type de X . L'écart type est le paramètre crucial ici. Mais il n'existe pas de règle empirique dans le cas où des données d'autres enquêtes ne sont pas disponibles. Le meilleur conseil dans ce cas est le suivant : être prudent et choisir une valeur pour l'écart type qui est une grande proportion de la valeur de départ supposée d'un indicateur. Par exemple, des écarts types qui sont de 50% à 80% de l'estimation utilisée de X devraient être adéquats dans la plupart des cas.

3.4 Effets de conception : Les deux équations de base ci-dessus comprennent « D » pour l'effet de conception. Cela permet de corriger la perte de l'efficacité de l'échantillonnage due à l'utilisation de grappes au lieu d'un seul

échantillonnage aléatoire (voir Chapitre 4). De fait, il s'agit du facteur utilisé pour augmenter la taille de l'échantillon pour une grappe donnée afin de produire des estimations d'enquête avec la même précision qu'un échantillon aléatoire simple. L'ordre de grandeur de D dépend de deux facteurs : (1) le degré d'analogie ou d'homogénéité des éléments au sein des grappes, et (2) le nombre d'unités de mesure devant être prises auprès de chaque grappe.

Dans l'idéal, l'on pourra obtenir une estimation de D pour les indicateurs intéressant l'étude grâce à une enquête précédente dans un contexte donné. Cela vous donnera une certaine idée de l'analogie ou de l'homogénéité des éléments de la grappe. Si cela n'est pas possible, il faudra utiliser des valeurs *typiques* venant d'enquêtes réalisées ailleurs. Malheureusement, de telles directives, souvent, ne sont pas disponibles et, par conséquent, une valeur implicite de 2 est généralement utilisée surtout pour les enquêtes anthropométriques et des vaccinations. En supposant que les tailles des grappes soient relativement petites (voir Section 4 ci-après pour une discussion plus approfondie des tailles des grappes), l'utilisation d'une valeur standard de $D = 2$ devrait compenser de manière adéquate l'utilisation de l'échantillonnage par grappe dans la plupart des cas.

3.5 Signification et puissance : La signification statistique (α) et la puissance statistique (β) sont analogues aux faux positifs et aux faux négatifs : la signification statistique évite de conclure à tort qu'un changement a eu lieu, alors que la puissance statistique évite d'arriver à la conclusion erronée que rien ne s'est passé suite au programme. Le

paramètre de la puissance, β , est le plus important des deux pour les évaluations du programme car il garantit qu'un programme n'est pas classé comme un échec alors qu'en fait, il a eu un résultat positif. A moins que les tailles de l'échantillon suffisent pour détecter de manière fiable les changements ou les différences entre les groupes de comparaison d'une taille spécifiée, l'utilité des enquêtes en tant qu'outil d'évaluation du programme est compromise. Une puissance insuffisante peut mener à la conclusion erronée selon laquelle il n'y a pas eu de changement significatif des indicateurs dans le temps ou des différences entre les groupes du projet et les groupes témoins alors que de fait, il y a eu de *réels* changements ou de *réelles* différences qui n'ont pas été détectés à cause de la taille insuffisante de l'échantillon utilisée. Pour garantir une puissance suffisante, il faudrait utiliser une valeur minimale de β de 0,80 et 0,90 est préférable si les ressources le permettent.

Pour α , ou le niveau de signification, le standard est de 95% dans la plupart des enquêtes. L'on suppose que c'est une proportion suffisante pour vérifier que tout changement observé n'a pas eu lieu par hasard. Mais si les ressources ne le permettent pas, ce paramètre pourrait être réduit à $\alpha = 0,90$. Ce chiffre plus faible ne comporte qu'un petit risque supplémentaire de conclure à tort qu'un changement a eu lieu ou que les indicateurs sont différents pour les groupes du projet et les groupes témoins. L'on ne conseille pas d'aller en dessous de ce niveau de signification et aucun équivalent pour un chiffre plus faible n'est donné sur la Figure 3-2.

3.6 Tenir compte de la non-réponse : La non-réponse est chose courante dans

les enquêtes. L'on encourage certes de faire tous les efforts possibles pour minimiser le niveau de non-réponse (voir Chapitre 4), mais il n'existe pas de limites pratiques à ce que l'on peut faire à ce propos.

Afin de vérifier que les tailles souhaitées sont atteintes dans le cadre des enquêtes, il faut tenir compte de la non-réponse lors du calcul de cette taille de l'échantillon. A cette fin, il faut généralement augmenter la taille de l'échantillon par le biais d'un facteur d'*assurance* de non-réponse. Cela peut varier quelque peu d'un contexte à l'autre, mais une marge de 10% devrait s'avérer adéquate dans la plupart des situations. Par conséquent, si la taille de l'échantillon calculée pour une enquête demande 1 000 ménages et une marge de sécurité de 10% pour la non-réponse incorporée à la conception de l'étude, la taille révisée de l'échantillon pour l'enquête sera de 1 100 ménages.

4. Déterminer le nombre de grappes et le nombre de sujets par grappe

Une fois déterminée la taille générale de l'échantillon, l'étape finale consiste à faire un plan de sondage pour déterminer le nombre de grappes et le nombre de ménages par grappe qui devraient être choisis. Trois aspects entrent en ligne de compte :

- Le premier est l'ordre de grandeur de l'effet de conception du sondage par grappe (D). Plus le nombre de ménages par grappe est petit, moins l'effet de conception est prononcé. Cela est dû au fait que les unités élémentaires au sein des grappes tendent généralement à avoir un certain degré d'homogénéité concernant les caractéristiques générales et peut-être les

comportements. Alors que le nombre de ménages augmente par grappe, la précision de l'échantillonnage se perd.

- Deuxièmement, le nombre de ménages dans une grappe donnée place une limite sur la grandeur éventuelle de l'échantillon par grappe. Le recensement ou autre matériel utilisé comme base d'échantillonnage devrait être examiné attentivement avant de décider de la taille de l'échantillon par grappe qui sera utilisée.
- Troisièmement, les ressources disponibles pour réaliser l'enquête sur le terrain dictent souvent ce qui est faisable. En effet, le transport et les indemnités journalières du personnel et des superviseurs sur le terrain constituent les coûts les plus importants de l'enquête et ces coûts tendent à varier plus ou moins directement en proportion avec le nombre de grappes devant être couvertes. Par conséquent, les coûts de terrain sont minimisés quand le nombre de grappes est petit.

Vu que les deux derniers aspects sont susceptibles de varier énormément d'une application et d'un contexte à l'autre, seules des directives générales sont données ici. Vu sous l'angle de la précision de l'échantillonnage, l'on préfère les grappes plus petites aux grappes plus grandes. Par conséquent, pour une taille d'échantillon fixée (par exemple, 600 ménages), un plan de sondage de 30 grappes avec 20 ménages dans chaque grappe sera préféré à un plan comportant 20 grappes et 30 ménages, formule à nouveau préférée à 10 grappes de 60 ménages. En général, une formule relativement sûre consiste à choisir un maximum de 40-50 ménages

par grappe. Bien sûr, si les ressources ne permettent pas des grappes de cette taille, la grappe peut être augmentée, mais il faut savoir que cela se fait au détriment de l'exactitude car le risque d'erreur d'échantillonnage sera plus grand.

L'utilisation de 30 grappes dans le cadre des enquêtes auprès de la population est une formule de plus en plus courante, mais de fait il n'existe aucune justification statistique montrant que 30 est un nombre minimum ou idéal. Il n'en sert pas moins de directive sommaire représentant un chiffre adéquat pour vérifier que les échantillons des membres du groupe cible sont suffisamment bien répartis dans les grappes pour que les estimations de l'enquête ne soient pas trop influencées par deux ou trois grappes.

En utilisant 50 ménages comme la taille standard de l'échantillon par grappe, le nombre de grappes est normalement calculé en divisant la taille de l'échantillon par 50. Par exemple, si la taille d'un échantillon est de 2 000 ménages, l'échantillon devrait être réparti sur plus de 40 grappes de 50 ménages chacune. Si la taille de l'échantillon est trop petite (par exemple, 1 000), les grappes seront insuffisantes (dans ce cas, 20) et la taille de l'échantillon ne sera pas suffisamment distribuée. Dans une telle situation, l'on conseille de prendre 34 grappes avec 30 ménages (même si cela signifie que la taille de l'échantillon est dépassée d'un petit chiffre).

Il vaut mieux que chaque grappe compte le même nombre d'éléments de l'échantillon. En effet, cela signifie plus ou moins la même charge de travail dans chaque grappe, d'où un contrôle opérationnel plus simple sur les travaux de terrain de l'enquête. Une seconde

raison est d'éviter les biais d'estimation en aidant à garantir un échantillon auto-pondéré. Cet aspect est discuté plus en détail dans les Chapitres 4 et 5.

5. Taille de l'échantillon nécessaire pour les enquêtes de suivi

Les méthodes permettant de déterminer la taille des échantillons pour les enquêtes par sondage décrites ci-dessus sont conçues pour tenir compte des conditions d'une série d'enquêtes de suivi entrant dans l'évaluation du programme. Mais dans certains cas, la taille de l'échantillon devra être élargie lors de la seconde série. Tel est notamment le cas quand la taille de l'échantillon pour la première série d'enquêtes n'a pas tenu compte des conditions liées à la mesure du changement dans le temps, par exemple, si on a utilisé la formule de taille de l'échantillon pour une enquête unique pour déterminer la taille de l'échantillon pour l'enquête initiale. Ou alors, une formule adéquate pour la taille de l'échantillon a été utilisée, mais les niveaux des indicateurs observés dans le cadre de l'enquête initiale étaient différents de ceux escomptés lorsque les calculs de la taille de l'échantillon ont été faits avant l'enquête. De manière plus

spécifique, dans les cas où l'indicateur a été mesuré comme une proportion, la valeur observée de l'enquête initiale s'est avérée bien plus proche de 0,50 qu'on ne l'avait pensé ; quand l'indicateur a été mesuré comme une moyenne ou un total, l'écart type de l'indicateur s'est avéré bien plus grand que prévu. Dans ces deux cas, la taille de l'échantillon utilisée pour l'enquête initiale serait trop petite pour répondre aux niveaux de précision nécessaires pour l'évaluation si cette taille est utilisée pour l'enquête de suivi.

Deux manières sont disponibles pour calculer l'augmentation de la taille de l'échantillon du suivi. La première consiste à utiliser l'une des formules existantes, mais c'est problématique car ces formules tiennent compte de la signification statistique et non pas de la puissance statistique. De fait, la solution recommandée est la suivante : (a) calculer une estimation révisée de la taille de l'échantillon nécessaire en utilisant l'Equation de Base 1 ou 2 en tenant compte des résultats de l'enquête initiale et (b) compenser toute insuffisance au niveau de la taille de l'échantillon de l'enquête initiale en augmentant davantage la taille de l'échantillon pour l'enquête de suivi. Cette méthode en deux étapes est montrée sur les Figures 3-7 et 3-8.

Figure 3-7 : Exemple de calcul de la taille de l'échantillon pour les enquêtes de suivi pour des indicateurs exprimés en tant que proportions

Supposons que l'on cherche à mesurer un accroissement de 10 points de pourcentage de la proportion de ménages adoptant un comportement correct du lavage des mains. Au moment de l'enquête initiale, l'on jugeait qu'environ 30% des ménages avaient de bonnes pratiques de lavage des mains. Par conséquent, P_1 et P_2 ont été fixés, respectivement, à 0,30 et 0,40. En utilisant la formulation pour les indicateurs exprimés en tant que proportions (Equation de Base 1) et en supposant des paramètres standard avec un niveau de signification de 95% et une puissance de 80%, une taille d'échantillon $n = 556$ est calculée.

Mais l'enquête initiale a révélé qu'une proportion bien plus importante de ménages (50%) pratiquait déjà un comportement correct au niveau du lavage des mains. La procédure recommandée ci-dessus demande que l'on change la taille de l'échantillon en utilisant les résultats de l'enquête initiale – par conséquent, $P_1 = 0,50$, tel que suit :

$$\begin{aligned}n &= D [(Z_\alpha + Z_\beta)^2 * (P_1 (1 - P_1) + P_2 (1 - P_2))] / (P_2 - P_1)^2 \\n &= 2 [(1,645 + 0,840)^2 * ((0,5)(0,5) + (0,6)(0,4))] / (0,6 - 0,5)^2 \\&= 2 [(6,175 * 0,49) / 0,10^2] \\&= 2 [(3,02575) / 0,01] = 2 (302,575) = 605,15 \\&\text{soit } 606 \text{ ménages par série d'enquêtes.}\end{aligned}$$

Pour compenser l'insuffisance au niveau de la taille de l'échantillon lors de l'enquête initiale, la différence entre les estimations initiales et les estimations revues de la taille de l'échantillon (dénotées, respectivement, n_1 et n_2) est ajoutée à la taille de l'échantillon devant être utilisée pour l'enquête de suivi. Par conséquent, la taille de l'échantillon pour l'enquête de suivi serait la suivante :

$$n = 606 + (n_2 - n_1) = 606 + (606 - 556) = 606 + 50 = 656$$

Figure 3-8 : Exemple de calcul de la taille de l'échantillon pour les enquêtes de suivi pour un indicateur exprimé en tant que moyenne

La Figure 3-5 a montré comment calculer la taille des échantillons pour des enquêtes initiales et de suivi visant à évaluer les effets d'un programme du Titre II sur la consommation de calories par jour et par habitant. Supposons qu'à la place de la consommation moyenne escomptée de calories par jour et par habitant de $X_1 = 1\,700$ et écart type de $et_1 = 1\,020$, des valeurs de $X_1 = 1\,800$ et $et_1 = 1\,280$ ont été observées dans le cadre de l'enquête initiale. Suivant la procédure recommandée, il faut recalculer la taille de l'échantillon de la manière suivante (en utilisant des paramètres standard d'un niveau de signification de 95% et une puissance de 80% et en recalculant et_2 en utilisant le ratio de X_1/et_1 observé dans l'enquête initiale) :

$$\begin{aligned}n &= D [(Z_\alpha + Z_\beta)^2 * (et_1^2 + et_2^2) / (X_2 - X_1)^2] \\n &= 2[(1,645 + 0,840)^2 * (1\,280^2 + 1\,536^2) / (2\,160 - 1\,800)^2] \\&= 2[(6,175)(3\,997\,696) / (360)^2] \\&= 2[24\,686\,672 / 129\,600] = 2(190,484) = 380,97, \\&\text{soit } 381 \text{ ménages par série d'enquêtes.}\end{aligned}$$

Pour compenser l'insuffisance de la taille de l'échantillon lors de l'enquête initiale, la différence entre les estimations initiales ($n=272$) et les estimations revues de la taille de l'échantillon (dénotées, respectivement, n_1 et n_2) est ajoutée à la taille de l'échantillon devant être utilisée pour l'enquête de suivi. Par conséquent, la taille de l'échantillon pour l'enquête de suivi serait la suivante :

$$n = 381 + (n_2 - n_1) = 381 + (381 - 272) = 381 + 109 = 490$$

4. Choisir l'échantillon

1. Vue générale de l'échantillonnage

1.1 Echantillonnage probabiliste/non probabiliste : Les méthodes d'échantillonnage sont rangées dans deux catégories : les méthodes formelles ou probabilistes et les méthodes informelles ou non probabilistes.

- Les méthodes d'échantillonnage formelles reposent sur la théorie du sondage probabiliste. Deux choses sont nécessaires à ce propos : (1) chaque unité d'échantillonnage doit avoir une probabilité de sélection connue et différente de zéro, et (2) le facteur aléatoire est le facteur contrôlant la sélection des unités de l'échantillonnage. L'échantillonnage tend également, dans la pratique, à être caractérisé par (1) l'utilisation de listes ou bases de sondage pour choisir l'échantillon, (2) des procédures de sélection de l'échantillon clairement définies, et (3) la possibilité d'estimer l'erreur d'échantillonnage à partir des données de l'enquête (discutée plus

dans le détail au Chapitre 5, Section 3).

- Les méthodes d'échantillonnage informelles comportent un certain nombre d'approches qui reposent sur d'autres principes que ceux de la probabilité. L'on cherche en général à faire des déductions et généralisations à un groupe plus grand de la population, mais les méthodes de sélection tendent à être plus subjectives. Dans la plupart des cas, l'on suppose que la personne choisissant l'échantillon connaît bien les dimensions sous-jacentes du phénomène de l'étude et, par conséquent, choisira l'échantillon de telle manière que ces phénomènes *sont couverts* adéquatement (à savoir sans biais). L'on espère que l'échantillon est suffisamment représentatif aux fins de l'enquête, mais on ne peut pas le savoir avec un degré mesurable de certitude. Les échantillonnages sur *quota* et *échantillonnage dirigé (échantillon choisi à dessein)* sont deux exemples de plusieurs formes d'échantillonnage informel ou non

probabiliste (voir ci-après, Sections 3.2 et 3.5).

L'on recommande vivement les méthodes d'échantillonnage probabiliste pour les évaluations des programmes du Titre II en dépit de leurs coûts légèrement plus élevés. Il est un fait que les deux types de méthodes d'échantillonnage peuvent produire les mêmes résultats, mais vu qu'ils se fondent sur la théorie statistique, les évaluations utilisant l'échantillonnage probabiliste ont un degré plus élevé de crédibilité et peuvent être défendues plus aisément que celles reposant sur des méthodes d'échantillonnage informel plus vulnérables aux questions concernant le caractère représentatif de l'échantillon face à la population ou le fait que les résultats peuvent être biaisés.

Les procédures d'échantillonnage décrites dans ce guide se fondent toutes sur les principes de la probabilité. Elles ont été adaptées pour tenir compte des types de difficultés que l'on rencontre souvent dans les contextes de pays en développement.

1.2 Sondage par grappe : Le sondage par grappe tend à être le type d'échantillonnage probabiliste le plus utilisé. Normalement, il est préféré à l'échantillonnage aléatoire simple, bien que cette méthode soit la méthode d'échantillonnage probabiliste la plus facile et la plus connue. L'échantillonnage aléatoire, qui suppose que les unités sont choisies individuellement et directement par le biais d'un processus aléatoire où chaque unité a la même chance (probabilité) d'être choisie, demande des listes complètes d'unités complémentaires (par exemple, ménages dans la région du projet). Vu que ces éléments sont

rarement disponibles et généralement trop chers pour être créés, l'échantillonnage aléatoire ne sera probablement pas utilisé comme méthode d'échantillonnage unique pour les évaluations du Titre II. Par contre, l'échantillonnage par grappe limite la portée de la construction de la base de l'échantillon et des travaux de terrain connexes à un ensemble ou échantillon de zones géographiques devant être couvertes et, par conséquent, permet de contrôler les coûts de terrain.

Une grappe est simplement une agrégation d'unités d'échantillonnage revêtant un intérêt pour une enquête particulière qui peuvent être définies sans ambiguïté et utilisées comme unités d'échantillonnage à partir de laquelle un sous-échantillon plus petit peut être choisi. Dans l'idéal, les grappes devraient répondre à quatre critères. (1) Elles devraient avoir des limites physiques relativement claires pour faciliter l'identification sur le terrain. (2) Elles devraient être situées relativement proches les unes des autres, sinon les coûts vont grimper, allant contre l'objectif principal de l'échantillonnage sur grappe. (3) Les grappes ne devraient pas comprendre un trop grand nombre de personnes ; cela aidera à minimiser le travail de développement des bases d'échantillonnage. (4) L'information sur la taille de la grappe devrait être disponible en principe avant la sélection des échantillons. Cela permettra d'utiliser des procédures de sélection des grappes conçues pour améliorer l'efficacité de l'échantillonnage ou la sélection de la *probabilité-proportionnelle-à-la-taille* (PPT) (voir Section 3.1 ci-après). (L'incapacité d'obtenir des mesures sur la taille de la grappe avant la sélection des

échantillons n'empêche pas l'utilisation de l'échantillonnage probabiliste.)

Le développement de la base d'échantillonnage suppose normalement deux étapes : (1) sélection d'unités de premier stade ou *unités primaires* et (2) sélection d'unités d'échantillonnage élémentaires au sein des unités primaires. Mais dans un grand nombre d'applications, les villages et/ou les rues de la ville/lots de maisons seront choisis lors d'une première étape et un échantillon de ménages de chacun lors de la seconde étape. Dans certains cas, des individus des ménages devront être choisis, ajoutant ainsi une troisième étape au processus.

Quand la population de l'unité de l'échantillon est jugée trop large, la sélection d'unités d'échantillonnage de *premier stade* ou *primaires* peut se faire dans le cadre de deux étapes. Dans de tels cas, la grappe choisie est divisée en grappes plus petites ou en deux grappes, l'une des grappes plus petites est choisie aléatoirement et le développement de la base d'échantillonnage et les opérations d'échantillonnage sont réalisées avec la grappe plus petite qui a été choisie. Le problème ici, c'est qu'un biais peut se présenter au niveau de la sélection de la grappe plus petite.

2. Bases d'échantillonnage

Une base d'échantillonnage est une liste des unités d'échantillonnage possibles. Pour les types d'enquêtes considérés dans ce chapitre, les bases d'échantillonnage sont des listes de zones d'énumération du recensement correspondant plus ou moins aux villages et lots de maisons d'une ville, à partir du dernier recensement de la population. Quand de telles listes n'existent pas, une liste des villages et des villes qui couvrent toute la

population ciblée par le projet suffira, mais il faudra probablement des étapes supplémentaires de sélection de l'échantillon pour produire des unités suffisamment petites pour qu'elles soient pratiques pour les opérations de terrain de l'enquête.

Deux autres options sont disponibles quand il n'existe pas une base d'échantillonnage adéquate. L'une consiste à développer une base initiale pour la région en question. Par exemple, si le projet ne dispose pas d'une liste de villages couvrant une province entière, un échantillon de circonscription de la province peut être compilé et une base d'échantillonnage des villages peut être formulée uniquement dans les circonscriptions de l'échantillon. Ce type d'échantillonnage par grappe à plusieurs stades est souvent utilisé pour compenser les inadéquations de la base d'échantillonnage. Il faut souvent développer des bases d'échantillonnage ou mettre à jour des travaux dans le cadre des enquêtes et il convient de le faire de sorte à minimiser le temps et les coûts.

La seconde option consiste à limiter l'enquête à la partie de l'univers de l'enquête pour laquelle existe une base d'échantillonnage. Toutefois, cela limite la capacité de généraliser les résultats de l'évaluation à la partie de la région du projet couverte par l'évaluation, d'où l'introduction d'un biais. Cette option ne sera choisie qu'en dernier recours.

3. Procédures de sélection de l'échantillon pour des programmes avec des niveaux élevés de couverture générale de la population

Dans certains cas, une proportion importante de la population dans la zone ciblée par le projet devra être *exposée* à

l'intervention qui est évaluée (c'est-à-dire qu'elle recevra ou du moins, aura l'occasion de recevoir les avantages du projet). Par exemple, les programmes visant à influencer les pratiques d'alimentation infantile ciblent généralement la population générale d'une zone géographique choisie pour une intervention. Dans de tels cas, les enquêtes auprès de la population représentent un moyen valide de mesurer les effets du programme.

Des directives sont données ci-après sur la manière de choisir les échantillons à chaque étape du processus d'échantillonnage : (1) choisir les grappes de l'échantillon, (2) choisir les ménages de l'échantillon, et (3) choisir les sujets individuels de l'enquête quand la nécessité se présente. Différentes variantes sont décrites pour chaque étape.

3.1 Procédure pour choisir les grappes de l'échantillon : Dans le cadre des projets avec couverture générale de la population, il y aura probablement un trop grand nombre de grappes figurant dans la base d'échantillonnage pour permettre un échantillonnage complet. Par conséquent, une première étape sera de réduire le nombre de grappes devant être choisies à un chiffre plus gérable. La procédure recommandée, l'échantillonnage systématique aléatoire suppose que l'on choisit une grappe de manière aléatoire, puis la i^{th} grappe par la suite dans la série. (1^{th} concerne le nombre de chaque grappe choisie ; à savoir, dans la Figure 4-2, la première, cinquième et huitième grappes sont choisies). On peut le faire de deux manières, suivant l'information disponible sur la taille des grappes au sein de la population ciblée. Les deux méthodes sont décrites ci-après.

Lorsque des mesures sont disponibles de la taille de la grappe : Le plan de la grappe à deux stades, le plus efficace du point de vue statistique est le suivant : (1) les grappes sont choisies avec probabilité-proportionnelle-à-la-taille (PPT) lors de la première étape de la sélection de l'échantillon et (2) un nombre constant de ménages est choisi de chaque grappe lors de la seconde étape.

Le terme probabilité-proportionnelle-à-la-taille (ou PPT) signifie que des grappes plus grandes ont une chance plus grande de sélection que les grappes plus petites. L'utilisation de la procédure de sélection PPT demande qu'une base d'échantillonnage de grappes avec des mesures de taille soit disponible ou soit développée avant la sélection de l'échantillon. Une mesure de la taille est simplement un compte ou une estimation d'une variable susceptible d'être corrélée au nombre de sujets intéressant l'enquête au sein d'une grappe donnée. Par exemple, la population totale d'un village et le nombre total de villages auront probablement une étroite corrélation avec le nombre de nourrissons/enfants de la grappe et, par conséquent, représentent de bonnes mesures de la taille pour des enquêtes ou les nourrissons/enfants seront des unités de mesure importantes. Des comptes exacts ne sont pas nécessaires : une approximation ou des estimations sommaires suffisent. Toute inexactitude sera corrigée lors de la seconde étape de la sélection de l'échantillon quand des nombres spécifiques de ménages seront choisis (voir Section 3.2 ci-après).

La Figure 4-1 indique les étapes de la sélection d'un échantillon de grappes utilisant l'échantillonnage systématique

avec probabilité proportionnelle à la taille. La Figure 4-2 applique ces étapes à un exemple où l'objectif est de choisir 40 grappes d'un total de 170 grappes.

Figure 4-1 : Etapes de la sélection d'un échantillon systématique aléatoire de grappes avec PPT

- (1) Préparer une liste des unités d'échantillonnage du premier stade (à savoir, grappes) avec une mesure correspondante de la taille pour chacune (voir la colonne 2 de la Figure 4-2).
 - (2) En commençant en haut de la liste, calculer la mesure cumulative de la taille et marquer ces chiffres dans une colonne à côté de la mesure de la taille pour chaque unité (voir colonne 3).
 - (3) Calculer l'intervalle d'échantillonnage (IE) en divisant la mesure cumulative totale de la taille pour le domaine ou strate (M) par le nombre planifié d'unités devant être choisies (a) – à savoir, $IE = M/a$.
 - (4) Choisir un nombre aléatoire (début aléatoire ou DA) entre 1 et (IE). Comparer ce nombre avec la colonne de la mesure cumulative de la taille. L'unité dans laquelle entre le chiffre (DA) de la mesure cumulative de la taille est la première unité de l'échantillon (voir colonne 4).
 - (5) Les unités subséquentes sont choisies en ajoutant l'intervalle d'échantillonnage (IE) au nombre identifié à l'étape (4) ; à savoir $DA + IE$, $DA + IE * 2$, $DA + IE * 3$, etc. (voir colonne 4).
 - (6) Cette procédure est suivie jusqu'à ce que la liste soit épuisée. Le nombre résultant d'unités devrait être plus ou moins égal au nombre ciblé de grappes.
-

Figure 4-2 : Exemple - sélection d'un échantillon systématique aléatoire de grappes avec PPT

Grappe No.	Taille – Nb de ménages	Taille cumulative	Nb d'échantillonnages	Grappe choisie
001	120	120	73	X
002	105	225		
003	132	357		
004	96	453		
005	110	563	503	X
006	102	665		
007	165	839		
008	98	937	934	X
009	115	1 052		
.	.	.		
.	.	.		
.	.	.		
170 (dernier)	196	17 219 (M dans l'équation)		
Taille cumulative totale		17 219		

Nb planifié de grappes = 40 (a dans l'équation)
 Intervalle d'échantillonnage = $17\,219/40 = 430,475$
 L'échantillonnage commence entre 1 et $430,475 = 73$
 Grappes choisies = 001, 005, 008,

Chaque fois que c'est possible, les grappes seront choisies avec probabilité-proportionnelle-à-la-taille dans les enquêtes par sondage. Une des raisons, c'est que cette procédure est relativement efficace du point de vue

précision de l'échantillonnage. La seconde raison est la suivante: si un nombre égal d'éléments est choisi dans chaque grappe lors de la seconde étape de la sélection de l'échantillon, le résultat final sera un échantillon où chaque ménage a la même probabilité générale de sélection ou *auto-pondération*. C'est un grand avantage lors de l'analyse de données (voir Chapitre 5 pour une discussion supplémentaire).

Quand les mesures de la taille des grappes ne sont pas disponibles : Une procédure légèrement différente devrait

être utilisée quand les mesures de la taille des grappes ne sont pas disponibles avant la sélection de l'échantillon. Dans le cadre de cette méthode, toutes les grappes auront la même chance ou probabilité de sélection, ou *probabilité égale*, plutôt qu'une probabilité étant reliée à leur taille. Les procédures pour choisir un échantillon de grappes avec probabilité égale sont décrites à la Figure 4-3 et un exemple est donné à la Figure 4-4. Dans l'exemple, l'objectif est à nouveau de choisir 40 grappes dans un total de 170 grappes.

Figure 4-3 : Etapes de la sélection d'un échantillon systématique aléatoire de grappes avec probabilité égale

- (1) Préparer une liste numérotée de sites ou grappes, de préférence arrangés géographiquement (par exemple, par zones d'une ville).
 - (2) Calculer l'intervalle d'échantillonnage (IE) en divisant le nombre total ou grappes dans le domaine (groupe cible) (M) par le nombre de grappes devant être choisies (a) – à savoir, $IE = M/a$.
 - (3) Choisir un nombre aléatoire (début aléatoire ou DA) entre 1 et (IE). La grappe de la liste numérotée correspondant à ce chiffre sera la première grappe de l'échantillon.
 - (4) Les unités subséquentes sont choisies en ajoutant l'intervalle d'échantillonnage (IE) au nombre identifié à l'étape (3) ; à savoir $DA + IE$, $DA + IE * 2$, $DA + IE * 3$, etc.
 - (5) Cette procédure est suivie jusqu'à ce que la liste soit épuisée.
-

Figure 4-4 : Exemple – sélection d'un échantillon systématique aléatoire de grappes avec probabilité égale

Grappe No.	Sélection	
001		Nb planifié de grappes = 40
002	X	Intervalle d'échantillonnage = $170/40 = 4,25$
003		Début d'échantillonnage entre 1 et $4,25 = 2$
004		Grappes choisies = 002, 006, 011,
005		
006	X	
007		
008		
009		
010		
011	X	
.		
.		
.		
170 (dernière)		

Notons qu'en choisissant les grappes de l'échantillon, il est important de garder les chiffres après la virgule dans l'intervalle d'échantillonnage. La règle est la suivante : quand la partie décimale du nombre de sélection de l'échantillon est inférieure à 0,5, la grappe du numéro inférieur est choisie et, quand la partie décimale du nombre de sélection de l'échantillon est 0,5 ou supérieure, c'est la grappe du numéro supérieur qui est choisie. Dans l'exemple ci-dessus, le nombre de sélection de l'échantillon pour la troisième grappe de l'échantillon était 10,5 et, par conséquent, la grappe 011 a été choisie pour l'échantillon.

On suppose qu'un nombre fixe de ménages doit être choisi dans chaque grappe de l'échantillon. Mais dans ce cas, vu que la probabilité de sélection de la grappe n'est pas basée sur le nombre de ménages que la grappe contient, la procédure nous donne des éléments de l'échantillon qui ont des probabilités générales différentes de sélection. En d'autres termes, l'échantillon n'est *pas auto-pondéré*. Cela complique la situation pendant l'analyse (voir Chapitre 5).

3.2 Procédures pour choisir les ménages de l'échantillon : Dans l'idéal, les ménages de l'échantillon seront choisis en créant une liste ou une base d'échantillonnage de tous les ménages situés au sein de chaque grappe et en choisissant un échantillon d'unités en utilisant soit un échantillonnage aléatoire simple ou un échantillonnage systématique. Mais la compilation de telles listes de ménages est probablement trop chère et prend trop de temps. Pour contourner cette difficulté, trois options sont possibles et elles sont décrites ci-

après : *segmentation* et deux variantes de la méthode *aléatoire-marche*.

Méthodes de segmentation : La méthode de segmentation consiste à diviser les grappes de l'échantillon en segments plus petits de taille plus ou moins égale en choisissant l'un au hasard dans chaque grappe et en interviewant tous les ménages du segment choisi. La taille du segment (à savoir, le nombre de ménages dans chaque segment) devrait être la même que le nombre cible de ménages de l'échantillon qui seront choisis par grappe (voir la Section 4 dans le

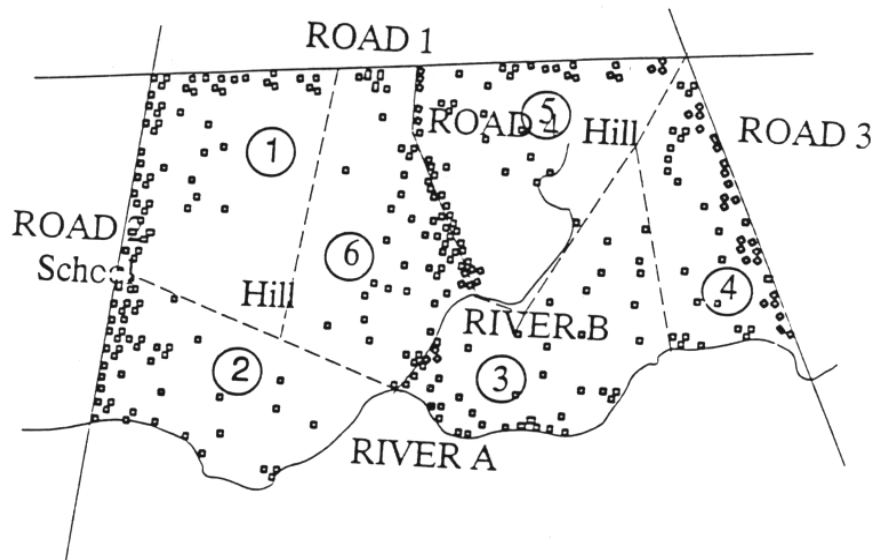
Chapitre 3). Par exemple, si l'on a déterminé que 30 grappes seront choisies pour une enquête donnée et que 50 ménages seront choisis par grappe (nous donnant une taille de l'échantillon de 1 500 ménages), la taille du segment ciblé dans le cadre de la méthode de segmentation serait de 50 ménages.

La Figure 4-5 propose les étapes lors de l'utilisation de la méthode de segmentation et la Figure 4-6 est une représentation graphique d'une grappe hypothétique qui a été créée en utilisant la méthode de segmentation.

Figure 4-5 : Etapes d'utilisation de la méthode de segmentation pour choisir les ménages de l'échantillon

- (1) Calculer le nombre de segments devant être créés. Diviser le nombre de ménages notés dans le dernier recensement par la taille du segment ciblé. Le résultat sera le nombre de segments qu'il faut créer sur le terrain. Par exemple, si le dernier recensement indique qu'il existait 250 ménages dans la grappe et que la taille du segment ciblé est de 40 ménages, il faudra créer six segments. (En faisant ce calcul, les chiffres décimaux des segments devraient être arrondis au chiffre entier le plus proche.)
 - (2) Mettre à jour la carte des grappes. En utilisant une carte de la grappe, vérifier/mettre à jour les frontières externes de la grappe et indiquer toute caractéristique interne pouvant être utile pour diviser la grappe en segments facilement reconnaissables.
 - (3) Compter et indiquer sur la carte l'emplacement des ménages situés dans la grappe. Cela représente une opération rapide pour que la grappe puisse être divisée en segments avec un nombre plus ou moins égal de ménages.
 - (4) En fonction de la carte des grappes, diviser la grappe en segments de taille égale. Le nombre de segments devant être utilisés est le nombre déterminé dans le cadre de l'Étape 1 ci-dessus.
 - (5) Choisir de manière aléatoire un segment.
 - (6) Interviewer tous les ménages situés dans les frontières du segment choisi aléatoirement.
-

Figure 4-6 : Exemple d'une grappe hypothétique qui a été divisée en six segments



Source : Fonds des Nations Unies pour l'Enfance. 1995. *Monitoring Progress Toward the Goals of the World Summit for Children: A Practical Handbook for Multiple-Indicator Surveys*. New York: UNICEF.

Les travaux de terrain pour une grappe donnée dans le cadre de la méthode de segmentation sont jugés terminés quand tous les ménages du segment choisi pour l'enquête ont été interviewés (quel que soit le nombre de sujets de l'étude effectivement trouvés).

Méthode aléatoire-marche : La méthode aléatoire-marche est utilisée dans le cadre des enquêtes sur grappes du PEV (programme élargi de vaccinations) et elle est donc assez bien connue. La méthode suppose la démarche suivante: (1) choisir aléatoirement un point de départ et une direction de déplacement au sein d'une grappe de l'échantillon, (2) réaliser une interview dans le ménage le plus proche, et (3) choisir continuellement le ménage

suivant le plus proche pour une interview jusqu'à ce que le nombre cible d'interviews soit atteint.

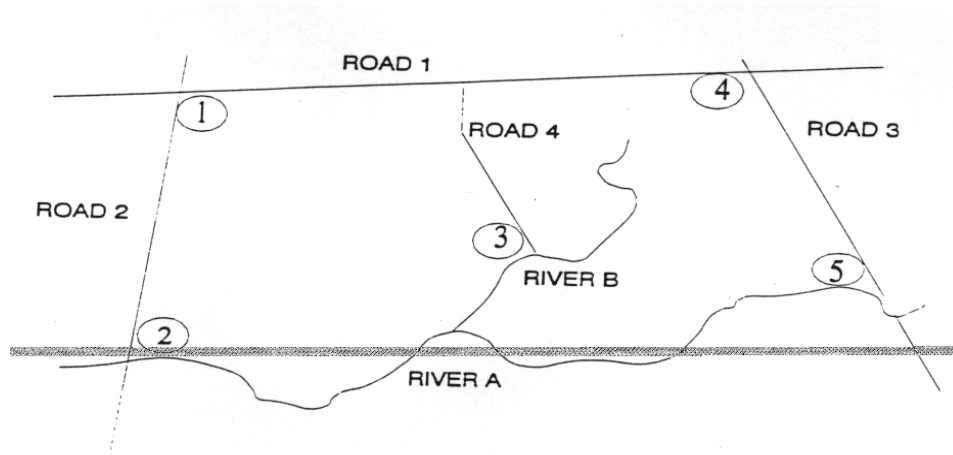
Cette approche peut être exécutée de deux manières, la seule différence étant la disponibilité d'une carte, ce qui détermine la sélection du point de départ. Un récapitulatif est fourni ci-après.

Choisir le point de départ avec une carte portant des limites : Quand une carte de la grappe de l'échantillon est disponible, on peut choisir un certain nombre de points de départ possibles à différents endroits facilement identifiables (voir Figure 4-7 pour un exemple) et à partir de là, un point de départ peut être choisi aléatoirement. L'avantage, c'est que le personnel de

supervision peut choisir le point de départ avant que ne commence le travail de terrain, réduisant tout risque de biais susceptibles de se présenter quand les

points de départ sont choisis sur la base de l'aspect pratique et non pas du facteur aléatoire.

Figure 4-7 : Carte d'une grappe d'échantillons hypothétiques montrant les points de départ possibles



Source : Fonds des Nations Unies pour l'Enfance. 1995. *Monitoring Progress Toward the Goals of the World Summit for Children: A Practical Handbook for Multiple-Indicator Surveys*. New York: UNICEF.

Choisir le point de départ en utilisant la méthode du PEV : Pour les enquêtes du PEV, l'on suppose qu'il n'existe pas de carte de la grappe et que le personnel de terrain choisira l'emplacement de départ. Pour cela, il va suivre des instructions qui recommandent les étapes suivantes (1) se rendre à un endroit central de la grappe et choisir une direction de déplacement au hasard en faisant tourner une bouteille, (2) partir en ligne droite dans cette direction et compter tous les ménages jusqu'à la limite de la grappe, et (3) choisir aléatoirement un chiffre entre 1 et le nombre de ménages comptés comme le point de départ de l'enquête.

Aucune des variantes de la méthode aléatoire-marche ne demande une mesure de la taille des grappes de

l'échantillon, mais une estimation du nombre de ménages situés dans chaque grappe de l'échantillon devrait être obtenue pour les évaluations du programme du Titre II quand l'une ou l'autre des approches est utilisée. Cela permet de calculer les probabilités d'échantillonnage (voir Chapitre 5 pour une discussion plus approfondie). Compter ou estimer le nombre de ménages n'est pas très cher et ne demande pas trop de temps. Dans la plupart des cas, un informant local connaissant bien les lieux peut fournir un chiffre raisonnable. Ou alors, un tour rapide de la grappe permet généralement d'arriver à un compte acceptable. Ce type de *compte rapide* est souvent utilisé dans le sondage à grappes.

Si l'on utilise la méthode aléatoire-marche, un *quota* devra être choisi, par exemple, l'agent de terrain devra continuer à contacter les ménages jusqu'à ce qu'un nombre prédéterminé de sujets de l'étude (par exemple, les enfants de moins de cinq ans) soient localisés. Le quota devrait être le nombre cible de *ménages* devant être choisis par grappe plutôt que le nombre des *sujets de l'étude* nécessaires pour les différents indicateurs (par exemple, enfants de moins de 24 mois, enfants ayant souffert d'un épisode diarrhéique les deux dernières semaines, etc.). Par exemple, si la taille de l'échantillon ciblée par grappe est de 50 ménages, on adoptera plutôt la procédure aléatoire-marche jusqu'à ce que ce chiffre de ménages soient interviewés. Les quotas de ménages sont suivis de préférence aux quotas des sujets de l'étude car il sera difficile de suivre ces derniers quand de multiples indicateurs seront mesurés. Tel que proposé ci-dessus, les quotas ne sont pas fixés pour la méthode de segmentation car les travaux de terrain dans une grappe donnée sont jugés achevés quand tous les ménages du segment ont été interviewés (quel que soit le nombre de sujets de l'étude effectivement trouvés).

Les trois méthodes décrites ci-dessus varient considérablement du point de vue exposition au risque de biais. La méthode de segmentation se rapproche le plus de la grappe conventionnelle à deux étapes et, par conséquent, est moins susceptible aux biais. Mais vu qu'il est nécessaire de faire une carte de la communauté pour utiliser la méthode, elle n'est pas toujours faisable dans tous les contextes. La meilleure méthode ensuite du point de vue biais est la variante de la méthode aléatoire-marche où le ménage de départ de l'échantillon

est choisi à partir d'une carte avec des limites. La méthode la moins préférée est la méthode aléatoire-marche du PEV.

3.3 Procédure pour choisir des sujets individuels de l'enquête : La recommandation générale pour la plupart des enquêtes, c'est d'inclure à l'échantillon tous les sujets présentant les conditions requises (par exemple, les enfants dans certaines fourchettes d'âge) au sein des ménages de l'échantillon. Mais pour certains types d'enquêtes (par exemple, quand des balances alimentaires sont utilisées) et dans des communautés où des familles élargies vivent ensemble, y compris toutes les femmes adultes et leurs enfants, la méthode prendrait trop de temps et s'avérerait trop chère. La procédure recommandée dans un tel cas est de choisir aléatoirement une femme adulte parmi celles trouvées dans un ménage donné de l'échantillon et de terminer l'enquête uniquement pour cette répondante et sa famille/enfants. Cela constitue une troisième étape dans le processus de sélection de l'échantillon.

Quand un sujet parmi plusieurs sujets possibles est choisi au sein d'un ménage, le calcul des poids de l'échantillon doit tenir compte de cela si l'on veut obtenir des estimations d'enquête non biaisées. C'est une démarche relativement simple : l'enquêteur doit uniquement noter sur le protocole d'enquête le nombre de répondants éventuels présentant les conditions requises dans chaque ménage de l'échantillon au sein duquel le répondant effectif de l'enquête a été choisi aléatoirement. Cela permet de calculer la probabilité de la sélection du répondant effectif de l'enquête. Cette probabilité est ensuite intégrée au calcul d'une probabilité globale de

l'échantillonnage (voir Chapitre 5, Figure 5-1[D]).

3.4 Traiter les problèmes opérationnels de l'échantillonnage :

Des problèmes au niveau de l'exécution peuvent se présenter, même dans le cadre des enquêtes les mieux planifiées. Il s'agit généralement des problèmes suivants : grappes inaccessibles, non-réponse et nombre insuffisant de ménages dans une grappe donnée.

Grappes inaccessibles : Parfois, il est impossible d'arriver jusqu'à la grappe de l'échantillon à cause du mauvais temps ou de routes impraticables. Dans un tel cas, la meilleure approche consiste à remplacer la grappe avec une autre grappe choisie aléatoirement possédant des caractéristiques analogues. Par exemple, si la grappe en question est située à l'Extrême Nord de la zone du projet, il faudrait la remplacer avec une autre grappe dans la même région en général, mais une grappe qu'il est possible d'atteindre pendant la période des travaux de terrain de l'enquête. Pour minimiser le risque de biais, les grappes de remplacement devraient être choisies parmi des grappes analogues. Le côté pratique ne devrait pas entrer en ligne de compte. Dans la mesure du possible, le personnel de supervision devrait prendre des décisions sur les grappes de remplacement.

Non-réponse de l'enquête : La non-réponse est un problème commun à toutes les enquêtes. Généralement, la non-réponse est due au fait que personne n'est présent dans les ménages de l'échantillon ou quand les sujets de l'enquête refusent d'être interviewés. Voici ce que l'on conseille de faire face à ces problèmes :

- **Personne à la maison** : Quand une personne ne répond dans un ménage choisi de la grappe, il faudrait demander aux voisins si (1) la maison est habitée et, si tel est le cas, (2) l'heure à laquelle les habitants sont généralement chez eux. Si la maison n'est pas habitée, aucune autre action n'est nécessaire. Si elle est habitée, l'on pourra faire au moins une autre visite et, peut-être, plusieurs autres visites (de préférence) à l'heure à laquelle les voisins ont indiqué que les habitants étaient généralement chez eux.
- **Refus** : Quand les membres d'un ménage cible refusent d'être interviewés, au moins une visite devra être faite, peut-être par un autre membre de l'équipe de terrain ou par le superviseur de l'équipe. Toutefois, la priorité des revisites devrait être pour ceux qui n'étaient pas à la maison.

Un certain degré de non-réponses est généralement intégré au calcul de la taille de l'échantillon (voir Chapitre 3, Section 3.7), mais quand ce facteur se présente, il peut biaiser les résultats de l'enquête. Cela est dû au fait qu'il existe souvent des différences systématiques entre les gens qui choisissent de répondre et ceux qui ne veulent pas et que ces différences peuvent être reflétées dans les indicateurs qui sont mesurés. La meilleure manière de traiter de tels biais possibles dus à la non-réponse est de minimiser, dans la mesure du possible, la non-réponse. Par conséquent, les plans opérationnels de terrain devraient prévoir un temps suffisant pour le suivi des non-répondants. Vu qu'un certain degré de non-réponse a déjà été anticipé lors des calculs sur la taille de l'échantillon, l'impact de la non-réponse

sur la taille finale de l'échantillon de l'enquête devrait être, dans la plupart des cas, tolérable, du moment que le niveau de non-réponse ne dépasse pas nettement le niveau escompté sur un nombre important de grappes.

La grappe a un nombre insuffisant de ménages pour répondre à la taille de l'échantillon ciblée : Si les calculs de la taille de l'échantillon ont été effectués correctement et si les tailles des grappes dans les bases d'échantillonnage disponibles ont été prises en compte, il devrait y avoir suffisamment de ménages pour répondre à la taille souhaitée de l'échantillon. De plus, si une *marge de sécurité* a été intégrée aux calculs de la taille de l'échantillon, tel que recommandé au Chapitre 3, on a généralement compensé les effets sur la taille finale de l'échantillon de l'enquête. Dans le cas où cette situation se présente, on recommandera aux équipes de terrain de ne pas choisir des ménages supplémentaires des grappes voisines. Elles devraient plutôt chercher à minimiser le nombre de ménages à non-réponse et ensuite, passer à la prochaine grappe qui leur est assignée.

3.5 Procédure pour choisir les échantillons dans les zones de comparaison : Généralement, les groupes de comparaison comprennent

des habitants d'une ou de plusieurs circonscriptions, municipalités ou autres unités administratives voisines comportant des caractéristiques analogues à celles du programme évalué. Le processus de sélection comprend généralement deux étapes. La première consiste à identifier des groupes qui répondent aux critères d'analogie. Le choix devra être fait *à dessein* (à savoir, les caractéristiques du groupe doivent être définies à l'avance et la sélection doit s'effectuer en fonction des critères convenus), à moins que plusieurs régions n'aient des profils analogues à la zone du programme et, dans ce cas, le choix peut être aléatoire. (Voir le Guide de Suivi et de l'Évaluation pour des conseils supplémentaires). Une fois défini l'univers de l'enquête pour la zone de comparaison, il reste à choisir un échantillon de grappes et de ménages pour représenter la zone de comparaison. Les procédures d'échantillonnage sont identiques à celles décrites ci-dessus pour les enquêtes auprès de la population en général.

La Figure 4-8 est un exemple des décisions d'échantillonnage pour un groupe de comparaison.

Figure 4-8 : Exemple de la conception d'un échantillon et de la sélection pour une zone de comparaison

Supposons qu'un programme visant à améliorer la production et la transformation agricoles a été réalisé dans deux districts d'un pays. Un prétest/post-test avec groupe de comparaison est utilisé aux fins d'évaluation du programme. Pour la zone de comparaison, tous les districts adjacents aux deux districts du projet sont jugés des candidats possibles. Deux sont éliminés car ils ont des caractéristiques socioéconomiques différentes des districts du projet et un autre est éliminé car c'est la cible d'un programme comparable financé par un autre bailleur de fonds. Parmi les districts adjacents restant, les deux districts qui sont les plus similaires aux deux districts du projet sont choisis.

Dans chaque district du groupe de comparaison, un échantillon de 30 grappes est choisi en utilisant une procédure de sélection systématique-aléatoire avec PPT, et 40 ménages sont choisis dans chaque grappe d'échantillon, utilisant la méthode de segmentation dans deux séries d'enquêtes : une enquête initiale et une enquête de suivi.

4. Procédures de sélection de l'échantillon pour des programmes avec une couverture de population limitée

4.1 Le problème d'échantillonnage :

Certains programmes du Titre II ont des niveaux assez limités de couverture de la population. Par exemple, la population bénéficiaire éventuelle d'un programme de crédit pour des mères célibataires avec des enfants de moins de cinq ans pourrait représenter moins de 5% des ménages dans la zone géographique ciblée pour un programme. Dans de tels cas, il n'est pas possible de mesurer l'impact du programme au niveau de la population générale. Il sera pratiquement impossible de détecter les changements dans les indicateurs à ce niveau, même si des changements de taille importante ont eu lieu pour les bénéficiaires du projet. Par conséquent, l'évaluation du programme cherche souvent à évaluer

les changements des indicateurs uniquement parmi les ménages bénéficiaires. Trois stratégies d'échantillonnage pour de telles situations sont présentées ci-après : limiter l'univers de l'enquête, dépistage et utilisation des bases de listes.

4.2 Autres stratégies l'échantillonnage :

Limiter l'univers de l'enquête : Une approche convient quand les bénéficiaires du projet sont très concentrés dans certaines parties de la région du projet (par exemple, dans les municipalités ou villages spécifiques). Dans de tels cas, l'univers de l'enquête peut être limité à des zones d'impact du programme et la base d'échantillonnage se limitera à un sous-ensemble de grappes. Grâce à une telle restriction, les stratégies d'échantillonnage décrites ci-dessus pourraient être utilisées sans modification supplémentaire. Cette

approche ne pourrait pas s'appliquer dans la plupart des cas car les bénéficiaires du programme sont généralement répartis sur l'ensemble de la région géographique ciblée par un projet donné.

Dépistage : Quand les bénéficiaires sont dispersés géographiquement sur l'ensemble de la région du projet, une approche utilisée couramment consiste à introduire une *procédure de dépistage*. Au début des travaux de terrain de l'enquête, l'on demande à d'éventuels ménages de l'échantillon s'ils ont bénéficié du projet. (Des questions à cet effet peuvent être rajoutées au début du questionnaire.) Les agents de terrain réalisent l'interview de l'enquête uniquement pour les ménages qui sont des bénéficiaires du projet, laissant de côté les ménages qui ne sont pas des bénéficiaires. Ou alors, une petite quantité d'information pourrait être réunie auprès de ménages non bénéficiaires pour faire des comparaisons entre bénéficiaires et non-bénéficiaires. Cette approche garantit que les ressources servent à réunir une information sur les bénéficiaires du projet.

Cette stratégie demande que l'on ajoute des ménages à la taille de l'échantillon ciblé. Par exemple, si seulement 20% des ménages dans une zone géographique ciblée doivent être des bénéficiaires du projet, un ménage sur cinq devra remplir les questionnaires. Dans ce cas, le nombre de ménages de l'échantillon devra être augmenté cinq fois.

Prenons un exemple : supposons que la taille de l'échantillon nécessaire pour une enquête donnée est de $n = 1\,500$ ménages, et que le plan de sondage consiste à avoir 30 grappes de 50 ménages chacune. Si l'on pense que

seuls 20% des ménages dans la zone ciblée par le projet été des bénéficiaires de ce projet, alors il faudra contacter un total de $n = 7\,500$ ($7\,500 = 1\,500 / 0,20$) ménages. Si l'on veut garder l'objectif des 30 grappes, cela nécessiterait que l'on contacte 250 ménages par grappe. Dans de nombreux cas, cela constitue la grappe entière et, dans ce cas, la stratégie d'échantillonnage consiste à interviewer tous les ménages des grappes de l'échantillon. Mais dans de nombreux contextes, les grappes dans les bases d'échantillonnage disponibles (à savoir, régions d'énumération du recensement ou autres unités administratives) sont plus petites que les 250 ménages, demandant que l'on choisisse un nombre plus grand de grappes. Par exemple, si la taille moyenne des grappes dans un contexte donné est de 100 ménages, il sera nécessaire de choisir un échantillon de 75 grappes et de contacter tous les ménages situés dans ces grappes ($7\,500 / 100 = 75$ grappes).

Utilisation des bases de listes : Quand des listes de bénéficiaires du programme sont disponibles, elles représentent une manière bien plus efficace de choisir un échantillon de ménages bénéficiaires que la formule du dépistage. Deux approches sont possibles ici. L'une consiste à choisir un échantillon de ménages bénéficiaires en utilisant un échantillonnage aléatoire simple ou systématique aléatoire. Toutefois, cela représente une opération de terrain difficile du point de vue logistique car les ménages de l'échantillon peuvent être très dispersés dans un grand nombre de villages. Autre solution, les bénéficiaires de la liste peuvent être regroupés géographiquement et l'échantillonnage sera réalisé avec les grappes de ménages

bénéficiaires qui en résultent. La base de sondage subséquente avec ces grappes peut être traitée exactement de la même manière qu'une base d'échantillonnage de grappes plus classiques (à savoir, zones d'énumération du recensement). Les procédures d'échantillonnage pour les enquêtes auprès de la population décrites ci-dessus peuvent ensuite être utilisées.

5. Questions de conception pour les enquêtes de suivi

5.1 Rétention ou remplacement des grappes de l'échantillon : L'une des questions clés de la conception d'une enquête sur plusieurs séries utilisant un échantillonnage sur grappes est de savoir s'il faut garder les mêmes grappes dans chaque enquête ou choisir un nouvel échantillon. D'un point de vue statistique, la stratégie préférée est de garder les mêmes grappes. En effet, les caractéristiques et comportements des individus dans les petites zones géographiques tendent à être corrélés dans le temps. L'effet de cette corrélation est de diminuer la variabilité des estimations de changement, facilitant ainsi la tâche consistant à détecter le changement *réel*. Le degré d'efficacité ainsi gagnée au niveau de l'échantillonnage varie d'un endroit à l'autre et d'un indicateur à l'autre mais on y gagnera toujours du point de précision de la mesure. Aussi la recommandation est-elle la suivante : il faut garder le même échantillon de grappes tout en choisissant un nouvel échantillon de ménages dans chaque grappe en suivant les procédures susmentionnées.

Cette stratégie comporte un danger : la tentation d'octroyer les ressources du programme de manière disproportionnée en faveur de régions dont on sait qu'elles

sont ciblées par l'évaluation du programme. Si le programme concentre ses ressources de cette manière, ces régions auront des résultats plus positifs que ceux notés sur toute la région du programme. Les responsables du programme doivent faire attention à ne pas déformer l'octroi de ressources en faveur des zones cibles de l'évaluation. Si les responsables de l'évaluation craignent ce type de biais, le recours le plus facile consiste à choisir un nouvel échantillon de grappes pour l'enquête de suivi. Certes, cela diminue, dans une certaine mesure, l'efficacité de l'échantillonnage pour l'évaluation du programme, mais cela réduit également le risque du biais, considération plus importante dans l'ordre plus grand des choses.

5.2 Effets des changements au niveau de la méthodologie de l'enquête : Autre problème : les effets *réels* du programme peuvent être déformés ou *subir l'interférence* de changements dans d'autres facteurs, d'une série à l'autre. Les *facteurs interférents* peuvent être divisés grosso modo en deux catégories : externes et internes.

Les *facteurs interférents externes* sont des changements des facteurs opérant dans la population étudiée sur lesquels les programmes et les responsables de l'évaluation n'ont peu, voire aucun contrôle. Les changements au niveau de la composition de la population suite aux flux migratoires, aux environnements économiques changeants, aux catastrophes naturelles, etc. sont des exemples de tels changements. Les évaluations tiennent généralement compte de tels facteurs en utilisant des groupes ou régions témoins, ce qui permet de séparer les effets réels du programme de ceux causés par des

facteurs hors programme et en utilisant des méthodes statistiques à plusieurs variables lors de l'analyse de données.

Les *facteurs interférents internes* sont des changements au niveau des facteurs qui sont internes à l'évaluation du programme et sur lesquels le personnel de terrain responsable de faire l'évaluation peut exercer un contrôle. En voici des exemples : changements de la définition de l'univers de l'enquête, formulation du questionnaire, méthodologie de l'échantillonnage et qualité ou caractère ponctuel (lié au caractère saisonnier) des activités de terrain de l'enquête. Surtout quand on utilise des tailles relativement petites de l'échantillon, les résultats de l'évaluation sont très sensibles à des changements même mineurs au niveau de la méthodologie et ou de l'exécution de l'enquête. La leçon ici est la suivante : la méthodologie et l'exécution de l'enquête doivent rester aussi constantes que possible.

5.3 Abandon au sein de l'échantillon :

Parfois, il faudra changer de fond en comble les plans faits en vue d'utiliser le même échantillon de grappes lors des

enquêtes initiales et de suivi à cause de catastrophes naturelles, de problèmes de sécurité etc. Dans de tels cas, la meilleure chose à faire consiste à suivre la procédure recommandée pour des grappes inaccessibles et de remplacer la grappe pour laquelle il n'est pas possible d'obtenir des données par une autre grappe choisie aléatoirement et comptant les mêmes caractéristiques (voir Section 3.4 ci-dessus). L'utilisation du choix aléatoire minimisera le risque du biais. Il vaut mieux avoir des caractéristiques analogues pour essayer de limiter la variabilité de l'échantillonnage entre les séries des enquêtes. Le personnel de supervision devrait prendre des décisions quant au choix des grappes de remplacement.

Toutefois, si la grappe est inaccessible temporairement quand l'enquête de suivi est programmée, ceux qui réalisent l'enquête devrait attendre que la grappe soit à nouveau accessible plutôt que choisir un remplacement du moment que cela n'entraîne pas un retard inacceptable pour les travaux de terrain de l'enquête.

5. Analyser les données

1. Vue d'ensemble

Une fois les données de l'enquête pour le suivi et/ou l'évaluation du programme réunies et saisies dans une base de données informatique, ce qu'il reste à faire, c'est d'analyser les données. Pour cela, il faut calculer les divers indicateurs mesurés dans les enquêtes et évaluations de l'ampleur et de la signification statistique des changements dans le temps et/ou différences entre les groupes de comparaison. Des directives pour l'analyse des données des évaluations des programmes du Titre II sont fournies dans le Guide du Suivi et de l'Évaluation. Dans cette section, nous nous concentrons sur deux questions liées à l'analyse très influencées par la manière dont est effectué l'échantillonnage : *poids de l'échantillonnage* et *calcul des erreurs types des estimations des enquêtes*.

2. Pondérer les données

Seul l'un des plans de sondage par grappe décrit dans le chapitre précédent résultera en échantillons auto-pondérés : c'est-à-dire que les grappes de l'échantillon sont choisies avec une probabilité-proportionnelle-à-la-taille

(PPT) et la méthode de segmentation est utilisée pour choisir les ménages de l'échantillon. Tous les autres aboutissent à des échantillons non auto-pondérés ou des échantillons où les ménages de l'échantillon ont des probabilités inégales de sélection. Cela doit être compensé au stade de l'analyse des données. Car sinon, cela entraîne un biais d'estimation.

Les poids compensent les probabilités inégales de sélection. La méthode standard pour corriger ces probabilités inégales consiste à appliquer des poids d'échantillonnage aux données de l'enquête pendant l'analyse en multipliant la valeur de l'indicateur par le poids. Le poids convenant pour chaque sujet de l'échantillon est simplement la réciproque de la probabilité de sélection de ce sujet ou l'inverse de la probabilité.

Equation de Base 3 :

Proportions

$$W_i = 1/P_i$$

CLE :

W_i = poids d'échantillonnage pour les éléments de la i^{th} grappe ; et

P_i = probabilité de sélection pour les éléments de la i^{th} grappe.

Les probabilités de sélection (P_i) doivent être calculées pour déterminer les poids de l'échantillonnage. La formule dépendra de la variante utilisée pour le plan de sondage par grappe de deux (ou trois) stades. La Figure 4-1 montre les calculs des probabilités d'échantillonnage pour quatre

combinaisons possibles : (A) PPT (premier stade)/segmentation (second stade); (B) PPT/aléatoire-marche ; (C) probabilité égale/aléatoire-marche ; et (D) PPT/segmentation, mais avec un répondant. (Les options non montrées sont les suivantes : probabilité égale/segmentée ; probabilité égale/aléatoire-marche, mais avec un répondant ; et probabilité égale/segmentation, mais avec un répondant. Les analystes peuvent adapter les options à leurs cas.) Le calcul de l'exemple A (probabilité-proportionnelle-à-la-taille/segmentation) ne sera jamais nécessaire car, tel que démontré, cette conception aboutit à un échantillon auto-pondéré.

Figure 5-1 : Procédures pour calculer les probabilités d'échantillonnage pour les éléments de l'échantillon (P_i) pour les plans d'échantillonnage par grappe à deux stades

A. PPT au premier stade, méthode de segmentation au second stade :

$$P_i = (m * M_i/M) * 1/S_i = m * C/M$$

CLE

- m = nombre de grappes d'échantillons choisies
- M_i = mesure de la taille pour la i^{th} grappe
- M = mesure totale de la taille pour l'univers de l'enquête
($M = \sum M_i$)
- S_i = nombre de segments créés dans la i^{th} grappe
- C = taille standard du segment (constante)

Notons que ce plan entraînant un échantillon auto-pondéré, l'application de poids d'échantillonnage pendant l'analyse n'est pas nécessaire.

B. PPT au premier stade, nombre constant d'éléments choisis au second stade en utilisant une méthode aléatoire-marche :

$$P_i = (m * M_i/M) * k/N_i$$

CLE

- m = nombre de grappes d'échantillons choisies
- M_i = mesure de la taille pour la i^{th} grappe
- M = mesure totale de la taille pour l'univers de l'enquête
($M = \sum M_i$)
- k = nombre constant de ménages choisis par grappe
- N_i = nombre total de ménages dans la i^{th} grappe, à savoir M_i mis à jour sur le terrain avec compte estimé ou effectif de ménages

C. Probabilité égale au premier stade, nombre constant d'éléments choisis lors du second stade en utilisant la méthode aléatoire-marche :

$$P_i = (1/m) * (k/N_i)$$

CLE

- m = nombre de grappes d'échantillons choisies
- k = nombre constant de ménages choisis par grappe
- N_i = nombre total de ménages dans la i^{th} grappe

D. Conception (B) ci-dessus, mais avec un répondant (par exemple, femme adulte) choisi par ménage de l'échantillon

$$P_i = (m * M_i/M) * (k/N_i) * 1/R_{ij}$$

CLE

- m = nombre de grappes d'échantillons choisies
- M_i = mesure de la taille pour la ith grappe
- M = mesure totale de la taille pour l'univers de l'enquête
(M = Σ M_i)
- k = nombre constant de ménages choisis par grappe
- N_i = nombre total de ménages dans la ith grappe
- R_{ij} = nombre total de répondants présentant les conditions requises dans le jth ménage

Un problème se présentera quand des poids d'échantillonnage seront appliqués à l'analyse de données effectuée avec des logiciels types (par exemple, Epi-Info, SPSS). L'application de poids aux logiciels standard va augmenter le nombre de cas de l'échantillon et, par conséquent, implique une taille plus grande de l'échantillon que celle utilisée dans la réalité.³ Par conséquent, les tests statistiques des différences et changements dans le temps se fonderont sur des tailles incorrectes de l'échantillon et des conclusions erronées quant aux effets des programmes peuvent en résulter. Par exemple, des changements ou des différences entre les groupes de comparaison qui n'étaient pas statistiquement significatifs sur la base de la taille de l'échantillon effectif sembleront significatifs en fonction du nombre pondéré de cas.

Pour compenser cela, on utilise souvent des *poids standardisés*. Les

poids standardisés assignent un poids à chaque observation de l'échantillon qui reflètent sa probabilité relative de sélection par rapport à d'autres observations de l'échantillon, mais qui ne changent pas la taille générale de l'échantillon de l'enquête. Les poids standardisés (w_i') pour les éléments de l'échantillon de la ith grappe sont calculés de la manière suivante :

$$w_i' = w_i n_i / \sum w_i n_i$$

Vu que chaque élément dans une grappe donnée comporte la même probabilité de sélection, chacun recevra également le même poids standardisé. La Figure 5-2 illustre le calcul des poids standardisés en utilisant des données d'enquêtes hypothétiques.

Pour utiliser les poids d'échantillonnage standardisés lors de l'analyse de données, il faut inclure une variable de poids appropriée dans le fichier des données d'enquête qui seront analysées. Les poids standardisés peuvent être calculés soit manuellement, soit avec un tableur, et saisis comme une variable lors de la saisie de données. Ou

³ Quand la valeur d'une variable ou d'un indicateur est multipliée par le poids de l'échantillon, le résultat concerne l'apparence voulant que la taille est plus grande qu'elle ne l'est en fait.

alors, les probabilités de sélection du premier et du second stade peuvent être saisies et les poids calculés en utilisant

les commandes appropriées SPSS ou Epi-Info.

Figure 5-2 : Exemple de calcul des probabilités de sélection, poids d'échantillonnage et poids d'échantillonnage standardisés – données hypothétiques

Dans cet exemple, les calculs des poids standardisés sont montrés pour les cinq premières grappes d'un échantillon de grappe choisie d'une enquête hypothétique.

CLE

- n_i = le nombre d'éléments d'enquête choisis dans la grappe i (qui est de 50 puisqu'un nombre constant d'élément a été choisi par grappe)
- P_i = probabilité générale de sélection pour les éléments de l'échantillon dans la grappe i
- w_i = poids d'échantillonnage pour les éléments de l'échantillon de la grappe i
- w_i' = poids d'échantillonnage standardisé pour les éléments de l'échantillon de la grappe i

Grappe No.	n_i	P_i	w_i	$w_i n_i$	w_i'
1	50	0,033	30,30	1515,00	0,0167
2	50	0,022	45,45	2272,50	0,0251
3	50	0,030	33,33	1666,50	0,0184
4	50	0,043	23,26	1163,00	0,0125
5	50	0,023	43,48	2174,00	0,0240
.					
.					
.					
Total	2 500			90 526,28	

3. Estimer les erreurs types

Tester la signification statistique des tendances ou changements observés demande que l'on estime l'ampleur de l'erreur d'échantillonnage associée aux estimations d'enquête appelées

couramment les *erreurs types*. L'estimation de l'erreur type peut être complexe suivant le plan de sondage utilisé pour collecter les données. Plus les plans de sondage deviennent complexes (par exemple, stratification,

sondage par grappe et multiple stades de la sélection de l'échantillon), plus les procédures pour estimer les erreurs types deviennent compliquées. L'estimation des erreurs types pour de tels plans n'entre pas dans la portée des évaluations des programmes du Titre II et ne peuvent pas être exécutées sans l'assistance d'un statisticien.

Malheureusement, les logiciels statistiques standard tels que SPSS et Epi-Info n'apportent pas une solution adéquate à ce problème. Les deux logiciels estiment les erreurs types des changements observés et/ou différences dans les groupes de comparaison par rapport aux indicateurs et effectuent des tests statistiques adéquats, mais les erreurs types produites par ces deux logiciels supposent qu'un échantillonnage aléatoire simple a été utilisé pour réunir les données de l'enquête. Vu qu'il est très probable qu'un échantillonnage par grappe a été utilisé dans le cadre des enquêtes servant à l'évaluation des programmes du Titre II, les erreurs types estimées produites par ces logiciels seront généralement sous-estimées. Le résultat est le suivant : certains changements ou certaines

tendances sont jugés des changements *réels* ou des tendances *réelles* alors que de fait, les changements sont trop petits pour être statistiquement significatifs.

Une des options serait de compenser pour la sous-estimation escomptée des erreurs types en rétrécissant les critères utilisés pour juger de la signification statistique à moins de créer des logiciels plus sophistiqués pour traiter du problème. Par exemple, au lieu d'utiliser $p < 0,05$ comme le point seuil pour juger qu'un changement ou une différence observé dans un indicateur est significatif, on pourrait utiliser $p < 0,04$ ou même $p < 0,03$. Ainsi, on risque moins de juger incorrectement qu'un changement observé est significatif sans ajouter toutefois la complexité des analyses statistiques des données de l'enquête ou de devoir utiliser un logiciel que le personne ne connaît pas. Mais, en fin de compte, quand des analyses intensives des données d'évaluation d'un programme sont nécessaires, il faut utiliser des logiciels qui sont capables de tenir compte de plans d'échantillonnage complexes (par exemple, STATA ou SUDAAN).

Annexe 1 : Liste des indicateurs génériques du Titre II

Catégorie	Niveau	Indicateur
Santé, nutrition et SMI	Impact	% d'enfants rabougris 20-59 mois (taille/âge score-Z)
		% d'enfants rabougris par groupe d'âge (taille/âge score-Z)
		% de nourrissons allaités dans les 8 heures suivant la naissance
		% de nourrissons de moins de 6 mois allaités uniquement
		% de nourrissons 6-10 mois recevant des aliments complémentaires
		% de nourrissons alimentés continuellement pendant la diarrhée
		% de nourrissons recevant des aliments complémentaires pendant les 2 semaines après la diarrhée
	Suivi annuel	% d'enfants concernés dans le suivi/promotion de la croissance
		% d'enfants vaccinés pour la rougeole à 12 mois
		% de communautés avec une organisation sanitaire communautaire
% d'enfants dans le programme de promotion de la croissance qui ont pris du poids ces 3 derniers mois		
Eau et assainissement	Impact	% de nourrissons souffrant ces 2 dernières semaines
		Litres d'eau de ménage utilisés par personne
		% de la population avec bon comportement de lavage des mains
		% de ménages ayant accès à un assainissement adéquat (également suivi annuel)
	Suivi annuel	% de ménages avec accès à une eau potable toute l'année
	% de structures eau/assainissement entretenues par la communauté	
Consommation alimentaire des ménages	Impact	% de ménages consommant une nourriture avec le minimum de vitamines/protéines recommandées
		Nombre de repas/goûters mangés chaque jour
		Nombre d'aliments différents/groupes alimentaires mangés
Productivité agricole	Impact	Rendement annuel des cultures ciblées
		Ecart de rendement (effectif versus potentiel)
		Variabilité de rendement dans des conditions différentes
		Valeur de la production agricole par ménage vulnérable
		Mois de provisions de céréales des ménages
		% de cultures perdues à cause des ravageurs ou de l'environnement
	Suivi annuel	Rendement annuel des cultures ciblées
		Nombre d'hectares où ont été adoptées les pratiques améliorées
Nombre d'installations de stockage construites et utilisées		
Gestion des ressources naturelles	Impact	Erosion des sols imputée
		Fertilité des sols imputée
		Rendements ou variabilité des rendements (également suivi annuel)
	Suivi annuel	Nombre d'hectares où sont utilisées les pratiques GRN
	Ensemencement/taux de survie	
Routes FFW/CFW	Impact	Marge de prix des intrants agricoles entre les régions
		Disponibilité des principaux intrants agricoles
		Coûts des transports des aliments de base par saison
		Volume des produits agricoles transportés par les ménages au marché
		Volume du trafic de véhicules, par type de véhicule
	Suivi annuel	Kilomètres de routes rénovées entre exploitation agricole et marché
		Mesures annuelles choisies des indicateurs d'impact

Annexe 2

Tailles de l'échantillon nécessaires pour les combinaisons choisies de P_1 et changements ou différences entre groupes de comparaison devant être détectés (pour $\alpha = 0,95$ et $\beta = 0,90$)

Changement/différence devant être détecté ($P_1 - P_2$)												
	Augmentation						Diminution					
P_1	0,05	0,10	0,15	0,20	0,25	0,30	0,05	0,10	0,15	0,20	0,25	0,30
0,10	1491	428	211	129	87	63	942	154	29	•	•	•
0,15	1970	540	257	152	101	71	1491	300	97	32	5	•
0,20	2382	634	295	171	112	78	1970	428	158	69	29	10
0,25	2724	711	326	186	120	83	2382	540	211	101	51	26
0,30	2999	771	348	197	125	86	2724	634	257	129	71	40
0,35	3204	814	364	203	128	87	2999	711	295	152	87	52
0,40	3341	840	371	206	128	86	3204	771	326	171	101	63
0,45	3410	848	371	203	125	83	3341	814	348	186	112	71
0,50	3410	840	364	197	120	78	3410	840	364	197	120	78
0,55	3341	814	348	186	112	71	3410	848	371	203	125	83
0,60	3204	771	326	171	101	63	3341	840	371	206	128	86
0,65	2999	711	295	152	87	52	3204	814	364	203	128	87
0,70	2724	634	257	129	71	40	2999	771	348	197	125	86
0,75	2382	540	211	101	51	26	2724	711	326	186	120	83
0,80	1970	428	158	69	29	10	2382	634	295	171	112	78
0,85	1491	300	97	32	5	•	1970	540	257	152	101	71
0,90	942	154	29	•	•	•	1491	428	211	129	87	63

Quelle est la confiance et la puissance pour les sous-groupes ?

Problème : Taille de l'échantillon, n , calculée sur la base d'un groupe cible (par exemple, enfants de moins de 23 mois), mais quelle est la confiance (α) et la puissance (β) pour les autres sous-groupes (par exemples, enfants 12-23 mois) ?

Hypothèses : $\alpha = 0,95$; $\beta = 0,8$;

L'ampleur de la différence, $p_2 - p_1$, est la même pour les deux sous-groupes

Pour tout sous-groupe plus grand (n est plus grand) que le groupe original, la puissance et la signification seront plus importantes et acceptables par définition.

Si le sous-groupe a n , les valeurs approximatives sont les suivantes :

Ratio des tailles de l'échantillon	α	β
0,9 (10% plus petite)	0,937	0,776
0,8	0,920	0,748
0,75	0,911	0,732
0,6	0,871	0,678
0,5 (moitié aussi grande)	0,834	0,634
0,33 (un tiers)	0,740	0,540
0,25 (un quart)	0,673	0,478
0,20 (un cinquième)	0,619	0,433
0,17 (un sixième)	0,576	0,399
0,14 (un septième)	0,541	0,372
0,13 (un huitième)	0,513	0,349
0,10 (un dixième)	0,465	0,314

La puissance et la signification sont trop approximatifs pour arriver à de solides conclusions pour les sous-groupes si le ratio des tailles de l'échantillon est de moitié ou moins.