

Boîte à outils pour l'évaluation d'impact

Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats

Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

Version 1.0. Juin 2012

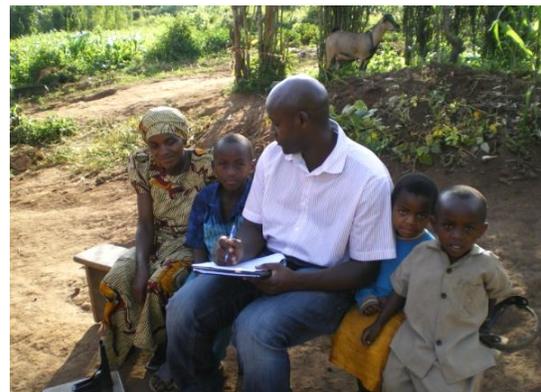
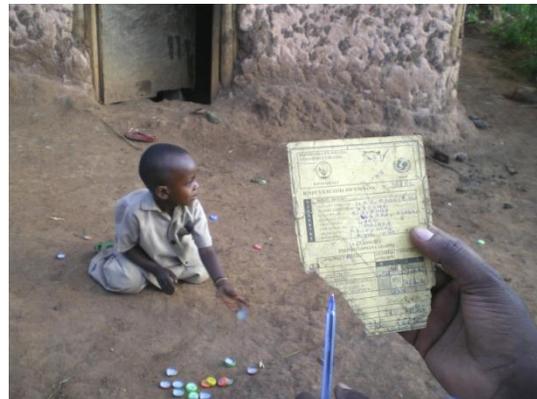


Table des matières

Remerciements.....	iii
Liste des Figures.....	v
Liste de Tableaux.....	vi
Liste des sigles et abréviations.....	vii
Mise en route : Qu'est-ce que la boîte à outils et comment ça marche ?.....	viii
Contexte général.....	viii
Qu'est-ce que la boîte à outils pour l'évaluation d'impact ?.....	viii
À qui la boîte à outils sert-elle ?.....	xi
Quelle est la portée de la boîte à outils ?.....	xi
Tour d'horizon des modules.....	xii
Aperçu des outils.....	xv
Tirer parti de l'expérience d'autres pays.....	xviii
Adaptation des outils aux besoins du pays.....	xix
Établissement de l'ordre de priorité des recommandations de la boîte à outils.....	xix
Module 1. Choix des questions d'évaluation.....	1-2
La promesse du financement basé sur les résultats et les lacunes de données.....	1-3
Qu'est-ce que l'évaluation d'impact ?.....	1-7
Pourquoi évaluer les programmes de FBR ?.....	1-8
Formulation des questions d'évaluation.....	1-9
Théorie du changement pour le financement de la santé basé sur les résultats.....	1-9
Questions d'évaluation d'impact.....	1-17
Questions de première génération : Le FBR fonctionne-t-il ?.....	1-18
Questions de deuxième génération : Comment améliorer le fonctionnement des FBR ?.....	1-22
Portefeuille d'évaluation financé par le HRITF.....	1-22
Module 2. Constitution de l'équipe d'évaluation d'impact.....	2-2
Enquêteur principal.....	2-6
Coordinateur de l'évaluation.....	2-7
Analyste des données.....	2-8
Expert du calcul de puissance.....	2-8
Expert de la qualité des données.....	2-9
Superviseur externe.....	2-10
Expert de la recherche qualitative.....	2-11
Expert de l'analyse des coûts.....	2-12
Participation de chercheurs locaux à l'évaluation d'impact.....	2-13
Module 3. Conception de l'évaluation d'impact.....	3-2
Contexte et justification.....	3-5
Cadre des résultats du FBR dans le secteur de la santé.....	3-5
Questions de recherche et pertinence stratégique.....	3-5
Indicateurs des résultats et des réalisations à l'étude.....	3-7
Stratégie d'identification.....	3-8
Échantillon.....	3-12
Besoins de données.....	3-20

Calendrier	3-22
Équipe d'EI	3-23
Plan de diffusion	3-23
Budget.....	3-24
Module 4. Préparation de la collecte des données.....	4-2
Planification des activités à l'aide d'un diagramme de Gantt	4-5
Définition de la propriété des données	4-5
Protection des sujets humains.....	4-6
Définition de la stratégie de saisie des données	4-13
Recrutement d'un bureau d'études	4-15
Comprendre et adapter les instruments d'enquête.....	4-21
Instruments d'enquête	4-27
Module 5. Conduite de la collecte des données	5-2
Cadre d'échantillonnage et échantillon.....	5-4
Définition de codes d'identification uniques.....	5-8
Identificateurs des groupes de traitement et de comparaison.....	5-13
Essai préalable des questionnaires.....	5-14
Gestion de la saisie des données	5-15
Risques sur le plan de la validité interne de l'EI pendant la collecte des données	5-18
Plan de travail sur le terrain.....	5-19
Recrutement et formation des membres des équipes de terrain.....	5-21
Pilotage des questionnaires.....	5-23
Gestion du travail sur le terrain	5-24
Préparation des rapports.....	5-24
Module 6. Stockage des données et accès à ces dernières	6-2
Pourquoi utiliser un catalogue de données ?	6-4
Types de données	6-5
Étapes du catalogage des données d'enquête	6-6
Catalogue de données FBR : Caractéristiques	6-7
Accès aux données.....	6-8
Module 7. Analyse des données et diffusion des résultats.....	7-2
Rapport de référence.....	7-5
Nettoyage des données	7-6
Création de variables aux fins d'analyse.....	7-7
Validation de la méthode d'évaluation	7-8
Analyse des impacts.....	7-12
Diffusion.....	7-14
Module 8. Suivi et documentation des programmes FBR.....	8-2
Suivi des interventions FBR.....	8-3
Suivi et documentation dans le contexte de l'évaluation d'impact	8-6
Sources de données complémentaires.....	8-9
References	1

Remerciements

La Boîte à outils pour l'évaluation de l'impact (EI) du financement basé sur les résultats (FBR) dans le domaine de la santé a été conçue par le HDNHE (santé, nutrition et population), en collaboration avec les équipes régionales FBR et le Groupe de recherche économique sur le développement (DECRG).

Les auteurs de la boîte à outils sont Christel Vermeersch (Banque mondiale, TTL), Elisa Rothenbühler (Banque mondiale) et Jennifer Sturdy (Millennium Challenge Corporation). Leurs points de vue exprimés ici ne coïncident pas nécessairement avec ceux de la Banque mondiale.

Les auteurs tiennent à remercier les nombreux membres du personnel et consultants de la Banque mondiale, ainsi que les homologues de différents pays qui ont contribué à l'élaboration de la Boîte à outils EI pour le FBR dans le domaine de la santé. Adam Ross (Banque mondiale), Sebastian Martinez (Banque interaméricaine de développement) et Paul Gertler (Université de Californie, Berkeley) ont contribué à la conception initiale de la boîte à outils. Des chefs d'équipe de projet et enquêteurs régionaux, notamment Tekabe Belay (Afghanistan), Rafael Cortez (Argentine), Adam Ross (Argentine), Paula Giovagnoli (Argentine), Luis Perez (Argentine), Christophe Lemièrre (Bénin), Gaston Sorgho et Paul Jacob Robyn (Cameroun/République centrafricaine), Jean-Jacques Frère (République démocratique du Congo), Juliette Seban (République démocratique du Congo), Dinesh Nair (Nigéria), Paulin Basinga (Rwanda) et Ashis Das (Zambie) ont partagé leurs expériences régionales de l'EI dans la rubrique « Pleins feux sur le pays ». Nous tenons tout particulièrement à remercier le Dr. Muhammad Pate, ministre d'État de la Santé de la République fédérale du Nigéria, qui a bien voulu nous faire connaître son point de vue sur la façon dont l'EI peut faire la différence. Edit Velenyi a participé à des entretiens et à la rédaction des rubriques « Pleins feux sur le pays ». Andres Martinez (Université du Michigan) a mis à notre disposition ses compétences en matière de calcul de puissance. Alvaro Canales, Victor Canales, Beatriz Godoy, Juan Muñoz et Mario Navarrete, de Sistemas Integrales, ont participé à l'élaboration des questionnaires, des procédures d'échantillonnage et d'enquête, des matériels de formation et des programmes de saisie de données CS-Pro. David Peters et Shivam Gupta ont dirigé l'équipe de chercheurs de l'Université Johns Hopkins qui a donné des conseils sur les indicateurs et mesures clés de santé maternelle et infantile, et établi une première version du guide sur les questionnaires et les indicateurs à l'intention des centres de santé. Ashis Das, Jed Friedman, Paul Jacob Robyn, Gaston Sorgho, Monique Vledder et Damien de Walque ont testé les questionnaires à l'intention des ménages et des centres de santé en Zambie, au Cameroun et en République centrafricaine et fourni à l'équipe une rétro-information extrêmement précieuse. Bianca Brijnath (Université Monash) a participé à l'élaboration des termes de référence pour les travaux qualitatifs, Logan Brenzel a fait de même pour le calcul des coûts des travaux et des questionnaires. Joe Naimoli (USAID) a contribué à l'élaboration du cadre conceptuel pour le suivi et la documentation des programmes FBR. Matthew Welch, Olivier Dupriez, Akiko Sagesaka et Mehmood Asghar ont aidé l'équipe à établir le catalogue de données FBR et les directives sur l'accès aux données. Petra Vergeer et Gyuri Fritsche ont donné des conseils sur l'exécution des programmes FBR. Alaka Holla, Tony Measham et Petra Vergeer ont formulé des commentaires sur les versions préliminaires de la boîte à outils. Darren Dorkin et Rama Lakshminarayanan ont apporté un appui constant à l'élaboration de la boîte à outils en tant que directeurs de programme pour le Fonds fiduciaire pour l'innovation dans le domaine de la santé (HRITF). Enfin, Matthew Mulligan a mis ses compétences de chercheur au service de l'équipe. Cette dernière est également très reconnaissante à l'égard des évaluateurs-experts - Margaret Grosh, Christophe Lemièrre, Mattias Lundberg, Adam Wagstaff et Feng Zhao - dont les commentaires ont été très instructifs. La

présente boîte à outils a été élaborée à l'aide d'un généreux financement accordé par le Fonds fiduciaire pour l'innovation dans le domaine de la santé (HRITF) financé par plusieurs bailleurs de fonds.

Citation recommandée :

Vermeersch C., Rothenbühler E. et J.R. Sturdy, 2012. Boîte à outils pour l'évaluation d'impact : Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats. Version 1.0. Banque mondiale, Washington, DC.®

Liste des Figures

Figure 1 : Structure de base de la boîte à outils	x
Figure 5 : Calculs de puissance ex-ante	3-18
Figure 6 : Calculs de puissance ex-post.....	3-19
Figure 7: Exemple de calendrier pour un document de conception de l'évaluation d'impact	3-23
Figure 8 : Copies d'écran des questionnaires de la boîte à outils pendant le processus d'adaptation nationale	4-26
Figure 9 : Capture d'écran d'un fichier de contrôle de l'identification	5-13
Figure 10 : Fichier d'identificateurs du lieu géographique et de l'assignation (traitement/comparaison)	5-14
Figure 11 : Gestion des données d'El.....	6-6
Figure 12 : Exemple de plan de rapport de référence	7-6
Figure 13 : Exemple de plan de rapport final.....	7-14

Liste de Tableaux

Tableau 1 : Liste de recommandations	xx
Tableau 2 : Interventions évaluées par pays	1-23
Tableau 3 : Résultats étudiés par pays.....	1-24
Tableau 4 : Temps de travail jugé nécessaire à un enquêteur principal.....	2-5
Tableau 5 : Temps de travail estimatif nécessaire à un coordinateur de l'évaluation.....	2-6
Tableau 6 : Temps de travail jugé nécessaire à un analyste des données.....	2-7
Tableau 7 : Temps de travail jugé nécessaire à un expert de la qualité des données.....	2-9
Tableau 8 : Temps de travail jugé nécessaire à un superviseur externe de la qualité des données.....	2-10
Tableau 9 : Temps de travail jugé nécessaire à un expert de la recherche qualitative.....	2-10
Tableau 10 : Temps de travail jugé nécessaire pour un expert de l'analyse des coûts.....	2-11
Tableau 11 : Besoins de données.....	3-21
Tableau 12 : Instruments d'enquête, participants et méthodes de vérification	4-11
Tableau 13 : Avantages et inconvénients de choisir un ou plusieurs bureaux d'études.....	4-15
Tableau 14 : Calendrier des produits escomptés et des paiements.....	4-20
Tableau 15 : Résumés des méthodes d'évaluation des connaissances des prestataires et de la qualité des soins fournis	4-33
Tableau 16 : Les rapports d'évaluation d'impact et leurs fonctions.....	7-4

Liste des sigles et abréviations

CCT	Transfert monétaire conditionnel
CD	Comité de déontologie
CE	Coordinateur de l'évaluation
CEI	Comité d'examen institutionnel
CEI	Comité d'éthique indépendant
CEP	Chef d'équipe de projet
CIG	Corrélation intra-grappe
CO	Contrat d'objectifs
Co-EP	Co-enquêteur principal
CS	Centres de santé
CTAO	Contrôles sur le terrain assistés par ordinateur
DC	Dublin Core
DDI	Initiative de Dublin pour les données
DECRG	Groupe de recherche sur l'économie du développement (service de la Banque) mondiale)
EDS	Enquêtes démographiques et sanitaires
EI	Évaluation d'impact
EP	Enquêteur principal
FBP	Financement basé sur la performance
FBR	Financement basé sur les résultats
HDNHE	Santé, nutrition et population (service de la Banque mondiale)
HRITF	Fonds fiduciaire pour l'innovation dans le domaine de la santé
ID	Identification
OMD	Objectifs du Millénaire pour le développement
OMS	Organisation mondiale de la santé
ONG	Organisation non gouvernementale
OSD	Opérateur de saisie des données
PA	Protocole d'accord
PDF	Portable Document Format
PMTCT	Prévention de la transmission de mère à enfant
PPT	Powerpoint©
RCA	République centrafricaine
SAN	Soins anténatals
SIGS	Système d'information de gestion pour la santé
SMI	Santé maternelle et infantile
STATA©	Programme d'analyse statistique
UPE	Unité primaire d'échantillonnage
VIH/SIDA	Virus de l'immunodéficience humaine/Syndrome de l'immunodéficience acquise

Mise en route : Qu'est-ce que la boîte à outils et comment ça marche ?

Contexte général

La boîte à outils a été élaborée avec l'aide financière du Fonds fiduciaire pour l'innovation dans le domaine de la santé (HRITF). Ce fonds fiduciaire a pour objectifs d'élaborer, de mettre en œuvre et d'évaluer des programmes pilotes viables de financement basé sur les résultats (FBR) capables d'améliorer la santé maternelle et infantile et d'accélérer ainsi les progrès vers la réalisation des OMD 1c, 4 et 5. Il veille en particulier à ce que chacun des projets nationaux de FBR prévoit une évaluation d'impact rigoureuse et bien conçue pour déterminer dans quelle mesure et dans quelles circonstances les programmes FBR sont efficaces et réalisables sur le plan opérationnel. Les évaluations sont essentielles pour produire les informations qui permettront d'étayer et d'améliorer le FBR dans les pays pilotes du HRITF et ailleurs. Le HRITF offre des dons aux pays pour la mise en œuvre de projets pilotes de FBR, d'activités axées sur le savoir et l'apprentissage, d'évaluations d'impacts et de travaux d'analyse¹.

Le programme de travail sur l'évaluation d'impact du financement basé sur les résultats comporte trois grands volets :

- Réalisation d'évaluations prospectives rigoureuses de l'impact des interventions de FBR sur l'accès aux services de santé, la qualité de la prestation de ces services, les dépenses de santé et les résultats sanitaires. Ces évaluations peuvent en outre se pencher sur la rentabilité et la complexité opérationnelle d'interventions de FBR de rechange.
- Coordination et normalisation, dans la mesure du possible, des méthodes d'évaluation au fil de multiples interventions de FBR afin de faciliter la comparaison de méthodes de rechange, d'évaluer la validité externe des impacts et de déterminer la faisabilité d'interventions similaires dans des contextes socioéconomiques et culturels différents.
- Synthèse et diffusion des acquis sous forme de matériel accessible et pertinent pour les décideurs nationaux et les autres parties prenantes.

Qu'est-ce que la boîte à outils pour l'évaluation d'impact ?

La boîte à outils pour l'évaluation d'impact est un guide pratique des méthodes de conception et de mise en œuvre des évaluations d'impact. Bien que plusieurs de ses éléments puissent servir à des évaluations générales, elle a pour but principal d'aider à évaluer l'impact des projets de financement basé sur les résultats (FBR) sur la santé maternelle et infantile. La boîte à outils est structurée comme suit :

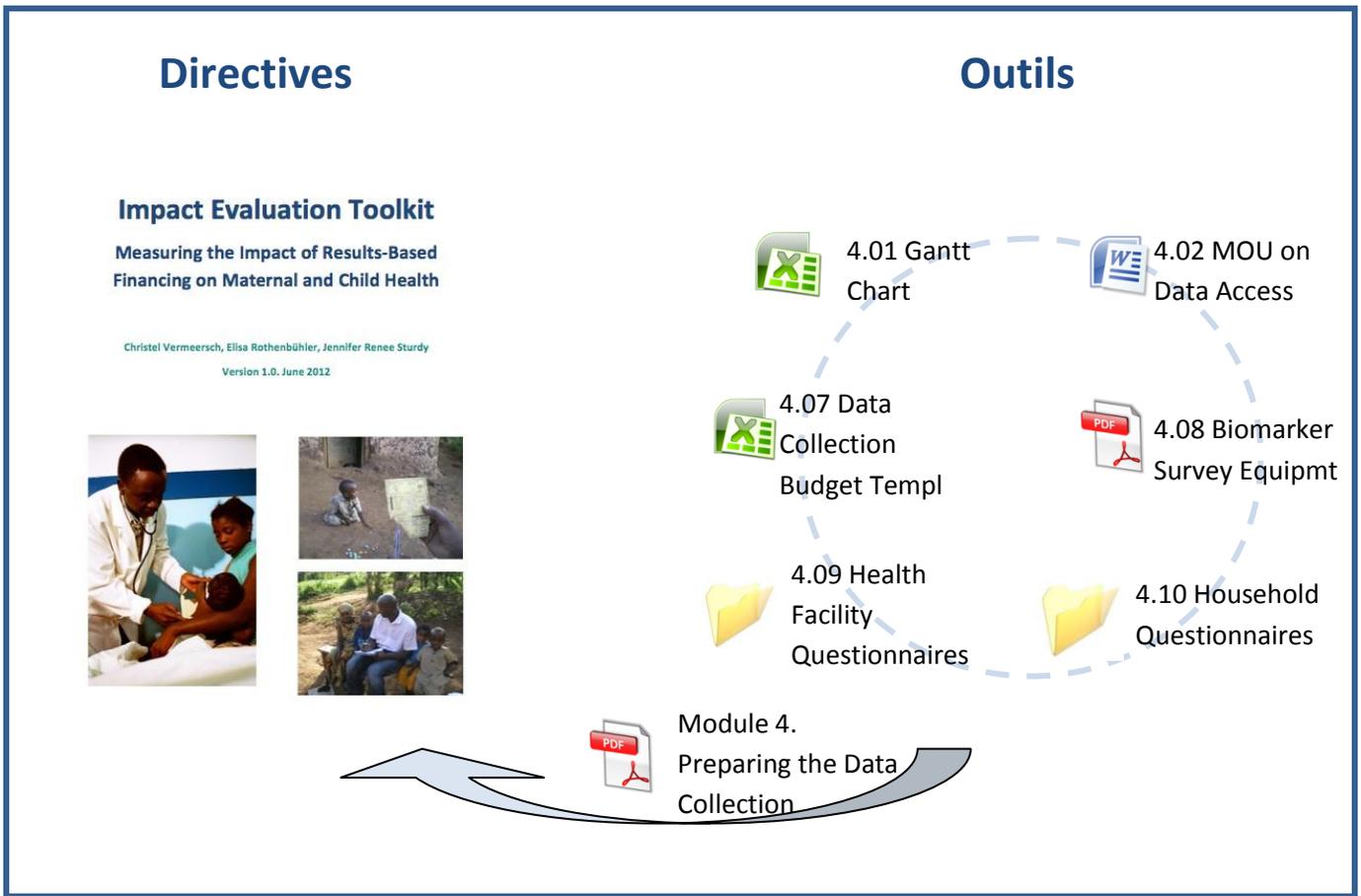
¹ Pour un aperçu du travail d'analyse, voir l'outil **0.01 – HRITF Analytical Work Program Overview (HRITF Aperçu analytique du programme de travail)**.

- Pour chaque étape du cycle d'évaluation d'impact (EI), la boîte à outils présente les meilleures pratiques dans un exposé narratif appelé ***Directives***. Chaque étape correspond à un module.
- Dans chaque module, la boîte à outils présente les ***outils*** techniques qui peuvent servir à la mise en œuvre des recommandations présentées dans les Directives : il y en a plus de 50, y compris les termes de référence des membres de l'équipe d'EI et des bureaux d'études, une liste d'indicateurs pertinents de la santé maternelle et infantile (SMI), des protocoles de recherche, des questionnaires, des manuels de formation et des programmes de cours destinés aux enquêteurs, du matériel de suivi du travail de terrain, des outils d'analyse des données, etc. Ces outils normalisés peuvent faciliter les comparaisons internationales des résultats des projets de FBR.

Si vous voulez utiliser au maximum de ses possibilités la boîte à outils pour l'évaluation d'impact, vous devez en outre accéder aux outils, les utiliser et les adapter conformément aux Directives suivantes. Pour accéder à la fois aux Directives et aux outils, veuillez visiter le site Web sur la boîte à outils pour l'évaluation d'impact :

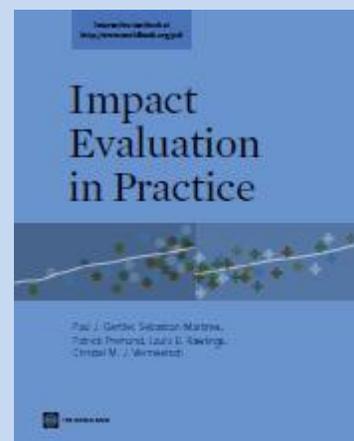
www.worldbank.org/health/impacetevaluationtoolkit

Figure 1 : Structure de base de la boîte à outils



La boîte à outils est le complément pratique du manuel intitulé « *L'évaluation d'impact en pratique* » (Gertler *et al.* 2011).

Le manuel aborde d'une manière plus approfondie la théorie de l'évaluation d'impact, tandis que la boîte à outils fournit les conseils et les outils pratiques qui serviront aux responsables des évaluations d'impact sur le terrain.



À qui la boîte à outils sert-elle ?

La boîte à outils est conçue pour aider les chefs d'équipes, les enquêteurs principaux, les chercheurs, les bureaux d'études, les parties prenantes gouvernementales et les autres membres des équipes nationales d'évaluation d'impact à concevoir et à mettre en œuvre les évaluations d'impact.

Quelle est le champ d'application de la boîte à outils ?

Financement basé sur les résultats et santé maternelle et infantile : La boîte à outils est conçue principalement pour la réalisation d'évaluations d'impacts des projets de FBR axés sur l'amélioration de la SMI. Elle utilise la définition du FBR proposée par le HRITF : tout programme conçu pour récompenser la production d'un ou plusieurs résultats à l'aide de mesures d'incitation financière ou autre, sous réserve d'un contrôle préalable de la production effective du ou des résultats en question. Les mesures d'incitation peuvent s'adresser aux prestataires de services (offre), aux bénéficiaires du programme (demande) ou aux deux. Il ne sera pas question dans l'immédiat de définitions plus larges du FBR comme celle associée aux groupes homogènes de malades.

Évaluations d'impact prospectives randomisées : La boîte à outils est conçue pour les équipes qui sont déjà convenues de procéder à une évaluation d'impact prospective. Bien que d'autres méthodes d'évaluation soient envisageables (par exemple, l'évaluation d'impact rétrospective) et risquent, dans certaines situations, de s'avérer préférables, de telles méthodes sortent du cadre d'application de la présente boîte à outils.

Bien que le quart des éléments de la boîte à outils aient été conçus spécialement aux fins du FBR et/ou de la SMI, le reste, soit **75 %, est destiné à un usage général**. Les praticiens qui ont une connaissance approfondie de leur domaine d'intérêt ont tout loisir d'adapter la portion spécialisée de la boîte à outils à un domaine différent.

Évaluations prospectives

Les évaluations prospectives sont prévues dès la conception du programme et font partie intégrante de sa mise en œuvre. Les données de l'enquête de base (ou enquête de référence) sont collectées avant la mise en place du programme tant pour le groupe de traitement que pour le groupe de comparaison.

L'évaluation d'impact en pratique, Gertler et al. 2011

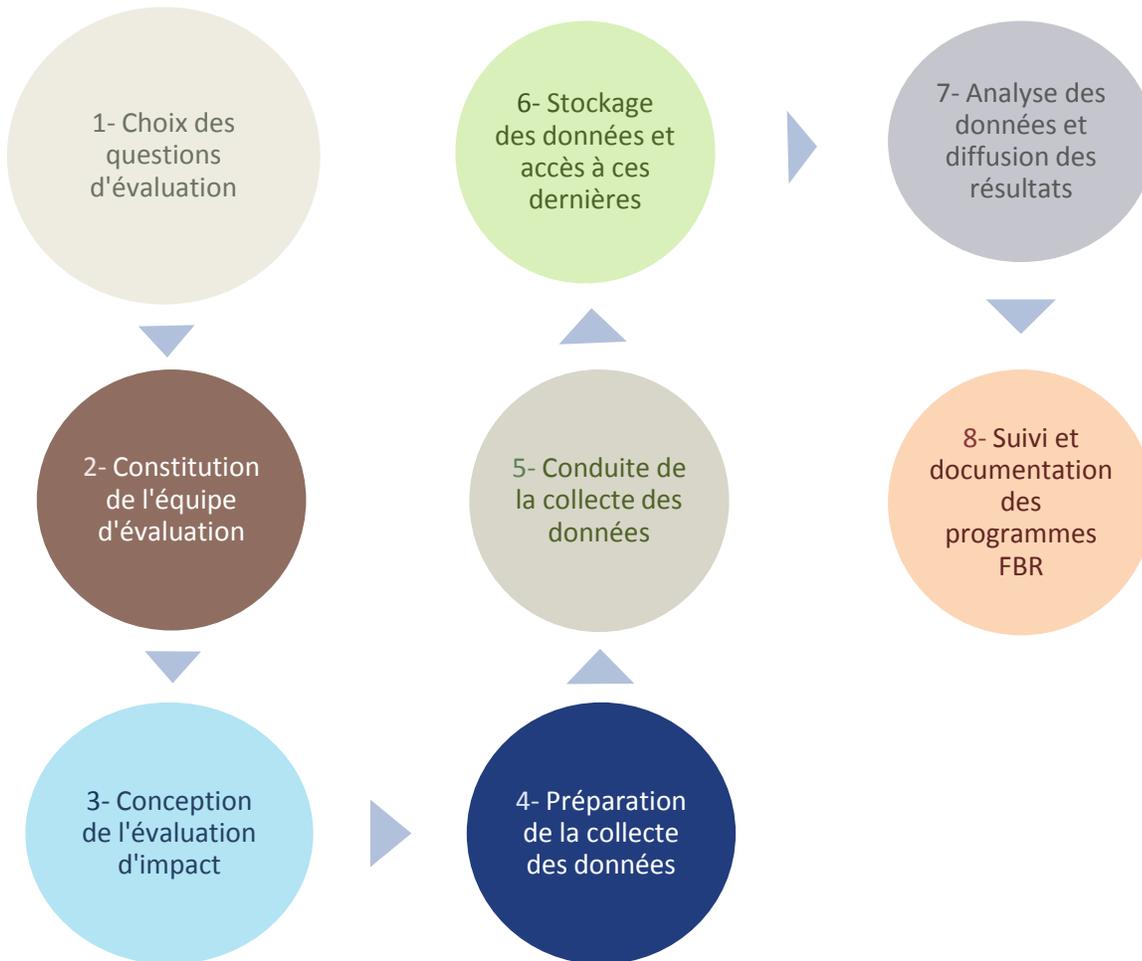
Tour d'horizon des modules

La boîte à outils se compose de huit modules qui portent chacun sur des étapes différentes du cycle de l'évaluation d'impact :

Module 1. Choix des questions d'évaluation. Chaque évaluation d'impact a pour but ultime d'**étayer la prise de décisions** qui permettront de renforcer les systèmes de santé et d'améliorer la santé des populations. Il est donc essentiel de définir des questions d'évaluation pertinentes à chaque pays et qui contribueront à l'élaboration d'une base de connaissances globale sur le FBR. Les questions stratégiques devraient non seulement permettre de vérifier i) que le FBR fonctionne, mais également de déterminer ii) pourquoi il fonctionne. La théorie du changement sur laquelle s'appuie l'intervention de FBR établit le cadre des questions stratégiques et servira à l'étape de la conception de l'évaluation d'impact.

Module 2. Constitution de l'équipe d'évaluation en tenant compte des compétences des membres et du temps qu'ils seront en mesure de consacrer à l'exercice de leurs fonctions. Chaque EI doit être dirigée par un enquêteur principal dévoué et compétent, assisté par un co-enquêteur ou par un solide coordinateur de l'évaluation. La collaboration avec des chercheurs locaux peut accroître la sensibilité culturelle et institutionnelle, la portée et la crédibilité de l'analyse et de la présentation des résultats. Elle peut en outre contribuer à renforcer les capacités locales de direction des travaux d'évaluation d'impact menés dans le pays. Enfin, les enquêteurs auront besoin d'évaluer (et, le cas échéant, de renforcer) les aptitudes et les capacités locales de collecte de données, et d'identifier les personnes qui pourront s'acquitter de tâches complémentaires comme l'analyse des coûts et la recherche qualitative.

Figure 2 : Les huit modules de la boîte à outils



Module 3. Conception de l'évaluation d'impact

- L'équipe d'EI apprendra à **articuler une chaîne de résultats pour l'intervention de FBR** en : i) définissant la ou les interventions de FBR à mettre en œuvre dans le pays ; ii) définissant la population qui sera visée par un programme pilote ; iii) utilisant le cadre de la chaîne de résultats pour déterminer les intrants, les extrants, les activités et les indicateurs de résultats qui serviront à l'évaluation d'impact ; iv) élaborant les questions d'évaluation primaires et les hypothèses y afférentes.
- L'équipe d'EI apprendra à **élaborer une stratégie d'évaluation**, définie dans un document de conception de l'évaluation d'impact décrivant d'une manière rigoureuse l'effet causal de l'intervention. Il s'agit pour cela de définir un groupe (ou des groupes) de traitement et de comparaison et de recueillir des données initiales et des données finales sur ces groupes, de définir les critères d'inclusion pour le cadre

d'échantillonnage, et d'effectuer des calculs de puissance afin de déterminer la taille appropriée des échantillons de l'étude.

Module 4. Préparation de la collecte des données

- L'équipe d'EI devra **établir le diagramme de Gantt de l'EI et du projet** pour veiller à assurer une coordination adéquate entre l'intervention et les activités d'EI. Elle devrait coordonner son travail avec celui de l'équipe de conception du projet pour veiller à assurer la conformité du concept opérationnel de l'intervention et de l'évaluation, et veiller à ce que l'évaluation cadre bien dans le concept opérationnel. Il convient par ailleurs de veiller à une coordination efficace pour faire en sorte que les mesures initiales des indicateurs soient prises avant le début de l'intervention, et ménager une période d'exposition suffisante pour la ou les interventions de FBR.
- L'équipe d'EI apprendra à **élaborer le protocole de recherche et veillera à obtenir l'autorisation éthique** de l'étude conformément aux exigences locales.
- L'équipe d'EI sollicitera des conseils concernant **l'embauche d'un bureau d'études** justifiant des capacités et de l'expérience nécessaires pour la gestion d'activités de collecte de données à grande échelle dans de multiples sites.
- L'équipe d'EI apprendra à **élaborer les instruments d'enquête et les procédures de terrain** nécessaires pour la collecte des données. La qualité de l'EI dépend au premier chef de la qualité des données recueillies, laquelle dépend directement des instruments d'enquête et des procédures de terrain utilisés. Les instruments d'enquête ont été mis au point par l'équipe focale Santé, Nutrition et Population (HNP) à l'intention des équipes nationales afin d'optimiser la coordination et la normalisation des mesures effectuées dans chacun des pays. Toutefois, les instruments ont besoin d'être adaptés pour tenir compte des environnements culturels et institutionnels locaux.

Module 5. Conduite de la collecte des données. L'équipe d'EI apprendra à **assurer l'exécution et le suivi** des activités de formation et de collecte et de saisie des données, **et la communication adéquate** des informations sur ces activités pour veiller à ce que le bureau d'études adhère aux plans et protocoles convenus. Elle prendra par ailleurs les mesures voulues pour respecter l'échéancier dont elle est convenue avec les parties prenantes gouvernementales.

Module 6. Stockage des données et accès à ces dernières. L'équipe d'EI apprendra à élaborer un **plan de documentation, de stockage des données et d'accès à ces dernières** en collaboration avec les équipes de conception de projet et d'EI afin de garantir la sécurité, la confidentialité et la documentation des données.

Module 7. Analyse des données et diffusion des résultats. L'équipe d'EI apprendra à élaborer un **plan d'analyse et de diffusion des données** afin de veiller à la diffusion en temps opportun de descriptions et de produits d'analyse.

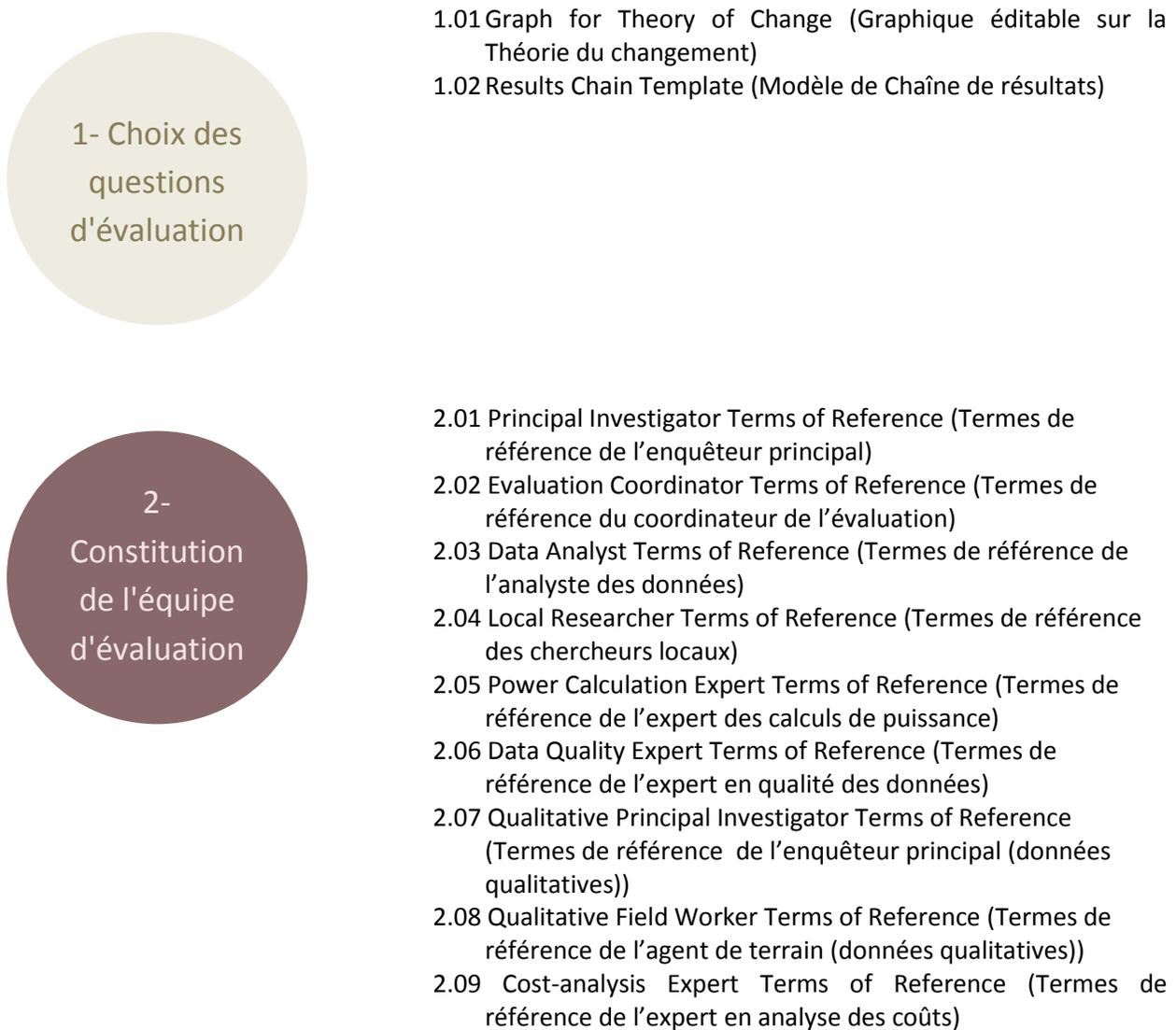
Module 8. Suivi et documentation des programmes FBR. L'équipe d'EI apprendra à élaborer un **plan de suivi et de documentation** pour assurer le contrôle de la mise en œuvre du projet, l'adhésion aux protocoles d'étude et d'assignation au groupe de traitement ou au groupe de comparaison, et pour recenser les sources de données

complémentaires comme les systèmes d'information de gestion de la santé (HMIS), et les données financières et administratives.

Aperçu des outils

Chaque module contient les outils utiles pour la mise en œuvre de l'étape correspondante de l'évaluation d'impact. Les outils mentionnés dans les Directives de la boîte à outils sont indiqués en **caractères gras rouges**.

Figure 3 : Outils





3- Conception de l'évaluation d'impact

- 3.01 RBF Output and Outcome Indicators (Indicateurs du FBR)
- 3.02 WHO Output and Outcome Indicators (Indicateurs de produits et de résultats (OMS))
- 3.03 IE Design Paper Template (Modèle de document de conception de l'EI)
- 3.04 IE Budget Template (Modèle de budget EI)
- 3.05 Ex ante Power Calculation Example (Exemple de calcul de puissance ex ante)
- 3.06 Power Calculations for Binary variables (Calculs de puissance binaires)
- 3.07 Power Calculation References (Références Calcul de puissance)



4- Préparation de la collecte des données

- 4.01 Impact Evaluation Gantt Chart (Graphique de Gantt pour l'évaluation d'impact)
- 4.02 Memorandum of Understanding on Data Access (Protocole d'accord sur l'accès aux données)
- 4.03 Research Protocol Example (Exemple de protocole de recherche)
- 4.04 Informed Consent Templates (Modèles de consentement éclairé)
- 4.05 Health Facility Survey Firm TOR (Termes de référence du bureau chargé des enquêtes formations sanitaires)
- 4.06 Household Survey Firm TOR (Termes de référence du bureau chargé des enquêtes ménages)
- 4.07 Data Collection Budget Template (Modèle de budget pour la collecte des données)
- 4.08 Consumables and Equipment for Biomarker Data (Produits consommables et matériel pour les données des biomarqueurs)
- 4.09 Health Facility Questionnaires (Questionnaires Formations sanitaires)
- 4.10 Household Questionnaires (Questionnaires Ménages)
- 4.11 Community Questionnaires (Questionnaires Communauté)
- 4.12 Costing Questionnaires (Questionnaires relatifs aux coûts)
- 4.13 Data Entry Program (Programme d'entrée des données)
- 4.14 Anemia Referral Guidelines (Directives relatives à l'orientation des cas d'anémie)
- 4.15 Anemia Referral Form (Formulaire relatif à l'orientation des cas d'anémie)
- 4.16 How to Translate Questionnaires (Comment traduire les questionnaires)
- 4.17 Institutional Review Board TOR (Termes de référence du Conseil d'examen institutionnel)
- 4.18 Certificate of Accurate Translation (Certificat d'exactitude de la traduction)

5- Conduite de la collecte des données

- 5.01 Interview Duration Tracking Sheet (Traçage de la durée des entretiens)
- 5.02 Enumerator Evaluation Form (Formulaire d'évaluation de l'enquêteur)
- 5.03 Survey Progress Report I (Word) (Rapport sur l'avancement de l'enquête I (Word))
- 5.04 Survey Progress Report II (Excel) (Rapport sur l'avancement de l'enquête II (Excel))
- 5.05a Household survey Field Manual (Manuel de terrain pour les enquêtes ménages)
- 5.05b Household survey Training Program (Programme de formation aux enquêtes ménages)
- 5.05c Household survey Training (PPTs de Formation aux enquêtes ménages)
- 5.06a Health Facility Survey Field Manual (Manuel de terrain pour les enquêtes formations sanitaires)
- 5.06b Health Facility Survey Training (Formation aux enquêtes pour les formations sanitaires)
- 5.07 Survey Training, CAR & Cameroon (Formation aux enquêtes, Exemples de RCA & Cameroun)
- 5.08 Health Facility Supervisor Checklist (Liste de contrôle du superviseur formations sanitaires)
- 5.09 Health Facility Arrival Checklist (Liste de contrôle d'arrivée en formation sanitaire)
- 5.10 Health Facility Supervisor Tracking Form (Formulaire de traçage pour le superviseur formation sanitaire)
- 5.11a Daily Listing of U5 Exit Interviews (Liste quotidienne des entretiens de sortie des enfants de moins de 5 ans)
- 5.11b Daily Listing of ANC Interviews (Liste quotidienne des entretiens de sortie des soins prénataux)
- 5.12 Cash Management Sheet (Fiche de suivi des décaissements)

- 6.01 Data Deposit Form – IE Micro-data Catalog (Formulaire de dépôt des données – Catalogue de micro-données EI)
- 6.02 Nesstar Data Storage Templates (Modèles de stockage des données Nesstar)
- 6.03 Login to Micro-data Management Toolkit (Accès à l'outil de la gestion des micro-données)
- 6.04 How to Access the Data Catalog and Data (Comment accéder au catalogue de données et aux données)

6- Stockage des données et accès à ces dernières

7- Analyse des données et diffusion des résultats

7.01 Household Baseline Report (Rapport de base d'enquête ménages)

7.01a Handbook Household Baseline Report (Manuel pour le Rapport de base d'enquête ménages)

7.01b Rwanda Household Baseline Outcome Indicators (Indicateurs enquêtes ménages de base au Rwanda)

7.01c Rwanda Household Baseline Report Do-Files (STATA) (Fichiers Do STATA enquêtes ménages de base au Rwanda)

7.01d Rwanda Household Baseline Ex-post Power Calculations (Calculs de puissance ex-post enquêtes ménages de base au Rwanda)

7.02 Health Facility Baseline Report (Rapport de base d'enquête Formations sanitaires)

7.02a Suggested detailed outline of health facility baseline report (Plan détaillé suggéré pour le rapport de base formations sanitaires)

7.03 Community Health Worker (CHW) Baseline Report (Rapport de base d'enquête Agents de santé communautaires (ASC))

7.03a Rwanda CHW Baseline Report Do-Files (STATA) (Fichiers Do STATA enquêtes ASC de base au Rwanda)

7.04 STATA ado file for Baseline balance table (Fichier Ado STATA pour tableau d'équilibre de l'enquête de base)

7.05 WHO Anthro calculation package (Pack de calcul OMS Anthro)

7.06 STATA training (Formation STATA)

7.07 STATA Training IE Design Validation (Formation STATA pour la Validation de la conception de l'EI)

8- Suivi et documentation des programmes FBR

8.01 Monitoring Indicators Rwanda Example (Modèle d'indicateurs de suivi pour le Rwanda)

8.02 Field Supervision Visit Templates Rwanda (Modèles de visites de supervision sur le terrain)

Tirer parti de l'expérience d'autres pays

La boîte à outils offre une série d'**études de cas nationales** illustrant les enjeux et les enseignements tirés d'évaluations d'impacts de programmes de FBR concrets. La plupart de ces documents ont été produits dans le cadre d'évaluations d'impacts financées par le HRITF, mais la boîte à outils propose également d'autres études

de cas intéressantes. Les études de cas ont été produites avec la collaboration de chefs d'équipes de projets et d'équipes d'évaluation d'impact. Les Directives de la boîte à outils ne contiennent que des extraits de ces études. On peut toutefois accéder aux documents complets dans la section du site Web concernant la Boîte à outils et intitulée « Pleins feux sur le pays ».

Adaptation des outils aux besoins du pays

Les **contenus propres aux pays** disponibles dans les outils sont mis en évidence pour permettre aux équipes d'EI de les adapter facilement à leur contexte national. Ces contenus sont indiqués en **caractères rouges** (dans les questionnaires) ou en caractères **surlignés en jaune** (dans la plupart des autres outils).

Établissement de l'ordre de priorité des recommandations de la boîte à outils

La boîte à outils a pour objectif d'établir des normes de qualité et de rigueur scientifique en proposant un ensemble complet de recommandations et d'outils. Cependant, les conditions réelles, les ressources financières, le dialogue national et le contexte influent sur la faisabilité de certaines recommandations. Le tableau 1 résume les principales recommandations formulées dans le cadre de la boîte à outils ; il vise à aider les équipes d'EI à établir un ordre de priorité entre ce que l'on peut considérer comme :

- Essentiel — c'est-à-dire ce qui doit impérativement être inclus ou envisagé lors de la conception et de la mise en œuvre des évaluations d'impact ;
- Important — c'est-à-dire ce qui devrait idéalement être inclus ou envisagé, en procédant le cas échéant à des changements ou ajustements ;
- Souhaitable — c'est-à-dire ce que les équipes d'EI sont invitées à inclure, mais dont elles pourraient faire l'économie le cas échéant.

Tableau 1 : Liste de recommandations

Module	Recommandations	Essentiel	Important	Souhaitable
1	<ul style="list-style-type: none"> La pertinence des questions stratégiques/d'évaluation retenues, tant sur le plan local que global, importe plus que le nombre de questions abordées. Il n'est pas possible d'examiner toutes les dimensions du FBR dans le cadre d'une seule et même évaluation d'impact ; l'équipe devra donc établir un ordre de priorité. 	✓		
1	<ul style="list-style-type: none"> Le contrôle de l'utilité du FBR constitue une première étape. Il est essentiel, pour pouvoir améliorer le programme de FBR et en assurer la viabilité, de bien cerner les raisons des échecs subis ou des succès remportés. 		✓	
2	<ul style="list-style-type: none"> Le(s) membre(s) de l'équipe chargé(s) principalement de la conception et de l'exécution du projet (par exemple, le chef d'équipe du projet) ne devraient pas jouer le rôle d'enquêteur principal. 	✓		
2	<ul style="list-style-type: none"> L'enquêteur principal et le coordinateur de l'évaluation jouent un rôle crucial dans la supervision du/des bureau(x) d'études. 	✓		
2	<ul style="list-style-type: none"> Les homologues locaux pour la recherche peuvent contribuer grandement au succès de l'évaluation d'impact parce qu'ils font intervenir leur connaissance du contexte local et favorisent l'appropriation du programme par le pays. 		✓	
2	<ul style="list-style-type: none"> Les équipes devraient évaluer la capacité locale à effectuer des enquêtes et déterminer si un appui technique devra être fourni pour assurer la qualité des données d'enquête. 		✓	
2	<ul style="list-style-type: none"> Un expert de la qualité des données peut aider à créer les conditions initiales appropriées pour assurer la qualité des données d'enquête avant que le bureau d'études se rende sur le terrain. Un superviseur local peut vérifier les processus d'assurance de la qualité des données durant l'exécution des enquêtes. 		✓	
2	<ul style="list-style-type: none"> Une analyse qualitative et une analyse du rapport coût-efficacité peuvent ajouter une richesse et une granularité considérables aux questions auxquelles répondra l'évaluation d'impact. 		✓	
2	<ul style="list-style-type: none"> Les évaluations d'impact nécessitent plusieurs séries de données complexes, et un bon analyste des données aidera l'équipe à gérer et analyser les données de façon rapide et fiable. 		✓	
2	<ul style="list-style-type: none"> Les calculs de puissance peuvent être confiés à l'enquêteur principal, mais un expert du calcul de puissance peut avoir davantage de temps et d'expertise à consacrer à cette tâche. 			✓
3	<ul style="list-style-type: none"> Il convient de concevoir une évaluation d'impact prospective avant ou pendant l'intervention. 	✓		

Module	Recommandations	Essentiel	Important	Souhaitable
3	<ul style="list-style-type: none"> Les équipes devraient élaborer un cadre des résultats pour le projet de FBR afin d'identifier les principaux canaux à travers lesquels les activités du programme de FBR influenceront sur les résultats et les réalisations clés. 	✓		
3	<ul style="list-style-type: none"> L'assignation aléatoire aux groupes d'intervention(s) et de comparaison est la stratégie d'identification recommandée aux fins de l'évaluation d'impact du FBR. 		✓	
3	<ul style="list-style-type: none"> Les équipes devraient évaluer les risques qui pèsent ou qui pèseront sur la validité interne de l'évaluation (par exemple, contamination, manque de puissance) et les contrôler au fil du temps. 	✓		
3	<ul style="list-style-type: none"> Les calculs de puissance constituent un élément important de la conception d'une évaluation d'impact. À défaut d'une puissance suffisante, l'évaluation d'impact pourrait ne pas être en mesure de répondre aux principales questions stratégiques. La taille de l'échantillon doit autoriser une puissance suffisante. 	✓		
3	<ul style="list-style-type: none"> L'échantillon doit être représentatif de la population qui, à terme, bénéficiera du programme. 	✓		
3	<ul style="list-style-type: none"> Lorsque les partenaires nationaux adoptent le concept de l'évaluation d'impact et reconnaissent l'importance de respecter les branches de l'étude, il devient plus facile de maintenir les groupes de traitement et de comparaison intacts jusqu'à la réalisation de l'enquête de suivi. 	✓		
3	<ul style="list-style-type: none"> Le choix des indicateurs de l'étude est crucial. Chacun des indicateurs retenus doit être mesurable à l'aide des instruments choisis de collecte de données. 	✓		
3	<ul style="list-style-type: none"> Les équipes peuvent se référer à l'ouvrage « <i>L'évaluation d'impact en pratique</i> »(Gertler <i>et al.</i>, 2011) pour un examen approfondi des stratégies d'identification de l'évaluation d'impact. 		✓	
3	<ul style="list-style-type: none"> Au moment de définir l'unité de randomisation, les équipes doivent faire les arbitrages requis entre la puissance de l'évaluation d'impact et le risque de contamination croisée entre les unités de randomisation. 		✓	
4	<ul style="list-style-type: none"> Le protocole de recherche doit contenir toutes les informations pertinentes concernant la protection des sujets humains, y compris des critères d'échantillonnage spécifiques et des protocoles relatifs au consentement éclairé et à la confidentialité des données. 	✓		
4	<ul style="list-style-type: none"> L'évaluation d'impact doit être approuvée par un comité d'examen institutionnel : l'enquêteur principal doit veiller à confier à un tel comité l'examen éthique et l'approbation de la recherche <u>avant</u> le début des activités sur le terrain. 	✓		
4	<ul style="list-style-type: none"> Un diagramme de Gantt du projet/de l'évaluation d'impact peut aider les équipes à coordonner leurs activités et les échéanciers à partir du projet ou de l'évaluation d'impact. 		✓	

Module	Recommandations	Essentiel	Important	Souhaitable
4	<ul style="list-style-type: none"> L'équipe d'évaluation d'impact doit convenir avec les partenaires de l'administration publique de la politique d'accès à l'information issue de l'évaluation d'impact. Un protocole d'accord écrit peut éviter les malentendus dans ce domaine. 		✓	
4	<ul style="list-style-type: none"> Le choix d'opter pour la saisie de terrain assistée par ordinateur (STAO) ou pour la saisie des données sur le terrain a une grande incidence sur le choix du bureau d'études et devrait être fait avant de recruter ce dernier. 		✓	
4	<ul style="list-style-type: none"> Le recrutement du bureau d'études est un processus de longue haleine qui prend d'ordinaire de 3 à 6 mois et qui devrait être entamé dès les premières étapes de la planification du projet. 	✓		
4	<ul style="list-style-type: none"> Selon la situation et l'expérience du pays, on pourra choisir de confier à un seul et même bureau les enquêtes sur les ménages et sur les établissements de santé, ou de recruter deux bureaux distincts. Nous recommandons en règle générale que l'équipe ait recours à cette fin à une procédure d'appel à la concurrence. 		✓	
4	<ul style="list-style-type: none"> L'équipe de gestion de l'enquête devrait être composée d'un responsable de projet, d'un administrateur local et d'un gestionnaire des données, et ce pour toute la durée de la préparation et de la mise en œuvre des activités de collecte des données. 	✓		
4	<ul style="list-style-type: none"> La conduite des négociations avec le bureau d'études nécessite une connaissance approfondie des contraintes budgétaires et temporelles qui auront une incidence sur la composition de l'équipe de terrain et sur la durée de l'enquête. 	✓		
4	<ul style="list-style-type: none"> Le bureau d'études devrait pouvoir compter dès les étapes initiales de la préparation de l'enquête sur les services d'un expert de la qualité des données, en particulier s'il s'agit d'une entreprise locale aux capacités limitées. 		✓	
4	<ul style="list-style-type: none"> La structure et la qualité des instruments d'enquête jouent un rôle essentiel dans l'assurance de la qualité et la comparabilité des résultats obtenus pour chaque pays. Nous recommandons que les équipes de projet utilisent comme point de départ les questionnaires d'enquêtes sur le FBR conçus pour les établissements de santé et les ménages. L'enquêteur principal de l'évaluation doit déterminer quels sont les modules appropriés ou non, et veiller à ce que les questionnaires permettent de cerner les principaux points d'intérêt. Les équipes ne doivent pas hésiter à procéder aux ajustements qu'elles estiment nécessaires. 		✓	
4	<ul style="list-style-type: none"> Les questionnaires de la boîte à outils ont été conçus pour être exhaustifs, et les équipes souhaiteront peut-être limiter le nombre de modules utilisés pour réduire le coût et le temps requis 		✓	

Module	Recommandations	Essentiel	Important	Souhaitable
4	<p>pour les administrer.</p> <ul style="list-style-type: none"> les enquêtes auprès des collectivités peuvent permettre d'évaluer les infrastructures et les réseaux d'aide qui les caractérisent. Elles peuvent aussi servir de compléments aux enquêtes auprès des ménages, en particulier lorsque ces dernières ont besoin d'être sensiblement écourtées. 			✓
5	<ul style="list-style-type: none"> L'équipe d'évaluation d'impact et le bureau d'études doivent définir le protocole pour identifier clairement les observations dans les bases de données et pour établir des liens entre ces dernières. 	✓		
5	<ul style="list-style-type: none"> L'équipe d'évaluation d'impact doit définir le protocole pour identifier les domaines de traitement et de comparaison au sein des bases de données. 	✓		
5	<ul style="list-style-type: none"> La qualité et la durée de la formation des équipes de terrain jouent un rôle essentiel dans l'assurance du succès de la collecte de données. 	✓		
5	<ul style="list-style-type: none"> Les bureaux d'études s'occupent de la collecte des données, mais il appartient à l'équipe d'EI de collaborer avec ces derniers pour veiller à ce que les résultats du travail de terrain soient communiqués en temps utile et d'une manière appropriée. 	✓		
5	<ul style="list-style-type: none"> Le protocole de recherche et les manuels d'enquête doivent contenir toutes les informations dont les bureaux d'études auront besoin pour veiller à ce que la collecte des données soit effectuée conformément aux plans établis et dans le respect des principes d'éthique reconnus. 	✓		
5	<ul style="list-style-type: none"> Il convient de veiller à la sécurité et à la confidentialité des données pendant les opérations de collecte et de saisie. Les équipes de terrain doivent signaler tout problème logistique ou tout accroc à la sécurité. 	✓		
5	<ul style="list-style-type: none"> L'équipe d'EI doit surveiller étroitement les opérations de collecte et de saisie des données, et pourrait décider de faire appel aux services d'un expert de la qualité des données pour mieux s'acquitter de cette tâche. 		✓	
5	<ul style="list-style-type: none"> Les bureaux d'études locaux risquent de manquer de moyens en matière de programmation de la saisie des données, de saisie et de gestion des données. La boîte à outils propose des formulaires de saisie des données pour le programme CS-Pro qui correspondent aux questionnaires d'enquêtes auprès des ménages et des établissements de santé. 		✓	
5	<ul style="list-style-type: none"> Il est préférable d'entrer les données à mesure que le travail sur le terrain progresse, au lieu d'attendre à la fin de ce travail. 			✓
6	<ul style="list-style-type: none"> Le chef d'équipe doit veiller à la planification et à la coordination d'une documentation complète des activités d'évaluation d'impact. <ul style="list-style-type: none"> ▶ Inclure des notes de synthèse, des protocoles de recherche, des questionnaires, des manuels de 	✓		

Module	Recommandations	Essentiel	Important	Souhaitable
	formation, etc., à jour. ▶ Déterminer quels sont les éléments d'information qu'il convient de supprimer pour assurer la confidentialité des répondants.			
6	<ul style="list-style-type: none"> L'enquêteur principal doit préparer une ou des fichiers d'identification permettant d'établir des liens entre les codes d'identification géographique et les codes d'identification de terrain. 	✓		
6	<ul style="list-style-type: none"> L'enquêteur principal doit déterminer qu'elles sont les variables qu'il faudra éviter de publier (par exemple, informations personnelles sensibles). 	✓		
6	<ul style="list-style-type: none"> Les fichiers confidentiels (fichiers d'identification et autres données non accessibles au public) doivent être stockés de façon sécurisée, de préférence dans une enclave de données. 		✓	
6	<ul style="list-style-type: none"> Les équipes d'évaluation d'impact doivent prévoir suffisamment de temps pour la documentation et le téléchargement des données vers l'amont, afin de garantir la continuité de l'accès aux données au sein de l'équipe, et la facilité du partage futur des données et du processus d'analyse. 		✓	
6	<ul style="list-style-type: none"> L'équipe d'évaluation d'impact doit se référer au protocole d'accord (ou à un accord similaire de partage des données) pour ses opérations de documentation, de stockage ou de partage des données. 		✓	
7	<ul style="list-style-type: none"> Les analystes doivent maintenir un registre de toutes les opérations de transformation et d'analyse statistique effectuées sur les données. Les données d'origine doivent absolument être conservées intactes. Toute modification doit être sauvee sous la forme d'une différente série de données. 	✓		
7	<ul style="list-style-type: none"> Avant de procéder à l'analyse des données de référence, l'analyste doit consulter les lignes directrices internationales et nationales concernant le calcul des indicateurs (par exemple, directives de l'OMS). 		✓	
7	<ul style="list-style-type: none"> L'analyste peut participer au dépistage des erreurs qui peuvent survenir pendant la collecte ou la saisie des données de référence. Cela peut permettre de procéder par la suite à des ajustements de la formation et de la supervision au cours des cycles ultérieurs de collecte des données. 			✓
7	<ul style="list-style-type: none"> Il faut du temps pour procéder au nettoyage et à l'analyse des données et à la diffusion des résultats. Il est donc utile de prévoir à l'avance les ressources humaines et financières qui seront requises à cette fin. 			✓
7	<ul style="list-style-type: none"> Les calculs de puissance ex-post font partie des contrôles de la validité interne de l'évaluation d'impact. Ils peuvent le cas échéant suggérer des moyens d'accroître la puissance au moment du 			✓

Module	Recommandations	Essentiel	Important	Souhaitable
7	<ul style="list-style-type: none"> suivi. L'analyse doit être élaborée en gardant à l'esprit les meilleurs moyens de diffuser les résultats et d'informer les décideurs. 			✓
7	<ul style="list-style-type: none"> Les données d'évaluation d'impact sont d'ordinaire très riches : bien que l'analyse de l'impact du FBR puisse constituer l'objectif principal, d'autres analyses peuvent également permettre d'informer les décideurs. 			✓
8	<ul style="list-style-type: none"> Le suivi et la documentation des activités du projet constituent un complément crucial de l'évaluation d'impact parce qu'ils fournissent des informations sur les interventions effectives sur le terrain et par conséquent sur l'intervention que l'on est en train d'évaluer. 		✓	
8	<ul style="list-style-type: none"> Les équipes d'évaluation d'impact souhaiteront que le programme identifie deux risques majeurs pour l'évaluation d'impact : 1) une compensation offerte au groupe de comparaison sous la forme d'une intervention ou d'un programme alternatifs ; et 2) l'imitation du traitement par le groupe de comparaison. 		✓	



Module 1

Choix des questions d'évaluation

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats
Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

www.worldbank.org/health/impacetevaluationtoolkit

Module 1. Choix des questions d'évaluation

Principales recommandations et outils disponibles pour ce module			
Recommandations	Essentiel	Important	Souhaitable
<ul style="list-style-type: none"> La pertinence des questions stratégiques/d'évaluation retenues, tant sur le plan local que global, importe plus que le nombre de questions abordées. Il n'est pas possible d'examiner toutes les dimensions du FBR dans le cadre d'une seule et même évaluation d'impact ; l'équipe devra donc établir un ordre de priorité. Le contrôle de l'utilité du FBR constitue une première étape. Il est essentiel, pour pouvoir améliorer le programme de FBR et en assurer la viabilité, de bien cerner les raisons des échecs subis ou des succès remportés. 	✓		
			
Outils			
<ul style="list-style-type: none"> 1.01 Editable Graph for Theory of Change (Graphique éditable sur la Théorie du changement) 1.02 Results Chain Template (Modèle de Chaîne de résultats) 			

Contenu du module

La promesse du financement basé sur les résultats et les lacunes de données.....	1-3
Qu'est-ce que l'évaluation d'impact ?	1-7
Pourquoi évaluer les programmes de FBR ?	1-8
Formulation des questions d'évaluation.....	1-9
Théorie du changement pour le financement de la santé basé sur les résultats	1-9
Questions d'évaluation d'impact	1-17
Questions de première génération : Le FBR fonctionne-t-il ?.....	1-18
Questions de deuxième génération : Comment améliorer le fonctionnement des FBR ?	1-22
Portefeuille d'évaluation financé par le HRITF.....	1-22

La promesse du financement basé sur les résultats et les lacunes de données

La stratégie Santé, nutrition et population (HNP, 2007) de la Banque mondiale a réaffirmé l'importance accordée par la Banque aux résultats et au renforcement des systèmes de santé. L'un des principaux objectifs de cette stratégie est d'établir des liens plus étroits entre la fourniture de crédits et les résultats grâce à un recours accru au financement basé sur les résultats (FBR).

Musgrove (2010) définit le FBR pour la santé comme « *tout programme conçu pour récompenser la production d'un ou de plusieurs extrants ou résultats à l'aide de mesures d'incitation financière ou autre, sous réserve de la vérification préalable de la production effective du ou des résultats en question. Les mesures d'incitation peuvent s'adresser aux prestataires de services (offre), aux bénéficiaires du programme (demande) ou aux deux [...] Il est impératif de vérifier que les résultats ont été atteints. Idéalement, la vérification pourrait être réalisée par un tiers neutre, même si le mandant paie les coûts associés, mais de nombreux dispositifs sont possibles. La vérification ex-ante (avant le paiement) peut être complétée par une évaluation ex-post.* »

Bien qu'il n'ait pas encore vraiment fait ses preuves, le FBR dans le secteur de la santé est considéré comme un outil puissant pour renforcer les systèmes de santé et accélérer les progrès vers les OMD liés à la santé (Levine et Eichler 2009, Cochrane 2012). Malgré les nombreuses preuves des impacts positifs des programmes de transferts monétaires conditionnels sur les résultats en matière de développement humain (Fiszbein et Schady 2009) et les nombreuses données sur l'impact des bons de services sanitaires axés sur la demande (Meyer et al. 2011), peu de données existent concernant l'impact sur les indicateurs sanitaires des pays à faible revenu des interventions de FBR axées sur l'offre ou des interventions axées sur la demande qui ne reposent pas sur des transferts monétaires conditionnels². De fait, seules quelques évaluations d'impact sur cohorte ont à notre connaissance été réalisées sur des programmes fournissant des incitations financières aux prestataires de soins dans les pays à revenu faible à intermédiaire, bien que quelques études présentent des résultats non expérimentaux prometteurs. Les encadrés ci-dessous présentent les résumés de quelques évaluations d'impact et examens des données.

Malgré le manque de données sur certains types d'interventions de FBR, de nombreux pays ont lancé ou développé ces interventions au cours des cinq dernières années. Le moment semble donc opportun pour investir dans des évaluations d'impact rigoureuses et bien conçues, afin de déterminer dans quelle mesure et dans quelles circonstances les politiques de FBR relatives à la santé sont efficaces et pratiquement réalisables. Un programme bien coordonné d'évaluation d'impact des FBR permettra aux pays et aux donateurs d'utiliser les données produites pour prendre des décisions stratégiques en connaissance de cause.

² Il existe toutefois un volume croissant de publications sur la rémunération au rendement (P4P) des soins médicaux aux États-Unis et au Royaume-Uni. Voir par exemple Fleetcroft et al (2012), Jha et al (2012), Lindenauer (2007), Doran et al (2006), et Perterson et al (2006).

Rémunération au rendement et amélioration de la prestation des interventions de santé dans les pays à revenu faible ou intermédiaire – Étude de Cochrane

« **Contexte** La rémunération au rendement suscite un intérêt croissant en tant que méthode permettant d'aligner les incitations proposées aux agents et prestataires de santé sur les objectifs de santé publique. Toutefois, peu de données rigoureuses existent à l'heure actuelle sur l'efficacité de ces stratégies pour améliorer les soins et la santé, en particulier dans les pays à revenu faible ou intermédiaire. De plus, la rémunération au rendement est une intervention complexe présentant des avantages et des inconvénients potentiels. Un examen des données sur l'efficacité arrive donc à point nommé, d'autant plus qu'il s'agit d'un domaine suscitant un intérêt croissant de la part des bailleurs et des gouvernements.

Objectifs Évaluer les données actuelles concernant les effets de la rémunération au rendement sur la prestation de soins de santé et les résultats sanitaires dans les pays à revenu faible ou intermédiaire.

[...]

Conclusions des auteurs La base de connaissances actuelle est trop limitée pour tirer des conclusions générales ; des études plus robustes et détaillées sont donc nécessaires. Le financement basé sur la performance ne correspond pas à une intervention uniforme, mais plutôt à un éventail d'approches. Ses effets dépendent de l'interaction de plusieurs variables, dont la conception de l'intervention (qui reçoit les paiements, l'importance des incitations, les cibles et les mécanismes de mesure), la quantité de financement supplémentaire, et d'autres éléments connexes, tels que l'assistance technique, et des facteurs contextuels, dont le paysage institutionnel dans lequel il est mis en œuvre. »

Witter et al. (2012)

Financement basé sur la performance – Le cas du Rwanda

« **Contexte** Il est urgent de disposer de données sur les meilleures méthodes pour accélérer les progrès vers la réalisation des objectifs du Millénaire pour le développement. Nous avons donc évalué les effets de la rémunération au rendement des prestataires de soins de santé sur l'utilisation et la qualité des services de santé infantile et maternelle dans les établissements de santé du Rwanda.

Méthodes 166 établissements ont été choisis au hasard au niveau du district, soit pour adopter la rémunération au rendement entre juin et octobre 2006 (groupe d'intervention ; n = 80), soit pour continuer d'utiliser le financement traditionnel fondé sur les moyens pour une période pouvant aller jusqu'à 23 mois après l'étude de référence (groupe témoin ; n = 86). La randomisation s'est faite à pile ou face. Nous avons réalisé une enquête sur les établissements et 2 158 ménages, dans le cadre de l'étude de référence puis 23 mois plus tard. Les principaux résultats mesurés ont été les visites de soins prénatals et les accouchements en milieu médicalisé, la qualité des soins prénatals, et les visites de soins préventifs et les vaccinations pour les enfants. Nous avons isolé l'effet incitatif de l'effet de la ressource en augmentant les budgets alloués aux établissements de référence de la moyenne de la rémunération au rendement versée aux établissements de traitement. Nous avons estimé une spécification de régression multivariée du modèle de double différence selon laquelle la régression du résultat d'une personne est effectuée par rapport à une variable fictive, pour indiquer si l'établissement a reçu une rémunération au rendement cette année-là, un effet d'établissement fixe, un indicateur annuel et une série de caractéristiques individuelles et ménagères. **Résultats** Notre modèle a estimé que les établissements du groupe d'intervention avaient observé une augmentation de 23 % du nombre d'accouchements en milieu médicalisé et une augmentation du nombre de visites de soins préventifs chez les enfants âgés de moins de 24 mois (56 %) et de 24 à 59 mois (132 %). Aucune amélioration n'a été observée concernant le nombre de femmes passant quatre visites de soins prénatals ou d'enfants recevant une vaccination complète. Nous estimons également une augmentation de 0.157 écart type (95 % CI 0.026 – 0.289) de la qualité prénatale, mesurée par rapport au respect des directives rwandaises relatives à la pratique clinique de soins prénatals.

Interprétation Le régime rwandais de rémunération au rendement a eu l'effet le plus notable sur les services qui avaient le plus fort taux de rémunération et exigeaient le moins d'effort de la part du prestataire de service. Les mesures d'incitation

financière fondées sur la rémunération au rendement peuvent améliorer l'utilisation et la qualité des services de santé maternelle et infantile, et pourraient être utiles pour accélérer les progrès vers les objectifs du Millénaire pour le développement liés à la santé maternelle et infantile. »

Basinga et al. (2011)

Financement basé sur la performance – Le cas du Rwanda (II)

« La présente étude examine l'impact des mesures d'incitation destinées aux prestataires de soins de santé du Rwanda sur les résultats en matière de santé infantile à l'aide d'un protocole quasi expérimental intégré au programme. Nous avons observé que le régime de rémunération au rendement avait un effet important sur le rapport poids/âge des enfants de 0 à 11 mois et sur le rapport hauteur/âge des enfants de 24 à 49 mois (0,53 et 0,25 écart type respectivement). Nous attribuons cette amélioration à l'augmentation de la quantité des soins prodigués aux enfants bien portants ainsi qu'à l'amélioration de la qualité des soins prénatals. Conformément aux théories économiques, les effets sont plus prononcés sur les aspects des services que les prestataires peuvent maîtriser et pour lesquels les avantages monétaires sont les plus élevés. Nous pensons que les modifications intervenues au niveau des efforts des prestataires sont essentiellement responsables des impacts observés. Nous observons une diminution de 20 % du déficit d'efficacité entre les connaissances et la pratique pour les soins prénatals. Enfin, nous avons constaté une forte complémentarité entre le programme de rémunération au rendement et la présence d'agents de santé hautement qualifiés dans les centres de santé. »

Gertler et Vermeersch (2012)

Sous-traitance des services santé – Le cas du Cambodge

« En 1999, le Cambodge a sous-traité la gestion des services publics de santé à des ONG dans cinq districts choisis de manière aléatoire. Les contrats spécifiaient des objectifs spécifiques concernant l'amélioration des services de santé maternelle et infantile. Les résultats ciblés ont progressé d'environ 0,5 écart type par rapport aux districts de référence. Les changements ont été faibles pour les résultats non ciblés. Le programme a augmenté les services disponibles 24 heures sur 24, réduit l'absence des prestataires de services, et augmenté les visites de supervision. Il est permis de penser qu'il a amélioré la santé. Ce programme reposait sur l'augmentation du financement de la santé publique mais a approximativement compensé la réduction des dépenses privées à mesure que les résidents des districts traités ont délaissé les vendeurs de médicaments non agréés et les guérisseurs traditionnels au profit des cliniques publiques. »

Bloom et al. (2006)

Inciter les villages à améliorer la santé et l'éducation – Le cas de l'Indonésie

« Le présent document rend compte d'une expérience menée dans plus de 3 000 villages indonésiens pour évaluer le rôle des mesures d'incitation dans l'amélioration de l'efficacité des programmes d'aide. Les villages situés dans un échantillon choisi au hasard d'un tiers des sous-districts ont reçu une subvention forfaitaire pour améliorer 12 indicateurs de santé maternelle et infantile et d'éducation, le montant de la subvention de l'année suivante dépendant de l'efficacité par rapport aux autres villages du sous-district. Les villages des autres sous-districts ont été affectés au hasard à un programme de subvention forfaitaire identique, mais sans lien avec les performances, ou à un groupe témoin. Nous avons observé que les villages bénéficiant de mesures d'incitation ont eu de meilleurs résultats en matière de santé que les autres villages, en particulier dans les provinces peu développées, mais que les mesures d'incitation semblent être sans effet dans le domaine de l'éducation. Nous n'avons pas observé de retombées négatives des mesures d'incitation sur les résultats non ciblés. Les mesures d'incitation semblent avoir entraîné une utilisation plus efficace des subventions forfaitaires et une augmentation de la main-d'œuvre des prestataires de santé, qui sont en partie payés à l'acte, mais qui ne sont pas des enseignants. On estime qu'entre 50 et 75 % de l'impact total du programme de subvention forfaitaire sur les indicateurs de santé peuvent être attribués aux mesures d'incitation. »

Olken et al. (2011)

Incitations liées à la performance du prestataire – Le cas des Philippines

« L'utilisation de mesures d'incitation financières pour améliorer la qualité des soins cliniques présente un attrait certain, mais peu d'études en ont rigoureusement évalué les avantages potentiels. L'incertitude entourant les évaluations de la qualité peut conduire à de piètres décisions de politique et, éventuellement, entraîner l'augmentation des coûts avec peu ou pas d'amélioration de la qualité, ou des occasions manquées en matière d'amélioration de la santé. Nous avons réalisé une expérience avec des médecins dans 30 hôpitaux des Philippines, qui a surmonté une grande partie des problèmes des études antérieures. Nous avons mesuré les résultats cliniques et avons ensuite examiné si de modestes primes correspondant à environ 5 % du salaire d'un médecin, ainsi que des mesures d'incitation systémiques augmentant la rémunération des hôpitaux et des groupes de médecins, entraînaient l'amélioration de la qualité des soins. Nous avons observé que les primes et les mesures d'incitation systémiques augmentaient de 10 % les scores du système de mesure de la qualité utilisé dans notre étude. Ces résultats suggèrent que lorsque l'on associe une mesure rigoureuse aux types d'incitations que nous avons étudiées, l'impact sur la qualité peut être supérieur à ce que nous avons préalablement observé. »

Peabody et al. (2011)

Examen systématique de l'impact des bons sur l'utilisation et la qualité des biens et services de santé dans les pays en développement

*« **Historique** : L'une des approches possibles pour apporter une assistance sanitaire aux pays en développement consiste à utiliser des programmes de bons de santé, dans lesquels des bons sont distribués à la population cible pour qu'elle puisse obtenir des biens/services de santé gratuits ou subventionnés. Théoriquement, les bons doivent bénéficier à des populations cibles spécifiques, améliorer la qualité et l'efficacité et, à terme, la santé des populations. Objectifs : Le principal objectif de cet examen systématique est d'évaluer si les programmes de bons ont à ce jour permis d'atteindre les résultats souhaités. Méthodes : Des recherches dans les bases de données bibliographiques, les revues spécialisées et les sites de plusieurs organisations ont été réalisées en septembre-octobre 2010 en utilisant des critères d'inclusion/exclusion explicites. Les autres stratégies de recherche utilisées comprenaient un référencement antérieur bibliographique, des recherches avec des mots-clés supplémentaires utilisant des informations spécifiques aux programmes, et la consultation d'experts du domaine. Une synthèse a été rédigée pour résumer de manière qualitative les variables de résultats quantitatifs identifiées selon cinq*

catégories (ciblage, utilisation, efficacité, qualité et incidence sur la santé). En utilisant la direction de l'effet des variables de résultats et la confiance dans les conclusions de l'étude, les résultats de chaque catégorie de résultats ont été agrégés et affectés à l'une des cinq catégories de conclusions préalablement définies : 1) preuves insuffisantes ; 2) preuves de l'absence d'effet ; 3) preuves conflictuelles ; 4) preuves raisonnables ; ou 5) preuves solides. Des analyses des sous-groupes et de la sensibilité ont également été effectuées. Une méta-analyse quantitative n'a pas été réalisée du fait de l'hétérogénéité des variables de résultats étudiées.

Résultats : Au total, 24 études évaluant 16 programmes de bons sanitaires ont été identifiées pour cet examen. Les données recueillies sur les 64 variables de résultats ont fourni les cinq principales conclusions suivantes : 1) il y a des preuves raisonnables que les programmes de bons ciblent efficacement les populations devant bénéficier de bons pour des biens/services de santé (sur la base de quatre programmes) ; 2) les preuves sont insuffisantes pour déterminer si les programmes de bons permettent de fournir des biens/services de santé plus efficacement que les autres stratégies de financement de la santé (sur la base d'un programme) ; 3) il y a des preuves solides que les programmes de bons augmentent l'utilisation des biens/services de santé (sur la base de 13 programmes) ; 4) il y a des preuves raisonnables que les programmes de bons améliorent la qualité des services de santé (sur la base de trois programmes) ; et 5) les données indiquent que les programmes de bons n'ont pas d'incidence sur la santé des populations (sur la base de six programmes) ; cette dernière conclusion n'a toutefois pas été jugée stable dans une analyse de sensibilité.

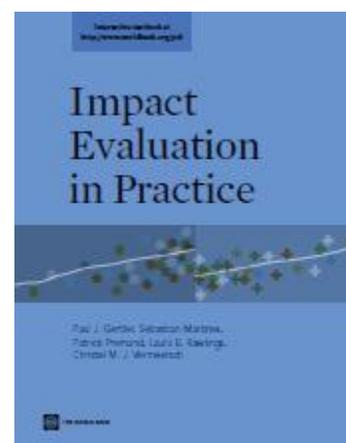
Conclusions : Les données montrent que les programmes de bons sanitaires ont réussi à augmenter l'utilisation des biens/services de santé, à cibler des populations spécifiques et à améliorer la qualité des services. Bien que ces résultats soient encourageants, les données analysées dans le cadre de cet examen ne montrent pas clairement que les programmes de bons améliorent la santé des populations concernées. La méthodologie utilisée dans cette analyse permet aux décideurs de synthétiser les conclusions d'études hétérogènes et donc d'inclure un plus grand volume de données qu'il ne serait possible dans une méta-analyse standard. Toutefois, l'utilisation des bons est relativement récente et le nombre d'études d'évaluation des bons publiées reste limité. À l'avenir, les études utilisant cette méthodologie pourront comparer les programmes de bons de santé aux autres techniques de financement et intégrer de nouvelles données sur ces programmes dans les évaluations en cours ; les outils de synthèse utilisés devront toutefois être validés. »

Meyer et al. (2011)

Qu'est-ce que l'évaluation d'impact³ ?

Les évaluations d'impact s'inscrivent dans la vaste tendance de l'élaboration de politiques fondée sur des données factuelles. Dans un contexte où les décideurs, les donateurs et la société civile demandent des résultats et exigent que les responsables de programmes rendent compte de la performance de leurs interventions, l'évaluation d'impact fournit des données solides et fiables qui indiquent si un programme donné a effectivement atteint les objectifs espérés. Globalement, les évaluations d'impact permettent de développer les connaissances sur l'efficacité des programmes.

L'évaluation d'impact est l'une des méthodes contribuant à l'élaboration de politiques fondée sur des données factuelles. Les autres méthodes comprennent



³ Cette section s'inspire directement du chapitre 1 de l'ouvrage intitulé *L'Évaluation d'impact en pratique* (Gertler et al. 2011). Veuillez vous reporter à cet ouvrage pour un examen plus détaillé de la question. Ce guide peut être téléchargé gratuitement à www.worldbank.org/ieinpractice.

le suivi, les évaluations des processus, les évaluations qualitatives et le calcul des coûts. Les évaluations d'impact sont différentes en ce qu'elles cherchent à identifier les modifications du bien-être qui peuvent être attribuées ou dues à une politique ou un programme particuliers. Contrairement au suivi et à l'évaluation, l'évaluation d'impact est généralement structurée autour d'une question du type : *Quel est l'impact (ou l'effet causal) d'un programme sur un résultat recherché ?* À la différence des comparaisons avant-après et des simples enquêtes sur la satisfaction des utilisateurs finals, l'évaluation d'impact entend isoler l'impact des programmes d'autres facteurs de confusion.

Pourquoi évaluer les programmes de FBR ?

Les évaluations d'impact sont particulièrement utiles pour les pays qui testent de nouvelles interventions innovantes, qui semblent prometteuses en théorie mais sur lesquelles nous manquons de données factuelles. Les décideurs qui veulent utiliser des données factuelles pour cautionner leurs politiques ont besoin d'informations sur plusieurs questions, par exemple : *Ce programme est-il efficace par rapport à la situation actuelle ? Parmi les nombreuses possibilités de mise en œuvre d'un programme de FBR, quelle est la plus efficace ?*

L'évaluation d'impact d'un programme de FBR fournit des données factuelles permettant de savoir si l'intervention étudiée a produit les résultats escomptés dans le contexte spécifique du pays. Les données des évaluations d'impact qui examinent divers mécanismes de FBR dans divers pays peuvent informer les gouvernements et les pays sur le moyen le plus efficace pour concevoir et utiliser ces mécanismes pour améliorer le fonctionnement du système de santé et les résultats de santé dans différents contextes. En outre, les évaluations d'impact peuvent déterminer si le FBR a des effets imprévus, par exemple encourager les prestataires à négliger les services qui ne sont pas inclus dans le mécanisme de FBR. Enfin, les évaluations d'impact peuvent aider à documenter les coûts associés à l'administration des systèmes de paiements fondés sur les résultats.

2012/03/15

Étude de cas nationale : Motivation de l'évaluation d'impact Nigéria : Diffuser les résultats pour mobiliser des financements⁴

Dr Pate : Dans mes fonctions antérieures, en tant que directeur exécutif de l'Agence nationale de développement des soins de santé primaires (NPHCDA), un organisme parapublic responsable de la prestation de soins primaires dans les 36 états du Nigéria, il était évident que davantage de ressources et d'innovation seraient nécessaires pour se concentrer sur les cibles des OMD liés à la santé dans un pays où la population est proche de 150 millions d'habitants et les indicateurs de santé maternelle et infantile (SMI) sont parmi les plus mauvais. Le NPHCDA étant mandaté pour offrir une couverture universelle pour tous les soins essentiels, alors que les cibles des OMD liés à la santé dépendent fortement de l'efficacité du système de prestation de soins primaires, nous savions que des innovations, qui augmenteraient l'efficacité des soins et de l'utilisation des ressources, seraient nécessaires pour améliorer les résultats en matière de santé maternelle et infantile. Toutefois, que nous recherchions une assistance pour augmenter les crédits budgétaires nationaux pour le secteur primaire ou des financements auprès des organismes de développement, nous devons répondre aux mêmes questions : *Pouvions-nous montrer des résultats ? Pouvions-nous montrer des impacts ? Pouvions-nous prouver*

⁴ Pour suivre l'intégralité de l'entretien avec le Dr Pate, voir M1_Nigeria_Motivation for IE.

que nous utilisons bien les crédits, qu'ils soient nationaux ou internationaux ? Nous avons vite compris que nous avions besoin de résultats crédibles pour bénéficier des crédits budgétaires publics et de l'assistance officielle au développement, qui comprenait des prêts et des dons. Le besoin de disposer de preuves et de résultats solides a augmenté pour les gouvernements comme pour les organismes donateurs du fait des pressions et difficultés budgétaires liées à la crise économique.

Formulation des questions d'évaluation

La première étape d'une évaluation consiste à définir le type de questions auxquelles l'évaluation répond, en élaborant une théorie du changement montrant comment le projet est censé atteindre les résultats espérés, en élaborant une chaîne de résultats, en formulant des hypothèses qui seront testées par l'évaluation et la sélection d'indicateurs de performance (Gertler et al. 2011).

Une théorie du changement est une description de la manière dont une intervention est censée produire les résultats espérés. Elle décrit la logique causale expliquant comment et pourquoi un projet, un programme ou une politique atteindra les résultats visés. L'existence d'une théorie du changement est fondamentale pour les évaluations d'impact étant donné l'importance qu'elles portent aux relations de cause à effet. La théorie du changement est l'une des premières étapes de la conception d'une évaluation, car elle contribue à la formulation des questions de recherche. Les théories du changement décrivent une série d'événements conduisant à un résultat. Elles énoncent les conditions et les hypothèses nécessaires pour que des changements se produisent ; elles mettent en évidence la logique causale sous-jacente au programme et inscrivent les interventions dans cette logique causale.

Gertler et al. 2011

Théorie du changement pour le financement de la santé basé sur les résultats

Les interventions de FBR s'inscrivent dans le cadre du système de santé d'un pays et varient donc selon les circonstances de chaque pays. Les évaluations d'impact doivent donc être adaptées à l'intervention et démarrer « dans la boîte noire » du programme de FBR de chaque pays. Les équipes nationales de conception de projets et d'évaluation d'impact devraient collaborer pour identifier les éléments des interventions de FBR, les questions de politique auxquelles une évaluation d'impact peut répondre, ainsi que les priorités du pays parmi ces questions, et comment l'EI peut contribuer à combler le déficit de connaissances internationales sur le FBR.

La présente Boîte à outils décrit des « théories du changement » possibles pour le FBR. Le glossaire de Musgrove définit différents mécanismes : le financement basé sur la performance (FBP), les contrats de performance, les transferts monétaires conditionnels, les transferts en nature et les bons. Nous n'examinerons ici que les théories du changement relatives à ces interventions de FBR plutôt que les assistances fondées sur les résultats ou le paiement à la livraison.

La théorie du changement sera déterminée par les principaux éléments de conception suivants :

- L'intervention est-elle axée sur l'offre ou sur la demande ?
- Les paiements sont-ils versés aux prestataires ou aux ménages ?
- Les paiements sont-ils versés à des personnes ou des groupes de personnes ?
- La performance est-elle mesurée au niveau du groupe ou des personnes ?
- Les paiements aux prestataires sont-ils liés à la quantité des services fournis ?
- Les paiements aux ménages sont-ils liés à l'utilisation des services de santé ?
- Les paiements aux prestataires sont-ils liés à la qualité des services fournis ?
- Le mécanisme de FBR a-t-il pour effet d'instaurer ou de renforcer le suivi et la surveillance des prestataires ou les retours d'expérience aux prestataires ?
- Le mécanisme de FBR augmente-t-il l'autonomie décisionnelle au niveau du prestataire ou à tout autre niveau ? Le FBR augmente-t-il la quantité totale de ressources disponibles aux prestataires de services (pour les interventions axées sur l'offre) ? Augmente-t-il la quantité totale de ressources disponibles à d'autres niveaux ?
- Comment les performances sont-elles mesurées et vérifiées ?
- Qui achète les services ?
- Quelle est l'importance du financement basé sur la performance fourni aux prestataires ou aux ménages ?
- Comment les prestataires et/ou ménages bénéficiaires sont-ils sélectionnés ? Quels sont les critères ?
- Des interventions parallèles sont-elles menées en même temps que le financement basé sur la performance, telles que la formation des prestataires ou l'information des communautés ?

La théorie du changement dépendra des principaux éléments de conception de chaque programme. Nous proposons ci-après une description des différents aspects d'une théorie du changement pour le programme de FBP au Rwanda, qui est un programme de FBR axé sur l'offre qui verse aux centres de santé des primes en fonction de la quantité et de la qualité des soins. Nous décrivons dans un premier temps comment les prestataires peuvent réagir à une formule de paiement contenant divers indicateurs quantitatifs et un indicateur qualitatif. Nous examinons ensuite un modèle de mesure du déficit d'efficacité entre les connaissances et la pratique des soins. Enfin, nous fournissons une représentation graphique d'une théorie du changement pour établir un lien entre les versements aux prestataires et la quantité. Des diagrammes semblables pourraient être produits pour décrire la théorie du changement d'autres éléments de conception du programme⁵. La théorie du changement pourrait être illustrée plus simplement en utilisant une simple chaîne de résultats établissant un lien entre les intrants et les activités, d'une part, et les extrants et les résultats intermédiaires et finaux, d'autre part⁶.

⁵ Pour obtenir une version modifiable du diagramme, voir l'outil **1.01 Graph for Theory of Change**

⁶ Pour toute information complémentaire, voir l'outil **3.01a RBF Indicators**. Pour obtenir un modèle Powerpoint© d'une chaîne de résultats, voir l'outil **1.02 Results Chain Template**.

Étude de cas nationale : Théorie du changement
Programme de FBP au Rwanda (rémunération au rendement)

D'après Gertler et Vermeersch (2012)

Systeme de paiement :

Le système de paiement [FBP au Rwanda] couvre 14 services de santé maternelle et infantile sur la base d'un score général d'évaluation de la qualité des établissements. La formule utilisée pour calculer le paiement à l'établissement i pour le mois t est la suivante :

$$Payment_{it} = \left(\sum_j P_j U_{jit} \right) \times Q_{it} \quad \text{with } 0 \leq Q_{it} \leq 1,$$

où P_j est le paiement par unité de service j (par exemple, accouchement en milieu médicalisé ou visites de soins préventifs pour les enfants), U_{jit} est le nombre de patients utilisant le service j dans l'établissement i pendant la période t , et Q_{it} est l'indice de qualité général de l'établissement i pendant la période t .

[...]

Modèle comportemental :

[...] nous utilisons un simple modèle comportemental pour estimer l'effet potentiel de l'introduction de la rémunération au rendement sur le comportement des prestataires de soins. Notre modèle envisage une clinique rurale dotée de quatre à six médecins sans remplaçants à proximité. Nous supposons, pour simplifier, qu'un établissement fonctionne comme un unique organe de décision que nous désignons prestataire. La fonction-objectif du prestataire joue un rôle essentiel à cet égard. Nous supposons que la santé des patients est importante pour les prestataires de soins au même titre que les revenus qu'ils tirent des services fournis à leurs patients. Nous tenons compte de cet aspect éthique des préférences en supposant que les prestataires traitent tous les patients qui se présentent à eux et leur fournissent des soins correspondant au moins au niveau minimal défini dans les normes éthiques.

Nous commençons par examiner le cas d'un établissement recevant un montant fixe pour les frais de personnel et ayant un budget fixe pour les autres frais, en supposant que ces derniers ne puissent être réaffectés aux frais de personnel. Les revenus du prestataire ne dépendent donc pas du nombre de patients reçus ni de la qualité des soins fournis. Le prestataire traite ainsi tous les patients qui se présentent à lui en leur fournissant un niveau minimal de soins.

Le programme de rémunération au rendement introduit une nouvelle dimension dans le problème d'optimisation du prestataire, en établissant un lien entre une partie des revenus de l'établissement, d'une part, et la prestation de certains services et la qualité des soins, d'autre part. Pour simplifier, nous supposons que l'activité du prestataire est axée sur deux types de services (soins prénatals et accouchements) et la qualité des soins. Compte tenu de la structure de base de la formule de rémunération au rendement, nous pouvons écrire la nouvelle fonction de profit suivante :

$$V = I + [P_1 U_1(\varepsilon_1) + P_2 U_2(\varepsilon_2)] Q(\varepsilon_q) - C(\varepsilon) \quad (3)$$

où I est le salaire fixe, P_i est la rémunération au rendement du service i , U_i est la quantité totale des services i fournis aux patients, Q est la qualité générale des soins, et $C(*)$ est le coût des efforts. Rappelons que les U_i figurent dans le tableau 1 et que Q est un indice déterminé à partir des éléments du tableau 2.

Le prestataire choisit les niveaux d'efforts ε_1 et ε_2 pour augmenter la quantité des services fournis au-delà des niveaux minimums pour traiter les patients et l'effort ε_q pour améliorer la qualité des soins au-delà des normes éthiques minimales⁷. Les fonctions de production des services $U_i(.)$ et la fonction de production de la qualité $Q(.)$ augmentent au niveau des efforts mais à un taux décroissant. Enfin, la fonction de coût des efforts $C(.)$ est une fonction des efforts totaux ($\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_q$) et est convexe.

Le prestataire détermine ensuite les niveaux d'efforts nécessaires pour maximiser les revenus, avec des niveaux d'efforts légèrement positifs. Dans le cas d'une solution intérieure, l'effort est alloué de telle sorte que les revenus marginaux des efforts soient identiques pour les trois types d'efforts et soient égaux aux coûts marginaux des efforts :

⁷ Ceci nous permet de normaliser à zéro les niveaux d'effort minimum.

$$P_1 U_1'(e_1) = P_2 U_2'(e_2) = P_1 U_1'(e_1) + P_2 U_2'(e_2) = Q'(e) = C'(e) \quad (4)$$

On notera que le rendement marginal de l'effort consacré à chaque service dépend de son propre prix mais également du prix de l'autre service, de même que le rendement marginal de l'effort consacré à la qualité dépend de ces deux prix. Il s'ensuit qu'une augmentation de l'un des deux prix se traduit toujours par une hausse du rendement des efforts fournis en matière de qualité. Quelle qu'en soit la cible, les efforts fournis augmentent le coût marginal des efforts car le coût des efforts est une fonction de l'effort total.

La quantité relative des efforts consacrés aux deux types de services satisfait à la condition suivante :

$$\frac{U_2'(e_2)}{U_1'(e_1)} = \frac{P_1}{P_2} \quad (5)$$

En d'autres termes, le rapport entre les rendements marginaux et les efforts pour fournir les services devrait être égal au rapport entre les taux de rémunération de ces services. Il s'ensuit que davantage d'efforts seront consacrés au service ayant le prix le plus élevé et la productivité marginale la plus haute en matière d'efforts.

Prévisions économiques :

Nous pouvons examiner les effets potentiels de l'introduction de la rémunération au rendement en termes d'analyse comparative statique de l'augmentation des prix, de sorte que le niveau d'origine de P et ε soient proches de zéro. Supposons une augmentation de P_1 , qui correspond à la rémunération du service 1. Ceci aura pour effet d'augmenter le revenu marginal de l'effort associé à la prestation du service 1 et à la fourniture d'un service de qualité, et incitera donc à consacrer davantage d'efforts au service et à la qualité. L'augmentation de l'effort ayant pour effet d'accroître le coût marginal de l'effort total, le prestataire réduira l'effort consacré au service 2. Il s'ensuit que l'augmentation de l'effort consacré au service 1 et à la qualité a pour contrepartie la réduction de l'effort consacré à l'autre service et la diminution du temps de loisirs. Alors que l'effort total augmente, sa répartition relative se fait au profit du service 1 et de la qualité et aux dépens du service 2. Si l'augmentation de prix est suffisamment importante, l'effort optimal consacré au service 2 descendra en dessous du minimum éthique, qui deviendra donc contraignant.

L'analyse statique comparative d'une seule modification de prix n'est toutefois pas exactement applicable à l'instauration d'un système de rémunération au rendement qui a pour effet de modifier tous les prix simultanément. Avant que le prix n'augmente, tous les niveaux d'efforts correspondent au minimum éthique. L'augmentation des prix des services entraînera un effort accru consacré à la qualité car l'augmentation de tout prix augmente le rendement marginal de l'effort consacré à améliorer la qualité. L'ampleur de l'effort consacré à un service sera la plus forte dans le cas des services pour lesquels la hausse relative du prix est la plus importante et la productivité marginale de l'effort est la plus élevée. De même, l'ampleur de l'effort sera la plus faible dans le cas des services où la hausse relative du prix et le rendement marginal de l'effort seront les plus bas. De fait, si, pour un service donné, la hausse relative du prix et la productivité marginale de l'effort sont suffisamment faibles, le prestataire ne consacra pas plus d'efforts à ce service malgré l'augmentation absolue du prix. Dans ce cas, l'effort fourni restera au niveau éthique minimum.

L'effet de l'instauration de la rémunération au rendement ne dépend donc pas uniquement des taux de rémunération relatifs mais également de la difficulté rencontrée pour augmenter le niveau des services. En règle générale, nous estimons que l'augmentation des services qui dépendent des choix des patients nécessite plus de travail que l'augmentation des services que le prestataire peut entièrement contrôler. Par exemple, il faut plus de temps pour convaincre une femme enceinte de se rendre dans une clinique pour recevoir des soins prénatals que pour lui injecter un vaccin antitétanique à la clinique. Par conséquent, même si les paiements étaient les mêmes pour la consultation d'un patient supplémentaire et l'injection d'un vaccin antitétanique, on pourrait s'attendre à une plus forte augmentation du nombre d'injections de vaccin antitétanique (qui relèvent essentiellement du prestataire) que du nombre de consultations dans l'établissement (qui dépendent largement de la volonté des patients). De surcroît, nous pensons que le démarrage des soins nécessite davantage d'efforts que leur poursuite. Par exemple, un prestataire devra consacrer d'importants efforts pour trouver des femmes enceintes dans la communauté, surtout pendant le premier trimestre de la grossesse, et leur prodiguer des soins prénatals. En revanche, il est relativement plus facile de tirer parti d'une visite de soins prénatals pour convaincre les femmes qui reçoivent déjà ces soins d'accoucher dans l'établissement.

L'analyse ci-dessus suppose que les prix des services contribuent de manière simple et linéaire à la fonction de profit, représentée par l'équation 2. Dans la réalité, le système de paiement est plus compliqué et les services répertoriés au tableau 1 sont composés de raisons primaires pour se rendre dans une clinique et de services fournis en cas de visite. Bien qu'ils soient tous des U_i , les services fournis pendant les visites contribuent également à l'indice de qualité Q . De plus,

le paiement P pour voir un patient dépend des services fournis pendant cette visite. Dans le cas des soins prénatals, les prestataires reçoivent 0,18 \$ pour chaque femme enceinte qui commence à recevoir des soins prénatals, puis 0,37 \$ au bout de quatre visites, 0,92 \$ pour un vaccin antitétanique et une prophylaxie antipaludique pendant une visite de soins prénatals, et 1,83 \$ s'ils estiment que l'accouchement présente des risques et demandent à la femme d'accoucher dans l'hôpital du district. Les paiements des soins prénatals dépendent donc du nombre de femmes enceintes recevant des soins et du nombre de visites, mais également du contenu des soins fournis pendant ces visites.

De fait, les taux de rémunération des visites sont beaucoup plus élevés si le contenu des soins fournis par le prestataire est de meilleure qualité. Comme on l'a vu plus haut, un prestataire recevra 0,55 \$ pour quatre visites de soins prénatals de faible qualité contre 1,47 \$ pour des soins de haute qualité. Si le prestataire détecte une grossesse à risque élevé et demande à la femme d'accoucher à l'hôpital, la rémunération de ces soins de haute qualité passe alors à 3,30 \$. Dans le cadre de la surveillance de la croissance, la rémunération du prestataire est de 0,18 \$ par visite plus 1,83 \$ si l'enfant est sous-alimenté et envoyé à l'hôpital pour traitement. Étant donné que 45 % des enfants rwandais de moins de cinq ans souffrent d'une malnutrition chronique modérée et que 19 % de ces enfants souffrent de malnutrition chronique grave⁸ (Institut National de la Statistique du Rwanda et ORC Macro 2006), la rémunération prévue pour une visite de surveillance de la croissance de haute qualité est relativement élevée. Globalement, le régime incitatif est axé sur le traitement d'un plus grand nombre de patients mais aussi sur la prestation de soins de plus haute qualité à un plus grand nombre de patients ; ces aspects sont représentés par le facteur d'échelle multiplicatif Q et par la rémunération directe du contenu des soins de santé représentée par les U_i .

Prévisions empiriques :

Cette discussion fournit un certain nombre de prévisions empiriques. Premièrement, l'augmentation de la rémunération sera plus efficace pour les services où la hausse relative des prix et le rendement marginal relatif de l'effort sont les plus élevés. Deuxièmement, l'augmentation de la rémunération n'entraînera pas nécessairement l'accroissement de tous les services. Elle pourra être sans effet sur les services où les taux de rémunération et le rendement marginal de l'effort sont faibles. Troisièmement, le taux de rémunération d'un service, qui détermine l'allocation des efforts, dépend non seulement du nombre de patients traités mais aussi du contenu des soins fournis lors d'une visite. Enfin, nous pensons que l'instauration de la rémunération au rendement augmentera la qualité Q , qui est le facteur multiplicatif dans la formule de rémunération.

⁸ La malnutrition chronique modérée (grave) correspond à une taille pour l'âge inférieure à -2 (-3 respectivement) écarts types de la médiane de la population de référence. (Institut National de la Statistique du Rwanda et ORC Macro 2006)

Étude de cas nationale : Déficit d'efficacité
Programme de FBP au Rwanda (rémunération au rendement)

D'après Gertler et Vermeersch (2012)

Une autre interprétation de la rémunération au rendement repose sur l'idée que les services fournis par les prestataires ne tirent pas pleinement parti de leurs capacités (connaissances). Cet argument est appuyé par le fait que les services cliniques fournis par les prestataires dans le cadre de soins prénatals sont sensiblement inférieurs à leur connaissance des procédures cliniques appropriées. À ce titre, il est bon de rappeler qu'en moyenne les prestataires connaissent 63 % des procédures appropriées mais n'en appliquent que 45 %. Il existe donc un écart de 18 % entre les connaissances et la pratique. Si l'on considère les connaissances des prestataires comme la limite de leurs possibilités de production, on peut interpréter l'écart entre les connaissances et la pratique comme une mesure de leur inefficacité technique. Les mesures incitatives de la rémunération au rendement ont pour objet de réduire l'inefficacité technique.

Nous présentons le déficit d'efficacité à l'aide d'un diagramme où l'axe des abscisses représente la part des soins prénatals, recommandés dans les Directives sur les bonnes pratiques cliniques et connus du prestataire, et l'axe des ordonnées représente la qualité fournie, c'est-à-dire la part effectivement fournie des soins prénatals recommandés dans ces directives. La ligne de 45° représente la limite des possibilités de production, où les prestataires fournissent des soins cliniques de qualité dans la limite de leurs connaissances. Les prestataires dont la qualité des soins est inférieure à leur niveau de connaissance se situent en dessous de cette ligne. La distance verticale entre la limite et le point de performance est une mesure de l'inefficacité technique.

Nous avons également inclus à la figure A les courbes de performance des prestataires figurant dans notre ensemble de données. Ces courbes sont des régressions bivariées non paramétriques de la qualité par rapport aux connaissances pour des groupes de traitement et de référence en fin de programme. On notera que ces deux courbes sont nettement inférieures à la limite des possibilités de production, ce qui implique des niveaux élevés d'inefficacité technique à tous les niveaux de qualification. De plus, bien que les courbes de performance suivent une pente ascendante, elles sont plus plates que la limite des possibilités de production. Cela implique que bien que les connaissances améliorent la performance, le déficit d'efficacité augmente avec les connaissances. Enfin, la courbe de performance du groupe de traitement se trouve au-dessus de celle du groupe de référence et a également une pente plus forte. Cela signifie que la rémunération au rendement réduit le déficit d'efficacité, et davantage encore chez les prestataires qualifiés.

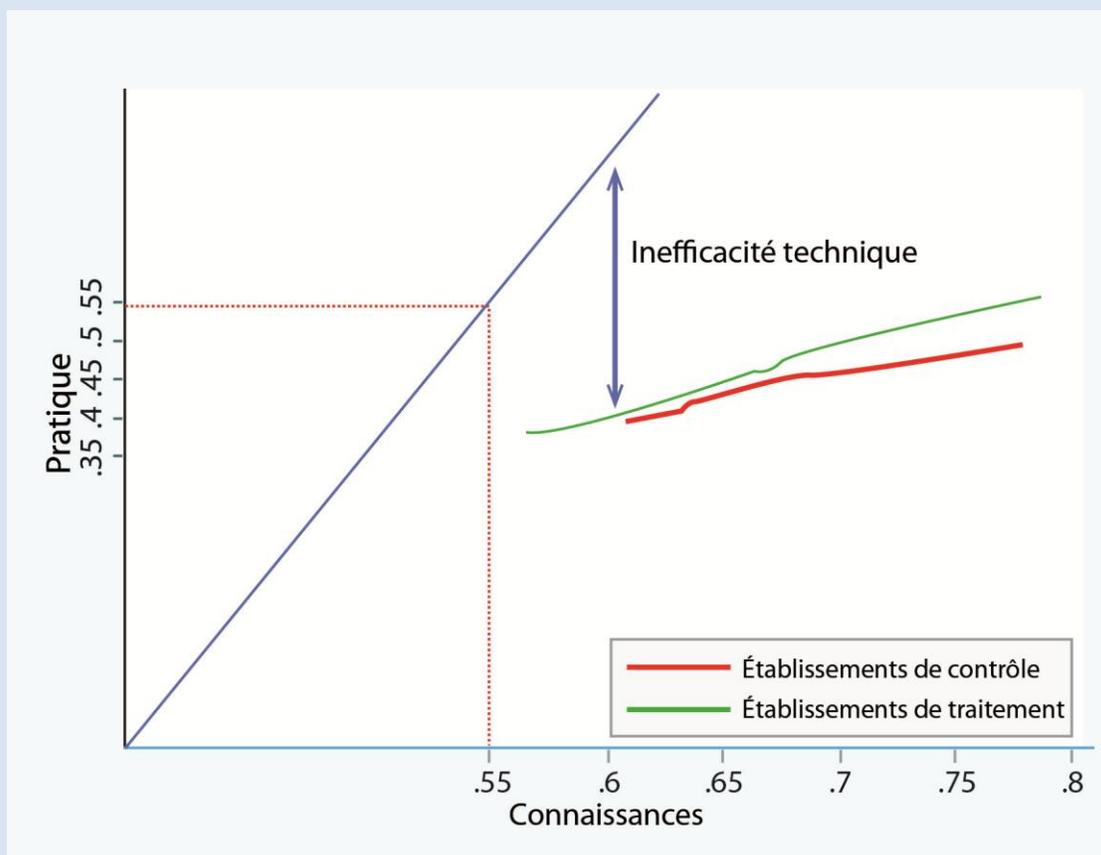
Nous allons maintenant estimer l'ordre de grandeur de l'impact de la rémunération au rendement sur le déficit d'efficacité. Nous mesurons le déficit d'efficacité comme la différence entre les services cliniques connus et les services cliniques fournis, tels que définis dans les Directives sur les bonnes pratiques cliniques. On observe que la rémunération au rendement réduit le déficit d'efficacité de 3,5 %, soit en moyenne environ 20 % de ce déficit (tableau 9, modèle 1). Si nous examinons les connaissances des prestataires, l'effet sur l'efficacité de la rémunération au rendement augmente légèrement pour atteindre 4 % (tableau 9, modèle 2). Dans ce modèle, plus les connaissances sont vastes et plus le déficit d'efficacité est important. En d'autres termes, bien que le développement des connaissances des prestataires améliore la qualité des soins, l'amélioration de la qualité est moins importante que celle des connaissances. Enfin, nous estimons que la rémunération au rendement a un effet beaucoup plus important sur l'efficacité des prestataires les mieux informés. Alors que nous n'observons aucune augmentation de l'efficacité des prestataires en dessous de la médiane des connaissances, l'efficacité augmente de 6 % chez les prestataires situés au-dessus de cette médiane (tableau 9, modèle 3).

Tableau 9 : Impact de la rémunération au rendement sur le déficit d'efficacité (Connaissances – Qualité)

	β	P	β	P	β	P
P4P (=1)	-0,035	0,00	-0,04	0,03	-0,02	0,24
Z (connaissances)			0,16	0,00	0,21	0,00
P4P * Connaissances de la moitié supérieure					-0,06	0,01
Nb d'observations	3 709		3 709		3 709	

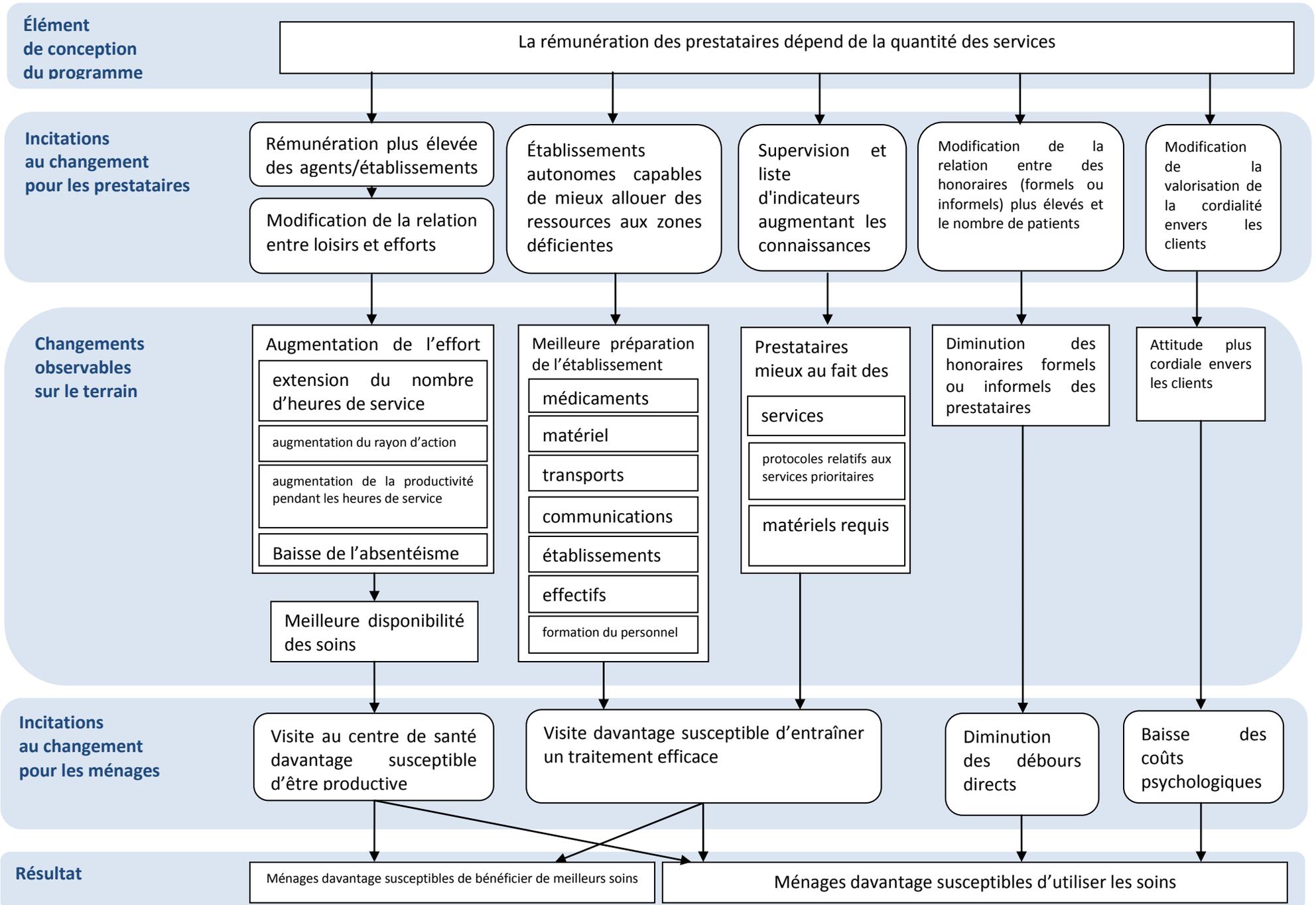
Notes : Les valeurs de P correspondent à des tests unilatéraux de l'hypothèse nulle que $\beta = 0$ et sont calculées avec un « WILD bootstrap » et 999 tirages.

Figure A : Déficit d'efficacité entre les connaissances et la pratique dans le suivi des soins prénatals en 2008



Notes : L'axe des abscisses représente les connaissances exprimées en pourcentage des éléments du protocole correctement identifiés par le prestataire pendant l'administration de la vignette. L'axe des ordonnées représente le pourcentage des éléments du protocole fournis dans le cadre de soins prénatals, tels que déclarés par les patients à la sortie de l'établissement et dans les enquêtes sur les ménages.

Figure 4 : Exemple de théorie du changement concernant la rémunération des prestataires



Étude de cas nationale : Modèle comportemental concernant la fourniture de résultats ciblés et non ciblés
Approche du Cambodge en matière de passation de marchés

Le modèle du Cambodge en matière de sous-traitance des services de santé établit un lien entre les mesures incitatives et huit résultats ciblés. Bloom et al. (2006) définissent un modèle décrivant comment un contrat de ce type peut affecter la prestation des services ciblés et la prestation des services non ciblés.

« Un cadre Holmstrom-Milgrom (1991) suggère que les contrats qui établissent un lien entre les mesures incitatives et les huit résultats ciblés produiront de meilleurs résultats sur ces mesures mais que leur influence sur les autres résultats variera selon que l'effort consacré à ces résultats ciblés complète ou remplace les résultats ciblés. Les deux scénarios sont possibles. Par exemple, il est possible que les mesures incitatives fournies aux sous-traitants les incitent à créer des incitations pour réduire l'absentéisme des agents de santé, et que cela constitue une action complémentaire à la fourniture d'autres types de soins. En revanche, il est possible que les établissements transfèrent des ressources vers les résultats ciblés aux dépens des soins non mesurés.

Nous allons préciser cette idée à l'aide d'un cadre Holmstrom-Milgrom (1991) simplifié. Supposons qu'il existe deux résultats sanitaires. L'agent peut contrôler deux types d'efforts qui nécessitent une lourde dépense. Supposons maintenant qu'un seul des deux résultats peut être sous-traité. Désignons les résultats C et NC et les types d'effort e_1 et e_2 et calculons-les comme suit

$$C = f(e_1, e_2) + \varepsilon$$

$$NC = g(e_1, e_2) + \eta$$

L'agent s'intéresse à la rémunération w ainsi qu'au coût de l'effort,

$$u(w, e_1, e_2) = w - c(e_1, e_2)$$

Les agents reçoivent un salaire linéaire en fonction du résultat sous-traité produit

$$w = \alpha + B.C$$

Les conditions de premier ordre de l'agent sont

$$\frac{dc}{de_1} = B \frac{df}{de_1} ; \frac{dc}{de_2} = B \frac{df}{de_2}$$

On notera que la fonction $g(e_1, e_2)$ n'apparaît pas dans les conditions de premier ordre. L'agent choisit l'effort uniquement en fonction du rapport entre le coût de l'effort et l'augmentation marginale du résultat C découlant de l'effort. L'augmentation de B entraînera généralement celle de C mais pourra augmenter ou diminuer NC. »

Bloom, E. et al (2006)

Questions d'évaluation d'impact

Les questions d'évaluation d'impact découlent directement de la théorie du changement associée à une intervention donnée. Bien que chaque intervention de FBR soit différente, plusieurs questions d'évaluation

apparaissent dans différentes évaluations du programme d'évaluation d'impact financé par le HRITF. Dans les sections qui suivent, nous classifions ces questions dans les catégories de « première » ou de « deuxième » génération, en notant qu'elles ne sont pas nécessairement séquentielles. Lors de la sélection des questions d'évaluation d'impact, il convient de garder à l'esprit que toutes les EI ne doivent pas nécessairement utiliser chaque question d'évaluation et qu'il est plus important que les questions soient pertinentes, aux échelons tant local qu'international.

Questions de première génération : Le FBR fonctionne-t-il ?

Comme indiqué ci-dessous, l'évaluation d'impact est censée répondre à la question : « Quel est l'impact (ou l'effet causal) d'une intervention sur les résultats étudiés ? » La première génération des questions de politique devant être abordée par les EI a pour objet de déterminer *si le FBR relatif à la santé fonctionne, dans quelle mesure et dans quels contextes*, ou en d'autres termes : « Le FBR fonctionne-t-il ? » Dans le cas présent, nous avons plusieurs types de **résultats étudiés** :

Quantité des services de santé fournis : La majorité des interventions existantes de FBR ont pour objet de développer l'utilisation des principaux services de santé maternelle et infantile en fournissant des primes aux prestataires et/ou aux usagers. Ces services comprennent généralement des interventions de médecine préventive : immunisation, soins prénatals, accouchements en milieu médicalisé ou distribution de moustiquaires. Nous pouvons mesurer la prestation de services au niveau du prestataire et de la population, que l'on évalue les interventions axées sur l'offre ou sur la demande. Par exemple, nous pouvons mesurer les indicateurs tels que le nombre de visites de soins prénatals, d'accouchements en milieu médicalisé et de visites de surveillance de la croissance en utilisant les données d'enquête sur les établissements de santé. Pour ce qui est des usagers, nous pouvons calculer la probabilité qu'une femme effectue quatre visites de soins prénatals, la probabilité qu'elle accouche dans un établissement de santé et la probabilité que la croissance d'un enfant ait été surveillée au cours des six derniers mois en utilisant les données des enquêtes sur les ménages.

Qualité des services fournis : On peut se demander si les primes versées aux prestataires pour augmenter la quantité des services fournis entraîneront la baisse de la qualité des services, en particulier en zone rurale où les ressources humaines et financières sont limitées. C'est pourquoi les primes ne sont généralement pas uniquement liées à la quantité des services mais également à leur qualité. Que ces primes soient ou non liées à la qualité, il est essentiel que l'EI vérifie si le mécanisme de FBR a une influence positive ou négative sur la qualité. Globalement, ce type d'information nous aidera à comprendre comment augmenter l'utilisation et la qualité des principaux services.

État de santé de la population : L'objectif ultime de tout mécanisme de FBR n'est pas uniquement d'augmenter la quantité et/ou la qualité des services, mais aussi et surtout d'améliorer l'état de santé de la population. La majorité des interventions de FBR sont censées avoir un impact direct sur l'état de santé maternelle et infantile des populations. Comme on l'a vu précédemment, bien qu'il ne soit pas toujours possible de mesurer le taux de

mortalité maternelle ou infantile, des indicateurs de résultats tels que l'état nutritionnel sont observables par des mesures anthropométriques et/ou la réalisation de tests d'anémie.

Gestion des ressources au centre de santé : Le FBR est fréquemment évoqué comme mesure de renforcement des systèmes de santé. C'est pourquoi l'une des principales questions de politique concerne l'influence que les interventions de FBR peuvent avoir sur la gestion des ressources financières et humaines, du matériel et des médicaments à différents niveaux du système de santé.

Services fournis hors FBR : Il est également important de mesurer les externalités positives ou négatives associées aux interventions de FBR. Il est à craindre que les prestataires fournissent davantage de soins correspondant à des services liés au FBR, afin d'augmenter les paiements associés, au détriment des autres services. C'est pourquoi les EI devraient saisir des informations sur les services hors FBR pour déceler toute modification de la qualité et de la quantité de ces services pouvant découler des interventions de FBR.

Équité de la prestation et de l'utilisation des services : Le FBR peut influencer de différentes manières l'équité de la prestation des services. Par exemple, le FBR en zone rurale ou isolée peut cibler les populations défavorisées et peut ou non réussir à augmenter l'accessibilité financière ou autre des soins pour ces populations. Le FBR peut avoir des effets en aval sur les débours directs (formels ou informels), qui pourraient influencer le type de personnes utilisant ces services. Plusieurs programmes de FBR prévoient une rémunération différentielle supérieure pour les services fournis aux populations pauvres et/ou isolées, et il serait important de savoir si ces paiements réussissent à compenser les coûts encourus par les prestataires pour atteindre ces populations. Les évaluations devraient, dans la mesure du possible, vérifier si un programme de FBR bénéficie aux pauvres de manière disproportionnée.

Effets du marché : Dans de nombreux cas, les ménages peuvent choisir leur prestataire de soins. C'est souvent le cas en zone urbaine, où les ménages peuvent choisir entre prestataires publics et privés. Mais même en zone rurale, les ménages de nombreux pays peuvent choisir leur prestataire public et peuvent dans certains cas avoir accès à un prestataire privé. Outre la dimension public-privé, plusieurs types de prestataires peuvent exister : médecins, pharmacies, comptoirs de médicaments, guérisseurs traditionnels, agents de santé communautaires, etc. Ces aspects sont importants dans l'analyse des impacts du FBR, et ce pour plusieurs raisons :

- Prenons par exemple le cas d'un programme de FBR axé sur l'offre, qui récompense expressément les prestataires publics pour les services fournis, ou le cas d'un programme axé sur la demande, qui distribue des bons aux femmes pour qu'elles se rendent chez des prestataires publics. Dans les deux cas, une augmentation de la quantité des services fournis par ces prestataires publics n'est pas suffisante pour prouver que les niveaux de service ont généralement augmenté. Il est possible que les patients aient délaissé des prestataires privés au profit de prestataires publics. On pourrait même imaginer des situations où la prestation de services diminue dans son ensemble, par exemple si des prestataires privés cessent leur activité et les prestataires publics ne récupèrent pas tous leurs patients. Les enquêtes sur les ménages

permettent de mesurer si la prestation de services augmente dans son ensemble, à condition de poser des questions sur l'utilisation des services fournis par tous les types de prestataires.

- Les effets du marché peuvent également avoir un impact sur la qualité globale des soins fournis. Imaginons un scénario avec des prestataires publics et privés, où les prestataires privés fournissent des services de meilleure qualité que leurs homologues publics. Si le FBR (axé sur la demande ou l'offre) incite les patients à délaisser les prestataires privés au profit des prestataires publics, et si la qualité des soins ne change pas, la qualité des soins reçus par les patients sera en moyenne moins bonne avec le FBR que sans. Pour vérifier si une telle situation se présente, il faudrait mesurer la qualité des soins au niveau de la population. On pourrait pour cela réaliser des enquêtes sur les ménages, bien que ce type de mesures de la qualité puisse être sujet à un biais de mémorisation. On pourrait également réaliser une enquête pour mesurer la qualité dans un échantillon représentatif d'établissements publics et privés ; il faudrait alors également pouvoir mesurer le nombre de patients de ces deux types d'établissements afin d'estimer le niveau moyen de qualité pour l'ensemble de la population. Les mesures peuvent devenir très compliquées lorsque de nombreux types de prestataires coexistent.

Effet de revenu ou effet d'incitation ? Le versement de paiements basés sur la performance aux prestataires de soins de santé ou de transferts subordonnés au comportement aux ménages peut avoir deux effets. Le premier est ce qu'on appelle l'effet de *ressource*⁹ et tient au fait que les paiements augmentent les ressources des prestataires ou des ménages. Le second est l'*effet d'incitation* et tient au fait que les paiements sont liés au comportement ou à la performance, contrairement aux sommes forfaitaires ou aux paiements inconditionnels. L'importance relative des effets de ressource et d'incitation est significative lors de l'élaboration des politiques : si le premier est très important par rapport au second, il serait probablement plus économique d'augmenter la quantité des ressources sans les lier à la performance ou aux comportements. On pourrait ainsi éviter les activités de vérification souvent coûteuses qui sont nécessaires lorsque l'on effectue des paiements fondés sur la performance ou le comportement.

L'existence des effets de ressource et d'incitation a des incidences sur l'évaluation d'impact. Imaginez par exemple que le groupe de traitement reçoive des paiements fondés sur la performance tandis que le groupe témoin ne reçoit rien. En comparant ces deux groupes, l'évaluation d'impact estime sans distinction l'effet de ressource et l'effet d'incitation. Il n'est pas possible d'estimer spécifiquement l'effet d'incitation. Il est donc difficile de savoir si le coût de la vérification est une dépense justifiée ou s'il aurait été préférable de distribuer les ressources sans imposer de conditions au niveau des performances. Imaginez maintenant que le groupe de traitement reçoive des paiements fondés sur la performance tandis que le groupe témoin reçoit la même quantité d'argent (en moyenne) sans toutefois établir de lien avec les performances. En d'autres termes, le groupe témoin reçoit en moyenne la même somme d'argent que le groupe de traitement mais cette somme ne dépend pas de la performance du groupe témoin. Dans ce cas, le groupe de traitement et le groupe de référence disposent en moyenne de la même quantité de ressources. Les différences entre les deux groupes en termes d'extrants ou de résultats sont alors attribuables à l'effet d'incitation et non pas à l'effet de ressource.

⁹ L'effet de ressource est parfois désigné effet de revenu.

Étude de cas nationale : Effet de ressource ou d'incitation
Programme de FBP au Rwanda (rémunération au rendement)

« Notre objectif étant d'évaluer l'effet du programme de primes incitatives (rémunération au rendement) indépendamment de l'effet de l'augmentation des ressources financières, la quantité des ressources destinées aux établissements de l'intervention et de référence devait rester constante. Les budgets traditionnels fondés sur les moyens et alloués aux établissements du groupe témoin ont été augmentés du montant moyen de la rémunération au rendement que les établissements du groupe d'intervention ont reçu tous les trois mois au cours de la période d'évaluation de 23 mois. »

Basinga et al. (2011)

Outre le fait de répondre aux questions susmentionnées, il est essentiel lorsque l'on détermine les effets du FBR de désagréger les impacts en fonction des caractéristiques des prestataires et des bénéficiaires. Celles-ci comprennent :

Caractéristiques du prestataire : Nous aimerions également désagréger les impacts des interventions de FBR en fonction des niveaux de formation et de connaissances, de l'autonomie, du type de fonctionnement (public/privé), etc. du prestataire. Nous pourrions ainsi déterminer si les impacts des interventions de FBR sont plus importants pour :

- Le personnel plus hautement qualifié
- Les prestataires plus autonomes
- Les établissements publics plutôt que privés.

Les impacts des programmes de FBR peuvent varier selon les types de prestataires, et ce pour diverses raisons. Par exemple, il est possible que les mesures de référence des connaissances soient plus faibles (ou plus élevées) pour les prestataires plus âgés, et que les prestataires ayant des connaissances plus limitées soient moins sensibles aux mesures incitatives. On pourra alors compléter l'intervention de FBR par une forme ou une autre de formation médicale continue pour les prestataires aux compétences plus limitées.

Caractéristiques de la population : Nous aimerions désagréger les impacts des interventions de FBR sur la population par âge, sexe, niveau de pauvreté et zone rurale/urbaine. Nous pourrions ainsi déterminer si les impacts des interventions de FBR sont plus importants pour :

- Les femmes les plus jeunes
- Les enfants les plus jeunes
- Les ménages les plus riches
- Les ménages ruraux ou urbains.

Questions de deuxième génération : Comment améliorer le fonctionnement des FBR ?

Alors qu'un nombre croissant de structures adopte le FBR, les gouvernements, agences internationales et partenaires d'exécution sont confrontés à des problèmes communs de conception et de mise en œuvre. Les parties concernées constatent qu'il ne suffit pas de savoir si les FBR fonctionnent, mais qu'il s'agit aussi de savoir comment maximiser leurs avantages. Les évaluations d'impact peuvent être conçues pour aborder une partie de ces questions fondamentales concernant la conception des FBR, notamment :

Quels sont les niveaux appropriés de rémunération ? Quel type de rémunération devrait être adopté (en espèces ou en nature) ? Quel est le montant le plus efficace au plan des coûts pour améliorer les résultats ? Quels sont les niveaux de rémunération appropriés pour chaque indicateur sélectionné ? Quels sont les indicateurs appropriés pour déclencher la rémunération ? Pouvons-nous proposer une formule pour déterminer le niveau de la rémunération par type de service ?

Qui devrait être visé par les interventions axées sur l'offre ? Les paiements devraient-ils être instaurés au niveau national ou infranational ? Les paiements devraient-ils être effectués au niveau de l'établissement ou du prestataire ? Les paiements devraient-ils être instaurés au niveau de l'hôpital ou des soins de santé primaire ?

Qui devrait être visé par les interventions axées sur la demande ? Les paiements devraient-ils être déterminés en fonction de critères socioéconomiques ? Devraient-ils être versés aux chefs de ménage, femmes, hommes ou enfants ? Les paiements devraient-ils être versés en espèces ou dans un compte bancaire ?

Comment réduire les erreurs de déclaration et la corruption ? Quelles sont l'intensité et la fréquence optimales pour vérifier et contre-vérifier les données ? Quelles sont les sanctions les plus efficaces pour combattre les déclarations incorrectes ou la corruption ?

Comment les connaissances des prestataires affectent-elles leurs réactions aux récompenses basées sur la performance ? Les prestataires hautement qualifiés réagissent-ils mieux que les prestataires moins qualifiés ? Le renforcement des capacités (par la formation, par exemple) améliore-t-il les réactions des prestataires ? Quel est le niveau optimal de renforcement des capacités ou de remise à niveau ?

Quels sont les principaux aspects structurels nécessaires à la réussite des FBR ? Quel est le niveau approprié d'autonomie concernant l'utilisation des fonds, le recrutement, la passation de marchés, etc. ? Quel est le meilleur régime de propriété (public, privé ou ONG) ?

Portefeuille d'évaluation financé par le HRITF

Les questions de recherche sur les FBR posées au niveau national contribuent à l'amélioration des connaissances mondiales. Les évaluations d'impact financières du HRITF sont conçues autour d'une méthodologie et d'un programme de recherche communs tout en étant adaptées aux spécificités, objectifs

opérationnels et intérêts stratégiques de chaque pays. Les résultats de divers pays et diverses approches du FBR créeront une évaluation unique et détaillée du FBR qui permettra d’explorer les multiples dimensions du FBR, sa mise en œuvre et les comportements et résultats qui en découlent.

Questions d’évaluation d’impact. Bien que toutes les évaluations d’impact aient pour objet de déterminer l’impact du FBR sur des indicateurs communs concernant les services et résultats sanitaires, chaque évaluation fournit également des informations complémentaires sur un aspect spécifique du FBR ou un type spécifique d’intervention de FBR. Certains pays comparent l’impact des paiements axés sur l’offre et la demande ; l’impact de niveaux d’incitation différentiels ; les aspects d’équité du FBR ; etc. Ce travail contribuera à combler le déficit de connaissances mondiales sur l’efficacité des FBR mais aussi sur les raisons de son efficacité ou inefficacité et les éléments contribuant à leur succès (échec). Tableau 1 présente les centres d’intérêt des évaluations des pays.

Résultats étudiés. La majorité des évaluations d’impact financées par le HRITF sont axées sur la santé maternelle et infantile. Dans ce cadre, les pays se concentrent sur des aspects spécifiques, tels que le planning familial, la prévention de la transmission de la mère à l’enfant (PTME) du VIH/sida ou des questions transversales telles que les débours directs pour les soins de santé ou la motivation du personnel. Certaines opérations appuyées examinent le FBR et les MNT dans le but de tirer des leçons applicables à la SMI et à la nutrition. Par exemple, l’évaluation de Karnataka est axée sur les paiements FBR pour le traitement des maladies cardio-vasculaires et du cancer au niveau des hôpitaux tertiaires. En outre, l’évaluation du programme turc est axée sur la détection et le traitement du diabète et de l’hypertension au niveau des pratiques familiales.

Tableau 3 ci-dessous présente les divers résultats étudiés pour chaque pays.

Tableau 2 : Interventions évaluées par pays

Évaluer l’impact de...	Pays
Paiements FBR axés sur l’offre	Afghanistan, Argentine, Bénin, Burkina Faso, Cameroun, Inde, Kirghizstan, Lesotho, Nigéria, RCA, RDC, Rwanda, Turquie, Zambie, Zimbabwe
FBR et formation des prestataires	Zimbabwe
Financement complémentaire	Zambie, Zimbabwe
FBR axé sur la qualité des soins	Afghanistan, Bénin, Cameroun, Kirghizstan, Nigéria, Zambie
Niveaux d’incitation différentiels	Argentine
Amélioration de la surveillance et de la supervision	Argentine, Cameroun, Kirghizstan, RCA
FBR axé sur les hôpitaux	Argentine, Inde, Kirghizstan
Paiements FBR axés sur la demande	Rwanda
FBR axé sur la communauté	Inde, Rwanda
Autre ou à déterminer	Burkina Faso, Laos, Libéria, Sri Lanka, Tadjikistan

Tableau 3 : Résultats étudiés par pays

Résultats étudiés	Pays
Soins maternels (qualité/utilisation)	Afghanistan, Argentine, Bénin, Burundi, Cameroun, Kirghizstan, Lesotho, Libéria, Nigéria, RCA, RDC, Rwanda, Tadjikistan, Zambie, Zimbabwe
Planning familial	Afghanistan, Cameroun, Éthiopie, Lesotho, RCA, Rwanda, Zambie, Zimbabwe
Soins de santé infantile (qualité/utilisation)	Afghanistan, Argentine, Bénin, Burundi, Cameroun, Kirghizstan, Lesotho, Libéria, Nigéria, RCA, RDC, Rwanda, Tadjikistan, Zambie, Zimbabwe
Qualité des soins	Afghanistan, Bénin, Burundi, Kirghizstan, Lesotho, Libéria, Zimbabwe
Débours directs	Afghanistan, Bénin, Inde, RDC, Tadjikistan, Zimbabwe
Tuberculose, paludisme, VIH/sida	Afghanistan, Bénin, Libéria, Nigéria, Zambie, Zimbabwe
Motivation du personnel	Bénin, RDC
Maladies non transmissibles	Inde (soins tertiaires), Turquie (prévention)



Module 2

Constitution de l'équipe d'évaluation d'impact

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats
Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

www.worldbank.org/health/impacetevaluationtoolkit

Module 2. Constitution de l'équipe d'évaluation d'impact

Principales recommandations et outils disponibles pour ce module			
Recommandations	Critique	Important	Utile
<ul style="list-style-type: none"> Le(s) membre(s) de l'équipe chargé(s) principalement de la conception et de l'exécution du projet (par exemple le Chef d'équipe du projet) ne devraient pas jouer le rôle d'enquêteur principal. 	✓		
<ul style="list-style-type: none"> L'enquêteur principal et le coordinateur de l'évaluation jouent un rôle crucial dans la supervision du/des bureau (x) d'études. 	✓		
<ul style="list-style-type: none"> Les homologues locaux pour la recherche peuvent contribuer grandement au succès de l'évaluation d'impact parce qu'ils font intervenir leur connaissance du contexte local et favorisent l'appropriation du programme par le pays. 		✓	
<ul style="list-style-type: none"> Les équipes devraient évaluer la capacité locale à effectuer des enquêtes et déterminer si un appui technique devra être fourni pour assurer la qualité des données d'enquête. 		✓	
<ul style="list-style-type: none"> Un expert de la qualité des données peut aider à créer les conditions initiales appropriées pour assurer la qualité des données d'enquête avant que le bureau d'études se rende sur le terrain. Un superviseur local peut vérifier les processus d'assurance de qualité des données durant l'exécution des enquêtes. 		✓	
<ul style="list-style-type: none"> Une analyse qualitative et une analyse du rapport coût-efficacité peuvent ajouter une richesse et une granularité considérables aux questions auxquelles répondra l'évaluation d'impact. 		✓	
<ul style="list-style-type: none"> Les évaluations d'impacts nécessitent plusieurs séries de données complexes, et un bon analyste des données aidera l'équipe à gérer et analyser les données de façon rapide et fiable. 		✓	
<ul style="list-style-type: none"> Les calculs de puissance peuvent être confiés à l'enquêteur principal, mais un expert du calcul de puissance peut avoir davantage de temps et d'expertise à consacrer à cette tâche. 			✓



Outils
<ul style="list-style-type: none"> 2.01 Principal Investigator TOR (Termes de référence de l'enquêteur principal) 2.02 Evaluation Coordinator TOR (Termes de référence du coordinateur de l'évaluation) 2.03 Data Analyst TOR (Termes de référence de l'analyste des données) 2.04 Local Researcher TOR (Termes de référence des chercheurs locaux) 2.05 Power Calculation Expert TOR (Termes de référence de l'expert des calculs de puissance)

Principales recommandations et outils disponibles pour ce module

- 2.06 Data Quality Expert TOR (Termes de référence de l'expert en qualité des données)
- 2.07 Qualitative Principal Investigator TOR (Termes de référence de l'enquêteur principal (données qualitatives))
- 2.08 Qualitative Field Worker TOR (Termes de référence de l'agent de terrain (données qualitatives))
- 2.09 Cost-analysis Expert TOR (Termes de référence de l'expert en analyse des coûts)

Contenu du module

Enquêteur principal.....	2-6
Coordinateur de l'évaluation	2-7
Analyste des données	2-8
Expert du calcul de puissance	2-8
Expert de la qualité des données	2-9
Superviseur externe	2-10
Expert de la recherche qualitative	2-11
Expert de l'analyse des coûts	2-12
Participation de chercheurs locaux à l'évaluation d'impact	2-13

Une importante caractéristique d'une évaluation crédible est l'absence de conflits d'intérêts au niveau des évaluateurs. Autrement dit, ceux-ci doivent être suffisamment séparés des exécutants du programme. Toutefois, il est souvent difficile de dissocier complètement une évaluation d'impact des règles opérationnelles du programme car ce sont les règles du programme qui déterminent entre autres d'où va venir le groupe de comparaison.

Face à cette difficulté, nous recommandons que la conception et l'exécution de l'évaluation d'impact et de l'analyse des données soient confiées à une équipe qui soit suffisamment distincte de celle qui est chargée de la conception et de l'exécution du projet. Ces équipes devront toutefois collaborer pour assurer que :

- Les questions prioritaires de politique générale pour les pays respectifs sont intégrées à l'évaluation d'impact
- Le calendrier des activités EI et du projet est approprié (par exemple, les travaux de base devraient être achevés avant le début de l'intervention)
- L'exécution de l'intervention va de pair avec la sélection de groupes de traitement et de comparaisons pour la stratégie d'évaluation d'impact.

Une équipe d'EI comprend généralement un personnel à temps plein et à temps partiel aussi bien local qu'international. Elle se compose généralement comme suit :

- Enquêteur principal (EP) et (le cas échéant) co-enquêteur principal (Co-EPI)
- Coordinateur de l'évaluation (CE)
- Analyste des données
- Expert(s) de la qualité des données et éventuellement un superviseur externe
- Expert du calcul de puissance

L'équipe d'EI pour le projet peut également comprendre :

- Un expert en recherche qualitative
- Un expert de l'analyse des coûts.

Les **termes de référence** de tous ces membres de l'équipe sont inclus dans la boîte à outils.

Il est utile d'inclure des collaborateurs locaux dans l'équipe d'EI lorsque les capacités nécessaires existent ou peuvent être renforcées. C'est ainsi que des universitaires locaux peuvent être désireux de participer à l'évaluation en tant que co-enquêteurs principaux et se perfectionner ainsi progressivement. La participation local peut accroître l'implication du pays, le savoir local et l'appropriation du programme, et il en résulte une situation où il y a tout à gagner.

Il est important de souligner que la communication entre les membres des équipes et la coordination entre les équipes du projet et d'évaluation d'impact par l'intermédiaire du chef d'équipe de projet sont cruciales pour le succès de l'EI.

Pleins feux sur le pays : Implications du renforcement des équipes et du compte rendu de l'exécution de l'évaluation d'impact

Projet d'appui à la réhabilitation du secteur de la santé en République démocratique du Congo (PARSS)

Aux premiers stades de l'évaluation d'impact et pour une part importante de la collecte des données de base, l'enquêteur principal (EP) de l'équipe d'EI et le Chef d'équipe de projet n'étaient pas basés en RDC, (...) des consultants internationaux étaient chargés d'une part de diriger l'échantillonnage et la randomisation et d'autre part, la préparation et l'exécution du recueil des données en effectuant seulement quelques missions sur le terrain et sans aucun contact avec l'Unité d'exécution du projet ni avec l'équipe de la Banque mondiale en RDC. En dehors de l'équipe, la communication entre les autorités provinciales et les équipes du projet et d'EI faisait également défaut. Du fait de la rotation de personnel au sein des équipes et du manque de présence sur le terrain, le bureau d'études n'a pas bénéficié de la formation, de la supervision, de la vérification et du contrôle de la qualité nécessaires durant les travaux sur le terrain.

[Face à ces problèmes], une fois achevée la collecte des données de base, (...) le nouveau Chef d'équipe de projet a chargé un nouvel enquêteur principal et un co-enquêteur principal d'analyser les données de base et de préparer l'enquête complémentaire. L'enquêteur principal et le co-enquêteur principal n'étaient pas basés en RDC, mais ils ont chargé un assistant de recherche (AR), basé à plein temps à Lubumbashi de gérer les activités courantes de l'EI et de comprendre la nature des problèmes pratiques à résoudre dans un contexte de sortie de conflit. L'AR était le point focal des activités sur le terrain et a donné régulièrement des informations en retour aux membres de l'équipe d'EI hors terrain. L'AR a également été capable de comprendre plusieurs problèmes liés à l'exécution du programme. Toute l'équipe a mis l'accent sur la communication entre les membres de l'équipe.

La présence continue sur place des membres de l'équipe d'EI est une condition clé du succès de l'EI. Le fait de disposer d'au moins un co-enquêteur principal, d'un coordinateur de l'évaluation ou d'un assistant de recherche sur le terrain, en particulier durant les activités de collecte des données, constitue une stratégie extrêmement précieuse.

Un problème courant rencontré durant les évaluations d'impact est le manque de communication entre l'équipe d'EI et l'équipe du projet. Ce problème risque manifestement de compromettre la validité de l'EI. Les **Chef d'équipes de projet de la Banque mondiale doivent** veiller à combler le déficit d'information entre les deux équipes, et **faciliter la collaboration entre les équipes opérationnelles et d'EI.**

Il est important de définir clairement le rôle de chaque membre de l'équipe. Les termes de référence devraient inclure toutes les activités qu'un membre de l'équipe est censé approuver, et les modalités de notification au Chef d'équipes de projet et/ou à l'équipe. (...) Les Chefs d'équipes de projet sont chargés de faire en sorte que chaque membre de l'équipe ait les capacités qu'exige l'exercice de son mandat, et de lui assurer la formation nécessaire dans le cas contraire.

Pour plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.

Dans la section suivante, nous examinons en détail les rôles et responsabilités qui devraient selon nous être ceux des différents membres éventuels de l'équipe d'EI :

Enquêteur principal

Le rôle de l'enquêteur principal est d'assurer la direction technique au niveau de la conception, de la méthodologie et de l'analyse de l'EI, ainsi que pour la gestion générale de l'étude. L'enquêteur principal adapte ensuite l'évaluation au contexte particulier du pays tout en gardant à l'esprit les objectifs du programme global d'EI du FBR.

L'enquêteur principal collabore avec le Chef d'équipe du projet et des homologues du gouvernement afin d'assurer que la conception et l'exécution de l'EI soient intégrées au déploiement de l'intervention FBR. Comme on l'a vu, pour qu'une évaluation soit crédible, les évaluateurs doivent être suffisamment séparés des exécutants du programme. Sauf circonstances vraiment exceptionnelles, nous recommandons donc vivement que le Chef d'équipe du projet ou un/d'autre(s) membre(s) de l'équipe principalement responsable(s) de la conception et de la supervision du projet ne jouent pas le rôle d'enquêteur principal pour l'évaluation d'impact.

Tableau 4 indique le temps de travail jugé nécessaire à un enquêteur principal.

Tableau 4 : Temps de travail jugé nécessaire à un enquêteur principal

Activités	Jours de travail
Conception de l'évaluation d'impact et discussions avec des homologues clés (y compris au moins une mission sur place)	30
Recueil, supervision et gestion des données de base	15
Analyse et diffusion des données de base	20
Suivi et activités complémentaires	15
Supervisions et gestion de la collecte des données de base (y compris au moins une mission sur place)	15
Analyse d'impact et diffusion	30
TOTAL	125 jours

Ce temps de travail estimatif part du principe que l'enquêteur principal collaborera avec un coordinateur de l'évaluation (voir plus bas) pour apporter son assistance au bureau d'études pour les tâches courantes durant la préparation de la collecte des données, le travail sur le terrain, la saisie et l'analyse des données. Dans certains cas, les enquêteurs principaux risquent de ne pas pouvoir consacrer eux-mêmes suffisamment de temps et d'attention à une évaluation et devront s'associer à un autre enquêteur en tant que co-enquêteur principal (co-EP).

Nous recommandons que l'enquêteur principal ait au moins les qualifications suivantes :

- Doctorat dans un domaine pertinent, de préférence économie ou politique de la santé.
- Minimum de 5 ans d'expérience de l'évaluation d'impact des projets

- Minimum de 5 ans d'expérience de la conception et de l'exécution des évaluations quantitatives d'impact à l'aide de modèles randomisés ou contrôlés d'une autre façon
- Expérience appropriée de la mesure des résultats dans le domaine de la santé au moyen d'enquêtes auprès des ménages
- Expérience appropriée de la conception et de la coordination du travail de terrain pour de grandes enquêtes auprès des ménages et des centres de santé
- Expérience appropriée de l'analyse des données quantitatives (ménages et centres de santé) à l'aide d'un logiciel d'analyse statistique (de préférence STATA)
- Expérience appropriée à la coordination de l'exécution du travail sur le terrain aux fins de l'évaluation d'impact
- Excellente aptitude à la communication par écrit en anglais, et en particulier à la rédaction de protocoles et de documents de recherche et de rapports descriptifs à l'intention d'un public varié
- Aptitude à faciliter la communication entre divers niveaux de gestion et à travailler de façon indépendante afin de respecter les délais
- Idéalement, l'enquêteur principal devrait avoir publié des évaluations dans des revues possédant un comité de lecture.

Coordinateur de l'évaluation

Le coordinateur de l'évaluation dirige les activités courantes liées à la conception de l'évaluation d'impact, ainsi qu'au recueil et à l'analyse des données. Cela exige généralement un investissement considérable en temps, estimé comme suit au tableau 5 :

Tableau 5 : Temps de travail estimatif nécessaire à un coordinateur de l'évaluation

Activités	Jours de travail
Conception et discussions sur l'évaluation d'impact avec des homologues clés (y compris au moins une mission sur place)	25
Supervision et gestion de la préparation de la collecte des données de base (sur place pour la préparation et toute la durée de l'enquête)	125
Analyse et diffusion des données de base	50
Suivi et activités complémentaires entre les enquêtes de base et finales	60
Supervision et gestion de la collecte des données fiables (sur place pour la préparation de l'enquête et toute sa durée)	125
Analyse d'impact et diffusion	75
TOTAL	460 jours

Nous recommandons en outre que le coordinateur de l'évaluation ait au moins les qualifications suivantes :

- Maîtrise ou diplôme équivalent dans un domaine pertinent comme la santé, la santé publique ou l'économie
- Expérience du logiciel d'analyse statistique (STATA)
- Expérience appropriée de la conduite, de la gestion et de la conception du travail de terrain et de la collecte de données pour des recherches empiriques

- Excellente aptitude à communiquer par écrit en anglais, en mettant l'accent sur les protocoles de recherche, les documents de recherche et les rapports descriptifs aux destinataires divers
- Connaissance des langues locales préférables
- Qualités d'organisation exceptionnelles, aptitude à faciliter la communication entre divers niveaux de gestion et à travailler de façon indépendante pour respecter les délais
- Une expérience antérieure de l'évaluation de l'impact des projets dans les pays en développement est très souhaitable.

Analyste des données

L'analyste des données est chargé d'aider l'enquêteur principal à mener à bien en temps voulu l'analyse des séries de données de base et finales. Une analyse des données de base est nécessaire pour valider la conception de l'évaluation, fournir à l'équipe du projet et aux partenaires un rapport descriptif des données, et des recommandations concernant des séries intermédiaires (le cas échéant) et finales. Pour les évaluations financées par le HRITF, l'analyse initiale des données de base en vue de valider la conception de l'évaluation est une étape capitale pour le déblocage de la seconde tranche de financement. De plus, des pressions considérables s'exerceront pour réaliser l'analyse d'impact après la collecte des données finales. L'enquêteur principal et le coordinateur de l'évaluation peuvent avoir les compétences nécessaires pour effectuer cette analyse, mais ils risquent de ne pas avoir suffisamment de temps pour trier et documenter les données, diriger les analyses et consigner les résultats. Le tableau 6 indique le temps de travail jugé nécessaire à un analyste des données.

Tableau 6 : Temps de travail jugé nécessaire à un analyste des données

Activités	Jours de travail
Analyse et diffusion des données de base	50
Analyse d'impact et diffusion	75
TOTAL	125 jours

Expert en calcul de puissance

Un expert du calcul de puissance détermine la taille de l'échantillon nécessaire ou l'effet minimum détectable de traitement pour répondre aux questions relatives à l'évaluation proposée. En bref, cet expert évaluera la taille minimum de l'échantillon nécessaire pour détecter une différence de résultats significative entre les groupes de traitement et de comparaison. Pour les études faisant l'objet d'un échantillon d'une taille donnée, cet expert estimera le plus petit effet de traitement statistiquement détectable. Le temps que nécessite un expert du calcul de puissance-échantillonnage est minime par rapport à celui qu'exigent d'autres fonctions, et si les données requises sont déjà disponibles, trois à cinq jours devraient être suffisants.

Le rôle d'expert du calcul de puissance peut être confié à l'enquêteur principal mais compte tenu de la haute technicité de cette tâche et du nombre limité de jours de travail nécessaires, nous recommandons de recruter un spécialiste.

Expert en qualité des données

Lorsque l'on effectue une enquête, un problème majeur consiste à faire en sorte que la qualité des données soit suffisante. Certains pays sont parfaitement capables de planifier et exécuter eux-mêmes la collecte des données de façon satisfaisante, tandis que d'autres n'ont que des moyens limités. L'expérience a montré que ces derniers recourent généralement à un appel d'offres pour charger un cabinet international de collecter des données. En pareil cas, les problèmes suivants se sont posés : i) les cabinets internationaux font appel à des sous-traitants et/ou à des employés locaux pour procéder à la collecte des données, ce qui signifie que la qualité des données dépend de celle de la formation des intéressés et de la supervision assurée par le cabinet international ; ii) les cabinets internationaux n'ont pas nécessairement l'expérience du pays où l'enquête est réalisée, ce qui peut entraîner des résultats médiocres ; et iii) les cabinets internationaux ont tendance à limiter la durée du séjour sur place de son personnel international pour des raisons d'économie, ce qui se traduit par une supervision et un appui technique réduits.

Pour les raisons citées plus haut, nous recommandons que l'équipe d'EI comprenne un consultant (soit un individu, soit un cabinet engagé spécifiquement à cette fin) pour fournir une assistance technique en vue d'assurer sur place la qualité des données et de la maximiser ainsi durant l'étude, notamment :

Examen des propositions techniques et financières d'un bureau d'enquête. Un expert de la qualité des données peut examiner et commenter les propositions techniques et financières d'un bureau d'études pour déterminer si celui-ci propose une méthodologie, une composition des équipes de terrain et un plan de travail appropriés, et si le budget est suffisant compte tenu des besoins liés à la collecte, à la saisie et à la gestion proposées des données. Il est préférable de faire examiner ces propositions avant de sélectionner et de retenir les services du responsable de l'exécution du fait qu'il est généralement très difficile de modifier ou de rallonger un budget une fois le contrat signé. Un examen approprié des propositions techniques et financières aide à atténuer les problèmes futurs tels qu'une sous-estimation des besoins budgétaires. Les questions relatives aux services du/des bureau(x) d'études sont examinées plus en détail au Module 4.

Conception, adaptation et mise à l'essai préalable des instruments d'enquête. Le premier outil pour assurer la qualité des données est un instrument d'enquête bien conçu caractérisé par un contenu et un formatage appropriés. Bien que des instruments d'enquête normalisés aient été élaborés pour évaluer le FBR, les équipes sous-estiment généralement le temps nécessaire pour adapter et mettre à l'essai les instruments d'enquête dans un pays donné. Si l'enquêteur principal et le coordinateur de l'évaluation ne disposent pas de suffisamment de temps, l'expert de la qualité des données peut prodiguer des conseils sur le développement et la mise à l'essai sur place d'instruments d'enquête clés. Les questions relatives à la conception, à l'adaptation et à la mise à l'essai préalable de/des instrument(s) d'enquête sont examinées plus en détail au



Module 4

Préparation de la Collecte des Données

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats
Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

www.worldbank.org/health/impactevaluationtoolkit

Préparation de la collecte des données.

Développement et adaptation d'un/de programme(s) de saisie des données. Les données recueillies sur le terrain devront être traitées à l'aide d'un programme de saisie de données (PED) et l'expert de la qualité des données peut donner des conseils sur l'élaboration d'un PED ou y participer. Un PED utile intégrera des mesures de la qualité des données telles que des vérifications d'écart et de cohérence afin de réduire les erreurs intervenant au point de saisie des données. Les questions liées à la conception du/des programme(s) de saisie des données sont examinées plus en détail au Module 4.

Développement et exécution du programme et des matériels de formation. Pour assurer la qualité des données, il est très important que les superviseurs, les équipes de terrain et le personnel chargé de la saisie des données reçoivent une formation suffisante et assurée efficacement. Les enquêteurs principaux et les coordinateurs des évaluations ne sont généralement pas très expérimentés en la matière, et nous recommandons donc que l'expert de la qualité des données participe à la formation des équipes chargées du recueil et de la saisie des données et la supervise. Les questions relatives à la formation sont examinées plus en détail au Module 5.

Supervision directe du recueil, de la gestion et de la saisie des données. D'une façon générale, la qualité des évaluations d'impact est ni plus ni moins celle des données collectées. Il est donc crucial de veiller à ce que les mesures de la qualité des données soient respectées durant le recueil et le regroupement (y compris l'acheminement des données obtenues sur le terrain) des données et leur saisie. Nous recommandons qu'un expert de la qualité des données supervise directement le recueil, la gestion et la saisie des données dès le début du travail sur le terrain. Les questions relatives à la collecte, à la gestion et à la saisie des données sont examinées plus en détail dans les Modules 5 et 6.

Le temps de travail jugé nécessaire pour la première série de données à collecter dépend du travail qui peut être géré par le bureau d'étude. Le tableau 7 indique le temps de travail jugé nécessaire pour un expert de la qualité des données.

Tableau 7 : Temps de travail jugé nécessaire à un expert de la qualité des données

Activités	Jours de travail
Examen de la proposition technique et financière du bureau d'études	5
Conception, adaptation et test préalable des instruments d'enquête	20-30
Élaboration et adaptation du/des programme(s) de saisie des données	5-20
Élaboration et exécution du programme et des matériels de formation	5-20
Supervision de la collecte, de la gestion et de la saisie des données	5-15
Total	40-90

Superviseur externe

L'expert de la qualité des données peut fournir une expertise technique durant la préparation et au premier stade du recueil des données, mais c'est généralement un consultant international qui ne séjournera pas dans le

pays pendant toute la durée de l'opération de recueil et de la saisie des données. Toutefois, la qualité des données n'est pas déterminée seulement aux stages préparatoires, mais dépend pour beaucoup de la collecte des données proprement dite. C'est pourquoi un superviseur externe peut être recruté localement pour jouer un rôle plus durable d'assurance de la qualité des données, en particulier lorsque les moyens du bureau d'études sont limités. Sous le contrôle de l'expert de la qualité des données (contrôle qui peut être exercé directement sur place ou ultérieurement à distance), le superviseur externe est chargé de contrôler de façon aléatoire l'opération de collecte et de saisie des données. Il permet de réagir rapidement aux problèmes concernant la qualité des données durant son travail sur le terrain et de maintenir durablement des normes strictes de qualité des données. Le superviseur externe relève à la fois de l'expert de la qualité des données et de l'enquêteur principal, et peut conseiller rapidement au bureau d'études des mesures correctives lorsque l'audit révèle des menaces pour la qualité des données.

Les termes de référence du superviseur externe peuvent être adaptés à celui de l'expert de la qualité des données. Le temps de travail jugé nécessaire pour la première série de données à recueillir est le suivant :

Tableau 8 : Temps de travail jugé nécessaire à un superviseur externe de la qualité des données

Activités	Jours de travail
Contrôles aléatoires du recueil, de la gestion et de la saisie des données (sur place)	40
Notification de la qualité des données et mise en œuvre de mesures correctives	15
Total	55

Expert en recherche qualitative

L'analyse quantitative peut permettre de déterminer si le FBR a fonctionné dans un certain contexte pour obtenir des résultats particuliers mais dans bien des cas, elle ne permet pas de déterminer pourquoi le FBR a fonctionné ou non. Dans ces circonstances, des données qualitatives peuvent fournir de plus amples détails sur le contexte précis et les perspectives des intéressés eux-mêmes, mieux comprendre les processus et offrir de nouvelles explications pour certains résultats. Idéalement, l'analyse qualitative devrait être incorporée dans l'EI à divers points du cycle du projet et prévue comme un élément de l'évaluation d'impact. Il ne s'agit pas ici d'examiner en détail les méthodes de recherche qualitative, mais un protocole et des instruments de recherche qualitative seront élaborés pour la seconde version de la boîte à outils.

Comme la recherche qualitative donne une vue globale d'une intervention dans le contexte de la société, de la culture et/ou d'un groupe particulier de gens, elle nécessite généralement la collecte de précieuses données dans une perspective d'« initié », ce qui exige une forte présence sur place. Dans bien des cas, il peut être plus rentable de charger un consultant local ou régional de ce type de travail. Selon la portée du travail qualitatif proposé, le temps de travail jugé nécessaire pour la première série de données à recueillir est le suivant :

Tableau 9 : Temps de travail jugé nécessaire à un expert de la recherche qualitative

Activités	Jours de travail
Mission d'évaluation du contexte et d'identification des méthodes et questions de recherche	10
Élaboration d'un protocole et d'outils de recherche qualitative	10
Recrutement et formation sur place de l'équipe de recherche qualitative	15
Mise à l'essai des instruments et des méthodes d'enquête	10
Gestion du recueil des données	20
Transcription, analyse et diffusion	30
Total	95

Expert en analyse des coûts

Une analyse quantitative peut nous indiquer si et dans quelle mesure une intervention FBR a fonctionné, mais pour pouvoir décider si une intervention mérite d'être menée à plus grande échelle, il faut également en considérer le coût. Avec les résultats de l'évaluation d'impact, l'analyse des coûts nous permet de calculer le rapport coût-efficacité, la faisabilité financière et la durabilité des interventions FBR. On dispose actuellement de peu d'informations sur les coûts et les besoins financiers à long terme des interventions FBR au niveau aussi bien de l'offre que de la demande ; nous recommandons par conséquent que ces données soient collectées et analysées dans le cadre des activités de collecte de données de l'EI. Les résultats de l'analyse des coûts peuvent servir à aider les décideurs et les exécutants du programme à :

- Comparer les coûts et les produits d'une/d'intervention(s) FBR à ceux d'une situation inchangée, ou d'autres investissements dans la santé ;
- Déterminer si une/des intervention(s) FBR constitue (nt) des investissements justifiés d'un point de vue économique ;
- Déterminer si l'exécution à plus grande échelle d'une/d' intervention(s) FBR est faisable d'un point de vue économique et financier ;
- Évaluer le coût et la faisabilité financière de plans FBR et les moyens éventuels de les poursuivre ; et
- Identifier les domaines dans lesquels d'éventuels gains d'efficacité sont possibles.

Notre propos n'est pas d'examiner ici de façon approfondie les questions d'analyse des coûts et de rapport coût-efficacité, mais un protocole et des instruments d'analyse du rapport coût-efficacité sont en cours d'élaboration pour la seconde version de la boîte à outils.

Le temps de travail jugé nécessaire pour collecter sur place la première série de données pour une analyse des coûts est le suivant :

Tableau 10 : Temps de travail jugé nécessaire pour un expert de l'analyse des coûts

Activités	Jours de travail
Mission/Appui à distance pour évaluer le contexte et identifier la méthodologie d'analyse des coûts	5
Développement/adaptation d'outils d'analyse des coûts	5
Recrutement et formation sur place de l'équipe chargée de l'analyse des coûts	10
Mise à l'essai des instruments et de la méthodologie d'enquête	5
Suivi de la collecte des données	5
Analyse et diffusion	15
Total	45

Participation de chercheurs locaux à l'évaluation d'impact

Comme il ressort abondamment de la présente boîte à outils, l'exécution et la gestion d'une étude d'évaluation d'impact à grande échelle posent de nombreux problèmes. Selon l'endroit où sont basés l'enquêteur principal et le coordinateur de l'évaluation, un moyen d'améliorer les chances de succès d'un projet EI est de collaborer avec des chercheurs locaux, qui peuvent être en mesure de :

1. *Renforcer l'appropriation locale et la présence de l'étude.* Bien que les représentants locaux du ministère de la Santé (MS) et les équipes chargées de la conception du projet puissent apporter un appui à l'étude, ils sont souvent éloignés de son exécution et de sa gestion effectives. Sachant qu'une évaluation d'impact dure généralement de 3 à 5 ans, il est crucial que les autorités et partenaires locaux fidèles à la conception et au calendrier de l'évaluation conception, ce à quoi les représentants locaux de l'équipe EI peuvent contribuer.
2. *Assurer la supervision directe et en temps opportun de l'assurance de qualité.* Toutes sortes d'activités tout au long du cycle du projet EI nécessiteront un engagement et/ou une supervision directe sur place. En s'associant avec des chercheurs locaux, l'équipe EI peut assurer une solide collaboration avec le MS et d'autres partenaires, ainsi qu'avec le bureau d'études.
3. *Assurer la sensibilité et la pertinence culturelles.* Les services de santé, nutrition et population ont élaboré toutes sortes de ressources pour les équipes d'EI dans le pays, notamment des questionnaires ainsi que des matériels et protocoles de formation. Toutefois, tous ces outils doivent être adaptés au contexte local du pays afin d'assurer la prise en compte de ses caractéristiques culturelles spécifiques et d'assurer ainsi que la conception et la méthodologie globales de l'EI soient adaptées au contexte de ce pays.
4. *Renforcer la capacité locale d'évaluation d'impact.* Les évaluations d'impact à grande échelle offre une excellente occasion de renforcer les capacités locales en ce qui concerne les méthodes d'évaluation d'impact, la gestion des enquêtes et le contrôle de la qualité des données en permettant de moins faire appel à des chercheurs internationaux. Les compétences acquises par les chercheurs locaux en ce qui concerne l'évaluation des effets du FBR valent aussi pour d'autres types d'évaluation et de recherche et pour l'élaboration de politiques fondées sur des éléments probants.

Pleins feux sur le pays : Évaluation de la capacité locale et nécessité de constituer l'équipe EI
Projet de financement basé sur la performance au niveau communautaire au Rwanda

La planification du Projet de financement basé sur la performance (PBF) au niveau communautaire au Rwanda a commencé par trois missions conjointes des équipes de conception et d'évaluation d'impact du projet dirigées par le Chef de l'équipe de projet. La coordination entre les équipes a permis un dialogue régulier entre la Banque mondiale, le ministère de la Santé et les partenaires de développement sur les objectifs de politique générale du projet communautaire PBF et les priorités connexes de recherche de son évaluation d'impact.

Le Chef de l'équipe de projet a formé une équipe composée comme suit : i) un enquêteur principal basé à Washington (États-Unis), chargé d'apporter un appui technique de haut niveau pour la conception de l'évaluation ; ii) un coordinateur basé à Washington (États-Unis), chargé d'apporter un appui technique pour la conception ainsi que de fournir un soutien intensif pour la gestion courante du temps et des produits de l'équipe d'évaluation et à faire de longs séjours sur place pour contribuer à la préparation et l'exécution du projet ; iii) deux chercheurs basés à Kigali (Rwanda), chargés de fournir un appui technique pour la conception de l'évaluation, en particulier pour la mise au point du questionnaire, la stratégie d'échantillonnage sur le terrain et l'assurance de qualité des données ; et iv) un bureau d'études basé à Kigali (Rwanda), chargé de gérer la collecte des données au niveau de la coopérative des agents de santé communautaires, de ces agents proprement dit et des ménages. Deux experts supplémentaires de l'assurance de la qualité des données ont été engagés pour deux missions précises consistant à apporter un appui technique pour l'élaboration du programme de saisie des données, la gestion du travail de terrain, l'acheminement et la saisie des données, à savoir une mission pour mettre à l'essai les questionnaires et la stratégie de gestion du travail de terrain, et une mission au début du recueil des données pour donner des conseils sur les processus en cours.

Tout au long des phases de préparation de base et d'exécution, la participation des deux chercheurs locaux a été cruciale pour faire face à plusieurs problèmes concernant la qualité de la collecte des données : i) appui et supervision considérables du bureau d'études durant l'adaptation et la traduction des questionnaires, ii) orientation importante donnée au cabinet chargé de la collecte des données sur l'échantillonnage et la gestion sur le terrain pour une enquête à grande échelle auprès des ménages, iii) maintenir des normes de qualité grâce à une supervision, des vérifications ponctuelles aléatoires et une communication avec les agents de terrain durant le travail de terrain.

Au cours de la préparation du projet, les chercheurs locaux ont pu poursuivre le dialogue avec le ministère de la Santé et les partenaires de développement, ce qui a permis de coordonner les activités entre l'exécution du projet et le recueil des données de base. Ils ont également représenté l'enquêteur principal de l'étude au niveau local.

Les deux experts de la gestion des données ont procédé à un renforcement sensible des capacités du directeur du cabinet chargé du recueil des données et influé ainsi sur la culture de gestion des données de l'organisation.

Pour de plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.



Module 3

Conception de l'évaluation d'impact

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats
Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

www.worldbank.org/health/impacetevaluationtoolkit

Module 3. Conception de l'évaluation d'impact

Principales recommandations et outils disponibles pour ce Module			
Recommandations	Essentiel	Important	Souhaitable
<ul style="list-style-type: none"> Il convient de concevoir une évaluation d'impact prospective avant ou pendant l'intervention. 	✓		
<ul style="list-style-type: none"> Les équipes devraient élaborer un cadre des résultats pour le projet de FBR afin d'identifier les principaux canaux à travers lesquels les activités du programme de FBR influenceront sur les résultats et les réalisations clés. 	✓		
<ul style="list-style-type: none"> L'assignation aléatoire aux groupes d'intervention(s) et de comparaison est la stratégie d'identification recommandée aux fins de l'évaluation d'impact du FBR. 		✓	
<ul style="list-style-type: none"> Les équipes devraient évaluer les risques qui pèsent ou qui pèseront sur la validité interne de l'évaluation (par exemple, contamination, manque de puissance) et les contrôler au fil du temps. 	✓		
<ul style="list-style-type: none"> Les calculs de puissance constituent un élément important de la conception d'une évaluation d'impact. À défaut d'une puissance suffisante, l'évaluation d'impact pourrait ne pas être en mesure de répondre aux principales questions stratégiques. La taille de l'échantillon doit autoriser une puissance suffisante. 	✓		
<ul style="list-style-type: none"> L'échantillon doit être représentatif de la population qui, à terme, bénéficiera du programme. 	✓		
<ul style="list-style-type: none"> Lorsque les partenaires nationaux adoptent le concept de l'évaluation d'impact et reconnaissent l'importance de respecter les branches de l'étude, il devient plus facile de maintenir les groupes de traitement et de comparaison intacts jusqu'à la réalisation de l'enquête de suivi. 	✓		
<ul style="list-style-type: none"> Le choix des indicateurs de l'étude est crucial. Chacun des indicateurs retenus doit être mesurable à l'aide des instruments choisis de collecte de données. 	✓		
<ul style="list-style-type: none"> Les équipes peuvent se référer à l'ouvrage « <i>L'évaluation d'impact en pratique</i> » (Gertler <i>et al.</i> 2011) pour un examen approfondi des stratégies d'identification de l'évaluation d'impact. 		✓	
<ul style="list-style-type: none"> Au moment de définir l'unité de randomisation, les équipes doivent faire les arbitrages requis entre la puissance de l'évaluation d'impact et le risque de contamination croisée entre les unités de randomisation. 		✓	



Outils
<ul style="list-style-type: none"> 3.01 RBF Indicators (Indicateurs du FBR) 3.02 WHO Output and Outcome Indicators (Indicateurs de produits et de résultats (OMS)) 3.03 IE Design Paper Template (Modèle de document de conception de l'EI) 3.04 IE Budget Template (Modèle de budget EI) 3.05 Ex-ante Power Calculation Example (Exemple de calcul de puissance ex ante) 3.06 Binary Power Calculations (Calculs de puissance binaires)

- 3.07 Power Calculation References (Références Calcul de puissance)

Contenu du module

Contexte et justification	3-5
Cadre des résultats du FBR dans le secteur de la santé	3-5
Questions de recherche et pertinence stratégique.....	3-5
Indicateurs des résultats et des réalisations à l'étude	3-7
Stratégie d'identification	3-8
Quels sont les éléments du programme de FBR qui devront être évalués ?	3-8
Comment l'équipe estimera-t-elle le contrefactuel ?	3-10
Choix de l'unité de randomisation.....	3-10
Échantillon	3-12
Quels sont les critères d'inclusion du cadre d'échantillonnage ?.....	3-12
Quelle est la taille requise de l'échantillon ?.....	3-15
Besoins de données	3-20
Calendrier	3-22
Équipe d'EI	3-23
Plan de diffusion	3-23
Budget.....	3-24

Il convient d'entrée de jeu d'insister sur l'importance de concevoir l'évaluation d'impact (EI) avant ou pendant la conception de l'intervention. Le fardeau de travail des chefs d'équipe occupés à la préparation du projet s'en trouve accru, mais on assure ainsi l'alignement de la conception de l'EI et de la sélection des groupes de traitement et de comparaison sur la conception et le déroulement prévu des phases du programme. Cette façon de procéder permet en outre de maintenir l'importance stratégique des efforts de préservation de la validité de l'EI : une fois les décisions politiques prises et publiées, il ne sera plus possible de négocier des changements pour les besoins de l'EI. Il est essentiel que les équipes d'EI et de projet collaborent pour pouvoir mener de front la conception de l'EI et de l'intervention.

Le présent module fournit un aperçu des éléments qu'il convient d'inclure dans le **document de conception de l'évaluation d'impact**¹⁰. Ce document décrit les éléments de base de l'évaluation, y compris le cadre des résultats, les questions de recherche, la stratégie d'identification, les données, le recrutement du personnel et le budget. La première version de ce document peut servir de note de synthèse aux fins de l'examen par les pairs et de l'approbation par la direction de la Banque mondiale (voir plus loin). Au-delà de cette étape, le document de conception devrait être régulièrement mis à jour pour refléter l'état d'avancement de l'évaluation, tout changement apporté aux méthodes ou les problèmes risquant d'influer sur l'EI.

Bien que le présent module ait pour objet l'évaluation d'impact prospective randomisée, il convient de mentionner que d'autres types d'évaluation sont possibles, parfois moins coûteux, et peuvent utiliser des données déjà disponibles. Le manuel *L'évaluation d'impact en pratique* (Gertler *et al.* 2011) examine cette question d'une manière plus détaillée. À ce propos, il convient de souligner que la conception et la faisabilité de l'EI dépendent étroitement de la conception, du calendrier et de la portée de l'intervention, du budget et des capacités techniques disponibles, et de l'évolution du dialogue politique à l'étape préparatoire. Les équipes doivent garder à l'esprit que les questions de recherche et la conception de l'EI sont interdépendantes, et qu'elles peuvent toutes deux évoluer (et le feront probablement) lors de l'étape préparatoire dans la foulée du dialogue politique et des évaluations ultérieures de la faisabilité de l'EI. Le présent module vise à aider les équipes d'EI à élaborer une EI prospective randomisée rigoureuse. Cependant, les équipes doivent faire preuve de pragmatisme et ajuster leur concept de manière à faire en sorte que l'EI reste faisable, rigoureuse, et instructive tout en s'adaptant au plan et à l'évolution de l'intervention.

Chaque section du présent module correspond à une section du document de conception de l'évaluation d'impact. Nous recommandons ce qu'elles devraient selon nous contenir.

¹⁰ Pour en savoir plus sur les éléments de conception technique (par exemple, la stratégie d'identification), prière de consulter *L'évaluation d'impact en pratique* (Gertler *et al.* 2011).

Aperçu du document de conception de l'EI

- Contexte et justification
- Cadre des résultats
- Questions de recherche et pertinence stratégique
- Indicateurs des résultats et des réalisations à l'étude
- Stratégie d'identification
- Échantillon
- Besoins de données
- Calendrier
- Équipe d'EI
- Plan de diffusion
- Budget

Contexte et justification

Cette section devrait répondre aux questions suivantes :

- Quels sont les principaux obstacles/enjeux nationaux dans la poursuite des OMD relatifs à la santé ?
- Qu'est-ce qui permet de penser que le FBR pourrait être utile, compte tenu du contexte national, pour accélérer les progrès dans la réalisation des OMD relatifs à la santé ?

L'examen du contexte devrait en outre se pencher brièvement sur toute information actuelle concernant le programme pilote national de FBR :

- Contexte : contextes propres au pays ou à d'autres pays semblables (Afrique subsaharienne, Amérique latine, etc.)
- Conception : bénéficiaires, indicateurs, niveaux d'incitation, etc.

Cadre des résultats du FBR dans le secteur de la santé

Cette section devrait décrire le cadre des intrants, des activités, des résultats et des réalisations intermédiaires qui conduira à l'atteinte des réalisations souhaitées du programme.

Questions de recherche et pertinence stratégique

L'EI devrait idéalement avoir pour but de produire des informations sur la façon d'utiliser les programmes de FBR pour accélérer les progrès vers la réalisation des OMD, non seulement dans le pays

où s'effectue cette évaluation, mais également à l'échelle mondiale. Cette section devrait répondre à deux questions ayant trait au cadre global du FBR :

- Que souhaitons-nous apprendre au sujet du FBR, dans le contexte où se déroule l'intervention ?
- Comment cela contribue-t-il à combler les lacunes globales des connaissances sur le FBR ?

Un projet de FBR peut compter plusieurs composantes — par exemple, évolution des incitatifs, augmentation du financement, resserrement du contrôle et du suivi — et l'équipe pourrait ne choisir d'évaluer qu'un sous-ensemble de ces composantes. Si tel est le cas, cette section devrait inclure également un bref résumé du projet de FBR et de ses composantes, de celles de ces composantes qui seront évaluées et du raisonnement qui a conduit au choix de ces dernières.

2012/02/21

Études de cas nationales : Définition des questions de recherche de l'évaluation d'impact en fonction des priorités stratégiques

Programme de financement basé sur les résultats dans les hôpitaux du Kirghizistan

Pendant les consultations menées auprès du ministère de la Santé du Kirghizistan, il a été convenu que la mesure de l'impact d'une augmentation du financement ne constituait pas en soi une question stratégique principale pour le ministère. C'est la raison pour laquelle l'EI vise à vérifier l'efficacité et la rentabilité du FBR ainsi que d'une de ses composantes : le renforcement du contrôle de la qualité des soins.

Les principales questions de recherche qui déterminent la méthode d'EI sont les suivantes :

- Le financement basé sur la performance (FBP) (y compris celui visant à renforcer le contrôle) dans les hôpitaux régionaux permet-il d'améliorer la qualité des soins ?
- Le renforcement du contrôle permet-il à *lui seul* d'améliorer la qualité des soins dans les hôpitaux régionaux ?
- Comment l'efficacité relative par rapport aux coûts du FBP (y compris celui visant à renforcer le contrôle) se compare-t-elle à celle du renforcement du contrôle utilisé seul ou au statu quo, en termes de qualité quantifiable des indicateurs de santé ?

Programme de financement basé sur les résultats au Zimbabwe

Le Zimbabwe souffre d'un grave handicap et fait face à de sérieux défis en ce qui a trait à la gestion des ressources humaines en santé — formation, financement, supervision et maintien en fonction du personnel. L'équipe d'EI du Zimbabwe examinera les rapports qui existent entre le FBR et le renforcement des compétences et des capacités dans les établissements de santé.

- Quel est l'effet causal de la décision d'adopter le FBR en même temps que l'on supprime la tarification des usagers sur l'utilisation des services de santé par la population prioritaire et les réalisations mesurées dans les districts visés par le FBR ?
- Quels sont les effets du renforcement des compétences et des capacités des infirmières de soins primaires sur les réalisations des services de santé prioritaires, l'utilisation de ces services et la qualité des soins dont bénéficient les patients, ainsi que sur la motivation des agents de santé dans les établissements de santé ruraux ?
- Quels sont les effets combinés du renforcement des capacités des infirmières de soins primaires, du FBR, et de la suppression de la tarification des usagers sur les réalisations précitées dans les établissements de santé ruraux ?

Programme de financement basé sur les résultats au Cameroun

On s'intéresse en particulier aux effets du FBR et des améliorations qu'il entraîne en matière de surveillance, de contrôle et d'évaluation sur la qualité des soins et l'état de santé de la population :

- Le programme de FBP augmente-t-il la couverture des services de SMI ?
- Le programme de FBP améliore-t-il la qualité des services de SMI offerts ?
- Les améliorations observées de la qualité ou de la couverture sont-elles le fait d'un renforcement du suivi, de l'évaluation et de la surveillance, ou du lien qui existe entre les paiements et les résultats ?
- Dans quelle mesure le renforcement de la surveillance et du suivi contribue-t-il à l'amélioration de la couverture et de la qualité des services de SMI en l'absence d'une amélioration de l'autonomie ou de ressources financières supplémentaires ?
- Le programme de FBP entraîne-t-il une baisse des coûts informels ou formels des services de santé ?
- Le programme de FBP augmente-t-il le volume des fonds disponibles au niveau opérationnel (c'est-à-dire, des établissements) ?
- Le programme de FBP améliore-t-il l'accessibilité physique et sociale des services de santé ? Cette accessibilité sera mesurée à l'aune des heures d'ouverture des établissements, de la disponibilité des services par le biais de la communication, de l'accès aux services de santé tel qu'il est perçu par la clientèle et de l'attitude des prestataires de soins de santé vis-à-vis de leurs patients, telle qu'elle est perçue par ces derniers.
- Le programme de FBP réduit-il le taux d'absentéisme des employés ?
- Le programme de FBP favorise-t-il les activités génératrices de demande des établissements de santé ?

Indicateurs des résultats et des réalisations à l'étude

Les questions de recherche permettent logiquement d'assurer la mesure des résultats ou des réalisations à l'étude. Ces indicateurs, qui se distinguent des indicateurs des paiements du FBR, visent à mesurer l'impact de l'intervention de FBR par rapport aux questions de recherche choisies. Ils doivent être clairement définis dès l'étape de la conception pour les raisons suivantes :

- la définition des indicateurs permet à l'équipe de veiller à ce que les questions de recherche soient effectivement mesurables ;
- elle permet à l'équipe de déterminer dès le départ quels sont les instruments qui permettront le mieux de mesurer ces indicateurs.

Nous encourageons les équipes à adopter les définitions internationales de ces indicateurs, lorsqu'elles existent, quitte à envisager le recours à des définitions nationales pour renforcer la pertinence au contexte national. Nous présentons ci-dessous une liste non exhaustive de références qui peuvent servir à la définition des indicateurs. Rappelons toutefois que certains de ces outils ne visent pas à évaluer l'impact du FBR, mais peuvent couvrir un éventail plus large d'enjeux liés à la prestation et à l'utilisation des services de santé.

- **Liste d'indicateurs** pour l'évaluation de la santé maternelle et infantile de l'Organisation mondiale de la santé (OMS), publiée annuellement.
- Protocole de l'OMS sur la prise en charge intégrée des maladies de l'enfant : pour en savoir plus, consulter le site Web de l'OMS à l'adresse suivante : http://www.who.int/maternal_child_adolescent/topics/child/imci/fr/index.html.
- Autres protocoles et lignes directrices de l'OMS.
- Lignes directrices et outils portant sur la prestation des services — par exemple, évaluation de la disponibilité des services et de l'état de préparation (*Service Availability and Readiness Assessment – SARA*) de l'OMS ; carte de la disponibilité des services (*Service Availability Mapping – SAM*) de l'OMS ; évaluation de la prestation des services (*Service Provision Assessment – SPA*) de Mesure DHS ; Programme intégré pour la santé maternelle et infantile (*Maternal and Child Health Integrated Program – MCHIP*) ; trousse de formation MamaNatalie et NeoNatalie.
- Protocoles nationaux sur les indicateurs nationaux pertinents.

La boîte à outils fournit une liste d'**indicateurs de FBR** proposés, explique la manière de les calculer et indique les instruments à utiliser à cette fin.

Stratégie d'identification

Pour assurer le succès de l'évaluation d'impact, chaque équipe devra élaborer une stratégie d'évaluation qui lui permettra d'identifier l'effet causal de l'intervention. Pour être efficace, cette stratégie devra inclure les groupes de traitement et de comparaison, ainsi que la collecte de données de référence et de données postérieures à l'intervention sur les groupes de traitement et de comparaison.

Dans la section du document de conception de l'évaluation d'impact consacrée à la « stratégie d'identification », les questions suivantes devraient être abordées :

Quels sont les éléments du programme de FBR qui devront être évalués ?

Nous recommandons que chacun des éléments à évaluer fasse l'objet d'une « branche de traitement » distincte. Des branches de traitement supplémentaires pourront servir à évaluer les interactions de diverses composantes. Il conviendra ainsi de préciser *combien de branches de traitement (ou de sous-traitement) devront être incluses dans l'évaluation d'impact*. Même lorsque les programmes englobent plusieurs interventions « complémentaires », il pourrait s'avérer utile de chercher à distinguer les effets séparés des différentes interventions afin de supprimer les composantes inutiles dont la mise en œuvre pourrait s'avérer coûteuse.

Exemple 1 : Un programme de financement basé sur la performance (FBP) peut inclure des paiements basés sur les résultats et la formation des agents de santé. On peut envisager d'évaluer séparément l'impact de chacune des composantes du programme (une branche de traitement portant sur les paiements basés sur les résultats et une autre portant sur la formation des agents de

santé), mais si on souhaite mesurer l'impact combiné des deux composantes, il conviendra de recourir à une troisième branche de traitement (paiements basés sur les résultats + formation des agents de santé) afin de pouvoir comparer l'effet de l'ensemble à celui des composantes individuelles.

Exemple 2 : Les autorités publiques qui mettent en place un programme d'incitation axé sur la demande souhaiteront peut-être savoir si les incitatifs en espèces donnent des résultats différents de ceux des incitatifs en nature. Deux branches de traitement permettront de comparer l'impact d'incitatifs en espèces à celui d'incitatifs en nature de même valeur.

2012/04/03

Étude de cas nationale : Branches de l'étude

Programme communautaire rwandais de financement basé sur la performance

Le programme rwandais de financement basé sur la performance compte deux types d'interventions :

- **Modèle 1 :** mesures d'incitation conditionnelles, en nature, versées directement aux femmes dans les situations suivantes :
 - **Soins prénatals :** femmes enceintes accompagnées au centre de santé ou orientées vers ce dernier avant ou pendant le 4^e mois de leur grossesse : valeur de 5 USD
 - **Accouchement en maternité :** femmes qui accouchent dans un établissement de santé : valeur de 7 USD
 - **Suivi postnatal :** mères et enfants recevant des soins postnatals dans un établissement de santé dans les dix jours suivant la naissance/le congé : valeur de 4 USD.
- **Modèle 2 :** mesures d'incitation, en espèces, versées aux coopératives d'agents de santé communautaire en fonction de leurs résultats mesurés à l'aide des indicateurs suivants :
 - **suivi nutritionnel :** pourcentage des enfants de 6 à 59 mois faisant l'objet d'un suivi (aiguillage approprié au centre de santé des enfants souffrant de malnutrition) ;
 - **soins prénatals :** pourcentage des femmes accompagnées au centre de santé ou orientées vers ce dernier avant ou pendant le 4^e mois de leur grossesse ;
 - **accouchements en maternité :** pourcentage des femmes accouchant dans un centre de santé ;
 - **planning familial :** pourcentage des nouveaux usagers du service orientés par des agents de santé communautaire au centre de santé ;
 - **planning familial :** pourcentage des utilisateurs réguliers de méthodes modernes de contraception au centre de santé.

Pour l'évaluation d'impact du programme communautaire rwandais de FBP, les autorités souhaitaient obtenir des informations sur l'impact du modèle d'incitation axé sur la demande et du modèle axé sur les agents de santé communautaire examinés séparément, ainsi que sur l'impact additif des deux modèles combinés.

Comment l'équipe estimera-t-elle le contrefactuel ?

Dans toute évaluation d'impact, un des objectifs importants de la détermination de l'effet de l'intervention consiste à estimer le contrefactuel, c'est-à-dire à déterminer ce qu'aurait été le résultat pour le groupe examiné en l'absence de tout traitement/intervention. Comme le contrefactuel n'est pas observable (il ne peut y avoir de groupe traité sans traitement), il faut trouver un groupe de comparaison qui permettra d'estimer ce qui serait arrivé au groupe de traitement en l'absence de traitement. La stratégie exacte de sélection du groupe de comparaison dépendra des règles de fonctionnement de l'intervention. Dans le cadre de ces règles, les membres du groupe de comparaison doivent être choisis pour conduire à une estimation exacte du contrefactuel — c'est-à-dire ce qui serait arrivé aux groupe de traitement en l'absence de programme. Le groupe de comparaison doit satisfaire à deux exigences : 1) les caractéristiques observées et non observées des groupes de traitement et de comparaison doivent être identiques en moyenne ; 2) les groupes de traitement et de comparaison doivent être capables de réagir de la même façon au traitement et être soumis aux mêmes chocs extérieurs au fil du temps. Lorsque ces conditions sont respectées, toute différence observée des résultats moyens mesurés dans chacun des deux groupes à la suite de la mise en œuvre du programme pourra être attribuée à l'intervention.

La répartition des unités entre le groupe de traitement et le groupe de comparaison devrait *dans la mesure du possible* être aléatoire, c'est-à-dire être réalisée à l'aide d'un système de tirage au sort. Il existe également d'autres stratégies d'identification comme la promotion aléatoire, le modèle de discontinuité de la régression et la double différence. D'autres stratégies ne sont pas considérées propices pour l'estimation du contrefactuel — par exemple, les comparaisons « avant et après » et la comparaison des unités participantes et non participantes. Prière de consulter *L'évaluation d'impact en pratique* (Gertler *et al.* 2011) pour un exposé détaillé des stratégies d'identification appropriées pour l'EI. La sélection aléatoire des sujets d'un programme n'est pas une panacée, mais elle constitue dans la plupart des cas la méthode la plus directe et la moins coûteuse en ce qui a trait aux besoins de données.

Choix de l'unité de randomisation

L'unité de randomisation est l'unité de base de l'assignation entre le groupe de traitement et le groupe de comparaison ; toutes les entités qui la composent partagent la même assignation. Par exemple, si un pays est subdivisé en provinces, districts et quartiers, et que l'équipe décide de déterminer l'assignation aléatoire à l'échelle du district, tous les quartiers, établissements de santé et agents de santé qui appartiennent à ce district partageront la même assignation au groupe de traitement ou au groupe de comparaison.

Dans les essais cliniques traditionnels comme ceux portant sur les médicaments, l'unité de randomisation équivaut souvent à l'unité d'analyse : les patients sont sélectionnés pour faire partie du groupe de traitement ou du groupe de comparaison (le patient fait donc ici office d'unité de randomisation). De plus, l'analyse est réalisée au niveau du patient (elle compare les patients du groupe de traitement à ceux du groupe de comparaison). Par contre, dans les évaluations des programmes de FBR, l'unité d'analyse est souvent intégrée dans l'unité de randomisation. Par exemple, on peut décider

d'analyser les données au niveau de l'agent de santé. Dans un tel cas, les agents de santé qui travaillent dans un établissement se verront attribuer le statut de leur établissement. Les patients et les ménages peuvent également être assimilés à des unités d'analyse.

2012/04/03

Étude de cas nationale : Unité de randomisation

Programme communautaire rwandais de financement basé sur la performance

L'unité de randomisation de cette évaluation était le secteur de la santé : une unité géographique de rang inférieur au district. Tous les établissements de santé, coopératives d'agents de santé communautaire et ménages appartenant à un secteur de la santé donné se voyaient assignés au même groupe (traitement ou comparaison) que celui du secteur de la santé auquel ils appartenaient. Les secteurs étaient répartis entre quatre branches d'études : i) mesures d'incitation axées sur la demande ; ii) mesures d'incitation visant les agents de santé communautaire ; iii) combinaison des deux types de mesures d'incitation précitées ; iv) groupe de comparaison.

Le choix des unités d'analyse et de randomisation influe grandement sur la puissance statistique et la validité de l'évaluation. Supposons par exemple que quatre districts participent au programme, et que ces districts sont choisis au hasard dans un groupe de huit districts. Même si le choix des districts du « groupe de traitement » est aléatoire, il reste très peu probable que les caractéristiques des districts des groupes de traitement et de comparaison seront équilibrées, puisque le nombre de districts examinés est simplement trop limité pour permettre l'établissement d'une moyenne représentative. Il est intéressant de noter que l'absence de comparabilité entre les districts de traitement et les districts de comparaison se manifesterait non seulement pour les indicateurs du niveau « district », mais également pour ceux du niveau « établissement de santé », même si chaque district compte plusieurs établissements de santé.

Lorsque l'évaluation ne peut inclure qu'un nombre limité de districts ou zones administratives, la puissance statistique sera habituellement insuffisante pour permettre une évaluation d'impact solide. Dans de tels cas, il peut être utile de vérifier la possibilité de réaliser l'assignation aléatoire à un niveau inférieur — par exemple, celui des établissements de santé. Il sera ainsi plus facile de comparer les caractéristiques des groupes de traitement et de comparaison. Toutefois, le passage à un niveau d'assignation inférieur risque de s'avérer difficile au plan opérationnel, et de provoquer un effet de contamination entre les unités traitées et non traitées. Par exemple, si les établissements de santé d'un district donné sont répartis au hasard entre le groupe de traitement et le groupe de comparaison, les administrateurs du district risquent d'avoir du mal à traiter d'une manière distincte tous les établissements de leur district. De plus, la randomisation effectuée à un niveau si bas risque de s'avérer politiquement difficile. Par exemple, les agents de santé pourraient se sentir lésés lorsqu'ils constateront les différences entre leurs établissements respectifs. Cette communication entre les agents de santé peut ainsi conduire à invalider l'hypothèse de l'indépendance des unités de traitement et de comparaison. Elle risque par ailleurs d'entraîner un effet de contamination entre les deux groupes.

Étude de cas nationale : Traitement et prévention des problèmes de conformité Financement basé sur les résultats dans le secteur de la santé en Zambie

La conformité de la mise en œuvre d'un programme au modèle d'évaluation est un enjeu crucial pour toutes les évaluations. Des problèmes ont été relevés pendant l'évaluation du programme de lutte contre le paludisme [en Zambie]. Pourtant, cela n'a pas permis d'éviter les problèmes de conformité pendant la mise en œuvre du programme de FBR. L'équipe d'EI a découvert avant l'enquête de référence qu'un des districts du groupe de comparaison avait été contaminé. Pour corriger ce problème, elle a décidé d'inclure un district supplémentaire dans chacun des trois groupes de traitement restants, ce qui a eu pour effet d'augmenter le budget d'EI. Bien que les programmes de grande envergure ne puissent être réalisés dans les conditions parfaitement contrôlées du laboratoire, certains facteurs qui conduisent souvent à une contamination pourraient être corrigés relativement facilement — par exemple, à l'étape de la coordination, de la communication ou de l'établissement des rapports.

Pour en savoir plus, consulter le dossier « Pleins feux sur le pays » de la boîte à outils.

Enfin, l'assignation aléatoire au niveau des établissements de santé pose une autre difficulté : les hôpitaux sont des institutions uniques qui ne se comparent pas facilement aux établissements de rang inférieur ; en règle générale, ils sont exclus de l'évaluation pour cette raison.

Le choix du niveau de randomisation devrait refléter la nature du programme soumis à l'évaluation ; ce choix appartient en fin de compte à l'équipe d'EI du pays, et doit être bien justifié et expliqué dans le document de conception.

Échantillon

Le document de conception doit répondre aux questions suivantes :

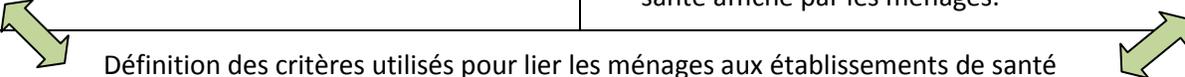
Quels sont les critères d'inclusion du cadre d'échantillonnage ?

Le cadre d'échantillonnage est l'ensemble d'unités à partir duquel l'échantillon sera prélevé. Il est important, surtout dans le cas d'un programme de FBR pilote, de veiller à ce que ce cadre soit représentatif de la population qui bénéficiera à terme de l'intervention et pour laquelle l'impact souhaité de l'intervention doit être mesuré. Il s'agit de veiller à ce que les résultats s'appliquent à la population à l'étude, c'est-à-dire d'assurer la « validité externe » de l'évaluation d'impact.

Dans la pratique, le cadre d'échantillonnage est une liste physique ou électronique d'unités qui peuvent se prêter à l'échantillonnage. Bien qu'il ne soit pas possible de dresser une telle liste à l'étape de la conception, il importe dès cette étape de définir les critères d'inclusion du cadre d'échantillonnage. Les critères d'inclusion jouent un rôle essentiel puisque pour des échantillons de même taille, ils peuvent conduire à de grandes variations de la puissance de l'étude. À l'étape de la conception, l'équipe d'EI devra définir les critères d'inclusion des cadres d'échantillonnage à de multiples niveaux.

- Par exemple, si l'assignation aléatoire des unités entre les groupes de traitement et de comparaison s'effectue au niveau du district, il conviendra d'établir d'abord la liste des districts.

- L'étape suivante consistera à déterminer le second niveau d'échantillonnage (par exemple, s'agira-t-il d'établissements de santé ou de ménages ? On peut par exemple échantillonner d'abord les ménages, pour ensuite déterminer les établissements de santé qu'ils utilisent ou qui se trouvent dans l'aire de dénombrement des ménages sondés. Inversement, on peut choisir d'abord les établissements de santé, puis échantillonner les ménages qui se trouvent dans leur zone d'intervention. Chaque méthode présente des avantages et des inconvénients, ainsi que des enjeux qui lui sont propres (voir le tableau ci-dessous). Il appartient à l'équipe d'enquête de les évaluer.

Établissements de santé échantillonnés d'abord	Ménages échantillonnés d'abord
<ul style="list-style-type: none"> • Il est plus facile d'établir la liste des établissements de santé. Une telle liste peut être obtenue auprès du ministère de la Santé ou de l'Institut national de la statistique. • Les ménages éloignés risquent d'être exclus de la zone d'intervention des établissements, ce qui pourrait influencer sur la représentativité des données. 	<ul style="list-style-type: none"> • L'établissement de la liste des ménages est plus coûteux et prend plus de temps. • Certains établissements de santé risquent d'être exclus si l'appariement des ménages et des établissements est fondé sur le comportement de recherche de services de santé, et si certains établissements ne sont pas fréquentés par les ménages échantillonnés. • L'échantillon risque d'inclure des établissements se trouvant hors de l'unité de randomisation si des ménages les fréquentent et si l'appariement des ménages et des établissements est fondé sur le comportement de recherche de services de santé affiché par les ménages.
 <p>Définition des critères utilisés pour lier les ménages aux établissements de santé</p>	
<ul style="list-style-type: none"> • Quelle est la zone d'intervention de l'établissement de santé ? Est-elle perçue de la même manière par l'établissement et par les ménages ? • Zone d'intervention réelle ou théorique : prend-on en compte le comportement réel de recherche de services de santé ou le comportement théorique ? • Combien y a-t-il de ménages par établissement, et combien y a-t-il d'établissements par ménage ? 	

2012/02/05

Étude de cas nationale : Cadre d'échantillonnage : par où commencer ?
Projet de financement basé sur les résultats au Bénin

La définition du cadre d'échantillonnage a soulevé quelques questions et donné lieu à des débats au sein de l'équipe [au Bénin]. L'équipe bénéficiait du soutien politique du ministère de la Santé, mais elle n'a pas obtenu l'autorisation de l'Institut national de la statistique pour utiliser les données de recensement. Il lui a donc fallu partir de zéro, et recenser elle-même les ménages admissibles à l'aide d'un travail préparatoire sur le terrain. Cette situation a conduit à une hausse importante des coûts de l'étude qui n'avait pas été prévue dans les termes de référence du bureau d'études ni dans le budget original d'EI. L'enquête de référence sur le terrain a également été retardée.

L'équipe a par ailleurs mis deux mois à débattre du lien entre les ménages et les établissements de santé : devait-elle procéder d'abord à une enquête auprès des ménages pour connaître qu'elles étaient les établissements de santé qu'ils fréquentaient le plus, ou se pencher d'abord sur les établissements de santé pour définir la zone d'intervention de chacun et localiser les ménages « attachés » à chacun d'eux ? Comme tous les établissements de santé ont été inclus dans l'enquête d'EI, il n'a pas été nécessaire d'échantillonner. L'équipe a opté pour la seconde solution, et défini des zones d'intervention d'un rayon de 5 kilomètres dans le Sud, et de 15 kilomètres dans le Nord du pays. Pour chaque établissement, dix ménages ont été retenus à partir d'un relevé initial des unités d'échantillonnage qui a permis d'établir le nombre de ménages compris dans chacune d'elles.

Pour en savoir plus, consulter la section « Pleins feux sur le pays » de la boîte à outils.

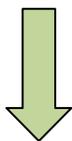
- Dès que l'échantillon d'unités de traitement et de comparaison est établi, l'équipe peut procéder au choix des critères d'inclusion pour les centres de santé, les agents de santé, les ménages, les mères, les enfants, etc., par ordre décroissant.

Exemple :

Critères de sélection pour le cadre d'échantillonnage des districts — tous les districts du pays.
Échantillon : 200 districts, choisis au hasard à partir du cadre d'échantillonnage.



Critères de sélection pour le cadre d'échantillonnage des établissements de santé — tous les établissements de santé, publics ou privés, situés dans un des districts de l'échantillon.



Critères de sélection pour les agents de santé — agents travaillant dans un des établissements de santé de l'échantillon et responsables des services de santé maternelle et infantile.

Critères de sélection pour les ménages : On pose comme hypothèse que le FBR améliorera la qualité des soins prénatals, et que cela conduira à une hausse du poids des bébés à la naissance. Dans ce cas, le poids à la naissance devient le « résultat à l'étude ». Les enquêtes devront donc servir à obtenir le poids à la naissance d'un échantillon suffisamment grand de bébés sur lequel, selon notre hypothèse, le FBR aura un impact. Supposons que le FBR démarre en janvier 2011 et que la fin de la période de collecte des données est fixée au mois d'août 2012, c'est-à-dire 20 mois plus tard. Pour être exposée aux répercussions du FBR pendant toute sa durée, une grossesse devra avoir débuté en janvier 2011 ou plus tard, et l'échantillon sera constitué de grossesses achevées dix mois ou moins avant la fin de l'enquête. Il convient donc de maximiser le nombre d'observations portant sur les grossesses qui se sont achevées il y a dix mois ou moins. Pour permettre une comparaison adéquate entre les groupes de traitement et de comparaison tant à la ligne de base qu'à la fin de l'étude, le critère d'inclusion des ménages pour les deux enquêtes sera donc le suivant : « Ménages ayant connu au moins une grossesse achevée il y a dix mois ou moins, et vivant dans la zone d'intervention d'un établissement de

santé inclus dans l'échantillon. » Ce critère pourrait s'avérer difficile à appliquer en pratique puisque les grossesses achevées ne sont plus observables. Un critère de rechange pourrait être formulé comme suit : « Ménages comptant au moins un enfant âgé de 0 à 10 mois. » Il convient cependant de mentionner qu'un tel critère exclura les ménages qui ont connu une grossesse achevée il y a dix mois ou moins, mais qui s'est soldée par le décès de l'enfant. Si le programme influe sur le taux de survie des enfants, le fait de compter sur la présence d'un enfant toujours vivant risquera de biaiser l'échantillon.

2012/04/10

Étude de cas nationale : Établissement du cadre d'échantillonnage Projet d'investissement en appui au secteur de la santé au Cameroun

L'étude en question est un essai randomisé par grappes de régions, avec comparaison avant-après. Les centres de santé de chacune des régions à l'étude sont répartis au hasard entre quatre groupes d'études : (T1) programme complet de financement basé sur la performance (FBP) ; (C1) ressources financières par habitant identiques à celles du FBP, mais non liées à la performance ; même surveillance et suivi et même autonomie de gestion que T1 ; (C2) pas de ressources supplémentaires, mais même surveillance et suivi que les branches du FBP, et T1 et C1 ; (C3) statu quo. Au total, environ 270 établissements de soins primaires sont inclus dans l'évaluation d'impact.

Les personnes qui vivent dans la zone d'intervention d'un établissement assigné à un groupe d'étude particulier (par exemple, C3) peuvent se rendre dans un établissement appartenant à un groupe différent (par exemple, T1). Pour minimiser la contamination entre les branches de l'étude, on a décidé d'assimiler les zones d'intervention des établissements à un cadre d'échantillonnage aux fins de la sélection aléatoire des ménages pour la composante de l'EI ayant trait à l'enquête auprès des ménages. Cependant, au Cameroun, l'unité de santé officielle la plus décentralisée au plan géographique est la « zone sanitaire » (« Health Area »), qui peut inclure un ou plusieurs centres de santé primaire, ou en être entièrement dépourvue. Ainsi, il n'existe pas de zone d'intervention officielle au niveau des établissements. Pire encore, les estimations sur les ménages pour les zones de dénombrement produites par le Bureau national du recensement ne sont ni actuelles, ni fiables.

Pour faire face à cette situation, les enquêteurs se sont rendus dans l'ensemble des 14 districts pilotes pour recueillir des informations ayant trait aux établissements de santé de chacune des zones sanitaires et de chacun des villages/secteurs servis par ces établissements. Les données recueillies incluaient la taille de la population de chaque localité, et la distance qui séparait ces dernières de l'établissement. Le bureau d'études chargé de l'enquête de référence a utilisé le cadre d'échantillonnage établi à partir de ce travail pour procéder à la première étape de l'échantillonnage des localités à l'intérieur des zones d'intervention de chaque centre de santé. Sur le terrain, les équipes d'enquête ont établi des listes de ménages de chacune des localités échantillonnées, où 16 ménages qui répondaient aux critères d'inclusion (au moins une grossesse ou une naissance au cours des deux années précédant l'enquête) ont été choisis au hasard.

Pour en savoir plus, consulter la section « Pleins feux sur le pays » de la boîte à outils.

Quelle est la taille requise de l'échantillon ?

L'échantillon est composé des unités que le bureau d'études cherchera à interviewer/sonder. Il est prélevé à partir du cadre d'échantillonnage. À l'étape de la conception, il n'est pas nécessaire de

disposer d'une liste des unités qui seront échantillonnées, mais il faut déjà savoir combien d'unités seront incluses dans l'échantillon, c'est-à-dire, la taille de l'échantillon.

L'équipe d'EI doit estimer la taille de l'échantillon dont elle aura besoin pour détecter un impact minimal acceptable sur les indicateurs à l'étude. Cette estimation est fondée sur ce qu'on appelle le « calcul de puissance statistique ». La puissance statistique correspond à la probabilité de détecter un effet s'il se produit. Comme il s'agit d'une probabilité, elle doit se situer entre 0 et 1, et nous souhaitons que l'étude présente la plus grande puissance statistique possible (c'est-à-dire, proche de 1, soit 80 à 90 %). Les études dont la puissance statistique est insuffisante risquent de ne pas pouvoir détecter l'effet d'un traitement, même s'il existe.

Le degré de puissance statistique d'une évaluation dépend de plusieurs paramètres, y compris la taille de l'échantillon disponible à chacun des niveaux de l'analyse. Pour les études dans lesquelles l'unité d'analyse correspond à l'unité de randomisation, il existe un rapport passablement direct entre le nombre total d'unités et le degré de puissance statistique : un échantillon plus grand donne une puissance statistique plus élevée. Cependant, dans le cas des études dans lesquelles l'unité d'analyse (par exemple, le ménage) est imbriquée, ou regroupée dans l'unité de randomisation (par exemple, les établissements de santé), ce qui arrive souvent avec les études d'évaluation des programmes de financement de la santé basés sur les résultats, le rapport entre la taille de l'échantillon et la puissance statistique est un peu plus embrouillé. On attachera en règle générale une importance plus grande au nombre d'unités de randomisation et au nombre d'unités de regroupement qu'au nombre d'unités d'analyse qu'elles renferment.

- Par exemple, une étude portant sur 100 établissements de santé comptant chacun dix patients présentera une puissance statistique plus grande qu'une étude portant sur dix établissements de santé comptant chacun 100 patients, même si le nombre total de ménages visés dans les deux cas s'établit à 1 000 (et en présumant que l'unité d'analyse est plus ou moins imbriquée dans l'unité de randomisation). Ainsi, cette imbrication (ou ce regroupement) des ménages dans les établissements de santé doit être prise en compte aux fins des calculs de puissance.
- De la même façon, si une évaluation d'impact est randomisée par exemple au niveau du district, et si les établissements de santé sont regroupés à l'intérieur de chaque district, une augmentation du nombre d'établissements de santé inclus dans l'échantillon ne produira pas une augmentation aussi prononcée de la puissance qu'une augmentation du nombre de districts. Dans les pays qui comptent un nombre limité de districts, il pourra donc être préférable, pour cette raison, d'opter pour une randomisation au niveau des établissements de santé, même si cela risque de poser d'autres problèmes (par exemple, la contamination).
- S'agissant des études avec regroupement en multiples étapes (par exemple, évaluation d'impact randomisée au niveau du district, avec établissements de santé regroupés dans chaque district et ménages regroupés dans la zone d'intervention de chaque établissement), l'augmentation du nombre d'unités de randomisation (par exemple, les districts) ou du nombre d'unités de regroupement (par exemple, les établissements de santé) produira une augmentation plus

prononcée de la puissance statistique qu'une augmentation du nombre d'unités d'analyse (par exemple, les ménages).

En résumé, la puissance augmentera plus rapidement avec le nombre de groupes qu'avec le nombre d'unités comprises dans chaque groupe. Il importe de noter que la proportion de la variabilité du revenu total observable entre les groupes, appelée « corrélation intra-classe » ou « corrélation intra-grappe » (CIG) dans le cas des résultats continus, évolue en sens inverse de la puissance statistique : une valeur de la CIG plus grande signifie moins de puissance statistique. Il est essentiel d'obtenir une estimation de la CIG pour pouvoir procéder aux calculs de puissance des évaluations d'impact randomisées en grappes.

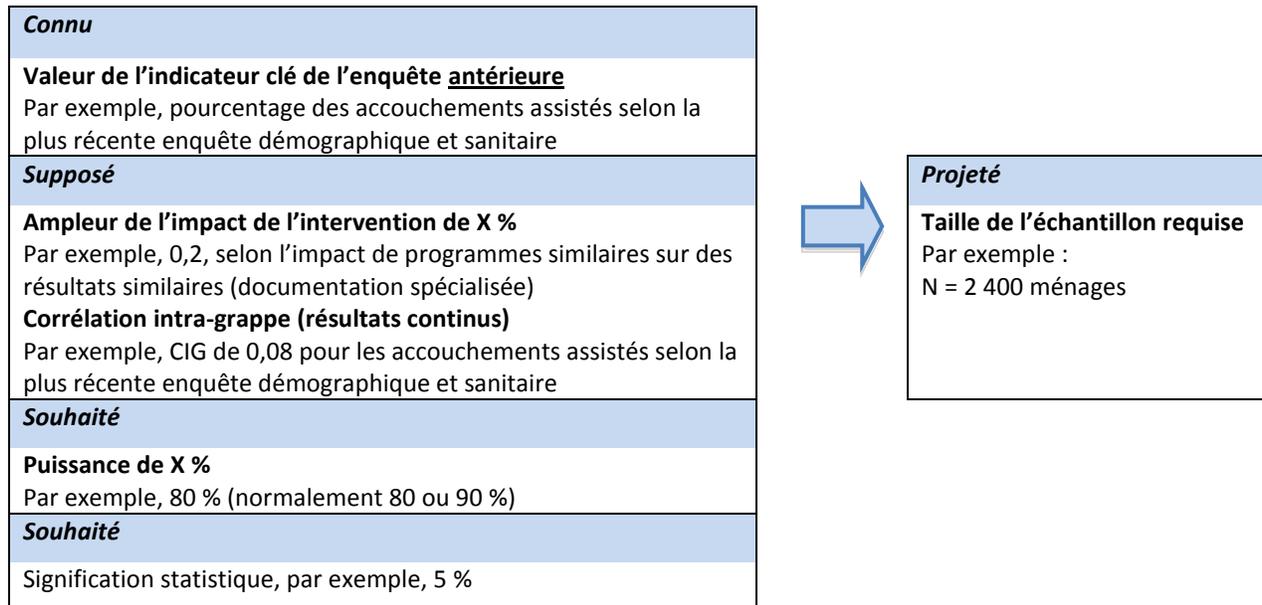
La puissance statistique d'une évaluation dépendra également de l'importance attendue de l'effet de l'intervention sur les variables de résultats à l'étude. Les effets plus petits sont naturellement plus difficiles à détecter, et les études qui visent à les détecter auront, toutes choses étant égales par ailleurs, une puissance statistique moindre que les études qui visent à détecter des effets de traitement plus importants. L'équipe d'évaluation est souvent impuissante face à l'effet de l'ampleur de l'intervention, mais cet effet doit être estimé sur la base de la connaissance préalable de la nature de l'intervention et du contexte dans lequel elle se déroule.

D'autres paramètres et options de plan d'étude jouent un rôle dans les calculs de puissance, notamment les valeurs de référence (par exemple, la moyenne et l'écart-type dans le cas des résultats continus) des indicateurs de résultats retenus, le nombre de branches de l'étude, l'état d'équilibre — ou de déséquilibre — du plan (par exemple, nombre égal d'unités dans chaque branche de traitement), la méthode d'échantillonnage aléatoire ou stratifiée (par exemple, 1 000 patients choisis au hasard ou dix patients choisis par établissement, dans 100 établissements), le choix de recourir à la randomisation par blocs (par exemple, regroupement des unités de randomisation par classes de pauvreté ou en fonction d'un autre critère afin de créer des branches plus facilement comparables) et le choix d'inclure les covariables dans les calculs de puissance.

À la lumière des études antérieures, nous recommandons que les équipes utilisent une puissance minimale de 0,9, et un niveau de signification de 0,05. Le document technique devrait également inclure les définitions des indicateurs et des sources de données utilisées pour procéder aux calculs de puissance, les estimations de la taille des échantillons à tous les niveaux d'observation (établissements de santé, agents de santé, ménages, patientes, etc.) et une description détaillée du plan d'étude adopté.

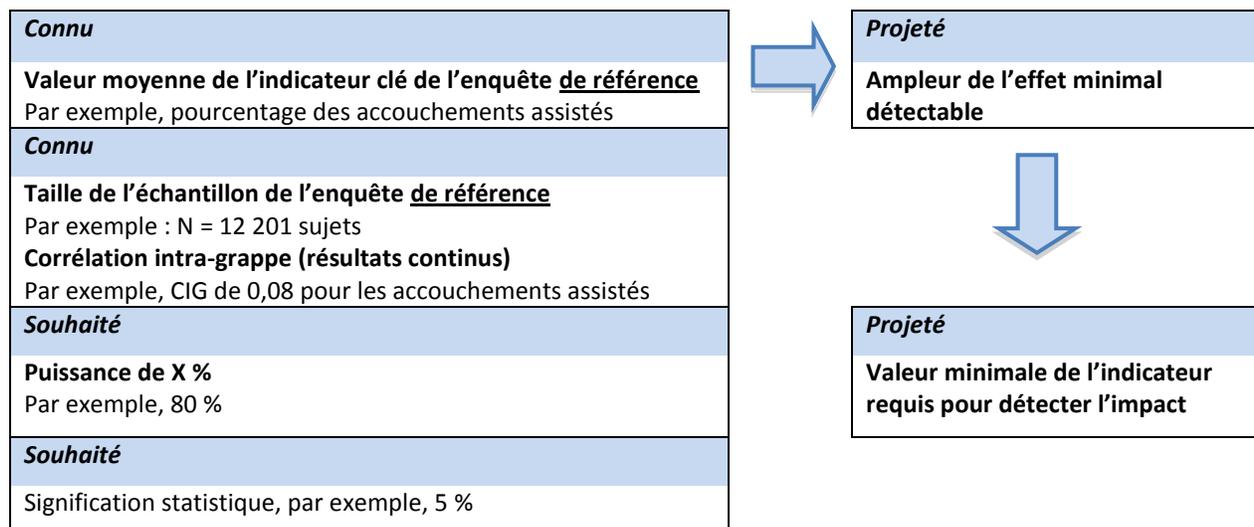
Il existe deux types différents de calculs de puissance, habituellement appelés « ex-ante » et « ex-post ». Les calculs de puissance ex-ante prennent en compte les valeurs des principales variables obtenues dans le cadre des études disponibles, l'ampleur de l'effet attendu de l'intervention, la puissance souhaitée de l'étude (80 ou 90 % en général) et le niveau de signification statistique souhaité (0,05 en général) pour calculer la taille requise de l'échantillon de l'étude de référence.

Figure 5 : Calculs de puissance ex-ante



On peut également utiliser les calculs de puissance « ex post » qui, malgré leur nom, n'ont pas besoin d'être effectués après l'étape de conception de l'étude. Pour certaines études, il est possible que certaines restrictions connues s'appliquent dès le départ quant à la taille des échantillons qui peuvent être utilisés. Il pourrait être instructif dans de tels cas d'aborder les calculs de puissance d'une manière qui permette de déterminer l'effet de l'intervention le plus petit qu'il sera possible de détecter statistiquement compte tenu de la taille de l'échantillon disponible. Dans le cas des études pour lesquelles la taille des échantillons pose de graves problèmes, cette approche aidera à décider de l'opportunité de procéder. Il est crucial d'éviter d'entamer des études dont on sait dès le départ que la puissance statistique sera faible.

Figure 6 : Calculs de puissance ex-post



2012/04/03

Étude de cas nationale : Branches de l'étude (suite)

Programme communautaire rwandais de financement basé sur la performance

Le Rwanda est divisé en quatre provinces, 30 districts et plus de 400 secteurs. En gardant à l'esprit que l'intervention se déroulerait au niveau des secteurs, les calculs de puissance ont démontré que l'évaluation d'impact nécessitait d'inclure 50 secteurs dans chacune des quatre branches de l'étude, soit un total de 200 secteurs. Les critères d'inclusion suivants ont servi à définir le cadre d'échantillonnage pour les secteurs de manière à : i) minimiser les décaissements sur le plan géographique en limitant le nombre de districts à 19 (sur un total possible de 30) ; ii) exclure 30 secteurs choisis pour bénéficier du traitement à la demande du ministère de la Santé ; iii) ne retenir que les secteurs qui comptaient au moins un centre de santé. Au total, 223 secteurs répartis dans 19 districts répondaient à ces critères. Comme l'étude n'en demandait que 200, l'équipe d'EI a utilisé le logiciel statistique STATA pour choisir au hasard les 50 secteurs à assigner à chacune des quatre branches de l'étude. Les 23 secteurs restants ont été mis en réserve.

Il convient de noter que la plupart des indicateurs ayant trait à la santé maternelle et infantile sont des résultats non continus dont la valeur est comprise entre zéro et un (par exemple, pourcentage des accouchements assistés par du personnel de santé qualifié). Les calculs de puissance ne sont habituellement pas effectués de la même manière pour les résultats continus et les résultats non continus.

La boîte à outils propose un exemple de **calcul de puissance** ex-ante ainsi qu'une **explication plus détaillée du calcul de puissance binaire** dans le cas de l'évaluation d'impact d'un programme FBR.

Nous recommandons que les équipes d'EI préparent un diagramme simple illustrant les différentes branches de l'étude en indiquant la taille proposée des échantillons à chacun des niveaux de l'intervention, comme dans l'exemple suivant :

Étude de cas nationale : Branches de l'étude (suite)
Programme communautaire rwandais de financement basé sur la performance

<p align="center">T1</p> <p align="center">Mesures d'incitation en nature seulement</p> <p align="center"><i>50 secteurs</i></p> <p align="center"><i>600 ménages</i></p> <p align="center"><i>1 200 agents de santé communautaire</i></p>	<p align="center">T2</p> <p align="center">Mesures d'incitation en espèces versées aux agents de santé communautaire</p> <p align="center"><i>50 secteurs</i></p> <p align="center"><i>600 ménages</i></p> <p align="center"><i>1 200 agents de santé communautaire</i></p>
<p align="center">T3</p> <p align="center">Combinaison des deux types de mesures d'incitation précitées</p> <p align="center"><i>50 secteurs</i></p> <p align="center"><i>600 ménages</i></p> <p align="center"><i>1 200 agents de santé communautaire</i></p>	<p align="center">C</p> <p align="center">Groupe de comparaison</p> <p align="center"><i>50 secteurs</i></p> <p align="center"><i>600 ménages</i></p> <p align="center"><i>1 200 agents de santé communautaire</i></p>

Besoins de données

La présente section répondra aux questions suivantes :

- Quel sera le type de données recueillies et analysées (quantitatives et/ou qualitatives) ?
- Quelles seront les sources des données ?
- À quel(s) niveau(x) les données seront-elles recueillies (établissements, ménages, etc.) ?
- Quelles seront les personnes interviewées (les répondants) ?
- Quelle sera la nature des données recueillies ?
- Quels seront les instruments d'enquête utilisés ?
- À quelle fréquence les données seront-elles recueillies ?

Bien qu'il existe une infinité de possibilités d'enquêtes réalisables aux divers niveaux de l'intervention, nous en décrivons cinq qui peuvent être réalisées au début et à la fin de l'opération de collecte de données sur les programme de FBR (voir ci-après).

Enquête auprès de l'établissement de santé. L'enquête auprès de l'établissement de santé mesure i) les principales caractéristiques de l'établissement, ce qui inclut d'ordinaire les effectifs, l'infrastructure, la disponibilité des services, la qualité des structures et des procédures, etc. ; ii) les caractéristiques des agents de santé telles que la formation, la rémunération, les heures de travail, les connaissances et les pratiques, la satisfaction et la motivation ; iii) la qualité des soins offerts par le biais d'entrevues de départ. Il convient de recueillir des données sur l'ensemble des établissements de santé compris dans l'échantillon d'évaluation, ce qui pourrait prendre de 1,5 à 2 jours par établissement.

Établissement d'une liste des ménages. L'équipe d'EI n'a souvent pas accès aux données d'un recensement récent pour établir son échantillon de ménages. Dans de tels cas, le bureau d'études devra établir une liste de tous les ménages compris dans les unités primaires d'échantillonnage (UPE) choisies à l'aide d'une courte enquête de présélection. Les enquêteurs doivent recueillir les informations de base sur chacun des ménages de l'UPE en ne consacrant à chacun de ces ménages que quelques minutes afin d'établir leur admissibilité, de consigner leurs informations démographiques de base et de noter leurs coordonnées en vue de contacts ultérieurs.

Enquête auprès des ménages. L'enquête complète réalisée auprès des ménages s'appuie habituellement sur un échantillon aléatoire de ménages situés dans les aires d'enquête choisies de l'échantillon d'évaluation. Elle devrait prendre en moyenne de 90 à 120 minutes par ménage. Les informations sont recueillies par le biais d'entrevues et d'observations directes. Ce type d'enquête est décrit d'une manière plus détaillée plus loin dans le présent module.

Indicateur(s) biométrique(s). Les indicateurs biologiques de la santé et du bien-être des ménages peuvent notamment inclure ce qui suit :

- grandeur et poids d'un sous-échantillon particulier des membres du ménage : femmes enceintes, enfants de moins de 5 ans ;
- taux d'hémoglobine (méthode de mesure de l'anémie) d'un sous-échantillon particulier des membres du ménage : femmes enceintes, enfants de moins de 5 ans ;
- test de dépistage du paludisme pour un sous-échantillon particulier des membres du ménage : femmes enceintes, enfants de moins de 5 ans.

Enquêtes auprès des collectivités. Les enquêtes réalisées auprès des collectivités s'adressent d'ordinaire aux dirigeants communautaires qui se trouvent dans les groupes échantillonnés et prennent environ 60 à 90 minutes. Les entrevues sont normalement organisées avec un ou deux dirigeants ou répondants principaux de chaque collectivité, et servent notamment à recueillir des informations sur les caractéristiques de la collectivité, les services, les infrastructures, l'accès aux marchés, les prix et les chocs subis par la collectivité.

Nous suggérons aux équipes de s'inspirer des exemples présentés au Tableau pour la collecte des données nécessaires à l'évaluation d'impact.

Tableau 11 : Besoins de données

	Type de données	Source des données	Niveau	Répondants	Description des données	Instrument d'enquête	Fréquence
Exemple 1	Quantitatives	SIGS	Établissement	S.O.	Couverture vaccinale mensuelle des enfants de moins de 5 ans	S.O.	Trimestrielle
Exemple 2	Quantitatives	Primaires	Établissement	Agents de santé	Charge de travail des employés, rémunération, motivation, connaissances, attitudes et pratiques	Questionnaire conçu pour l'établissement	Deux (évaluation de base et évaluation de suivi)
Exemple 3	Qualitatives	Primaires	Établissement	Femmes enceintes	Groupe de femmes enceintes, pour comprendre le comportement de recherche de services prénatals, périnatals et postnatals.	Questionnaire conçu pour un groupe d'étude	Deux (évaluation de base et évaluation de suivi)

Calendrier

Pour veiller à assurer la compatibilité de la conception opérationnelle du plan d'intervention et du plan d'évaluation, l'équipe d'EI doit s'engager tôt dans le processus de planification du programme. Le document de conception doit se pencher sur les questions suivantes :

- Quand les données de références seront-elles recueillies (données primaires) ou extraites (données secondaires) ?
- Quand les données de suivi seront-elles recueillies (données primaires) ou extraites (données secondaires) ?
- Dans quelle mesure les autorités publiques et les autres parties prenantes principales se sont-elles engagées à adhérer au calendrier convenu de manière à faire en sorte i) que la collecte des données de référence soit achevée avant le début de toute activité sur le terrain, et ii) que le groupe de traitement soit soumis au traitement pendant une période de temps assez longue, entre le début de l'intervention et l'enquête finale, pour permettre une mesure de l'impact sur les résultats ?

Figure 7 : Exemple de calendrier pour un document de conception de l'évaluation d'impact

Phase	An 1				An 2				An 3				An 4			
	T1	T2	T3	T4												
Conception du programme																
Conception de l'évaluation d'impact																
Préparation de l'évaluation																
Collecte des données de référence																
Début de l'intervention de FBR																
Exposition au traitement du FBR																
Documentation et stockage des données de référence																
Analyse de référence et rapport																
Préparation de l'évaluation																
Collecte des données finales																
Documentation et stockage des données finales																
Analyse d'impact et rapport																

Un calendrier général suffira pour les besoins du document de conception et de la note de synthèse de l'EI, mais l'équipe d'EI devrait élaborer un diagramme de Gantt plus détaillé énumérant l'ensemble des activités principales qui seront réalisées dans le cadre de l'EI et indiquant les échéanciers correspondant à chacune. Le module 4 porte sur le diagramme de Gantt de l'EI.

Équipe d'EI

Le document de conception doit également décrire en détails la composition de l'équipe d'EI (voir 0). En résumé, les informations suivantes devraient être incluses :

- Qui dirigera l'évaluation d'impact (enquêteur principal) ?
- Qui fournira l'aide technique requise à l'enquêteur principal (co-enquêteur principal, coordinateur de l'évaluation, expert de la qualité des données, assistant de recherche, etc.) ?
- Quelle personne ou quelle organisation peut faire office de collaborateur local de la recherche aux fins de la conception et de la gestion de l'étude ?
- Qui dirigera les activités de recherche complémentaires (par exemple, analyse des coûts, recherches qualitatives) ?

Plan de diffusion

L'équipe d'EI doit également prévoir comment les résultats de l'analyse de référence et ceux de l'analyse d'impact initiale seront diffusés.

Au minimum, elle devrait prévoir les activités de diffusion suivantes :

- rapport descriptif de base ;
- document pour collations de travail et autres moyens de communication ;
- rapport d'analyse d'impact ;
- atelier d'analyse d'impact ;
- collations de travail pour l'analyse d'impact.

Budget

Le budget de l'évaluation constitue un élément important de la conception de cette dernière, et nous recommandons qu'il énumère les dépenses attendues des éléments suivants de l'évaluation d'impact :

- *Équipe d'évaluation d'impact.* Le budget doit inclure le temps consacré par l'ensemble des membres du personnel et par les consultants à la gestion de l'évaluation d'impact, y compris pour la conception, la mise en œuvre et l'analyse.
- *Collecte des données.* L'équipe doit définir l'ensemble des besoins en matière de collecte de données primaires et secondaires, et présenter un budget pour satisfaire à ces besoins (au minimum, données de référence et de suivi), y compris les besoins en matière de données qualitatives et/ou d'analyse des coûts, le cas échéant.
- *Déplacements.* Le budget doit également inclure l'ensemble des coûts prévus des déplacements des superviseurs du projet, y compris pour les billets d'avion et les frais d'hébergement et de séjour.
- *Assistance technique.* Le budget doit inclure le temps nécessaire pour les déplacements et les services des prestataires d'assistance technique (par exemple, conception d'instruments d'enquête, contrôle de la qualité des données, élaboration du programme de saisie des données).
- *Plan de diffusion.* Le budget doit inclure tous les coûts liés aux déplacements ou aux services logistiques d'au moins un représentant sur le terrain, pour l'enquête de référence et le suivi, ainsi que tous les coûts liés à la préparation des documents écrits.
- *Divers.* Le budget doit inclure tous les coûts supplémentaires liés à l'évaluation d'impact, par exemple, ceux liés à l'examen institutionnel du protocole de recherche.

La boîte à outils contient un **modèle de budget d'EI** qui pourra servir à la préparation du budget de l'EI.

Étude de cas nationale : Relever le défi du financement de la collecte des données en faisant preuve de pragmatisme

Intervention de financement de la santé basé sur les résultats en Zambie

Étant donné les faibles densités de population qui caractérisent un certain nombre de districts en Zambie, le piètre état des routes en général, et particulièrement pendant la saison des pluies, le prix élevé du carburant et la volatilité des prix due aux fluctuations du taux de change, le transport a occupé une place importante dans le budget d'EI. Bien que le transfert des connaissances acquises dans le cadre des enquêtes sur le paludisme et sur le financement de la santé basé sur les résultats ait contribué à contenir les coûts, en réduisant par exemple les coûts de mise au point de produits et de formation, les coûts variables du transport sont restés problématiques. L'équipe d'EI était en partie impuissante à corriger la situation, mais certaines mesures telles que le contrôle de la qualité des données à la fin de chaque jour ont aidé à réduire les dépenses consacrées aux déplacements. Les informations sur la charge de travail peuvent aider les responsables de la planification de la saisie des données sur le terrain à réduire les coûts de transport [coûts des visites supplémentaires requises lorsque le nombre de patients interviewés dans les établissements de santé est insuffisant].

Pour en savoir plus, consulter la section « Pleins feux sur le pays » de la boîte à outils.

Étude de cas nationale : Ajustement du calendrier et du budget
Projet de financement basé sur les résultats au Bénin

Le Bénin s'emploie à mettre en œuvre un programme de financement basé sur les résultats dans 34 districts. L'intervention consiste à combiner diverses mesures telles que l'augmentation de l'autonomie de gestion, la hausse des bonus forfaitaires versés aux employés ou la hausse des bonus de FBR. La combinaison de ces mesures permet de définir cinq branches d'étude : i) autonomie accrue + bonus de FBR ; ii) autonomie accrue plus bonus forfaitaires ; iii) bonus de FBR seulement ; iv) bonus forfaitaires seulement ; v) statu quo (groupe de comparaison).

Pendant la phase préparatoire, le groupe de comparaison (statu quo) a été ajouté à la dernière minute, principalement parce que l'équipe d'EI craignait que les différences entre les quatre branches initiales de l'étude soient trop petites. Ce changement a entraîné une hausse soudaine du budget et une prolongation de la période requise pour réaliser l'enquête de référence. L'équipe a révisé son calendrier d'exécution et sollicité une hausse des ressources budgétaires pour réaliser l'étude et l'enquête requises.

Par ailleurs, alors que l'équipe se préparait à mettre en œuvre l'intervention et l'EI, d'autres donateurs multilatéraux se sont montrés intéressés à financer les interventions de FBR dans le pays. L'intégrité du groupe de comparaison a été menacée aux étapes précoces de l'EI, et l'équipe a dû négocier le report d'autres activités des donateurs pour protéger ce groupe le temps qu'il fallait.

Pour en savoir plus, consulter la section « Pleins feux sur le pays » de la boîte à outils.



Module 4

Préparation de la Collecte des Données

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats
Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

www.worldbank.org/health/impacetevaluationtoolkit

Module 4. Préparation de la collecte des données

Principales recommandations et outils disponibles pour ce module			
Recommandations	Essentiel	Important	Souhaitable
<ul style="list-style-type: none"> Le protocole de recherche doit contenir toutes les informations pertinentes concernant la protection des sujets humains, y compris des critères d'échantillonnage spécifiques et des protocoles relatifs au consentement éclairé et à la confidentialité des données. 	✓		
<ul style="list-style-type: none"> L'évaluation d'impact doit être approuvée par un comité d'examen institutionnel : l'enquêteur principal doit veiller à confier à un tel comité l'examen éthique et l'approbation de la recherche <u>avant</u> le début des activités sur le terrain. 	✓		
<ul style="list-style-type: none"> Un diagramme de Gantt du projet/de l'évaluation d'impact peut aider les équipes à coordonner leurs activités et les échéanciers à partir du projet et de l'évaluation d'impact. 		✓	
<ul style="list-style-type: none"> L'équipe d'évaluation d'impact doit convenir avec les partenaires de l'administration publique de la politique d'accès à l'information issue de l'évaluation d'impact. Un protocole d'accord écrit peut éviter les malentendus dans ce domaine. 		✓	
<ul style="list-style-type: none"> Le choix d'opter pour la saisie de terrain assistée par ordinateur (STAO) ou pour la saisie des données sur le terrain a une grande incidence sur le choix du bureau d'études et devrait être fait avant de recruter ce dernier. 		✓	
<ul style="list-style-type: none"> Le recrutement du bureau d'études est un processus de longue haleine qui prend d'ordinaire de trois à six mois et qui devrait être entamé dès les premières étapes de la planification du projet. 	✓		
<ul style="list-style-type: none"> Selon la situation et l'expérience du pays, on pourra choisir de confier à un seul et même bureau les enquêtes sur les ménages et sur les établissements de santé, ou de recruter deux bureaux distincts. Nous recommandons en règle générale que l'équipe ait recours à cette fin à une procédure d'appel à la concurrence. 		✓	
<ul style="list-style-type: none"> L'équipe de gestion de l'enquête devrait être composée d'un responsable de projet, d'un administrateur local et d'un gestionnaire des données, et ce pour toute la durée de la préparation et de la mise en œuvre des activités de collecte des données. 	✓		
<ul style="list-style-type: none"> La conduite des négociations avec le bureau d'études nécessite une connaissance approfondie des contraintes budgétaires et temporelles qui auront une incidence sur la composition de l'équipe de terrain et sur la durée de l'enquête. 	✓		
<ul style="list-style-type: none"> Le bureau d'études devrait pouvoir compter dès les étapes initiales de la préparation de l'enquête sur les services d'un expert de la qualité des données, en particulier s'il s'agit d'une entreprise locale aux capacités limitées. 		✓	
<ul style="list-style-type: none"> La structure et la qualité des instruments d'enquête jouent un rôle 		✓	

Principales recommandations et outils disponibles pour ce module

Recommandations	Essentiel	Important	Souhaitable
<p>essentiel dans l'assurance de la qualité et la comparabilité des résultats obtenus pour chaque pays. Nous recommandons que les équipes de projet utilisent comme point de départ les questionnaires d'enquêtes sur le FBR conçus pour les établissements de santé et les ménages. L'enquêteur principal de l'évaluation doit déterminer quels sont les modules appropriés ou non, et veiller à ce que les questionnaires permettent de cerner les principaux points d'intérêt. Les équipes ne doivent pas hésiter à procéder aux ajustements qu'elles estiment nécessaires.</p> <ul style="list-style-type: none"> • Les questionnaires de la boîte à outils ont été conçus pour être exhaustifs, et les équipes souhaiteront peut-être limiter le nombre de modules utilisés pour réduire le coût et le temps requis pour les administrer. • Les enquêtes auprès des collectivités peuvent permettre d'évaluer les infrastructures et les réseaux d'aide qui les caractérisent. Elles peuvent aussi servir de compléments aux enquêtes auprès des ménages, en particulier lorsque ces dernières ont besoin d'être sensiblement écourtées. 		✓	✓



Outils
<ul style="list-style-type: none"> • 4.01 Impact Evaluation Gantt Chart (Diagramme de Gantt pour l'évaluation d'impact) • 4.02 Data Access Memorandum of Understanding (Protocole d'accord sur l'accès aux données) • 4.03 Research Protocol Example (Exemple de protocole de recherche) • 4.04 Informed Consent Templates (Modèles de consentement éclairé) • 4.05 Health Facility Survey Firm TOR (Termes de référence du bureau chargé des enquêtes formations sanitaires) • 4.06 Household Survey Firm TOR (Termes de référence du bureau chargé des enquêtes ménages) • 4.07 Data Collection Budget Template (Modèle de budget pour la collecte des données) • 4.08 Consumables and Equipment for Biomarker Data (Produits consommables et matériel pour les données des biomarqueurs) • 4.09 Health Facility Questionnaires (Questionnaires Formations sanitaires) • 4.10 Household Questionnaires (Questionnaires Ménages) • 4.11 Community Questionnaires (Questionnaires Communauté) • 4.12 Costing Questionnaires (Questionnaires relatifs aux coûts) • 4.13 Data Entry Program (Programme d'entrée des données) • 4.14 Anemia Referral Guidelines (Directives relatives à l'orientation des cas d'anémie) • 4.15 Anemia Referral Form (Formulaire relatif à l'orientation des cas d'anémie) • 4.16 How to Translate Questionnaires (Comment traduire les questionnaires) • 4.17 Institutional Review Board TOR (Termes de référence du Conseil d'examen institutionnel) • 4.18 Certificate of Accurate Translation (Certificat d'exactitude de la traduction)

Contenu du module

Planification des activités à l'aide d'un diagramme de Gantt	4-5
Définition de la propriété des données	4-5
Protection des sujets humains.....	4-6
Principes fondamentaux de la protection des sujets humains.....	4-7
Formation sur les sujets humains	4-8
Protocole de recherche.....	4-8
Consentement éclairé.....	4-10
Protection de la confidentialité des participants.....	4-12
Obtention de l'autorisation éthique	4-13
Définition de la stratégie de saisie des données	4-13
Recrutement d'un bureau d'études	4-15
Effectifs des bureaux d'études	4-17
Calendrier des produits escomptés et des paiements	4-19
Budget.....	4-20
Comprendre et adapter les instruments d'enquête.....	4-21
Directives universelles en matière de format.....	4-22
Adaptation des instruments d'enquête de FBR.....	4-23
Traduction et rétro-translation	4-27
Instruments d'enquête	4-27
Questionnaires sur les établissements de santé	4-27
Évaluation des établissements :.....	4-28
Entretien avec les agents de santé	4-29
Entretien avec les patients à la sortie de l'établissement	4-31
Mesure de la qualité des soins.....	4-32
Questionnaires destinés aux ménages	4-35
Sujets au niveau des ménages	4-36
Modules sur la santé maternelle et infantile.....	4-37
Questionnaires destinés aux communautés.....	4-38

Ce module présente un aperçu des diverses étapes nécessaires à la préparation de la collecte des données de référence. Nous couvrons toutes les étapes devant être suivies entre l’approbation de la note conceptuelle et l’entrée en fonction du bureau d’études. Ces étapes comprennent : i) la planification de l’évaluation et de l’enquête de référence ; ii) la définition du protocole de recherche et l’obtention des autorisations éthiques ; iii) la définition de la stratégie de saisie des données ; iv) le recrutement d’un bureau d’études pour l’enquête de référence ; v) la compréhension et l’élaboration des questionnaires.

Planification des activités à l’aide d’un diagramme de Gantt

Pour qu’une enquête de référence reflète précisément la situation « avant le traitement », il est essentiel que l’intégralité de la collecte des données soit achevée avant le début de l’intervention de FBR, en particulier dans les groupes de traitement sélectionnés. Nous recommandons de partir de la date désirée pour le début de l’intervention et de planifier en conséquence toutes les activités devant être entreprises pour mener à bien l’enquête de référence. On pourra pour cela utiliser le modèle du **diagramme de Gantt**. En se fondant sur les autres enquêtes ayant été réalisées, on peut formuler quelques conseils concernant la durée des principaux processus :

Conseil 1 : Prévoir suffisamment de temps pour recruter les bureaux d’études – Du fait de la taille du budget, le recrutement des bureaux d’études sera presque toujours conclu par appel à la concurrence, ce qui peut prendre **entre quatre et six mois**.

Conseil 2 : Prévoir suffisamment de temps pour l’approvisionnement en matériel et produits consommables – Le matériel nécessaire pour les mesures anthropométriques et l’identification des biomarqueurs (pour l’anémie ou le paludisme) n’est pas disponible dans tous les pays. Bien que ces équipements puissent être acquis et gérés par les bureaux d’études, ce processus peut prendre **entre deux et trois mois**.

Conseil 3 : Prévoir suffisamment de temps pour obtenir les autorisations éthiques – Certains comités d’examen institutionnel (CEI) ont des calendriers fixes pour examiner les demandes nationales d’autorisation éthique. Dans la majorité des cas, l’enveloppe-recherche, y compris le protocole de recherche, les questionnaires et les modèles de consentement éclairé, doit être soumise plusieurs semaines avant que l’étude ne puisse être examinée par le CEI, qui fixe généralement ce délai. Les équipes devraient prévoir **entre deux et trois mois** pour obtenir les autorisations éthiques.

Définition de la propriété des données

La définition de la propriété des données peut être compliquée, car les grandes enquêtes font généralement intervenir de nombreuses parties ayant un intérêt dans les données :

- Les pouvoirs publics, parce que les données ont été collectées sur leur territoire.

- Les enquêteurs, en raison du temps et des efforts intellectuels investis dans la collecte des données.
- Les bailleurs de fonds, qui peuvent chacun avoir leurs propres idées concernant la propriété des données.
- Les bureaux d'études, qui peuvent insister pour que leur contrat leur réserve les droits sur les données.

Tous ces exemples montrent qu'il est très important de bien comprendre les questions relatives à l'accès aux données et à leur propriété. Un **protocole d'accord** peut être établi à cette fin entre les pouvoirs publics et la Banque ou un autre bailleur de fonds avant le début de la collecte des données. Ce protocole d'accord devrait couvrir les aspects suivants :

- Qui sera propriétaire des données collectées ?
- Quels sont les accords sur l'accès aux données ? Pendant combien de temps les données seront-elles indisponibles, et dans quel délai seront-elles disponibles dans le cadre d'une utilisation autorisée ou accessibles au public ? Quelles seront les conditions associées à une utilisation autorisée des données pendant leur délai d'indisponibilité ? Pour plus de détails sur les différents types d'accès aux données, voir le 0 sur le stockage des données.
- Accords sur le stockage et la protection des données.
- Qui gèrera les données ?

Les enquêteurs et les bureaux d'études ne devraient jamais être les « propriétaires » des données collectées dans le cadre d'une évaluation d'impact. Il est toutefois possible de prévoir des dispositions concernant les enquêteurs dans l'accord sur l'accès aux données, de sorte qu'ils soient suffisamment motivés pour investir du temps et de l'énergie dans la collecte des données, tout en veillant à ce qu'elles soient disponibles pour un usage ultérieur dans un délai raisonnable.

Les modalités de la disponibilité des données après leur collecte et les politiques d'accès aux données sont examinées au 0.

Protection des sujets humains

Les évaluations d'impact, bien qu'elles soient liées à l'exécution des projets et aux interventions gouvernementales, sont aussi une activité de recherche à laquelle participent des « sujets humains ». Les ménages, médecins, infirmiers et administrateurs qui répondent aux questionnaires peuvent subir des préjudices si les informations qu'ils fournissent sont mises à la disposition du public sans avoir pris des précautions suffisantes. À ce titre, les évaluations d'impact doivent être menées en respectant les bonnes pratiques de « protection des sujets humains ».

Exemples d'infractions à la protection des sujets humains : i) dans le cadre d'une enquête, une femme peut se trouver en danger si son mari l'entend fournir des informations confidentielles concernant ses méthodes de planning familial ; ii) la sécurité des ménages peut être compromise si les revenus ou les

avoirs financiers d'une famille spécifique peuvent être identifiés à partir de données accessibles sur internet ; iii) l'étude n'a pas fourni d'informations sur les risques associés aux tests biométriques ; iv) un participant ayant demandé de quitter l'étude au milieu de l'enquête est obligé par l'enquêteur de terminer l'entretien.

Principes fondamentaux de la protection des sujets humains

L'enquêteur principal et les autres enquêteurs sont responsables de veiller à la protection des droits et du bien-être des sujets humains participant aux travaux de recherche, conformément au code de déontologie ou à la législation du pays. En l'absence de directives nationales de déontologie, l'enquêteur devrait s'inspirer de la déclaration d'Helsinki adoptée par la vingt-neuvième Assemblée médicale mondiale à Tokyo (octobre 1975) et de l'Article 7 du Pacte international relatif aux droits civils et politiques, adopté par l'Assemblée générale des Nations Unies, le 16 décembre 1966.

Critères de l'OMS : Les critères de base recommandés par l'OMS pour évaluer les projets de recherche auxquels participent des sujets humains comprennent :

- les droits et le bien-être des personnes participant aux travaux de recherche devraient être protégés de manière adéquate ;
- un consentement libre et éclairé devrait être obtenu ;
- l'équilibre entre les risques et les avantages potentiels associés devrait être évalué et jugé acceptable par un groupe d'experts indépendants de l'institution ; et
- toute exigence nationale particulière devrait être respectée.

Aux États-Unis, les trois principes suivants constituent le fondement des directives sur l'éthique de la recherche sur des sujets humains¹¹ :

- Respect de la personne : Comment les chercheurs obtiendront-ils le consentement éclairé de leurs sujets ?
- Intention bénéfique : Comment les chercheurs s'assurent-ils que la recherche 1) ne porte pas préjudice aux sujets et 2) maximise les avantages potentiels tout en minimisant les préjudices potentiels.
- Justice : Comment les chercheurs s'assurent-ils que les avantages et les fardeaux de la recherche sont équitablement partagés ?

Ces principes sont fondés sur le rapport Belmont (1979) sur les Principes éthiques et directives concernant la protection des sujets humains dans le cadre de la recherche, disponible à <http://www.hhs.gov/ohrp/policy/belmont.html>.

¹¹ Ces principes sont énoncés dans le rapport Belmont, qui a été rédigé en 1974 par la Commission nationale pour la protection des sujets humains dans le cadre de la recherche biomédicale et comportementale.

Formation sur les sujets humains

Les Instituts nationaux de la Santé des États-Unis (NIH) recommandent que tous les enquêteurs principaux, co-enquêteurs principaux et coordinateurs de la recherche reçoivent une formation sur la protection des sujets humains et suivent des cours annuels de mise à jour. La formation en ligne des NIH est très instructive et ne prend qu'une heure. Elle est disponible à <http://phrp.nihtraining.com/users/login.php> et www.ohsr.od.nih.gov.

Veillez noter que les fonds du HRITF destinés aux enquêtes de référence et de suivi ne peuvent être transférés aux équipes que si l'enquêteur principal proposé, le co-enquêteur principal et le coordinateur de la recherche peuvent prouver qu'ils ont récemment suivi une formation sur les sujets humains. Le cours en ligne des NIH comprend un test, à l'issue duquel est délivré un certificat qui peut être utilisé à cette fin.

Le programme CITI est une autre source de formation sur les sujets humains. Le CITI offre des cours internationaux sur les CEI dans plusieurs langues, contre le versement de 75 dollars (par personne). Voir www.citiprogram.com.

Protocole de recherche

Le **protocole de recherche** décrit l'objet de l'étude, l'objectif de l'évaluation, les méthodes et procédures ainsi que les mesures prises par les chercheurs pour veiller à la protection des sujets humains. C'est donc l'un des plus importants documents écrits du dossier d'évaluation.

Le protocole de recherche est utilisé par différentes parties concernées par l'évaluation :

- Les **équipes d'EI et de projet** l'utilisent comme référence sur les méthodes employées.
- Le **bureau d'études** l'utilise pour orienter les méthodes de terrain.
- Les **analystes** y ont recours pour comprendre comment les données ont été produites.
- Les **comités d'éthique/comités d'examen institutionnel** (CEI) y ont recours pour déterminer s'ils peuvent approuver les travaux de recherche proposés.

La structure d'un protocole de recherche peut varier en fonction des exigences du comité d'éthique/comité d'examen institutionnel, mais comprend généralement les éléments suivants :

1. Objet de l'étude :

Cette section décrit l'intervention proposée, les lacunes existantes concernant l'impact de cette intervention et

comment celle-ci se propose de les combler. Cette section sera semblable à l'introduction et à l'exposé des motifs du document de conception de l'EI.

2. Objectifs de l'évaluation, questions stratégiques et méthodologie : Cette section devrait détailler les objectifs de l'EI et les questions stratégiques abordées dans l'évaluation, ainsi que la méthodologie sélectionnée pour isoler l'impact causal de l'intervention proposée sur les résultats escomptés.

3. Plan de l'analyse : Cette section devrait détailler les méthodes économétriques qui seront utilisées pour réaliser l'analyse des impacts.

4. Sélection des sujets : Cette section développe la section sur l'échantillon figurant dans le document de conception de l'EI et présente des chevauchements avec le **Plan d'échantillonnage de terrain** (voir 0). Elle devrait définir les éléments suivants :

- **Nombre de sujets :** Il s'agit du nombre de sujets nécessaires aux niveaux des établissements, des prestataires, des ménages et des individus, conformément aux besoins de l'étude.
- **Critères d'inclusion :** Les critères d'inclusion sont utilisés pour décider si une unité doit ou non être incluse dans la base d'échantillonnage. Les critères d'inclusion jouent un rôle essentiel car ils peuvent conduire à d'importantes variations de la puissance de l'étude, même avec des échantillons de taille égale.
- **Définition de l'échantillon en fonction de l'âge, du sexe et de la race :** L'équipe d'EI doit fournir tous les détails nécessaires sur les critères d'inclusion et justifier toute exclusion de l'étude.

Qu'est-ce qu'un comité d'éthique ?

Un **comité d'examen institutionnel (CEI)**, également désigné **comité d'éthique indépendant** ou **comité d'éthique (CE)** est un comité officiellement désigné pour approuver, surveiller et examiner les travaux de recherche biomédicale et comportementale impliquant des êtres humains.

L'objet d'un examen du CEI est de s'assurer, par anticipation mais aussi à intervalles réguliers, que des mesures appropriées sont prises pour protéger les droits et le bien-être des sujets humains participant à des travaux de recherche. Les CEI s'efforcent de veiller à la protection des sujets en examinant les protocoles de recherche et les documents connexes. L'examen d'un protocole par un CEI évalue le caractère éthique des travaux de recherche et de leurs méthodes, encourage la participation volontaire et informée de sujets potentiels capables de faire de tels choix (ou, si cela n'est pas possible, l'autorisation informée donnée par un mandataire approprié) et s'attache à maximiser la sécurité des sujets. (Wikipedia)

De nombreuses universités ont leur propre comité d'examen institutionnel et de nombreux pays sont également dotés d'un comité d'éthique national.

- **Protection des sujets vulnérables** : L'équipe d'EI devrait déterminer quelle partie de l'échantillon est jugée vulnérable et préciser les mesures qu'elle prendra pour minimiser les risques pesant sur ces sujets. L'équipe d'EI devrait surtout s'attacher à minimiser le risque qu'une pression soit exercée sur les sujets démunis, analphabètes, en bas âge, ou définis comme vulnérables, pour qu'ils participent à l'étude.
- **Confidentialité et autonomie** : Cette section devrait indiquer les mesures que l'équipe chargée de l'enquête prendra pour protéger la confidentialité et l'autonomie des personnes concernées.

5. Méthodes et procédures : Cette section devrait fournir des détails sur les informations collectées (y compris les données requises sur les biomarqueurs tels que l'anémie et le paludisme et les mesures anthropométriques) et les procédures de terrain utilisées pour sélectionner les ménages et les personnes sondés. Elle devrait également préciser comment le bureau d'études recueillera et gèrera ces données, notamment les procédures utilisées pour stocker les données et veiller à leur confidentialité sur le terrain, lors de la saisie des données et à l'issue de l'étude.

6. Évaluation des risques/avantages : L'équipe d'EI devrait évaluer tout risque (direct ou indirect) ou avantage (direct ou indirect) potentiel pour les participants à l'étude. Ces risques peuvent comprendre la perte de confidentialité et ceux associés aux tests sur l'anémie et le paludisme ainsi que les mesures anthropométriques. Les risques pour les participants à l'étude devraient être minimales dans la majorité des évaluations d'impact des FBR. Les avantages peuvent comprendre une faible compensation pour le temps passé (en conformité avec les traditions locales) et l'évaluation de l'état de santé des enfants découlant des résultats instantanés des tests sur l'anémie et le paludisme.

7. Identification et recrutement des sujets : Cette section devrait indiquer comment l'équipe de terrain évaluera la capacité des participants potentiels à comprendre les procédures de l'étude et les instruments de l'enquête. Si une équipe de terrain estime qu'un participant potentiel a une capacité ou une compréhension de l'étude limitée, le ménage concerné devrait être remplacé par le suivant dans la liste des ménages admissibles, dressée conformément aux critères d'inclusion et à la stratégie d'échantillonnage.

8. Documentation : Le protocole de recherche devrait inclure en annexe tous les documents ou outils pertinents, notamment mais non exclusivement : i) les questionnaires, ii) les formulaires de consentement éclairé, iii) les fiches de santé infantile, et iv) la description des avantages ou de la compensation prévus pour les participants, le cas échéant.

L'OMS a publié des directives sur la rédaction d'un protocole de recherche faisant intervenir des sujets humains, disponible (en anglais) à http://www.who.int/rpc/research_ethics/guide_rp/en/index.html.

Consentement éclairé

Le consentement éclairé est l'un des éléments fondamentaux pour protéger les droits des sujets humains participant à une étude ou une intervention. Il nécessite que les participants aient une bonne

compréhension de l’objet, des procédures, des risques et des avantages de l’étude. Par défaut, le **consentement éclairé** d’un adulte¹² nécessite un document écrit (formulaire) comprenant une section sur les méthodes utilisées pour protéger sa confidentialité, une section sur son droit de refuser ou d’annuler à tout moment sa participation à l’étude, une explication des risques et avantages potentiels, les coordonnées de l’équipe d’enquête ou de l’enquêteur pour le cas où les participants aimeraient les contacter, et un espace pour que le participant puisse enregistrer son consentement exprès écrit et y apposer sa signature. L’équipe de terrain garde un exemplaire du formulaire de consentement éclairé pour archivage et donne un exemplaire au participant à l’issue de l’entretien. Contrairement aux adultes, les *mineurs*¹³ ne peuvent pas consentir à participer à une enquête ; ils peuvent uniquement donner leur **assentiment après** avoir obtenu l’**autorisation** écrite de leur parent ou tuteur. L’assentiment et l’autorisation sont vérifiés de la même manière formelle et écrite que le consentement. Tableau 4 dresse la liste des personnes susceptibles de participer aux activités de collecte des données d’évaluation d’impact sur les FBR ainsi que les méthodes de vérification requises.

Tableau 4 : Instruments d’enquête, participants et méthodes de vérification

Instrument d’enquête	Participant	Vérification nécessaire
Évaluation des établissements de santé	Directeur et/ou administrateur de l’établissement de santé	Consentement éclairé
Agent de santé	Agent(s) de santé du service étudié	Consentement éclairé
Entretien avec les patients à la sortie de l’établissement	Patients du service étudié	Consentement éclairé Autorisation d’un tuteur adulte + assentiment de l’enfant, en cas d’examen de l’enfant
Principaux modules sur les ménages	Chef de ménage et/ou son conjoint. Comprend les réponses indirectes de tous les membres du ménage aux questions concernant l’éducation, le marché du travail et la santé.	Consentement éclairé
Modules sur la santé de la femme	Toutes les femmes du ménage âgées de 15 à 49 ans	Consentement éclairé
Modules sur la santé infantile	Mère des enfants de 0 à 5 ans et/ou personne qui en a la charge	Autorisation d’un tuteur adulte + assentiment de l’enfant, en cas d’examen de l’enfant

L’équipe d’EI peut demander deux dérogations à l’obligation d’obtenir le consentement ou l’assentiment exprès écrit des participants. Ces dérogations peuvent être obtenues auprès du Comité d’éthique.

¹² Tels que définis par la législation nationale, généralement des personnes âgées de 18 ans ou plus.

¹³ Tels que définis par la législation nationale, généralement des personnes âgées de moins de 18 ans.

Dérogation au consentement exprès écrit de potentiels participants adultes admissibles : Pour l'étude principale, les équipes d'EI peuvent demander que le consentement exprès et écrit sous la forme d'une signature soit remplacé par un consentement oral documenté. Si possible, l'enquêteur devrait documenter tout consentement oral et l'accompagner de la signature du superviseur (en tant que témoin). Les équipes d'EI peuvent utiliser les arguments suivants pour justifier l'utilisation de consentements oraux documentés. Premièrement, la majorité des activités de l'étude ne présente pas de risque plus important que ceux que l'on rencontre dans la vie quotidienne et, deuxièmement, l'équipe d'EI veillera à ce que le personnel de terrain ait reçu une formation approfondie concernant la présentation du consentement oral et de l'étude. Chaque famille recevra alors un exemplaire du formulaire de consentement de sorte à pouvoir contacter les agents chargés de l'étude en cas de question (les familles analphabètes pourront demander l'aide des membres alphabètes de leur communauté). Troisièmement, la validité scientifique de l'étude pourrait être compromise si des participants analphabètes refusent d'y participer, au motif qu'ils ne veulent pas signer un document qu'ils ne peuvent pas lire, et si le nombre de consentants analphabètes est inférieur à celui d'alphabètes. Cet aspect est important car l'alphabétisation est un élément-clé dans un grand nombre de processus causaux que l'étude entend évaluer, tels que l'utilisation des soins de santé et la qualité des soit reçus.

Dérogation à l'assentiment écrit des enfants : L'équipe d'EI peut demander une dérogation à l'assentiment écrit des enfants de moins de cinq ans lorsque l'enquête pose des questions sur leur santé alors que l'on ne peut pas raisonnablement supposer que ces enfants comprennent les risques et les avantages associés à leur participation à l'enquête. L'autorisation du parent devrait être suffisante pour ces sections. Pour les modules d'enquête nécessitant un contact direct avec les enfants (par exemple, mesure anthropométrique, mesure d'hémoglobine et tests de dépistage du paludisme), le Comité d'éthique peut lever l'obligation d'obtenir l'assentiment écrit des enfants. Toutefois, si des enfants refusent d'être mesurés ou de se soumettre à des tests réalisés par des agents de terrain, ceux-ci doivent respecter leurs souhaits et utiliser le code « refus » lors du remplissage du questionnaire. Veuillez noter que la participation d'un mineur nécessite impérativement une autorisation parentale.

Des **modèles de consentement éclairé** sont proposés aux équipes.

Protection de la confidentialité des participants

Toutes les informations fournies pendant l'entretien sont strictement confidentielles. Bien que les résultats de l'étude puissent être publiés à des fins scientifiques, ils devraient être rédigés de sorte à empêcher l'identification de la personne ou du ménage concerné. À cette fin, chaque sujet participant à l'étude devrait recevoir un numéro d'identification unique (ID) et tous les noms et identifiants devraient être effacés de la base de données utilisée à des fins de recherche. L'affectation de ces numéros est examinée au Module 5.

Obtention de l'autorisation éthique

Autorisation nationale : Il incombe à l'enquêteur principal d'identifier toutes les institutions exigeant l'autorisation ou l'approbation de l'étude au niveau national, en particulier lorsqu'elles nécessitent la collecte de données sur les ménages ou les personnes. L'équipe devra généralement obtenir l'autorisation éthique auprès du comité d'éthique ou du comité d'examen institutionnel du pays concerné. Le site ci-dessous (en anglais) répertorie les comités nationaux et internationaux chargés de la protection des sujets humains <http://www.hhs.gov/ohrp/international/intlcompilation/hspcompilation-v20101130.pdf>

L'équipe d'EI devrait également collaborer avec le ministère de la Santé concerné pour veiller au respect de toutes les exigences nationales relatives aux travaux de recherche sur des sujets humains. Dans de nombreux pays, le CEI exigera des enquêteurs principaux et du bureau d'études qu'ils accompagnent l'étude des preuves du respect des directives en matière de recherche.

Autorisation aux États-Unis : Bien que la Banque mondiale ne soit pas dotée de son propre CEI, le respect des normes internationales par le protocole d'EI peut être assuré de deux manières. Premièrement, les enquêteurs principaux attachés à des établissements universitaires (tels que l'Université Johns Hopkins et l'Université de Californie-Berkeley) doivent obtenir leur autorisation auprès de leur CEI pour participer à l'étude. Deuxièmement, l'équipe d'EI peut demander à un CEI indépendant agréé aux États-Unis de procéder à l'examen du protocole de recherche. Un **modèle de termes de référence du CEI** est proposé aux équipes.

Le processus d'autorisation éthique prévoit la soumission du protocole de recherche, des questionnaires, du consentement éclairé et des autres documents de l'étude au CEI, leur examen par ce dernier et la révision du protocole et des documents en conformité avec les éventuelles recommandations du comité. Ce processus prend généralement entre deux et trois mois, bien que cela dépende fortement de la fréquence des réunions du CEI. L'équipe d'EI devrait prévoir de détenir toutes les autorisations requises avant d'entamer la phase pilote de l'enquête avec l'équipe de terrain.

Définition de la stratégie de saisie des données

La définition d'une stratégie de saisie des données est une étape essentielle lors de la sélection du bureau d'études. Les deux principales stratégies de saisie des données sont : la saisie de terrain assistée par ordinateur (STAO) et la saisie centralisée des données. Dans la première, chaque équipe de terrain comprend un opérateur de saisie de données, chargé de numériser les données sur le terrain à l'aide d'un ordinateur portable. Dans la seconde, les questionnaires sur support papier sont envoyés à un site de saisie centralisée des données où une équipe d'opérateurs numérise les données de toutes les équipes de terrain.

Pour les évaluations financées par le HRITF, la stratégie de saisie des données recommandée est la STAO car les enquêtes sur la mesure des niveaux de vie ont montré qu'elle permet d'optimiser la qualité

globale des données collectées. Toutefois, cette stratégie de saisie des données a des incidences sur le budget et le calendrier de l'enquête. L'équipe d'EI devrait donc décider si elle veut adopter la STAO *avant* de recruter le bureau d'études.

La STAO nécessite du matériel, des logiciels, une organisation efficace et une planification réaliste du travail de terrain. Elle nécessite en particulier i) une équipe motivée et spécialisée et ii) la réalisation des activités de terrain dans un esprit d'équipe. Les éléments suivants sont importants lorsque l'on compare la STAO à la saisie centralisée des données :

Quelles sont les ressources nécessaires à la STAO ? La mise en œuvre efficace de la STAO nécessite un opérateur de saisie doté d'un ordinateur portable au sein de chaque équipe de terrain. Le bureau d'études devra probablement se procurer des ordinateurs portables, ce qui aura une incidence sur le budget et le calendrier de l'enquête. Le bureau d'études pourra également envisager de ne pas avoir recours à un opérateur de saisie et de former les enquêteurs pour qu'ils saisissent eux-mêmes les données. Le superviseur de l'équipe sera également responsable de la saisie des données, de l'analyse des rapports d'erreur et de la décision d'effectuer de nouvelles visites pour corriger ces erreurs. La formation des superviseurs d'équipe aux responsabilités associées à la STAO peut être difficile dans un premier temps, et la supervision centrale des équipes de STAO doit être beaucoup plus active, technique, compétente et motivée que celle des équipes de terrain. Les superviseurs centraux doivent savoir mieux que quiconque comment interpréter les rapports quotidiens produits par la STAO.

Combien de temps doit-on prévoir pour la formation à la STAO ? L'expérience montre qu'un bureau d'études devrait prévoir entre trois et cinq jours pour former les opérateurs de saisie, bien que cela puisse dépendre des capacités locales. La formation à la STAO devrait être axée sur la préparation des opérateurs de saisie et des superviseurs d'équipe à la saisie des données, à l'analyse des rapports d'erreur et à l'organisation de nouvelles visites pour corriger ces erreurs si nécessaire.

Quelles sont les différences dans la qualité des données entre la STAO et la saisie centralisée des données ? Le principal avantage de la STAO est que les valeurs aberrantes et les incohérences sont corrigées sur le terrain, en les comparant directement à la réalité des ménages plutôt qu'en faisant des suppositions comme cela peut être le cas avec la saisie centralisée des données. Le laborieux processus de « nettoyage centralisé des données » est généralement inévitable et peut menacer la pertinence des données aux fins de l'élaboration de politiques. Car le « nettoyage des données » a pour objet de veiller à leur cohérence interne sans pour autant vérifier qu'elles reflètent la réalité du terrain. La STAO permet de déceler les incohérences, immédiatement et sur le terrain, mais aussi d'évaluer directement la performance des agents de terrain et de détecter de manière précoce les comportements indésirables.

Quels sont les gains de temps imputables à l'utilisation de la STAO comparée à la saisie centralisée des données ? Dans le cas de la saisie centralisée des données, le bureau d'études devra généralement passer entre deux et trois mois à saisir et nettoyer les données après leur collecte. Avec la STAO, les bases de données définitives sont disponibles dès la fin de la collecte des données ; toutes les données sont disponibles dès la livraison du fichier de la dernière grappe sondée. La STAO peut toutefois légèrement prolonger la durée des activités de terrain, notamment lorsque de nombreuses visites

doivent être renouvelées. Les avantages liés à la qualité supérieure des données l'emportent toutefois sur les coûts associés à la prolongation des activités de terrain.

Recrutement d'un bureau d'études

L'équipe d'EI demandera au bureau d'études d'exécuter les activités suivantes de collecte des données : i) enquête sur les établissements de santé ; ii) liste des ménages ; iii) enquête sur les ménages ; iv) mesures biométriques ; et v) enquête communautaire. En fonction des capacités locales disponibles, l'équipe d'EI pourra se trouver dans la nécessité de recruter différents bureaux d'études pour exécuter ces différentes activités. Par exemple, le bureau d'études le mieux qualifié pour réaliser une petite enquête sur les établissements de santé n'aura pas nécessairement les capacités ou l'expérience nécessaires pour réaliser une grande enquête sur les ménages. Les équipes d'EI devraient donc envisager une procédure d'adjudication séparée pour ces deux activités, de sorte à pouvoir identifier le bureau d'études le mieux qualifié pour chaque enquête. Il peut toutefois être avantageux de confier toutes les enquêtes à un même bureau d'études. Tableau 5 résume les avantages et les inconvénients de ces deux approches concernant les enquêtes sur les ménages et les établissements de santé.

Tableau 5 : Avantages et inconvénients de choisir un ou plusieurs bureaux d'études

Plusieurs bureaux d'études	Un bureau d'études
<ul style="list-style-type: none"> + Expérience spécifique et compétence plus pointue dans la réalisation de l'enquête considérée (par exemple, expérience des enquêtes de grande envergure sur les ménages ; expérience du contexte médical pour les enquêtes sur les établissements de santé) + Besoins de formation réduits car le bureau d'études est déjà spécialisé dans les enquêtes de grande envergure ou les enquêtes en contexte médical - Coordination plus difficile des deux enquêtes et des coûts de transaction - Coûts accrus - Besoins de supervision potentiellement plus importants 	<ul style="list-style-type: none"> + Meilleure coordination des activités et synergies entre les enquêtes sur les ménages et les établissements de santé + Formation et renforcement des capacités réalisés par l'équipe d'EI pour le premier type d'enquête contribuant à la qualité du second type d'enquête - Formation et renforcement des capacités plus intenses au début - Importants effectifs de terrain nécessaires pour réaliser les enquêtes simultanément + Faibles coûts de transport et de transaction en général + Appariement des ménages et des établissements de santé plus facile lors de la phase d'échantillonnage

Étude de cas nationale : Potentiel et défis associés au recrutement d'un unique bureau d'études
Projet de financement basé sur les résultats au Bénin

L'équipe a invité des bureaux d'études à participer à un appel d'offres concernant la réalisation des deux enquêtes sur les ménages et les établissements de santé. Il a été décidé de confier ces deux enquêtes à un même bureau d'études pour minimiser les coûts de transaction et de transport. L'adjudicataire était un bureau d'études local, sélectionné pour des raisons financières mais aussi pour renforcer les capacités locales. Ce bureau était doté de personnel qualifié mais manquait d'expérience en matière d'enquêtes de grande envergure. L'équipe focale de la Banque mondiale et son partenaire en matière d'assurance de la qualité (Sistemas) ont appuyé et formé les agents du bureau d'études pendant une phase pilote afin de veiller à la qualité des données. [...] L'équipe a supervisé la collecte des données sur les ménages à l'aide de contrôleurs externes. Ce travail a permis de collecter des données sur les ménages de bonne qualité.

L'enquête sur les établissements de santé n'a toutefois pas bénéficié de la même supervision. Le contenu de cette enquête était par ailleurs très technique et les enquêteurs durent administrer 15 instruments aux établissements de santé. La mesure de l'absentéisme a également nécessité deux visites, l'une annoncée, l'autre inopinée, ce qui a augmenté la complexité et le coût de l'enquête. Enfin, le bureau d'études n'avait pas une grande expérience des enquêtes sur les établissements de santé. L'équipe craint donc que la qualité des données sur les établissements de santé ne soit pas aussi bonne que celle sur les ménages et envisage de collecter de nouvelles données lors du suivi si nécessaire.

Pour de plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.

Étude de cas nationale : Choix lors du recrutement d'un bureau d'études
Financement basé sur les résultats dans le secteur de la santé en Zambie

[En Zambie], l'une des principales questions soulevées dans le cadre du recrutement du bureau d'études concernait l'existence de bureaux ayant une expérience dans la gestion de la collecte de données à grande échelle et de bons antécédents en matière de qualité des données fournies [...]. À l'issue d'un appel à concurrence géré dans le cadre du système de passation de marchés de la Banque, l'équipe a sélectionné un bureau d'études local ayant de bons antécédents. Toutefois, la réalisation simultanée sur plusieurs sites d'enquêtes sur les établissements et les ménages nécessitant une main-d'œuvre et une capacité excédentaire importantes, l'équipe a choisi de réaliser séparément l'enquête sur les établissements [...] et l'enquête sur les ménages. L'équipe de la Zambie a contribué au renforcement des capacités du bureau d'études local.

Bien que l'adoption d'une démarche visant à renforcer les capacités locales doive à terme produire des économies au niveau des programmes, elle n'est pas sans risque. Pour minimiser les risques potentiels, des contrôles de qualité et fiduciaires ont été intégrés dans cette démarche. Par exemple, les termes de référence du bureau d'études chargé de réaliser l'enquête sur les ménages stipulaient des objectifs d'étape devant être atteints avant le décaissement des fonds et des niveaux de qualité des données minimum avant de lui confier une enquête de suivi. Ces mesures incitatives ont joué un rôle essentiel dans la gestion de l'évaluation.

Les enquêtes sur les établissements et les ménages étant réalisées par deux entités différentes, il a donc été possible de fournir un soutien au bureau d'études concerné dans le cadre de la mise en œuvre des enquêtes sur le financement de la santé basé sur les résultats [...].

Pour de plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.

Comme indiqué précédemment, le processus de recrutement d'un bureau d'études par appel à la concurrence est long : la préparation et l'évaluation des soumissions peuvent prendre deux à trois mois, auxquels s'ajoutent deux à trois mois supplémentaires pour la négociation et la signature du contrat.

La liste ci-dessous répertorie les principaux critères à prendre en compte lors du recrutement d'un bureau d'études :

- Sélection du bureau d'études :
 - La collecte des données fera-t-elle l'objet d'un appel à la concurrence ? Dans l'affirmative, veillez à ce que le calendrier prévoie suffisamment de temps pour l'appel d'offres, l'examen des soumissions et la conclusion du contrat avec le bureau retenu.
 - Compte tenu des besoins de l'étude, lancerez-vous des appels à la concurrence distincts pour chaque activité de collecte de données (établissements de santé, ménages et données qualitatives) ?
- STAO ou saisie centralisée des données :
 - Comptez-vous saisir les données sur le terrain ? Dans l'affirmative, précisez-le dans les termes de référence car ce type de saisie a une incidence sur les qualifications, le matériel, la formation et le budget requis.
- Liste des ménages (ou des établissements de santé) :
 - Une liste existe-t-elle déjà ? Dans la négative, précisez-le dans les termes de référence car cet aspect a de fortes incidences sur le budget et le calendrier.
- Comité d'éthique :
 - Quelles sont les exigences locales en matière de travaux de recherche et quelles sont les obligations auxquelles doit satisfaire le bureau d'études ?
 - Le comité d'éthique prélève-t-il des droits ? Qui en assumera le coût ? S'il s'agit du bureau d'études, précisez-le dans ses termes de référence.
 - Qui doit présenter l'étude au comité (équipe d'EI et/ou bureau d'études) ? S'il s'agit du bureau d'études, précisez-le dans ses termes de référence.
- Supervision :
 - Quels sont les mécanismes d'établissement de rapports et de supervision que le bureau d'études devra utiliser ?
- Calendrier :
 - Quelles sont les contraintes de temps relatives à la collecte des données ?
 - Combien de temps faut-il à chaque équipe de terrain pour collecter des données dans une grappe spécifique ?

Effectifs des bureaux d'études

Nous recommandons que la collecte des données auprès des établissements de santé et des ménages soit envisagée comme deux activités distinctes, quel que soit le nombre de bureaux d'études retenus. Pour chaque enquête, le ou les bureaux d'études devront être disposés à former une équipe composée de responsables, d'enquêteurs et d'opérateurs de saisie.

Responsables des bureaux d'études. Il est important de noter que plusieurs responsables seront nécessaires pour superviser les différents aspects du travail, mais qu'ils devront tous participer à temps plein à toutes les phases de l'enquête. Les principaux cadres responsables sont :

- Responsable de projet : planifie, supervise et gère l'intégralité de l'enquête avec l'assistance des responsables de terrain et des données. Le responsable de projet doit être établi dans le pays pendant toute la durée de l'enquête.
- Responsable de terrain : planifie, supervise et gère le travail de terrain.
- Responsable des données : planifie, supervise et gère la saisie des données, la détection des erreurs, le traitement et la consolidation des données.

Il arrivera souvent que les bureaux d'études demandent une assistance technique substantielle à l'équipe d'EI et à l'expert de la qualité des données (voir Module 2). Il est important de veiller à ce que le bureau d'études comprenne dès le début que l'expert de la qualité des données pourra participer à tous les aspects de la collecte et de la saisie des données si l'équipe d'EI le juge nécessaire.

Des systèmes d'établissement de rapports périodiques entre le bureau d'études et l'équipe d'EI, et entre les équipes chargées des enquêtes, devront également être mis en place.

2012/02/22

Étude de cas nationale : Supervision des bureaux d'études et du travail de terrain **Financement basé sur les résultats dans le secteur de la santé en Zambie**

Par exemple [en Zambie], les termes de référence du bureau d'études stipulaient la soumission de rapports sur la mise en œuvre des enquêtes à mi-parcours et en fin de programme mais aussi lorsque des problèmes étaient rencontrés. Le besoin de rapports en flux tendus est apparu dans le cadre de l'évaluation du programme antipaludique lorsque les équipes de terrain se sont heurtées à l'hostilité des communautés en raison des perceptions et croyances liées aux prélèvements sanguins nécessaires dans le cadre du volet biomédical. Outre les exigences stipulées dans les termes de référence au sujet des rapports devant être fournis par le bureau d'études, un coordinateur d'EI reçoit des mises à jour de l'équipe de terrain toutes les semaines environ, ou plus fréquemment si nécessaire.

[Au sein des équipes chargées des enquêtes,] les questionnaires des enquêtes sur le financement de la santé basé sur les résultats sont contrôlés tous les jours et les enquêteurs font le point tous les soirs pour réduire les anomalies et la quantité de données manquantes, et donc les nouvelles visites, ce qui a également pour effet de réduire les frais de transport.

Pour de plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.

Équipes de terrain. La structure recommandée pour chaque équipe de terrain comprend :

1 superviseur d'équipe de terrain : Le superviseur de l'équipe de terrain devrait être chargé d'aider les enquêteurs à résoudre tous les problèmes rencontrés sur le terrain. Les enquêteurs devraient par conséquent transmettre toutes les questions, observations et réclamations au superviseur de l'équipe de terrain. Ceci permet de régler tous les problèmes de manière méthodique, de recevoir des réponses dans des délais raisonnables et d'aider le responsable central des activités de terrain à communiquer avec le superviseur de terrain plutôt qu'avec chaque membre de l'équipe. Le superviseur de l'équipe de

terrain devrait également être chargé d'évaluer le travail des enquêteurs de son équipe : il devra aléatoirement observer les entretiens avec les membres des ménages et se rendre à nouveau dans certains ménages choisis au hasard. Le superviseur de terrain devrait contrôler la présence d'erreurs dans les questionnaires à mesure de leur remplissage et veiller à ce que les enquêteurs retournent voir les ménages pour lesquels des données incomplètes ou erronées ont été collectées. Nous pensons que les superviseurs d'équipes de terrain jouent un rôle important dans la collecte de données de haute qualité et que leur performance devrait être évaluée par les responsables de projet en fonction de la qualité des données.

2-4 enquêteurs : Selon le contexte national et la nature des données, le bureau d'études pourra se trouver dans l'obligation de former des équipes comprenant des enquêteurs des deux sexes. La composition de l'équipe de terrain devrait être déterminée par les exigences en matière de collecte des données et la nature des questions sexospécifiques figurant dans le questionnaire.

1 anthropomètre/collecteur de biomarqueurs : Les équipes de terrain pourront également inclure une personne qualifiée et formée dans la collecte de données anthropométriques et/ou biologiques. Dans certains contextes et pays, cette personne devra avoir une formation d'infirmier ou de médecin.

1 opérateur de saisie : **Si le système STAO est utilisé pour la saisie des données**, un opérateur de saisie sera généralement nécessaire. Nous recommandons que cette personne soit chargée de saisir les données à partir des questionnaires, dès qu'ils lui parviennent du terrain, et de contrôler immédiatement les données incohérentes ou improbables. Le superviseur de l'équipe de terrain devrait vérifier ces contrôles et décider si le ménage ou l'établissement devrait faire l'objet d'une deuxième visite.

Les contraintes budgétaires et temporelles seront des facteurs importants lors de la détermination du nombre d'équipes de terrain nécessaires. Lorsque le nombre d'équipes de terrain est faible, les responsables de l'enquête auront peu de mal à les superviser et à s'assurer que les mesures de qualité des données sont correctement effectuées ; les équipes de terrain doivent toutefois être en nombre suffisant pour pouvoir réaliser l'enquête dans les délais impartis. Comme mentionné précédemment, il est essentiel que la collecte des données de référence ait lieu avant le début de l'intervention sur les groupes de traitement et que la collecte des données en fin de programme soit achevée avant que l'intervention ne soit intégrée dans le groupe de référence.

Calendrier des produits escomptés et des paiements

Pour minimiser les risques de trésorerie pendant la réalisation des enquêtes, les produits escomptés du bureau d'études devraient être clairement liés aux activités requises et suivre un calendrier prévoyant le décaissement des fonds à des étapes critiques de la préparation et de la mise en œuvre des activités d'évaluation, tout en minimisant le risque que le bureau abandonne le travail avant de l'avoir terminé. Tableau 6 fournit un exemple de calendrier des produits escomptés et des paiements devant figurer dans les **termes de référence du ou des bureaux d'études** :

Tableau 6 : Calendrier des produits escomptés et des paiements

Produit	Date	Paiement
Signature du contrat	Mois 0	10 %
Produit 1 : 1.1 Diagramme de Gantt détaillant toutes les activités proposées, les produits escomptés et les échéances 1.2 Protocole de recherche et formulaire de consentement éclairé adaptés 1.3 Justificatifs de l'obtention de l'autorisation éthique, de la couverture d'assurance et des permis nécessaires pour réaliser l'enquête	Mois 2	20 %
Produit 2.1 : Questionnaires adaptés en anglais Traduction initiale des questionnaires en langue locale	Mois 2	
Produit 2.2 : Rapport de pré-test précisant le calendrier des modules, les observations des enquêteurs et superviseurs et les modifications devant être apportées au questionnaire Questionnaires définitifs en langue locale Questionnaires définitifs correspondants en anglais	Mois 3	20 %
Produit 3 : Plan d'échantillonnage écrit approuvé par l'équipe d'évaluation	Mois 3	
Produit 4 : Protocole écrit sur la collecte des données sur les biomarqueurs approuvé par l'équipe d'EI	Mois 3	
Produit 5 : Plan de travail de terrain écrit approuvé par l'équipe d'EI	Mois 3	20 %
Produit 6 : 6.1 Protocole écrit de saisie des données pour les opérateurs de saisie détaillant le programme 6.2 Programme définitif de saisie des données adapté aux questionnaires locaux 6.3 Dictionnaire de l'ensemble de données avec toutes les variables désignées et définies	Mois 3	
Produit 7 : Liste des agents recrutés avec leur qualification	Mois 4	
Produit 8 : Approvisionnement et formation 8.1 Fournitures (matériel anthropométrique, GPS, collecte de données sur les biomarqueurs) 8.2 Matériels didactiques et manuels opérationnels 8.3 Rapport contenant les résultats des évaluations des enquêteurs	Mois 4	20 %
Produit 9 : Rapport final et données de l'étude pilote transférés avec succès à l'équipe d'évaluation	Mois 5	
Produit 10 : Rapport final écrit du responsable de projet sur la collecte des données de référence	Mois 9	30 %
Produit 11 : Bases de données et rapport final sur la fourniture des données	Mois 10	
Produit 12 : Fourniture en temps opportun des rapports bihebdomadaires du responsable de projet	10 mois	

Budget

Plusieurs facteurs influencent le **budget du ou des bureaux d'études** :

- Le niveau des salaires et des avantages
- Le recours à du personnel étranger (salaires et voyages)
- La taille de l'échantillon et la répartition géographique
- La durée anticipée des entretiens (établissements de santé, ménages et/ou communautés)

- La composition des équipes de terrain
- Le nombre d'équipes de terrain
- La durée du travail de terrain
- Le besoin d'acquérir du matériel (anthropométrie, identification des biomarqueurs).

Pour qu'un bureau d'études puisse estimer avec précision le budget de l'enquête dans sa proposition financière, les termes de référence utilisés dans la demande de propositions devraient contenir toutes les informations nécessaires pour identifier les besoins relatifs à la collecte des données ainsi que les facteurs précités. Le facteur temps peut dans certains cas être décisif car les données doivent être collectées avant la mise en œuvre ou la transposition à plus grande échelle de l'intervention de FBR. Les termes de référence devraient alors préciser les contraintes temporelles de sorte que le bureau d'études puisse estimer la composition et le nombre des équipes de terrain requises pour collecter les données dans les délais prescrits. En outre, les termes de référence devraient indiquer tous les matériels nécessaires à la collecte des données et préciser qui sera responsable de leur acquisition. Dans de nombreux pays, les matériels recommandés pour les mesures anthropométriques et l'identification des biomarqueurs doivent être obtenus à l'étranger, ce qui a pour effet d'augmenter les délais et les coûts encourus par le bureau d'études. Il peut toutefois arriver que ces matériels puissent être obtenus gratuitement auprès de donateurs ou de programmes nationaux (programme national antipaludique, par exemple). Quels que soient leurs origines et leurs prix, les coûts et le calendrier associés doivent figurer dans les termes de référence.

L'expérience montre que la majorité des budgets de collecte des données sont sous-estimés. Nous recommandons de :

- Définir les termes de référence aussi précisément que possible et fournir une description claire de toutes les activités requises de collecte de données
- Fournir un modèle de budget au bureau d'études pour éviter l'omission de postes de dépenses importants
- Demander à un expert de la collecte des données d'examiner les propositions techniques et financières pour détecter toute lacune importante
- Prévoir un poste de dépenses accessoires de 10 % pour éviter les dépassements de budget
- Vérifier soigneusement la méthodologie, le calendrier et le budget proposés pour l'enquête sur les ménages, le cas échéant. Les bureaux d'études ont malheureusement tendance à sous-estimer le budget de ce type d'activités.

Comprendre et adapter les instruments d'enquête

Les **instruments d'enquête FBR** et le **programme de saisie des données** correspondant ont été élaborés pour saisir une quantité de données considérable au niveau des établissements de santé, des agents de santé, des ménages et des personnes. Cette section fournit un aperçu du format de base de ces questionnaires, des recommandations en vue de leur adaptation et traduction, ainsi que des descriptions détaillées des instruments d'enquête (échantillons, participants, calendrier et contenu).

Remarque : Chaque évaluation d'impact du réseau de FBR a un cadre, des questions et un contexte qui lui sont propres. Il n'est donc pas possible de préparer un ensemble unique ou uniforme d'instruments de collecte de données. Les instruments de FBR sont fournis aux équipes à titre de ressource ; nous espérons que l'utilisation par les équipes d'instruments communs permettra de développer les savoirs collectifs sur la mesure des impacts des FBR et de mettre en commun les expériences des uns et des autres.

Directives universelles en matière de format

Le format des questionnaires joue un rôle déterminant dans la qualité des données ; nous recommandons donc que les questionnaires soient conçus conformément à des directives universelles en matière de format. Nous conseillons vivement que toutes les modifications apportées aux questionnaires soient conformes à ces directives, de sorte à éviter les incohérences entre modules, évaluations nationales ou équipes.

Texte en MAJUSCULES et minuscules. Les instructions suivantes devraient figurer dans le manuel de l'enquêteur et être mises à la disposition du bureau d'études et des enquêteurs.

- Texte en minuscules : Le texte en minuscule doit être lu à haute voix et est utilisé pour formuler les questions de l'entretien et énumérer toutes les réponses possibles lorsqu'il est précisé que l'enquêteur doit les lire à haute voix.
- Le TEXTE EN MAJUSCULES ne doit pas être lu à haute voix par l'enquêteur ; il est utilisé pour lui donner des instructions, le cas échéant, ou pour énumérer les réponses possibles, qui ne doivent pas être lues à haute voix par l'enquêteur et doivent uniquement être consignées si elles sont mentionnées par la personne interrogée.

Codes de réponse. Nous recommandons que tous les codes de réponse utilisent un format à deux chiffres. Le tableau ci-dessous donne un exemple de codes utilisés dans un questionnaire.

(13.36) Qui avez-vous vu pour les soins prénatals lors de cette grossesse ? (SI PLUS D'UN, CHOISIR LE PRESTATAIRE DE SOINS PRIMAIRES)	
DOCTEUR EN MÉDECINE	01
INFIRMIÈRE/SAGE-FEMME	02
AGENT DE SANTÉ COMMUNAUTAIRE	03
TECHNICIEN DE LABORATOIRE	04
PHARMACIEN	05
GUÉRISSEUR TRADITIONNEL	06
GUÉRISSEUR SPIRITUEL	07
ASSISTANT TRADITIONNEL À LA NAISSANCE	08
MEMBRE DE LA FAMILLE	09
AMI/VOISIN	10
AUTRE (SPÉCIFIEZ)	96

Bien que nous recommandions que toutes les réponses utilisent un code à deux chiffres, la réponse Oui est généralement représentée par le chiffre 1 et la réponse Non par le chiffre 2.

Réponses « Ne sait pas » et « Refus ». La majorité des questions du questionnaire ne prévoit pas d'option pour une réponse du type « Ne sait pas ». En effet, l'expérience montre que les enquêteurs auxquels cette option est proposée dans le cadre d'enquêtes sur les ménages insistent moins pour obtenir une réponse des participants. Il peut bien entendu arriver que le participant ne connaisse pas la réponse à une question ; l'enquêteur devra alors écrire en toutes lettres « Ne sait pas » sur le questionnaire. De même, lorsque les participants refusent de répondre à une question, nous recommandons que les enquêteurs écrivent « Refus » sur le questionnaire. Les opérateurs de saisie utiliseront des codes prédéfinis pour consigner ces réponses au moment de la saisie des données. Le programme de saisie de données sur les FBR utilise -7 (moins 7) pour « Ne sait pas » et -8 (moins 8) pour « Refus ».

Nous recommandons que l'équipe d'EI identifie, pendant la phase pilote ou le pré-test, les questions produisant un nombre exceptionnellement élevé de réponses de ce type et i) révise le libellé de ces questions ou ii) supprime ces questions. Un grand nombre de réponses de ce type peut être directement corrélé à la qualité des entretiens et, in fine, de la collecte des données par le bureau d'études.

Réponse « Autre ». Nous recommandons fortement d'utiliser le code 96 pour la réponse « Autre » par souci de cohérence entre l'ensemble des questions et modules et avec les enquêtes antérieures. En outre, toutes les réponses « Autre » devraient être consignées par les enquêteurs en utilisant scrupuleusement les mots et la diction de la personne interrogée. Les modules de saisie des données prévoient un espace pour consigner les « autres réponses » ; ces réponses devraient être examinées par les agents d'assurance de la qualité pour vérifier s'il s'agit en fait de réponses prévues, auquel cas le code correspondant devrait être utilisé.

Adaptation des instruments d'enquête de FBR

2012/02/22

Étude de cas nationale : Adaptation et mise à l'essai des questionnaires Financement basé sur les résultats dans le secteur de la santé en Zambie

[...] Chaque projet étant unique, le contenu des produits liés à l'enquête a dû être adapté. L'équipe sur le financement de la santé basé sur les résultats a utilisé des instruments élaborés pour un projet de lutte contre le paludisme en Zambie pour l'enquête sur les ménages, et des instruments élaborés par l'équipe d'EI de la Zambie avec le soutien de l'équipe d'évaluation de la Banque pour l'enquête sur les établissements de santé. Les essais de terrain, qui ont duré environ six mois, ont permis d'identifier les principaux ajustements devant être apportés. En particulier, le questionnaire, initialement administré en quatre heures, a été réduit à une heure et demie en supprimant les sections superflues ou difficiles à administrer des livres socioéconomiques et sanitaires, y compris les questions sur la biométrie et l'activité physique. Malgré le temps consacré à ce travail d'adaptation, les coûts ont été inférieurs à ceux qu'aurait occasionnés la préparation d'un nouvel instrument.

Pour de plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.

Le contenu manifestement spécifique à un pays apparaît en rouge dans les instruments d'enquête sur les FBR. Les équipes d'EI peuvent remanier plus profondément les instruments d'enquête sur les FBR pour les adapter au contexte local mais doivent être conscientes du risque d'incohérence au sein de leurs questionnaires et avec les données collectées dans d'autres pays.

Inclusion de questions pertinentes pour répondre aux questions de recherche. En cas d'adaptation des questionnaires, il est important de garder à l'esprit que les instruments ont pour objet de mesurer l'impact de l'intervention sur les extrants et les résultats étudiés. Il est donc impératif de veiller à ce que :

- Les instruments les plus appropriés sont utilisés pour mesurer les indicateurs étudiés.
- Dans chaque instrument, les questions et réponses qui permettront de calculer les indicateurs étudiés sont toutes incluses et administrées aux personnes appropriées et concernent le calendrier approprié (voir la liste des indicateurs sur les FBR donnée au 0).

Dans le reste de cette section, nous formulons des recommandations pour trois types de modifications que les équipes nationales pourraient devoir apporter au questionnaire : i) ajout de questions spécifiques au pays ; ii) ajout et/ou suppression de réponses aux questions existantes ; et iii) suppression de questions sans intérêt. Ces modifications devraient être codées de sorte à indiquer qu'il s'agit d'ajouts locaux au questionnaire.

Ajout de questions spécifiques au pays. En cas d'adaptation des instruments d'enquête au contexte national, les questions peuvent être codées à l'aide d'une nomenclature à deux lettres spécifique au pays, par exemple :

AF – Afghanistan
BJ – Bénin
ZR – République démocratique du Congo
GH – Ghana
KG – Kirghizistan
RW – Rwanda
ZM – Zambie
ZW – Zimbabwe

Pour une liste complète des pays, voir <http://cds.worldbank.org/Pages/CntryGroup.aspx>.

Par exemple, pour ajouter trois questions après la question 10 du questionnaire, l'équipe du Rwanda pourra utiliser les codes RW10A, RW10B et RW10C. Ainsi, la numérotation des questions ultérieures ne sera pas modifiée et le programme de saisie des données et les fichiers de STATA pourront être utilisés.

Les questionnaires comprennent des sauts automatiques de questions qui permettent à l'enquêteur d'omettre des questions jugées sans intérêt sur la base des réponses préalablement reçues. Lorsque les équipes nationales ajoutent des questions au questionnaire, nous recommandons qu'elles vérifient que les sauts de questions figurant dans la section où ces ajouts ont été apportés ne sont pas perturbés et que les questions seront posées à tous les participants auxquels elles s'adressent. Il sera généralement plus simple pour les équipes d'ajouter de nouvelles questions dans leurs propres sections, à la fin du module/questionnaire, pour éviter les problèmes associés à la modification de sections existantes.

Ajout de codes pour des réponses spécifiques à un pays. Lorsque les modifications concernent des réponses aux questions, des codes manifestement différents de ceux des réponses standard existantes devront être affectés à ces nouvelles réponses. Les codes des réponses prévues dans la version internationale ne devraient pas être modifiés, même si certaines réponses sont supprimées. Il s'ensuit que les codes des réponses ne sont pas nécessairement séquentiels. En règle générale, les nouveaux codes utilisés pour des réponses spécifiques aux pays utilisent des chiffres supérieurs à 50. Les codes qui dénotent des conventions de nommage différentes dans différents pays peuvent être écrasés. Par exemple, si le code utilisé pour les médecins est 01 alors qu'ils sont désignés docteurs, l'équipe utilisera le code 01 pour les docteurs. L'utilisation de normes de codage semblables permet aux équipes de veiller à ce que les données de différents pays soient facilement comparables.

Par exemple : Pour ajouter deux nouvelles réponses possibles spécifiques à la Zambie pour la question 14.13 « Où avez-vous cherché à faire soigner la maladie de [VOTRE ENFANT] ? », on utilisera les codes 58 et 59 : Kantemba – 58 et Vendeur de médicaments – 59.

(14.13) Où avez-vous (le répondant) cherché à faire soigner la maladie de... ?	
HÔPITAL GOUVERNEMENTAL	01
CLINIQUE GOUVERNEMENTALE	02
DISPENSAIRE GOUVERNEMENTAL	03
HÔPITAL PRIVÉ	04
CLINIQUE PRIVÉE	05
PHARMACIE	07
GUÉRISSEUR TRADITIONNEL	09
GUÉRISSEUR PAR LA PRIÈRE/D'ÉGLISE	10
AGENT DE SANTÉ COMMUNAUTAIRE	11
KANTEMBA	58
VENDEUR DE MÉDICAMENTS	59
AUTRE, SPÉCIFIEZ	96

Suppression de questions sans intérêt. Les questionnaires standard contiennent un certain nombre de questions optionnelles. Il est par ailleurs possible que d'autres questions ne présentent pas d'intérêt compte tenu du contexte national ou ne correspondent pas aux questions de recherche concernant le pays. Nous recommandons que l'enquêteur principal décide si les questions sont pertinentes ou appropriées et, dans le cas contraire, les supprime du questionnaire. Lorsqu'une question est supprimée, nous recommandons de conserver, dans la mesure du possible, la numérotation des questions suivantes figurant dans le questionnaire standard, par souci de cohérence entre les pays.

Étude de cas nationale : Adaptation des questionnaires d'enquête Crédits d'investissement pour le programme de santé publique du Nigéria

Pour préparer son enquête de référence, l'équipe du Nigéria a utilisé les questionnaires du HRITF en les adaptant au contexte local et aux questions de recherche. Ce travail a nécessité une collaboration et des échanges entre les membres de l'équipe, notamment entre l'enquêteur principal, le coordonnateur de l'EI et l'assistant de recherche. Il est donc utile de pouvoir repérer les modifications apportées. L'équipe du Nigéria a utilisé le système suivant pour attirer l'attention sur les modifications apportées aux questionnaires dans le cadre du processus d'adaptation :

- Les cellules surlignées en jaune doivent être supprimées
- Les cellules surlignées en vert ont été modifiées par rapport au questionnaire original du HRITF
- Les cellules surlignées en bleu ont été ajoutées au questionnaire original du HRITF
- Les cellules surlignées en rouge méritent une attention particulière
- L'équipe a utilisé des commentaires pour donner des précisions sur ces modifications et attirer l'attention sur les questions en suspens.

Les copies d'écran ci-dessous illustrent le processus d'adaptation des questionnaires.

Figure 5 : Copies d'écran des questionnaires de la Boîte à outils pendant le processus d'adaptation nationale

(1) Informations générales					
(A) Généralités		NOTER LA RÉPONSE			
PERSONNE INTERROGÉE : RESPONSABLE DE L'ÉTABLISSEMENT DE SANTÉ OU SON ADJOINT SI ABSENT OU INDISPONIBLE					
(1.01)	Cet établissement est-il sous votre responsabilité aujourd'hui ?	OUI	1	Coordinateur Terrain: Peut être supprimé- il est plus probable que les enquêteurs trouvent les OICs eux-mêmes lors de l'enquête	
		NON	2		
(1.02)	Êtes-vous habilité à représenter cet établissement ?	OUI	1	Coordinateur Terrain: Supprimé car pas pertinent (surligné en jaune): (1) Responsable clinique (2) Assistant infirmier (3) CDE Changé les noms pour (surligné en vert): (1) Administrateur d'hôpital (2) Technologue Santé environnementale (3) Technicien Pharmacie (4) Technologiste de laboratoire Ajouté (surligné en bleu): (1) RCC (2) EASC (3) EASCJ	
		NON	2		
(1.03)	Quel est l'intitulé de votre poste dans cet établissement ?	Médecin ou responsable médical	01	Coordinateur Terrain: Devrions-nous mettre une catégorie distincte pour (1) Responsable infirmier en chef pour CSP plus grands et hôpitaux généraux? OUI (2) catégorie combinée infirmier et sage-femme OUI (3) responsables dossiers médicaux OUI (4) assistants de santé OUI (même si improbable que (3) et (4) répondraient au questionnaire)	
		Clinicien	02		
		Secrétaire d'Hôpital	03		
		Infirmier	04		
		Sage-femme	05		
		Pharmacien	06		
		Responsable Santé environnementale	07		
		Aide-soignant	08		
		Assistant Pharmacien	09		
		Technologue/Scientifique de Laboratoire	10		
		Technicien/Assistant Laboratoire	11		
		Employé journalier classifié	12		
		Responsable Santé Communautaire (RCC)	51		
		Employé additionnel Santé Communautaire (EASC)	52		
		Employé additionnel Santé Communautaire Junior	53		
		Autre, spécifier:	96		
(1.04)	Cet établissement est-il un hôpital de district, un centre de santé ou un poste de santé ?	Hôpital général	01	Coordinateur Terrain: Renommé de l'hôpital de district MAIS: Qu'en est-il des hôpitaux de campagne? AJOUTER UNE CATEGORIE	
		Centre de Santé Primaire Secondaire	51		

(4) Liste du personnel de l'établissement						
PERSONNES CONCERNÉES : TOUS LES MEMBRES DU PERSONNEL, QU'ILS SOIENT TEMPORAIRES OU PERMANENTS, SOIGNANTS OU NON, À COMMENCER PAR LE RESPONSABLE DE L'ÉTABLISSEMENT						
PERSONNE INTERROGÉE : RESPONSABLE DE L'ÉTABLISSEMENT OU EMPLOYÉ LE MIEUX INFORMÉ						
Coordinateur Terrain: A vérifier						
(4.01)	(4.02)	(4.03)	(4.04)	(4.05)	(4.05)	
C O D E D · I D E N T I F I C A T I O N	LISTEZ LES NOMS COMPLETS DE TOUT LE PERSONNEL TRAVAILLANT DANS LA FORMATION. POUR CHAQUE EMPLOYÉ, POSEZ TOUTES LES QUESTIONS DE CETTE SECTION, PUIS PASSEZ À L'EMPLOYÉ SUIVANT. S'IL Y A PLUS DE 15 EMPLOYÉS, UTILISEZ UN QUESTIONNAIRE DIFFÉRENT.	SEXE DE [NOM]	CODE D'ID. DU RÉPONDANT	Âge de [NOM]	Quel est le diplôme le plus élevé obtenu par [NOM] ?	Quel poste occupe [NOM] dans cet établissement ?
					Professions médicales/paramédicales	Professions médicales/paramédicales
					Médecin ou responsable n° 01	Responsable Santé Communautaire (f 51
					Clinicien 02	Employé additionnel Santé Communautaire 52
					Secrétaire d'Hôpital 03	Employé additionnel Santé Communautaire 53
					Infirmier 04	Responsable Dossiers Médicaux 54
					Sage-femme 05	Coordinateur Terrain: clinique ou non?
					Pharmacien 06	Autres
					Responsable Santé environnementale 07	Assistants Santé >1 year 14 ► 0.00
					Aide-soignant 08	Assistants Santé <1 year 15 ► 0.00
					Assistant Pharmacien 09	Assistant social 16 ► 0.00
					Technologue/Scientifique d 10	Conseiller 17 ► 0.00
					Technicien de laboratoire 11	Personnel administratif 13 ► 0.00
Employé journalier classifié 12	Autre poste non médical/paramédical 16 ► 0.00					
Autre poste médical/paramédical 13	Coordinateur Terrain: Changé de personnel auxiliaire					
NOM ET PRÉNOM	HOMMES 01 FEMMES 02	ANNÉES				
01						
02						

Traduction et rétro-translation

Lorsque les instruments anglais ont été adaptés au contexte local, le bureau d'études devrait les traduire dans le ou les langues locales requises. Pour vérifier que la traduction est fidèle, chaque version en langue locale peut être rétro-traduite par une équipe de traduction indépendante qui n'a pas participé à la traduction initiale. Cette activité devrait figurer dans les termes de référence et le budget du bureau d'études et être assortie d'un calendrier. Nous estimons que la traduction et la rétro-translation prennent chacune entre deux et trois semaines.

Une **macro Excel de traduction** est fournie dans la boîte à outils pour faciliter la traduction des questionnaires formatés.

Instruments d'enquête

Questionnaires sur les établissements de santé

Les questionnaires sur les établissements de santé ont été conçus pour fournir des données primaires sur la prestation de services, les structures opérationnelles, la qualité des processus, les ressources humaines et les infrastructures. Une enquête sur les établissements de santé implique la visite de tous les établissements recensés dans le plan d'échantillonnage et la collecte de données connexes. Des

méthodes spécialement conçues, telles que l'examen des dossiers, l'observation des relations prestataire-client et l'utilisation de patients normalisés, peuvent ajouter une valeur considérable à l'évaluation des établissements. De plus, les données collectées à partir de l'examen des dossiers et des inventaires sur le personnel peuvent être utilisées pour valider les statistiques administratives courantes sur le volume de services fournis et la disponibilité et la répartition géographique des ressources humaines, telles que les données disponibles dans le SIGS. Elles peuvent également augmenter les coûts et la complexité de la collecte des données.

Évaluation des établissements :

Échantillon. L'évaluation des établissements devrait être réalisée pour tous les établissements de santé de l'échantillon tel que défini dans le plan d'échantillonnage.

Personne interrogée. La personne interrogée lors de l'évaluation d'un établissement de santé devrait être son directeur ou responsable. Bien que certaines sections puissent prescrire un suivi auprès des responsables des services comptables, de la pharmacie et du laboratoire, les enquêteurs devraient entamer l'évaluation de l'établissement avec son directeur. Celui-ci peut ensuite déterminer si la participation d'autres agents est nécessaire pour des sections spécifiques.

Durée. Nous estimons que l'évaluation d'un établissement de santé prend environ **un jour**. L'entretien devrait être programmé en tenant compte des heures de service, pour minimiser les perturbations occasionnées pour les patients, et de la disponibilité du directeur. Il est vivement recommandé que les équipes de terrain contactent l'établissement de santé à l'avance pour s'assurer de la présence de son directeur le jour de l'entretien.

Contenu. L'instrument des enquêtes sur les établissements de santé comprend 15 sections :

1. *Généralités et précautions universelles* : Cette section présente un aperçu de l'infrastructure, des heures de service, des services d'orientation des patients et des aspects financiers des établissements. Dans le contexte des FBR, ces données sont importantes pour comprendre comment la direction des établissements réagit à un régime incitatif (agrandissement des chambres des patients, extension des heures de service pour augmenter l'utilisation, réallocation du budget pour améliorer la qualité de l'établissement, etc.).
2. *Administration et gestion* : Cette section collecte des données sur la gestion de l'établissement : relations avec la communauté, plan d'activités, encadrement, évaluation interne et externe et planification budgétaire. Elle donne également une idée de l'autonomie de l'établissement et de ses pouvoirs en matière de fourniture de médicaments et de décisions de gestion.
3. *Ressources humaines* : Cette section présente un aperçu des ressources humaines de l'établissement, y compris des récentes recrues, de la rotation du personnel, des postes vacants, de la formation et de la collaboration avec les agents de santé communautaires. Elle fournit également une liste du personnel avec les caractéristiques des principaux agents.
4. *Liste* : Cette section collecte des données sur tous les agents de l'établissement, telles que leur niveau d'études, leurs qualifications et leur charge de travail.

5. *Services de laboratoire* : Cette section collecte des données sur les services de laboratoire disponibles au sein de l'établissement.
6. *Services* : Cette section collecte des données sur les services (vaccination, soins prénatals, accouchements, soins postnatals, traitements antituberculeux et antipaludiques) offerts par l'établissement de santé et les protocoles connexes suivis.
7. *SIGS* : Cette section collecte des données de base du SIGS concernant la composition de la population desservie et des patients.
8. *Utilisation des services de santé* : Cette section collecte des données sur l'utilisation des services de santé à partir des données du SIGS.
9. *Participation financière des usagers* : Cette section collecte des informations sur la participation financière des usagers, la transparence concernant les honoraires, les exceptions, les pouvoirs en matière de fixation des honoraires, ainsi que la manière dont les recettes qui en découlent sont utilisées au sein de l'établissement. L'un des impacts potentiels des FBR est la réduction des débours directs des patients afin d'augmenter les taux d'utilisation. Ces données compléteront les informations sur la population pour définir l'impact sur la participation financière des usagers.
10. *Encadrement* : Cette section collecte des données sur le type d'encadrement fourni par le directeur de l'établissement de santé.
11. *Autorité* : Cette section collecte des informations sur le niveau d'autonomie décisionnelle au niveau de l'établissement de santé.
12. *Observation directe* : Les enquêteurs consignent dans cette section leurs observations concernant l'état général de l'établissement et l'affichage des honoraires et des protocoles nationaux.
13. *Matériel* : Cette section collecte des données sur le matériel disponible et opérationnel de l'établissement, en particulier le matériel nécessaire pour fournir les principaux services de santé maternelle et infantile. Ces données sont utilisées pour déterminer la qualité structurelle de l'établissement.
14. *Approvisionnement en médicaments* : Cette section collecte des données sur les stocks de médicaments de base (utilisation générale, paludisme, planning familial, tuberculose, obstétrique, vaccination) au cours des 30 derniers jours. Ces données sont également utilisées pour déterminer la qualité structurelle de l'établissement.
15. *Zone de desserte* : Cette section collecte des données nominatives supplémentaires sur les villages inclus dans la zone de desserte.

Entretien avec les agents de santé

Échantillon. L'échantillon dépendra des objectifs de l'EI et en particulier du type de services étudiés (soins prénatals, infantiles, adultes, etc.). Par exemple, si l'évaluation entend mesurer l'impact sur la qualité des soins prénatals, l'enquête devra inclure au moins un agent de santé fournissant ce type de service.

Personne interrogée. Si un seul agent fournissant le service étudié est présent le jour de l'entretien, cet agent devra être sélectionné pour l'entretien. Par contre, si plusieurs agents fournissant le service étudié sont présents le jour de l'entretien, l'équipe de terrain devra en choisir un au hasard. Nous recommandons que les équipes de terrain vérifient si certains services sont offerts le ou les jours prévus pour l'entretien, pour s'assurer qu'au moins un agent travaillant dans le domaine étudié sera présent le jour J.

Durée. Nous estimons que l'entretien avec l'agent de santé prend entre 45 et 60 minutes. L'entretien devrait être programmé de sorte à minimiser les perturbations occasionnées pour les patients.

Contenu. L'entretien avec l'agent de santé comprend 13 sections :

1. *Informations générales* : Cette section collecte des données démographiques de base sur le poste, l'expérience et les responsabilités de l'agent de santé.
2. *Formation* : Cette section collecte des données sur la formation reçue par l'agent de santé dans les principaux domaines sanitaires au cours de la dernière année (ou plus), ainsi que sur ses besoins de formation.
3. *Heures travaillées* : Cette section présente le détail des heures et des jours travaillés, ainsi que les raisons des absences.
4. *Salaires* : Cette section donne des détails sur les salaires mensuels et la régularité des paiements ainsi que les possibilités d'emploi et les salaires associés hors du secteur sanitaire.
5. *Autre compensation* : Cette section donne des détails sur la valeur des autres formes de compensation : voyages, hébergement, primes d'isolement et autres.
6. *Supervision* : Cette section collecte des informations sur les activités de supervision au sein de l'établissement de santé, les retours d'information du superviseur à l'agent de santé et les contributions du superviseur à l'environnement de travail.
7. *Emploi secondaire* : Cette section collecte des informations sur tout emploi secondaire occupé par l'agent de santé, ainsi que sur les revenus supplémentaires issus de ce travail.
8. *Bien-être* : Cette section collecte des informations sur le bien-être général de l'agent de santé.
9. *Satisfaction (optionnel)* : Cette section collecte des données sur la satisfaction de l'agent de santé concernant divers aspects de ses conditions de travail.
10. *Motivation personnelle (optionnel)* : Cette section collecte des données sur la motivation de l'agent de santé concernant divers aspects de ses conditions de travail.
11. *Innovation (optionnel)* : Cette section collecte des informations sur la capacité de l'agent à s'adapter aux changements pouvant intervenir dans l'établissement et/ou dans la communauté située à proximité de l'établissement.
12. *Vignettes* : Les vignettes mesurent les connaissances grâce à des études de cas. Elles décrivent une situation où un patient se présente avec des symptômes ou une affection spécifiques. Le prestataire est invité à dire ce qu'il ferait pour s'occuper du patient. L'accent est mis sur l'évaluation des connaissances médicales de l'agent de santé.
13. *Vignette* pour un protocole de soins prénatals.

Pour plus d'informations sur les méthodes de mesure de la qualité des soins à l'aide de vignettes, voir la section ci-dessous sur la mesure de la qualité des soins.

Remarque : L'entretien avec l'agent de santé durera probablement plus de 45 minutes s'il examine l'intégralité des 13 sections précédemment décrites. Les équipes ne devraient pas hésiter à modifier le contenu des instruments de collecte de données et à insérer ou supprimer des modules de sorte que le questionnaire final corresponde aux questions de recherche proposées.

Entretien avec les patients à la sortie de l'établissement

Échantillon. Comme pour l'entretien avec l'agent de santé, les services étudiés détermineront les types de patients devant être sondés à la sortie de l'établissement (soins prénatals, soins infantiles ou autres). En règle générale, les enquêteurs sondent entre 8 et 12 patients par service étudié.

Personne interrogée. Si le nombre de patients présents le jour de l'entretien est inférieur ou égal à 12, les enquêteurs peuvent tous les sonder. Toutefois, si le nombre de patients recevant le ou les services étudiés est important (supérieur à 15), l'équipe de terrain devrait sélectionner au hasard entre 8 et 12 patients. Il est vivement recommandé que les équipes de terrain vérifient si certains services sont offerts le ou les jours prévus pour l'entretien, pour s'assurer que le nombre de patients correspondant aux services étudiés sera suffisant le jour J.

Durée. L'entretien avec les patients à la sortie de l'établissement devrait durer environ **20 à 30 minutes**.

Contenu. Les deux modèles d'entretien avec les patients comprennent chacun huit sections :

1. *Identification à la sortie* : Cette section collecte des données sur l'établissement de santé (pour les contrôles par recoupement), le niveau d'études et la situation de famille du patient (ou de la personne qui en a la charge).
2. *Traitement et conseils* : Cette section collecte des données essentielles à l'analyse de la qualité des soins. Ces données reflètent les efforts déployés par le prestataire dans le cadre des consultations de santé maternelle et infantile et fournissent une liste de contrôle de toutes les questions posées, des examens et tests de laboratoire réalisés, des médicaments et des conseils fournis pendant la consultation.
3. *Temps et dépenses* : Cette section collecte des données supplémentaires sur la qualité des soins, notamment les temps d'attente et le temps passé avec le prestataire. Elle collecte également des informations sur les honoraires versés pendant la consultation.
4. *Satisfaction* : Cette section collecte des données sur la satisfaction du patient concernant les divers aspects de sa consultation.
5. *Sécurité et confiance* : Cette section collecte des données sur les sentiments de sécurité et de confiance du patient concernant les divers aspects de sa consultation et les alentours de l'établissement de santé.
6. *Ménages 1 et 2* : Cette section collecte des données sur la catégorie socioéconomique du patient, y compris ses biens fonciers et avoirs financiers et la structure de son ménage.

7. *Agent de santé communautaire* : Cette section collecte des informations sur la présence d'agents de santé communautaires et les services qu'ils fournissent à la communauté. Elle ne s'applique qu'à certains contextes.
8. *Assistante traditionnelle à la naissance (optionnel)* : Cette section collecte des informations sur la présence d'assistantes traditionnelles à la naissance et les services qu'elles fournissent à la communauté. Elle ne s'applique qu'à certains contextes.

Remarque : L'entretien avec le patient durera probablement plus de 30 minutes s'il examine l'intégralité des sections précédemment décrites. Les données collectées sur la satisfaction des patients indiquent en moyenne une satisfaction très élevée, ne présentent pas de grande variabilité et éliminent implicitement les patients qui ne se rendent pas à l'établissement du fait de leur insatisfaction concernant les services fournis. Cette section est donc optionnelle. Par ailleurs, en fonction des questions de recherche, l'équipe pourra envisager de réduire le volume de données collectées aux sections 5 (sécurité et confiance) et 6 (ménages 1 et 2).

Mesure de la qualité des soins

Plusieurs méthodes ont été utilisées pour mesurer les connaissances des prestataires et la qualité des soins fournis (Franco, Daly et al. 1997 ; Hermida, Nicholas et al. 1999 ; Peabody, Luck et al. 2000 ; Bessinger et Bertrand 2001 ; Franco, Franco et al. 2002 ; Leonard et Masatu 2005 ; Leonard et Masatu 2006 ; Das et Hammer 2007).

Tableau 7 présente un récapitulatif de ces méthodes, en comparant leurs validité, fiabilité, faisabilité et coûts relatifs.

Tableau 7 : Résumés des méthodes d'évaluation des connaissances des prestataires et de la qualité des soins fournis¹⁴

Méthode	Forces	Faiblesses	Fiabilité	Faisabilité	Coûts relatifs
Connaissances du prestataire					
Entretien avec le prestataire — connaissances, attitudes	Mesure les connaissances pratiques et les perceptions du prestataire	Effet Hawthorne – Modification du comportement ou de la performance de l'agent de santé découlant de sa prise de conscience qu'il participe à une étude expérimentale ; Ne mesure pas la pratique	Fiable pour les perceptions et les connaissances du prestataire (mais utilise un grand nombre de questions mal formulées)	Peut être réalisé sur de grands échantillons et dans un large éventail de cas, y compris des événements rares	Faible
Vignette	Mesure les connaissances pratiques et la prise de décision	Effet Hawthorne – Modification du comportement ou de la performance de l'agent de santé découlant de sa prise de conscience qu'il participe à une étude expérimentale ; Ne mesure pas la pratique	Variable en fonction de l'enquêteur ; la fiabilité peut être faible (mais bonne dans certaines circonstances)	Dépend de la qualité de l'enquêteur, utilise de petits échantillons et une faible variation au niveau des cas, quelques événements rares	Nécessite des enquêteurs ayant des qualifications et une formation d'un niveau plus élevé
Prestation de services de santé					
Observation des soins	Mesure la qualité fournie	Effet Hawthorne – Modification du comportement ou de la performance de l'agent de santé découlant de sa prise de conscience qu'il participe à une étude expérimentale ; Mesure les soins optimums fournis	Bonne – Problèmes limités entre évaluateurs mais mal documentée	Nécessite un nombre de cas important et/ou des affections courantes ; plus intrusive que les autres méthodes	Formation limitée (par ex. 1 semaine) d'agents de santé alphabétisés ou d'étudiants en santé ; coûts de supervision

¹⁴ Gupta et Peters (2010). Voir l'outil « **3.01a Indicateurs RBF** »

Méthode	Forces	Faiblesses	Fiabilité	Faisabilité	Coûts relatifs
Entretien de sortie – Interrogatoire du malade et examen physique	Effet Hawthorne plus faible que l'observation si l'étude n'est pas portée à la connaissance du prestataire	Ne mesure pas la séquence de l'examen physique	Mauvaise mémorisation du patient ou de la personne qui en a la charge : faible corrélation avec les antécédents documentés	Peut être appliquée à un grand échantillon mais plus adaptée aux affections courantes	Formation limitée (par ex. 1 semaine) d'agents de santé alphabétisés ou d'étudiants en santé ; coûts de supervision
Entretien de sortie – Conseils, perceptions et caractéristiques des patients	Mesure ce que le soignant comprend ; Mesure les perceptions immédiates ; Peut mesurer les principales caractéristiques (par ex. richesse) qui peuvent être comparées aux paramètres statistiques d'une population	Les perceptions sur le lieu des soins ne correspondent pas nécessairement à celles mesurées ultérieurement	Bonne – problème limité entre évaluateurs ; certaines questions ont montré une bonne corrélation avec les antécédents	Peut être appliquée à un grand échantillon pour les grands établissements, mais plus adaptée aux affections courantes	Formation limitée (par ex. 1 semaine) d'agents de santé alphabétisés ou d'étudiants en santé ; coûts de supervision
Clients simulés (pseudopatients)	Effet Hawthorne le plus faible avec un enquêteur « anonyme »	Faible vraisemblance pour les enfants et les femmes sur le point d'accoucher Impossible pour les femmes enceintes et les enfants malades (problèmes éthiques)	Variable selon l'acteur	Faible taille de l'échantillon ; éventail très limité des affections étudiées	Surcoûts possibles dus au niveau plus élevé des qualifications et de la formation
Examen des dossiers des patients	Si les dossiers sont bien tenus, peut refléter les soins devant être fournis ; préférable pour le diagnostic et le traitement (s'ils sont systématiquement consignés)	Les dossiers sont rarement bien tenus ; le peu d'information sur les tâches réalisées reflète les dires des prestataires au sujet des soins plutôt que les soins effectivement fournis	Mauvaise, en particulier lorsque la tenue des dossiers est hors normes	Non intrusive et rapide mais la qualité des dossiers est rarement suffisante pour qu'ils puissent être utilisés à des fins autres que la quantification du volume des services	Peu coûteuse

La boîte à outils pour l'évaluation d'impact contient trois types de vignettes :

1. Les vignettes d'études de cas (préparées par Shivam Gupta et David Peters, de l'Université Johns Hopkins) : Dans ces vignettes, l'enquêteur évoque le cas d'un patient présentant des symptômes particuliers et demande au prestataire d'énumérer toutes les actions et prescriptions qu'il jugerait utiles pour fournir le meilleur traitement possible à ce patient. L'enquêteur ne fournit généralement pas d'autres informations au prestataire après lui avoir présenté le cas. Toutefois, dans l'une des vignettes, le cas se présente en deux parties. Cette vignette se trouve sur la feuille 12 du questionnaire de l'agent de santé.
2. Vignettes sur le protocole de soins prénatals : Cette vignette contrôle les connaissances du prestataire concernant le protocole à suivre lors de la première visite prénatale d'une femme enceinte. Le scénario ne fournit pas d'information sur des symptômes autres que la grossesse visible à l'œil nu. Cette vignette se trouve sur la feuille 13 du questionnaire de l'agent de santé.
3. Vignettes Banque mondiale-ISERDD (créées par Jishnu Das et Jeffrey S. Jammer, en collaboration avec l'Institut de recherche socioéconomique sur la démocratie et le développement (Institute for Socio-Economic Research on Democracy and Development, ISERDD) (Das et Hammer, 2004 ; Das et Hammer, 2007) : Dans ces vignettes, l'enquêteur évoque le cas d'un patient présentant des symptômes particuliers. Il demande ensuite au prestataire de le soigner comme s'il était le patient. Le prestataire peut poser des questions à l'enquêteur-patient pour établir un diagnostic et peut également proposer des examens. L'enquêteur-patient répond aux questions du prestataire et lui donne les résultats des examens proposés. Le prestataire est censé établir un diagnostic et proposer un traitement. L'enquêteur peut poser de nouvelles questions au prestataire sur le traitement proposé. Veuillez noter que ces vignettes ne figurent pas dans le questionnaire de l'agent de santé mais plutôt dans le **dossier du questionnaire destiné aux établissements de santé**.

Outre les vignettes, la boîte à outils comprend un module d'observation directe, également élaboré par Jishnu Das et Jeffrey S. Jammer, en collaboration avec l'Institut de recherche socioéconomique sur la démocratie et le développement (ISERDD) (Das et Hammer, 2004 ; Das et Hammer, 2007). Ce module très court permet de consigner des renseignements de base sur les activités du prestataire de santé pendant la visite, en précisant si des questions sur les antécédents ont été posées, si le prestataire a réalisé des examens physiques, demandé des tests ou prescrit des médicaments. Ce module a été élaboré dans le contexte indien, où la majorité des visites sont très brèves. Les équipes devraient envisager de l'adapter à leur contexte national.

Questionnaires destinés aux ménages

Les enquêtes sur les ménages permettent à l'équipe d'EI de mesurer les résultats au niveau de la population. L'un des principaux risques associés à un projet de FBR est que les prestataires exagèrent les résultats obtenus. Bien qu'un FBR opérationnel doive être doté de systèmes de vérification efficaces, l'utilisation de données primaires indépendantes recueillies au niveau de la population permet à l'évaluation d'impact de contrôler si les effets du programme de FBR sont ressentis au niveau de la population cible.

Sujets au niveau des ménages

Les équipes de terrain devraient remplir un questionnaire destiné aux ménages pour chaque ménage sélectionné pour l'enquête dans le plan d'échantillonnage.

Personne interrogée. La principale personne devant répondre à ce questionnaire est le membre du ménage le mieux informé. Il peut s'agir du chef de ménage et/ou de son conjoint. Cette personne peut toutefois demander l'aide d'autres membres du ménage sur des questions spécifiques.

Durée. On estime que l'entretien avec les ménages prendra entre **60 et 90 minutes** en fonction du nombre de modules inclus. La personne interrogée pouvant avoir d'autres responsabilités, l'équipe de terrain pourra prévoir deux visites et revenir plus tard dans la journée ou le lendemain pour terminer l'entretien.

Contenu. L'enquête sur les ménages a été élaborée avec le concours de plusieurs équipes d'EI et comprend les sections suivantes :

1. *Liste* : Cette section collecte des données démographiques de base (âge, date de naissance, situation de famille et niveau d'études des parents) pour tous les membres du ménage.
2. *Éducation* : Cette section collecte des données sur les études achevées, l'assiduité et le temps consacré pour tous les membres du ménage âgés de 5 ans et plus.
3. *Travail* : Cette section collecte des données sur les emplois primaire et secondaire (rémunérés ou non) ainsi que les autres sources de rémunération (assurance, allocations de chômage ou de retraite) de tous les membres du ménage âgés de 12 ans et plus.
4. *Logement* : Cette section collecte des données sur tous les aspects du logement : plancher, toit et murs, eau et assainissement, combustibles et loyer.
5. *Biens* : Cette section collecte des données sur les avoirs financiers et la valeur des biens fonciers et matériels du ménage, sans oublier les animaux.
6. *Revenus* : Cette section collecte des données sur toutes les sources de revenu possibles (investissements, loyers, bourses, envois de fonds de l'étranger, héritages, etc.).
7. *Consommation* : Cette section collecte des données sur toutes les catégories de consommation (alimentation, produits consommables et durables, etc.) rapportées à la semaine, au mois et à l'année.
8. *Mortalité* : Cette section collecte des données sur tous les décès intervenus dans le ménage au cours des 12 derniers mois, en précisant leur cause.
9. *État de santé et utilisation* : Cette section collecte des données sur la morbidité de tous les membres du ménage. Ces données sont complétées par une section plus détaillée figurant dans les questionnaires sur la santé maternelle et infantile.
19. *Coordonnées* : Cette section collecte des données sur les moyens de contacter le ménage pour obtenir des informations complémentaires ou pour réaliser une enquête de suivi.

Remarque : L'entretien avec le ménage durera probablement plus de 90 minutes s'il examine l'intégralité des neuf sections précédemment décrites. En cas de contraintes de temps importantes, nous proposons que l'équipe envisage en premier lieu de réduire les données collectées aux sections 5 (biens), 6 (revenus), 7 (consommation) et/ou 8 (mortalité). On gardera toutefois à l'esprit que les informations sur les biens, les transferts et la consommation fournissent des indicateurs précieux sur la pauvreté, qui sont particulièrement importants pour réaliser une analyse de l'équité.

Modules sur la santé maternelle et infantile

Sujets. Toutes les femmes âgées de 15 à 49 ans faisant partie des ménages sélectionnés pour l'étude devraient être interrogées dans le cadre des modules sur la santé maternelle et au moins un entretien sur la santé maternelle devrait être effectué dans chaque ménage. Tous les enfants âgés de 0 à 5 ans faisant partie des ménages sélectionnés pour l'étude.

Personne interrogée. La ou les personnes interrogées dans le cadre des modules sur la santé maternelle sont la ou les femmes du ménage âgées de 15 à 49 ans. L'équipe ne devrait pas accepter de réponses indirectes, c'est-à-dire qu'un membre du ménage réponde pour le compte d'un autre. La personne interrogée concernant les enfants de 0 à 5 ans est celle qui en a la charge.

Durée. Nous estimons que l'entretien sur la santé maternelle prendra **60 minutes** pour chaque femme du ménage. La personne interrogée pouvant avoir d'autres responsabilités, l'équipe de terrain pourra prévoir deux visites et revenir plus tard dans la journée ou le lendemain pour terminer l'entretien. Nous estimons que l'entretien sur la santé infantile prendra **30 minutes** par enfant. La personne interrogée pouvant avoir d'autres responsabilités, l'équipe de terrain pourra prévoir deux visites et revenir plus tard dans la journée ou le lendemain pour terminer l'entretien.

Contenu. L'instrument sur la santé maternelle comprend les sections suivantes :

10. *Activités de la vie quotidienne* : Cette section collecte des données sur les activités de la vie quotidienne et mesure la capacité physique d'une personne à réaliser des activités quotidiennes courantes.
11. *Santé mentale* : Cette section collecte des données pour évaluer la santé mentale d'une femme et tout traitement récent pour dépression et/ou anxiété.
12. *Antécédents de grossesse* : Cette section collecte des données sur les grossesses antérieures, y compris les naissances vivantes, les fausses couches et les morts nés, ainsi que sur tous les enfants vivants ou non de la femme interrogée.
13. *Soins anténatals et postnatals* : Cette vaste section collecte des données sur les soins prénatals, obstétricaux et postnatals reçus au cours des deux années précédentes (cette durée peut être modifiée en fonction des objectifs de l'évaluation d'impact nationale). Dans la majorité des projets de FBR, les soins prénatals, obstétricaux et postnatals constituent d'importants indicateurs de succès. Cette section collecte des données sur l'utilisation des services et la qualité des soins (mesurée par le respect du protocole national par le prestataire).

14. *Santé génésique* : Cette section collecte des données sur le désir d'une femme d'avoir plus d'enfants ainsi que sur ses antécédents et pratiques actuelles en matière de contraception.
15. *Vaccination* : Cette section collecte des données sur les vaccinations de l'enfant, qu'elles aient été administrées dans l'établissement de santé ou dans le cadre de campagnes de santé communautaire.
16. *Anthropométrie* : Cette section collecte des données permettant de déterminer l'état nutritionnel de l'enfant en mesurant sa taille et son poids. Ces données sont utilisées pour calculer la valeur du Z de l'enfant. Dans certains contextes, ces données sont également collectées pour les femmes.
17. *Autres biomarqueurs* : Cette section collecte des données issues d'autres tests d'identification des biomarqueurs, tels que le paludisme et l'anémie, qui peuvent être réalisés dans le cadre de l'enquête sur les ménages. Dans certains contextes, ces données sont également collectées pour les femmes.
18. *Agents de santé communautaires* : Cette section collecte des données sur la satisfaction de la femme concernant les services prodigués par l'agent de santé communautaire. Elle peut être adaptée ou supprimée en fonction du contexte national. Dans certains pays, l'adoption du FBR a augmenté la collaboration entre les établissements de santé et les agents de santé communautaires afin d'encourager la demande de services essentiels.

Remarque : L'entretien sur la santé maternelle durera probablement plus de 60 minutes s'il examine l'intégralité des neuf sections précédemment décrites. En cas de contraintes de temps importantes, nous proposons que l'équipe envisage en premier lieu de réduire les données collectées aux sections 10 (activités de la vie quotidienne), 11 (santé mentale) et 18 (satisfaction). La section 14 (santé génésique) pourra également être réduite en fonction des objectifs de l'évaluation. En fonction des contraintes temporelles et budgétaires, l'équipe d'EI pourra décider de collecter des données sur les biomarqueurs pour un sous-échantillon d'enfants et/ou de femmes, ou de supprimer complètement les sections 16 et 17. La suppression ou la réduction de la section 12 (antécédents de grossesse) serait délicate car les informations collectées permettent de déterminer qui interroger à la section 13 (soins anténatals et postnatals). La section 13 est essentielle pour calculer les indicateurs sur les soins maternels.

Questionnaires destinés aux communautés

Sujets. Les équipes peuvent réaliser un module auprès des communautés de chaque village retenu pour l'étude.

Personnes interrogées. Les personnes interrogées dans le cadre des modules communautaires doivent être des responsables communautaires ou des administrateurs locaux ayant une bonne connaissance de leur communauté.

Durée. L'entretien réalisé dans les communautés devrait prendre **45 minutes**.

Contenu. L'instrument du module communautaire comprend dix sections :

1. *Observation directe* : Cette section demande à l'enquêteur de collecter des données provenant de l'observation directe de la communauté, notamment sur l'assainissement, la propreté, les logements et la topographie.
2. *Composition* : Cette section collecte des données sur la composition du groupe interrogé au cas où plusieurs responsables ou représentants communautaires doivent participer à l'entretien.
3. *Démographie* : Cette section collecte des données sur le nombre d'habitations et de membres de la communauté, ainsi que sur la religion, la langue et l'appartenance ethnique les plus répandues.
4. *Services de base* : Cette section collecte des données sur l'accès de la communauté aux services de base : établissements de santé, écoles, routes, marchés, eau et l'assainissement, etc.
5. *Capital social* : Cette section collecte des données sur les organisations communautaires et leur fréquentation, ainsi que les droits des femmes en matière de biens fonciers et de succession.
6. *Activités économiques* : Cette section collecte des données sur les principales activités économiques des membres de la communauté.
7. *Chocs externes* : Cette section collecte des données sur les principaux chocs externes subis par la communauté au cours des dix dernières années : inondations, tremblements de terre, sécheresses, maladies, etc.
8. *Programmes* : Cette section collecte des données sur les programmes sociaux menés dans la communauté au cours des trois dernières années : santé, éducation, eau et assainissement, accès au crédit, etc.
9. *Prix* : Cette section collecte des données sur les prix des produits alimentaires et les coûts des services de santé et de l'éducation.
10. *Coûts* : Cette section collecte des données sur le coût du projet de FBR pour la communauté.

Remarque : Bien que les données communautaires permettent aux équipes d'étudier des caractéristiques observables supplémentaires dans le cadre de leur analyse, les groupes de traitement et de référence devraient présenter des caractéristiques équilibrées du fait de la randomisation. Ce module est inclus à titre de référence mais ne sera pas toujours nécessaire.



Module 5

Conduite de la collecte des données

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats
Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

www.worldbank.org/health/impacetevaluationtoolkit

Module 5. Conduite de la collecte des données

Principales recommandations et outils disponibles pour ce module			
Recommandations	Essentiel	Important	Souhaitable
<ul style="list-style-type: none"> • L'équipe d'évaluation d'impact et le bureau d'études doivent définir le protocole pour identifier clairement les observations dans les bases de données et pour établir des liens entre ces dernières. • L'équipe d'évaluation d'impact doit définir le protocole pour identifier les domaines de traitement et de comparaison au sein des bases de données. • La qualité et la durée de la formation des équipes de terrain jouent un rôle essentiel dans l'assurance du succès de la collecte de données. • Les bureaux d'études s'occupent de la collecte des données, mais il appartient à l'équipe d'EI de collaborer avec ces derniers pour veiller à ce que les résultats du travail de terrain soient communiqués en temps utile et d'une manière appropriée. • Le protocole de recherche et les manuels d'enquête doivent contenir toutes les informations dont les bureaux d'études auront besoin pour veiller à ce que la collecte des données soit effectuée conformément aux plans établis et dans le respect des principes d'éthique reconnus. • Il convient de veiller à la sécurité et à la confidentialité des données pendant les opérations de collecte et de saisie. Les équipes de terrain doivent signaler tout problème logistique ou tout accroc à la sécurité. • L'équipe d'EI doit surveiller étroitement les opérations de collecte et de saisie des données, et pourrait décider de faire appel aux services d'un expert de la qualité des données pour mieux s'acquitter de cette tâche. • Les bureaux d'études locaux risquent de manquer de moyens en matière de programmation de la saisie des données, de saisie et de gestion des données. La boîte à outils propose des formulaires de saisie des données pour le programme CS-Pro qui correspondent aux questionnaires d'enquêtes auprès des ménages et des établissements de santé. • Il est préférable d'entrer les données à mesure que le travail sur le terrain progresse, au lieu d'attendre à la fin de ce travail. 	<p style="text-align: center;">✓</p>	<p style="text-align: center;">✓</p> <p style="text-align: center;">✓</p>	<p style="text-align: center;">✓</p>



Outils
<ul style="list-style-type: none"> • 5.01 Interview Duration Tracking Sheet (Traçage de la durée des entretiens) • 5.02 Enumerator Evaluation Form (Formulaire d'évaluation de l'enquêteur) • 5.03 Survey Progress Report I (Word) (Rapport sur l'avancement de l'enquête I (Word)) • 5.04 Survey Progress Report II (Excel) (Rapport sur l'avancement de l'enquête II (Excel)) • 5.05a Household survey Field Manual (Manuel de terrain pour les enquêtes ménages) • 5.05b Household survey Training Program (Programme de formation aux enquêtes ménages) • 5.05c Household survey Training PPTs (PPTs de Formation aux enquêtes ménages) • 5.06a Health Facility Survey Field Manual (Manuel de terrain pour les enquêtes formations sanitaires) • 5.06b Health Facility Survey Training PPTs (Formation aux enquêtes pour les formations sanitaires) • 5.07 Survey Training, CAR & Cameroon Examples (Formation aux enquêtes, Exemples de RCA & Cameroun) • 5.08 Health Facility Supervisor Checklist (Liste de contrôle du superviseur formations sanitaires)

- 5.09 Health Facility Arrival Checklist (Liste de contrôle d'arrivée en formation sanitaire)
- 5.10 Health Facility Supervisor Tracking Form (Formulaire de traçage pour le superviseur formation sanitaire)
- 5.11a Daily Listing of Under 5 Exit Interviews (Liste quotidienne des entretiens de sortie des enfants de moins de 5 ans)
- 5.11b Daily Listing of ANC Exit Interviews (Liste quotidienne des entretiens de sortie des soins prénataux)
- 5.12 Cash Management Sheet (Fiche de suivi des décaissements)

Contenu du module

Cadre d'échantillonnage et échantillon.....	5-4
Établissements de santé	5-4
Agents de santé	5-5
Ménages.....	5-5
Définition de codes d'identification uniques.....	5-8
Codes d'identification géographique des bases de données brutes	5-9
Codes d'identification de terrain pour la collecte et le partage des données.....	5-11
Fichier de contrôle de l'identification : Mise en correspondance des codes géographiques et des codes de terrain	5-12
Identificateurs des groupes de traitement et de comparaison.....	5-13
Essai préalable des questionnaires.....	5-14
Gestion de la saisie des données.....	5-15
Risques sur le plan de la validité interne de l'EI pendant la collecte des données	5-18
Plan de travail sur le terrain	5-19
Gestion des équipes de terrain.....	5-20
Collecte de données auprès des établissements de santé	5-20
Collecte de données auprès des ménages.....	5-21
Recrutement et formation des membres des équipes de terrain.....	5-21
Programme de formation et matériaux.....	5-22
Pilotage des questionnaires.....	5-23
Gestion du travail sur le terrain.....	5-24
Préparation des rapports.....	5-24

Ce module présente un aperçu des diverses étapes de la collecte des données. Il aborde toutes les étapes qui doivent être franchies après l'entrée en jeu du bureau d'études : i) définition du cadre d'échantillonnage et de l'échantillon ; ii) définition des identificateurs uniques ; iii) définition des identificateurs des groupes de traitement et de comparaison ; iv) essai préalable des questionnaires ; v) planification et gestion de la saisie des données ; vi) planification du travail de terrain ; vii) recrutement et formation des membres des équipes de terrain ; viii) test pilote ; ix) gestion du travail de terrain ; x) préparation du rapport.

Cadre d'échantillonnage et échantillon

Lors de l'étape de la conception de l'évaluation d'impact (voir Module 3), l'équipe d'EI devrait avoir défini :

- i) l'unité de randomisation : le niveau auquel le traitement doit être appliqué, et les groupes de comparaison formés ;
- ii) les critères d'inclusion : utilisés pour décider si une unité doit ou non être incluse dans le cadre d'échantillonnage. Les critères d'inclusion sont critiques puisqu'ils peuvent conduire à d'importantes variations de la puissance de l'étude, même avec des échantillons de taille égale. Les critères d'inclusion définissent la population à l'étude ;
- iii) le nombre d'unités comprises dans chacune des branches de l'étude et la taille de l'échantillon entier à chaque niveau (établissements de santé, agents de santé, femmes ou enfants).

Pendant la préparation de la collecte des données, l'équipe d'EI devra définir *d'abord, le cadre d'échantillonnage, et ensuite l'échantillon*. Rappelons que le cadre d'échantillonnage est la liste des unités à partir de laquelle l'échantillon sera prélevé. Il devrait être représentatif de la population à l'étude. Il s'agit essentiellement d'une liste physique ou électronique complète des unités qui pourront éventuellement faire partie de l'échantillon. L'échantillon est un sous-ensemble d'unités prélevées dans la liste du cadre d'échantillonnage ; il contient les unités qui devront être interviewées par le bureau d'études.

Dans la plupart des évaluations du FBR, la randomisation sera effectuée à un niveau géographique (par exemple, districts, secteurs, départements, etc.). Les cadres d'échantillonnage de l'ensemble des unités de rang inférieur (par exemple, établissements de santé, agents de santé, etc.) seront limités aux unités géographiques qui appartiennent à une branche de l'évaluation — traitement ou comparaison.

Établissements de santé

Si le ministère de la Santé est en mesure de fournir une liste complète des établissements de santé, l'équipe d'EI devrait pouvoir en extraire le cadre d'échantillonnage en utilisant les critères d'inclusion

préalablement établis — par exemple, « Tous les établissements publics ou privés situés dans les districts qui appartiennent à une des branches de l'évaluation ». L'équipe d'EI et le bureau d'études procéderont ensuite au prélèvement de l'échantillon à partir du cadre d'échantillonnage. Dans beaucoup de pays, il sera nécessaire d'inclure l'ensemble des établissements de santé compris dans le cadre d'échantillonnage afin d'assurer à l'évaluation une puissance suffisante. Dans certains cas, il pourrait arriver que les groupes de traitement et de comparaison renferment un nombre d'établissements de santé plus élevé que ce qui est requis pour obtenir la puissance suffisante. L'équipe d'EI pourra alors choisir de prélever un échantillon aléatoire.

2012/04/03

Étude de cas nationale : Cadre d'échantillonnage et échantillon
Programme communautaire rwandais de financement basé sur la performance

Dans le cas du programme communautaire rwandais de FBP, le *cadre d'échantillonnage* des établissements de santé incluait tous les établissements situés dans les 200 secteurs retenus aux fins de l'évaluation. Pour maximiser la puissance, l'*échantillon* d'établissements de santé incluait l'ensemble des établissements compris dans le cadre d'échantillonnage.

Agents de santé

Le bureau d'études procédera d'abord à la définition du cadre d'échantillonnage des agents de santé. Par exemple, si le critère d'inclusion retenu est « Agents de santé chargés de la prestation des soins prénatals le jour de l'entrevue de l'établissement », il devra établir la liste de l'ensemble de ces agents de santé. Il lui faudra ensuite déterminer la manière de prélever l'échantillon — par exemple, « Choisir au hasard deux agents chargés de la prestation des services prénatals le jour de l'entrevue de l'établissement de santé ».

2012/04/03

Étude de cas nationale : Critères de sélection pour le cadre d'échantillonnage
Programme communautaire rwandais de financement basé sur la performance

Le programme communautaire rwandais de FBP prévoit l'assignation de quatre agents de santé communautaire à chaque village : i) un coordonnateur de la santé maternelle et infantile ; ii) un coordonnateur des affaires sociales ; iii) un coordonnateur de la lutte contre le VIH/sida ; iv) un coordonnateur de l'administration. Le critère d'inclusion utilisé pour le cadre d'échantillonnage des agents de santé communautaires en poste dans les villages des 200 secteurs retenus était le suivant : « Agent de santé communautaire responsable de la prestation des services de santé maternelle et infantile, et un autre agent choisi au hasard, sans égard au sexe ou au niveau du poste ».

Ménages

La plupart des évaluations utiliseront pour les ménages un critère de sélection du type suivant : « Ménages comptant des enfants de moins de X ans situés dans la zone d'intervention d'un établissement de santé échantillonné ». Malgré la simplicité de ce critère, l'établissement du cadre d'échantillonnage au niveau des ménages peut prendre passablement de temps puisqu'il s'agit d'établir

la liste de tous les ménages de la zone d'intervention. Or, dans la plupart des cas, ces ménages sont nombreux et il n'existe habituellement aucune liste aisément disponible.

2012/04/03

Étude de cas nationale : Critères de sélection pour le cadre d'échantillonnage
Programme communautaire rwandais de financement basé sur la performance

Dans le cas du programme communautaire rwandais de FBP, le critère d'inclusion du cadre d'échantillonnage pour les ménages situés dans des villages appartenant à l'un des 200 secteurs à l'étude était le suivant : « Le ménage qui a connu la naissance la plus récente (au moins un enfant de 0 à 4 mois) ».

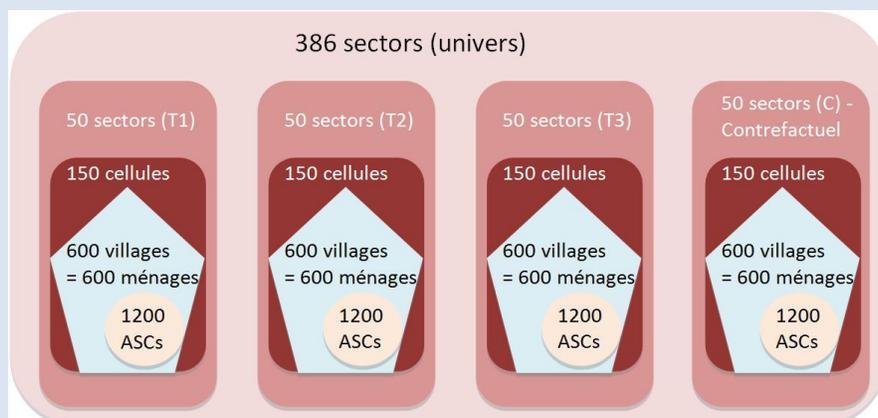
Au lieu d'établir la liste de tous les ménages situés dans la zone d'intervention d'un établissement donné, on peut également procéder à un échantillonnage aléatoire en deux étapes, en procédant comme suit :

i) déterminer les unités primaires d'échantillonnage (UPE) qui constituent la zone d'intervention de l'établissement de santé. Il devra s'agir de préférence de la plus petite unité géographique — par exemple, un village ou un îlot de recensement ;

ii) choisir au hasard les UPE de chaque zone d'intervention d'établissement de santé. Lorsque la zone d'intervention peut être définie à partir des données disponibles au *niveau central*, il est possible d'établir la liste des UPE comprises dans chaque zone d'intervention. Cependant, beaucoup de pays ne disposent pas de telles données au niveau central, et il faut alors obtenir les informations sur la zone d'intervention auprès des employés de *chacun des établissements de santé*. Dans un tel cas, il pourrait s'avérer nécessaire de procéder à l'échantillonnage des ménages dans la foulée des enquêtes auprès des établissements de santé : le bureau d'études peut être chargé d'établir la liste de toutes les UPE comprises dans la zone d'intervention de chaque établissement et de transmettre cette information à l'équipe d'EI, qui pourra ensuite prélever un échantillon aléatoire d'UPE à partir de cette liste.

Étude de cas nationale : Sélection aléatoire des unités primaires d'échantillonnage
Programme communautaire rwandais de financement basé sur la performance

Dans le cas du programme communautaire rwandais de FBP, les calculs de puissance ont démontré qu'il fallait retenir 12 ménages dans chacun des 50 secteurs de chacune des branches de l'étude => 50 secteurs*12 ménages = 600 ménages par branche de l'étude. La sélection des UPE a été effectuée au niveau central puisque l'équipe d'EI a pu obtenir la liste de toutes les cellules et villages compris dans chacun des secteurs en s'adressant au Bureau de la santé communautaire du ministère de la Santé. À partir de ces données, trois cellules ont été sélectionnées dans chacun des 200 secteurs afin de minimiser les décaissements sur le plan géographique à l'intérieur des secteurs. Quatre villages ont ensuite été sélectionnés dans chacune des cellules. On a ainsi obtenu un échantillon aléatoire de 12 villages par secteur.



iii) Établir la liste de tous les ménages compris dans les UPE échantillonnées — la « liste des ménages ».

Si le pays dispose de données de recensement récentes qui peuvent être liées aux UPE des établissements de santé, l'équipe d'EI sera en mesure d'en extraire la liste des ménages de chaque UPE échantillonnée, et d'établir ainsi le cadre d'échantillonnage des ménages. Cependant, il est rare que de telles informations soient disponibles, et le bureau d'études se verra donc souvent contraint d'établir la liste de tous les ménages qui répondent aux critères d'inclusion et qui se trouvent dans les UPE échantillonnées. Les listes complètes de ménages devraient contenir les informations de base sur les ménages et sur leurs membres — y compris leur âge et leur sexe, pour veiller à obtenir un cadre d'échantillonnage à partir duquel pourra être sélectionné un ménage admissible conforme aux critères d'inclusion. La liste de ménages peut être établie avant ou pendant la collecte des données sur les ménages, mais il importe dans un cas comme dans l'autre de contrôler correctement le processus pour veiller à ce que cette liste soit représentative de la population à l'étude. Nous avons également recommandé de ne pas confier à la même équipe l'enquête auprès des ménages et l'établissement de la liste des ménages. Le cumul de ces deux responsabilités risque de pousser l'équipe à exclure de la liste les ménages compris dans l'UPE, mais qui sont difficiles d'accès, ce qui peut entraîner une sous-représentation de ces ménages dans l'échantillon retenu pour l'enquête.

Étude de cas nationale : Établissement des listes de ménages
Programme communautaire rwandais de financement basé sur la performance

Dans le cas du programme communautaire rwandais de FBP, l'équipe d'EI n'avait pas accès à des données de recensement récentes. Elle comptait au départ établir une liste des ménages de chacun des villages sélectionnés. Cependant, des contraintes de temps et d'argent ont empêché l'équipe et le bureau d'études de procéder de cette façon. Au lieu d'établir une liste complète de tous les villages aux fins du dépistage des naissances les plus récentes (0 à 4 mois), le responsable des études de terrain a rencontré les chefs et les agents de santé communautaire de chacun des villages pour recenser les ménages qui avaient connu une naissance au cours des quatre derniers mois (en moyenne, on dénombre deux naissances par mois par village). Cette stratégie d'échantillonnage était vulnérable aux risques de l'aléa moral, et obligeait le responsable du projet et l'équipe d'EI à exercer une surveillance rigoureuse.

iv) Choisir au hasard les ménages à partir de la liste. Lorsque le bureau d'études aura achevé de dresser la liste des ménages compris dans chacune des UPE, l'équipe d'EI sera en mesure de choisir au hasard dans cette liste les ménages qui feront l'objet de l'enquête.

Définition de codes d'identification uniques

Pendant la collecte des données, le bureau d'études recueillera divers éléments d'information auprès des établissements de santé, des collectivités, des ménages et des particuliers, et les équipes de terrain se rendront à divers endroits, souvent à plusieurs reprises. Par ailleurs, plusieurs personnes interviendront dans le traitement des questionnaires, des données et des échantillons biologiques prélevés, et quelques années après la collecte des données de référence, un autre bureau d'études (ou le même) souhaitera retourner aux mêmes endroits pour recueillir une nouvelle série de données. De plus, les analystes souhaitent être en mesure de lier les diverses sources de données. Par exemple, l'information sur un ménage n'est utile que dans la mesure où on peut déterminer dans quelle zone d'intervention il se trouve. En conséquence, les données devront être organisées de telle manière que quiconque est appelé à les utiliser sera en mesure de retracer correctement l'origine de chaque réponse, sans toutefois violer les dispositions de confidentialité.

L'organisation adéquate des données d'enquêtes (ou autres) s'appuie sur l'utilisation cohérente de codes d'identification pendant toute la durée des opérations de collecte et de saisie des données. Il est ainsi possible de retracer l'origine de toutes les informations et de les relier, peu importe leur provenance. L'établissement du protocole des codes d'identification relève normalement de l'enquêteur principal, qui doit d'ordinaire collaborer à cette fin avec le bureau d'études. À défaut d'un protocole clair en cette matière, l'enquêteur principal ne sera pas en mesure d'achever l'élaboration des questionnaires, et le bureau d'études ne sera pas en mesure de planifier le travail de terrain. Pendant les études de base sur le terrain, le bureau d'études appliquera le protocole des codes d'identification

pour étiqueter tous les questionnaires imprimés, les échantillons biologiques, etc. Les mêmes codes devront être utilisés lors des enquêtes de suivi.

Il existe deux types de codes d'identification, chacun ayant son utilité propre.

- Les codes d'identification géographique permettent d'indiquer l'emplacement exact de chacune des unités incluses dans la base de données. Ils doivent rester confidentiels puisqu'ils ne protègent pas l'anonymat des répondants.
- Les codes d'identification de terrain permettent d'identifier précisément chaque unité sur laquelle des données sont recueillies, qu'il s'agisse d'une personne, d'un ménage, d'un village ou d'un établissement de santé, mais ils ne fournissent pas d'information sur l'emplacement physique exact des unités. On peut les utiliser dans les bases de données qui sont partagées et utilisés aux fins des analyses.

Les deux types de codes précités sont utilisés dans les questionnaires imprimés originaux.

Voici quelles sont les caractéristiques des bons codes d'identification :

- identification précise de l'unité d'observation : toutes les informations relatives à une personne doivent porter le même code d'identification, qu'il s'agisse de ses réponses à un questionnaire ou d'échantillons sanguins ;
- codes numériques, c'est-à-dire, ne contenant aucune lettre ni aucun caractère spécial ; les codes ne doivent contenir que des chiffres.

Codes d'identification géographique des bases de données brutes

Les bases de données brutes utilisent habituellement des codes d'identification géographique qui permettent d'identifier chaque emplacement. Chaque unité géographique (région, province, district, village...) devrait posséder son propre code, composé selon un ordre hiérarchique, de l'unité géographique la plus vaste à la plus petite, de manière que même les lieux les plus petits possèdent leur propre code. Le code peut être fondé sur les informations de santé existantes ou sur les données de recensement afin de maximiser la compatibilité entre les bases de données d'évaluation d'impact et les sources de données existantes.

Étude de cas nationale : Définition des codes géographiques
Programme communautaire rwandais de financement basé sur la performance

Le Rwanda est géographiquement divisé en provinces, districts, secteurs, cellules et villages. Pour les besoins de l'évaluation du programme communautaire de FBP, le secteur fait office d'unité de randomisation, et l'échantillon compte un total de 200 secteurs. Les enquêtes recueillent des données sur les villages, ainsi que sur les ménages et sur les agents de santé. La plus petite unité géographique est le village, et la base de données contient donc des codes d'identification géographique correspondant aux provinces, districts, secteurs, cellules et villages.

Geo-Codes									
province	prov_code	district	dist_code	sector	sect_code	cell	cell_code	village	village_code
SUD	2	NYANZA	201	Busasamana	20101	Gahondo	2010101	Kamatovu	201010103
SUD	2	NYANZA	201	Busasamana	20101	Gahondo	2010101	Karama	201010104
SUD	2	NYANZA	201	Busasamana	20101	Kavumu	2010102	Akirabo	201010201
SUD	2	NYANZA	201	Busasamana	20101	Kibinja	2010103	Kabuzuru	201010301
SUD	2	NYANZA	201	Busasamana	20101	Kibinja	2010103	Ngorongari	201010304
SUD	2	NYANZA	201	Busasamana	20101	Kibinja	2010103	Rugari A	201010307
SUD	2	NYANZA	201	Busasamana	20101	Kibinja	2010103	Rugari B	201010308
SUD	2	NYANZA	201	Busasamana	20101	Nyanza	2010104	Gatunguru	201010406
SUD	2	NYANZA	201	Busasamana	20101	Nyanza	2010104	Kavumu	201010408
SUD	2	NYANZA	201	Busasamana	20101	Nyanza	2010104	Kivumu	201010410
SUD	2	NYANZA	201	Busasamana	20101	Nyanza	2010104	Rubona	201010414
SUD	2	NYANZA	201	Busasamana	20101	Rwesero	2010105	Mwima	201010507

Les enquêteurs principaux ont besoin de connaître l'emplacement géographique des unités échantillonnées aux fins de l'enquête (district, établissement de santé, etc.) afin de pouvoir produire les données de panel, mais **ces codes d'identification peuvent facilement compromettre la confidentialité des réponses s'ils sont partagés**. Supposons par exemple, qu'une base de données contient des informations sur les ménages, et qu'il est possible de déterminer le village dans lequel se trouve un ménage donné. Supposons par ailleurs que le ménage en question comprend cinq enfants et possède 17 vaches, et que cette information est disponible dans la base de données. Si le village ne compte que quelques ménages comptant cinq enfants et possédant 17 vaches, n'importe quel utilisateur de la base de données sera en mesure d'identifier facilement le ménage qui a répondu à l'enquête. Il s'agira d'une violation de la confidentialité du ménage.

En conséquence, l'enquêteur principal ne devrait pas partager des bases de données incluant des identificateurs géographiques, et seules les bases de données brutes (qui ne sont accessibles qu'aux enquêteurs principaux) devraient contenir des codes d'identification géographique.

Codes d'identification de terrain pour la collecte et le partage des données

Les codes d'identification de terrain sont des codes uniques et simplifiés correspondant à chaque unité d'observation, mais ils ne contiennent pas d'informations géographiques comme tel. En plus de protéger la confidentialité des répondants, ces codes facilitent la collecte des données et leur codage pour le bureau d'études.

Supposons par exemple qu'une équipe d'EI souhaite mener une étude englobant 400 centres de santé et les zones d'intervention correspondantes. Une telle étude se composera en fait de 400 enquêtes portant chacune sur un centre de santé et les répondants correspondants. Il faudra dans ce cas utiliser un code numérique à trois chiffres puisqu'un code à deux chiffres ne permettrait de couvrir qu'un maximum de 99 enquêtes. Chaque enquête visera un centre de santé, et les établissements peuvent ainsi être numérotés de la même manière que les enquêtes. À chaque centre de santé/enquête correspondront normalement plusieurs niveaux supplémentaires de collecte des données — par exemple, agents de santé travaillant dans l'établissement, patients sortant de l'établissement et ménages se trouvant dans la zone d'intervention de l'établissement — qui devront tous être liés au code d'identification du centre de santé. On pourra d'ordinaire attribuer à ces unités supplémentaires un code de deux chiffres. Par exemple, s'agissant de l'enquête n° 232 :

- Centre de santé : un seul, portant le numéro d'identification 232.
- Agents de santé : si trois de ces agents sont convoqués pour une entrevue, les questionnaires seront numérotés 232-01, 232-02 et 232-03.
- Patientes de soins prénatals : si dix de ces patientes sont interviewées à leur sortie de l'établissement, leurs questionnaires seront numérotés 232-01 à 232-10.
- Patients de moins de 5 ans : si huit gardiens de patients de moins de 5 ans sont interviewés à la sortie de l'établissement, les questionnaires seront numérotés 232-01 à 232-08.
- Ménages : si 12 ménages sont sondés dans la zone d'intervention de l'établissement, les questionnaires correspondants seront numérotés 232-01 à 232-12.

Grâce à ce système simplifié de codification, le bureau d'études sera en mesure de suivre facilement le déroulement de chacune des enquêtes. Il pourra ainsi constituer des trousse individuelles d'informations et de mesures sur les enquêtes achevées. S'agissant de l'exemple ci-dessus, la trousse contiendra les données de l'évaluation du centre de santé en question (formulaire F1), des questionnaires administrés aux trois agents de santé (formulaire F2), des questionnaires administrés aux dix patientes de soins prénatals (formulaire F3), des questionnaires administrés aux huit gardiens d'enfants de moins de 5 ans (formulaire F4) et des questionnaires administrés aux 12 ménages (formulaire HH).

Étude de cas nationale : Définition des codes d'identification de terrain
Programme communautaire rwandais de financement basé sur la performance

L'évaluation du programme communautaire rwandais de FBP couvrait 200 secteurs correspondant à autant d'enquêtes. Dans chaque secteur, l'équipe d'EI a recueilli des données auprès d'un établissement de santé, de 12 villages par établissement et de deux agents de santé communautaire par village. Le nombre total d'entrevues par secteur se répartissait donc comme suit : 1 établissement de santé ; 12 ménages ; 24 agents de santé communautaires. Trois codes d'identification de terrain ont été générés :

- field_id1 : numéro à trois chiffres identifiant chacun des secteurs et établissements de santé ;
- field_id2 : numéro à trois chiffres identifiant chacun des villages et des ménages choisis dans la zone d'intervention de l'établissement de santé ;
- field_id3 : numéro identifiant chacun des agents de santé communautaires interviewés dans chaque établissement de santé.

Fichier de contrôle de l'identification : Mise en correspondance des codes géographiques et des codes de terrain

Le fichier de contrôle de l'identification établit la liste des codes d'identification de terrain et des codes géographiques correspondants ; il est d'ordinaire présenté en format Excel. Ce fichier est normalement préparé par le bureau d'études, en collaboration avec l'enquêteur principal, avant le début de l'enquête de référence.

Le fichier de contrôle de l'identification joue un rôle essentiel au moment de l'enquête de suivi, lorsque les enquêteurs jugent utile de retourner sur le terrain pour conduire des entrevues dans les établissements et les zones d'intervention qui ont fait l'objet de l'enquête de référence. S'ils disposaient uniquement de codes d'identification de terrain, ils seraient incapables d'établir les correspondances voulues entre les séries de données et les établissements sondés ; c'est-à-dire de déterminer l'emplacement physique de chacun des établissements décrits dans la base de données de référence. Le fichier de contrôle de l'identification leur permet d'établir ce lien, et d'utiliser pour l'enquête de suivi les numéros d'identification de l'enquête de référence.

Il convient de rappeler que le fichier de contrôle de l'identification est un document sensible puisqu'il permet de déterminer l'emplacement physique des répondants des ensembles de données, ce qui constituerait une infraction aux règles de protection de la vie privée. L'enquêteur principal est responsable de veiller au stockage de ces informations dans des conditions de sécurité adéquates. Voir Module 6 pour en savoir plus sur l'utilisation d'une enclave de données aux fins de la protection de la sécurité de ce type de données sensibles.

Figure 9 : Capture d'écran d'un fichier de contrôle de l'identification

Codes d'identification de terrain			Codes d'identification géographique									
field_id1	field_id2	field_id3	province	prov_code	district	dist_code	secteur	sect_code	cellule	cell_code	village	village_code
267	111	001	SUD	Sud	NYANZA	361	Busasamana	36101	Gahondo	3610101	Kamatovu	361010103
267	111	002	SUD	Sud	NYANZA	361	Busasamana	36101	Gahondo	3610101	Kamatovu	361010103
267	112	001	SUD	Sud	NYANZA	361	Busasamana	36101	Gahondo	3610101	Karama	361010104
267	112	002	SUD	Sud	NYANZA	361	Busasamana	36101	Gahondo	3610101	Karama	361010104
267	113	001	SUD	Sud	NYANZA	361	Busasamana	36101	Kavumu	3610102	Akirabo	361010201
267	113	002	SUD	Sud	NYANZA	361	Busasamana	36101	Kavumu	3610102	Akirabo	361010201
267	114	001	SUD	Sud	NYANZA	361	Busasamana	36101	Kibinja	3610103	Kabuzuru	361010301
267	114	002	SUD	Sud	NYANZA	361	Busasamana	36101	Kibinja	3610103	Kabuzuru	361010301
267	115	001	SUD	Sud	NYANZA	361	Busasamana	36101	Kibinja	3610103	Ngorongari	361010304
267	115	002	SUD	Sud	NYANZA	361	Busasamana	36101	Kibinja	3610103	Ngorongari	361010304
267	116	001	SUD	Sud	NYANZA	361	Busasamana	36101	Kibinja	3610103	Rugari A	361010307
267	116	002	SUD	Sud	NYANZA	361	Busasamana	36101	Kibinja	3610103	Rugari A	361010307
267	117	001	SUD	Sud	NYANZA	361	Busasamana	36101	Kibinja	3610103	Rugari B	361010308
267	117	002	SUD	Sud	NYANZA	361	Busasamana	36101	Kibinja	3610103	Rugari B	361010308
267	118	001	SUD	Sud	NYANZA	361	Busasamana	36101	Nyanza	3610104	Gatunguru	361010406
267	118	002	SUD	Sud	NYANZA	361	Busasamana	36101	Nyanza	3610104	Gatunguru	361010406
267	119	001	SUD	Sud	NYANZA	361	Busasamana	36101	Nyanza	3610104	Kavumu	361010408
267	119	002	SUD	Sud	NYANZA	361	Busasamana	36101	Nyanza	3610104	Kavumu	361010408
267	120	001	SUD	Sud	NYANZA	361	Busasamana	36101	Nyanza	3610104	Kivumu	361010410
267	120	002	SUD	Sud	NYANZA	361	Busasamana	36101	Nyanza	3610104	Kivumu	361010410
267	121	001	SUD	Sud	NYANZA	361	Busasamana	36101	Nyanza	3610104	Rubona	361010414
267	121	002	SUD	Sud	NYANZA	361	Busasamana	36101	Nyanza	3610104	Rubona	361010414
267	122	001	SUD	Sud	NYANZA	361	Busasamana	36101	Rwesero	3610105	Mwima	361010507
267	122	002	SUD	Sud	NYANZA	361	Busasamana	36101	Rwesero	3610105	Mwima	361010507

Note : Dans le cas de l'évaluation d'impact du programme communautaire rwandais de FBP, le code field_id1 identifie chacun des secteurs, mais n'inclut aucune information géographique. Le code field_id2 identifie chacun des villages ou chacun des ménages, puisqu'un seul ménage a été interviewé par village, mais il n'inclut aucune information géographique. Le code field_id3 identifie chacun des agents de santé communautaires, puisque deux de ces agents ont été interviewés dans chaque village, mais n'inclut aucune information géographique.

Prière de noter que les codes d'identification de terrain et les codes d'identification géographique présentés dans ce tableau ont été modifiés pour protéger l'identité des sources d'information.

Identificateurs des groupes de traitement et de comparaison

L'équipe d'EI devra faire en sorte qu'il soit possible de déterminer à quel groupe (traitement ou comparaison) chaque zone géographique est assignée. Il conviendra donc d'établir une liste des zones géographiques et de leur assignation.

Par ailleurs, les analystes devront aussi pouvoir déterminer à quel groupe correspondent chacun des ensembles de données. Il conviendra donc d'établir une liste des codes d'identification et des groupes (traitement ou comparaison) auxquels ils correspondent.

Il convient de noter que le fichier de correspondance des informations géographiques et des informations d'assignation (traitement/comparaison) et le fichier de correspondance des codes d'identification de terrain et des informations d'assignation devraient être distincts. Si les informations géographiques, les codes d'identification de terrain et les informations d'assignation étaient toutes comprises dans un seul et même fichier, il serait possible d'utiliser ce fichier pour localiser les répondants, ce qui serait contraire aux règles de confidentialité.

Figure 10 : Fichier d'identificateurs du lieu géographique et de l'assignation (traitement/comparaison)

province_code	province_name	district_code	district_name	sector_code	sector_name	group_code	group_name
2	SUD	201	NYANZA	20102	Busoro	1	Phase 1 Demand
2	SUD	201	NYANZA	20110	Rwabicuma	1	Phase 1 Demand
2	SUD	201	NYANZA	20101	Busasamana	2	Phase 1 CPBF Incentive
2	SUD	201	NYANZA	20103	Cyabakamyi	2	Phase 1 CPBF Incentive
2	SUD	201	NYANZA	20105	Kigoma	3	Phase 1 Demand+CPBF
2	SUD	201	NYANZA	20107	Muyira	3	Phase 1 Demand+CPBF
2	SUD	201	NYANZA	20108	Ntyazo	4	Control
2	SUD	201	NYANZA	20109	Nyagisozi	4	Control

Noter que les codes d'identification de terrain sont exclus de ce fichier.

Essai préalable des questionnaires

Il appartient d'ordinaire à l'équipe d'EI de procéder à l'essai préalable des questionnaires, en collaboration avec un groupe choisi de gestionnaires et d'enquêteurs du bureau d'études. Cet essai a notamment pour objectifs de :

- veiller à ce que les questionnaires soient adaptés au contexte local ;
- réviser la traduction des questionnaires pour veiller à ce qu'elle soit adaptée au contexte local ;
- déterminer les ajustements requis afin de réduire le risque de voir les personnes interviewées répondre « je ne sais pas » ou refuser de répondre ;
- tester l'organisation et la gestion du travail de terrain — par exemple, la répartition du travail entre les enquêteurs chargés d'administrer le questionnaire destiné à l'établissement de santé, et le rôle du chef d'équipe ;
- tester la collecte des données biométriques et ajuster les procédures et l'équipement, le cas échéant ;
- évaluer l'aptitude des enquêteurs à administrer les questionnaires ;

- estimer le temps qu'il faudra pour administrer chaque questionnaire sur le terrain. L'équipe d'EI et le bureau d'études doivent connaître la durée moyenne de chaque composante de collecte de données afin de planifier correctement le travail sur le terrain. Cf. **Traçage de la durée des entretiens**;
- documenter les changements et les adaptations suggérées par l'équipe.

Gestion de la saisie des données

La présente section examine un certain nombre de questions importantes ayant trait à la saisie des données.

Démarche. Comme mentionné plus haut, nous recommandons de procéder à la saisie des données pendant l'exécution du travail sur le terrain à l'aide de la méthode de STAO, au lieu d'attendre la fin du travail de terrain.

Logiciel. Le logiciel de saisie des données (de préférence, CSPro) doit être programmé afin de permettre le contrôle de la portée, le cas échéant, et de dépister les fautes de frappe qui donnent des valeurs impossibles à certaines réponses (par exemple, saisie du chiffre 7 lorsqu'on demande de donner un chiffre compris entre 1 et 5 ; inscription de 150 à une question sur l'âge ; etc.). Le logiciel de gestion des données devrait également être programmé pour signaler les incohérences internes entre les questions. En cas de message d'erreur, un enquêteur ou le chef de l'équipe de terrain devrait d'abord vérifier les données et, si nécessaire, reprendre l'entrevue du ménage pour éclaircir les réponses contradictoires. Par exemple, le logiciel devrait signaler les problèmes d'incohérence suivants :

- différence d'âge inférieure à 12 ans entre la mère et l'enfant ;
- nombre d'années de scolarité supérieur à l'âge du sujet moins quatre ans ;
- nombre total de pièces à usage exclusif supérieur au nombre total de pièces de l'habitation entière ;
- incohérences quant au sexe des sujets — par exemple, soins prénatals dans le cas d'un sujet de sexe masculin.

Les **modules de saisie des données CSPro** sont mis à la disposition des équipes nationales qui peuvent les adapter et les utiliser à leur guise. En suivant les directives d'adaptation des questionnaires fournies ci-dessus, les équipes nationales devraient être en mesure d'optimiser l'utilisation des programmes de saisie des données disponibles et de limiter les coûts supplémentaires de leur adaptation.

Matériel. L'équipe d'EI doit évaluer le matériel que le bureau d'études propose d'utiliser pour la collecte des données. La saisie et la gestion des données avec des ordinateurs périmés, vieux ou peu fiables accroît les risques de sécurité pour les données. Il est utile de procéder à cette évaluation avant de passer les contrats avec les bureaux d'études (voir la section consacrée au recrutement du bureau d'études).

Cependant, il n'est pas nécessaire d'utiliser les modèles d'ordinateurs de bureau ou portatifs les plus récents. Le logiciel CS-Pro peut tourner sur des ordinateurs moins coûteux qui présentent les caractéristiques suivantes : au moins 1 Go de mémoire vive, processeur cadencé à environ 1,5 GHz, environ 100 Go d'espace libre sur le disque dur, écran à résolution normale (au moins 1266 x 768) et de dimensions raisonnables (pas moins de 14 po pour les ordinateurs portatifs de terrain ; dimensions recommandées de 17 à 19 po pour les ordinateurs utilisés dans un centre de saisie) ; au moins trois ports USB ; capacités de réseautage de base. Le bureau d'études devrait par ailleurs avoir accès à un autre ordinateur plus puissant (ordinateur de bureau ou portatif) qui pourra servir de concentrateur de données.

Sécurité antivirus. L'équipe d'EI doit évaluer le logiciel de protection antivirus du bureau d'études avant le début du travail de collecte des données. Les systèmes mal protégés posent un risque accru de sécurité pour les données. Nous recommandons que les ordinateurs soient équipés de programmes de sécurité antivirus légers et à mise à jour rapide comme NOD32 ou Kaspersky, de préférence aux logiciels lourds comme Norton or McAfee qui exigent des mises à jour régulières.

Réseaux. Le concentrateur de données doit pouvoir communiquer avec tous les ordinateurs portatifs de terrain et les ordinateurs de bureau utilisés au centre de saisie des données. Nous recommandons d'utiliser un réseau privé de communication sans fil (wifi) de manière à éviter les protocoles de réseautage complexes. Dans la plupart des cas, un simple mappage du disque dur (lecteur C) de chaque ordinateur permettra au concentrateur de se connecter à distance à tous les ordinateurs du réseau en leur attribuant une lettre (M / N / ... / Z). Cette méthode permet au concentrateur d'utiliser jusqu'à 14 postes de saisie. Nous suggérons d'éviter les réseaux câblés Ethernet, à moins qu'ils ne soient déjà disponibles, puisqu'ils sont coûteux et difficiles d'entretien.

Boîtes étiquetées pour questionnaires imprimés. Le bureau d'études doit établir un protocole de stockage et de gestion des questionnaires imprimés remplis sur le terrain. Ces questionnaires peuvent par exemple être conservés dans des chemises ou des boîtes de classement (en fonction du volume) et étiquetés à l'aide du code d'identification de terrain. Grâce à cette stratégie, l'équipe peut déterminer facilement si la collecte des données auprès d'un établissement de santé ou d'une « zone de santé » est achevée. Les paquets ou les boîtes peuvent être entreposés sur des tablettes faciles d'accès, classés par ordre croissant à l'aide du code de l'EI du FBR. La pièce où sont entreposés les questionnaires devrait être protégée en permanence contre tout accès non autorisé.

Traitement ultérieur des données. Après la saisie des données dans les fichiers de données brutes, le bureau d'études doit procéder à un examen final au cours duquel il comparera les données transmises par les superviseurs de terrain à celles inscrites dans les registres de saisie. Nous recommandons de procéder au minimum aux vérifications suivantes :

- Vérification des fichiers de données brutes pour veiller à ce que le nombre de questionnaires remplis (indiqué dans les rapports du superviseur) corresponde au nombre de registres figurant dans les fichiers de données. Par exemple, si le superviseur fait état de sept entrevues de départ de patientes de soins prénatals pour le centre de santé numéro 232, le fichier de données

d'entrevues de départ des patientes de soins prénatals devrait compter sept registres pour cet établissement.

- Inspection des données brutes pour veiller à ce qu'il n'y ait pas de fichiers corrompus.
- Comparaison des codes de chaque variable de la série de données aux codes utilisés dans le questionnaire pour déceler les valeurs aberrantes. On estime qu'un pourcentage de valeurs aberrantes de 2 % est acceptable, mais qu'un pourcentage de 10 % pose un grave problème.
- Vérification manuelle de toutes les « autres » entrées pour veiller à ce que la description ne corresponde pas à une des réponses pré-codées. Cette tâche peut exiger l'inspection visuelle de plusieurs milliers de cas, et peut facilement demander de six à huit jours ; il faudra souvent examiner les questionnaires originaux pour contrôler l'exactitude des réponses inscrites.

Étiquetage des bases de données. Nous recommandons l'utilisation du protocole suivant pour l'étiquetage et l'organisation des fichiers de données.

``country'_'x'_'nn'`

Country : Préfixe indiquant le pays où les données ont été recueillies. Par exemple, le préfixe « NG » correspond au Nigéria.

x : Instrument d'enquête utilisé avec le codage suivant, par exemple :

- HH – Ménage
- F1 – Établissement de santé
- F2 – Agents de santé

nn : Numéro de séquence du fichier pour chaque formulaire ; par exemple :

- 00 pour les variables au sujet desquelles il n'y a qu'une réponse ; par exemple variables sur les actifs des ménages (une réponse par ménage) ou variables sur les caractéristiques de l'établissement (une réponse par établissement) ;
- 01 pour les variables qui reviennent souvent dans un questionnaire ; par exemple, celles figurant dans la liste des membres du ménage (un questionnaire contient des informations sur l'éducation et la santé de chacun des membres du ménage).

2012/04/03

Étude de cas nationale : Identification cohérente des bases de données
Crédits d'investissement pour le programme de santé publique du Nigéria

- NG_HH_00 contiendrait des données sur les ménages provenant du questionnaire destiné aux ménages ;
- NG_HH_01 contiendrait des données sur les membres du ménage ; c'est-à-dire des informations portant sur les membres individuels du ménage.

Double entrée. Lorsqu'on a recours à un service centralisé de saisie des données, nous recommandons que le bureau d'études utilise un système à double entrée (ou à « entrée de contrôle ») pour valider et mesurer l'exactitude des données saisies par les divers opérateurs. Le choix de la proportion du travail faisant l'objet de ce type de contrôle est flexible : l'équipe d'EI et le bureau d'études peuvent s'entendre au départ sur le nombre de fichiers qui devraient faire l'objet d'une double entrée dans chacune des zones d'enquête. Ce contrôle devrait d'ordinaire porter sur le fichier de données dont la qualité est jugée la meilleure. Si la comparaison des données primaires et des données de contrôle laisse constater plus de 10 à 15 erreurs de saisie environ par questionnaire, le bureau d'études devra prendre les mesures qui s'imposent — par exemple, augmenter le pourcentage des questionnaires qui feront l'objet d'une double entrée ou, dans les cas les plus graves, remplacer l'opérateur de saisie des données. Noter que la saisie des données à l'aide du système de STAO est fondée sur une approche entièrement différente, et n'exige pas normalement le recours au système de double entrée.

Risques sur le plan de la validité interne de l'EI pendant la collecte des données

Lorsqu'on passe de l'étape de la conception de l'EI à celle de la collecte des données, il est essentiel de porter attention aux risques possibles qui pèsent sur la validité interne de l'évaluation. Pour assurer la validité interne de l'EI, il importe de maintenir une estimation précise du contrefactuel (estimation de la situation qui aurait existé en l'absence du programme de FBR)¹⁵. Les activités de collecte des données risquent de nuire à la validité interne de l'évaluation si le bureau d'études réserve des traitements différents au groupe de traitement et au groupe de comparaison. Une différence de traitements introduit en effet un biais qui peut conduire à l'invalidation de l'évaluation, comme le montrent les exemples suivants.

- **Exemple 1 :** L'équipe de collecte de données n° 1 est la plus expérimentée et a fait preuve d'un niveau de compétence plus élevé à la formation que l'équipe n° 2. L'équipe n° 1 est chargée de recueillir les données auprès du groupe de traitement, tandis que l'équipe n° 2 s'occupe du groupe de comparaison. Après la communication des données de référence, l'équipe d'EI procède à des épreuves de comparaison des moyennes et constate que les résultats moyens du groupe de traitement sont significativement inférieurs à ceux du groupe de comparaison. Un examen plus attentif permet à l'équipe d'EI de découvrir que l'équipe n° 1 était plus habile que l'équipe n° 2 à repérer les cas de surdéclaration ou de réponses inexactes. Ainsi, les données recueillies auprès du groupe de traitement sont plus exactes que celles recueillies auprès du groupe de comparaison, et il n'est pas possible de comparer les moyennes des variables de résultats des deux groupes.

¹⁵ Prière de consulter le 0 et le manuel « L'évaluation d'impact en pratique » pour en savoir plus sur l'utilité de l'assignation aléatoire entre les groupes de traitement et de comparaison pour l'estimation exacte du contrefactuel.

- **Exemple 2** : Pendant la planification du travail sur le terrain, le bureau d'études décide de commencer la collecte des données auprès du groupe de traitement, avant de passer au groupe de comparaison. La collecte des données du groupe de traitement débute en janvier, tandis que celle du groupe de comparaison débute en avril. Après la collecte des données, les épreuves de comparaison des moyennes laissent constater des taux de prévalence du paludisme beaucoup plus élevés ainsi qu'une utilisation moindre des principaux services de santé chez les patients du groupe de traitement que chez ceux du groupe de comparaison. Vérifications faites, l'équipe d'EI découvre que la période de janvier à mars correspond à la saison des pluies, tandis que la période d'avril à juin correspond à la saison sèche, ce qui a de toute évidence influé sur les principaux indicateurs de santé. Bien que les données aient été recueillies avec la même exactitude dans les deux groupes, elles restent impossibles à comparer.
- **Exemple 3** : Pendant l'enquête de référence, l'équipe de projet demande au bureau d'études d'expédier une enveloppe de documents aux établissements du groupe de traitement. La directrice d'un établissement apprend en consultant les documents qui lui ont été transmis que son établissement a été choisi pour participer à un programme de FBR dans le cadre duquel les membres de son personnel recevront des primes d'encouragement fondées sur leur performance. Elle s'empresse de communiquer cette information aux membres de son personnel et leur demande de collaborer avec les enquêteurs afin de « faire bonne impression » auprès des administrateurs du FBR. Les agents de santé des établissements du groupe de traitement ont ainsi tendance à se montrer plus coopératifs et plus patients avec les enquêteurs que ceux des établissements du groupe de comparaison. Cette différence de comportement rend les données recueillies auprès des deux groupes impossibles à comparer.

Ces exemples illustrent les points qu'il convient de garder à l'esprit pendant la collecte des données :

- i) le calendrier de la collecte des données ne devrait pas favoriser les établissements d'un des groupes par rapport à l'autre ;
- ii) les équipes de terrain assignées aux deux groupes ne devraient pas être différentes au plan de la compétence, de la surveillance, etc. ;
- iii) les équipes d'enquête devraient être tenues à l'écart des opérations de mise en œuvre et de contrôle de l'intervention de FBR. Dans la mesure du possible, les équipes de terrain ne devraient pas être informées de la répartition des établissements entre les deux groupes de traitement, et ne devraient pas chercher à savoir à quel groupe les divers établissements ont été assignés.

Plan de travail sur le terrain

Le plan de travail sur le terrain est un document préparé par le bureau d'études dans lequel ce dernier décrit comment il assurera la gestion des équipes de terrain et comment il conduira les enquêtes auprès des établissements de santé et des ménages. Nous décrivons ci-dessous quelques-uns des éléments de ce plan.

Gestion des équipes de terrain

Le plan de travail fournit des détails sur la composition des équipes, leurs rôles et leurs responsabilités, ainsi que sur le calendrier de travail, les résultats attendus et les aspects logistiques.

- **Composition.** Le plan doit préciser combien d'équipes seront déployées sur le terrain, et indiquer la composition de chacune de ces équipes (superviseur, enquêteurs, personnel de saisie des données, etc.).
- **Rôles et responsabilités.** Comme chaque enquête comporte de multiples éléments, nous recommandons que le plan explique en détails quels sont les membres des équipes responsables de la réalisation de chacun des modules.
- **Calendrier.** Le plan doit inclure un **Fichier de contrôle de l'échantillonnage** dûment rempli indiquant les dates précises des entrevues réalisées par l'équipe de terrain auprès de chacune des unités d'observation (établissement de santé, ménage). On ajoutera également des détails sur les résultats attendus, puisque cela influe sur le temps que l'équipe compte passer à chaque endroit.
- **Résultats attendus.** Le plan doit inclure une description du travail que chacune des équipes compte accomplir chaque jour avec chaque type de questionnaire.
- **Aspects logistiques.** Les superviseurs des équipes de terrain sont responsables de la logistique du travail sur le terrain, y compris la gestion du matériel (questionnaires imprimés, stylos, matériel de collecte de marqueurs biologiques, etc.). Ils doivent en outre s'occuper de l'organisation des déplacements et du transport — par exemple, réservations d'hôtels, gestion des véhicules (essence, pneus, etc.). Le plan devrait inclure un exposé des exigences quotidiennes ou hebdomadaires de la gestion logistique de chacune des équipes de terrain.
- D'autres aspects propres aux activités de collecte de données auprès des établissements de santé et des ménages doivent aussi être pris en compte lors de l'élaboration du plan de travail sur le terrain.

Collecte de données auprès des établissements de santé

S'agissant de la collecte des données auprès des établissements de santé, le plan de travail doit prendre en compte les aspects suivants :

- **Heures d'accueil des patients.** Les équipes de terrain doivent s'enquérir des heures d'accueil des patients puisque dans la plupart des cas, les entrevues avec les prestataires devront être organisées hors de cette période. Par ailleurs, les entrevues de départ réalisées auprès des patients ne sont possibles que pendant les heures d'accueil.
- **Calendrier de prestation des services.** Les équipes de terrain doivent s'enquérir des jours au cours desquels les services à l'étude sont offerts dans les établissements de l'échantillon. Par exemple, si un établissement n'offre les soins prénatals que le lundi, le mercredi et le vendredi, il faudra éviter d'organiser les entrevues auprès des prestataires de soins prénatals et les entrevues de départ des patientes le mardi ou le jeudi. Ce genre de détails peut avoir d'énormes répercussions sur la planification du travail.

- **Horaires.** Dans certains pays, les horaires des établissements de santé sont uniformes et peuvent donc être intégrés dans le plan global de travail sur le terrain. Cependant, dans la plupart des cas, les jours et les heures de prestation de services sont déterminés par les établissements. En conséquence, les équipes de terrain doivent obtenir cette information avant de pouvoir estimer le nombre de jours qu'elles devront consacrer à l'ensemble des entrevues conduites dans les établissements de santé. En règle générale, les équipes d'EI peuvent s'attendre à ce que les équipes de terrain aient besoin de 1,5 à 2 jours par établissement.

Collecte de données auprès des ménages

Le travail de collecte de données auprès des ménages devrait prendre en compte les aspects suivants :

- **Disponibilité des répondants.** Les répondants aux entrevues réalisées auprès des ménages sont le chef du ménage et les soignants principaux, dont la disponibilité pour participer à une entrevue de deux à trois heures sera limitée. Idéalement, l'équipe de terrain devrait être composée de manière à permettre la réalisation d'entrevues parallèles auprès de plusieurs répondants d'un même ménage pour ainsi limiter le temps consacré à chaque ménage.
- **Durée des entrevues.** Le questionnaire destiné aux ménages est long et complexe, et son administration devrait prendre de deux à trois heures par ménage. L'administration du questionnaire entier en une seule séance risque de nuire à la qualité des réponses à cause de la fatigue du répondant. Il est donc recommandé de prévoir plus d'une visite par ménage. La première visite devrait servir à la collecte des données initiales, la seconde étant consacrée à l'entrevue d'autres membres du ménage absents lors de la première visite et à la correction des incohérences relevées par l'opérateur de saisie des données. Tel que mentionné plus haut, cette procédure s'applique à la fois à la STAO et à la saisie centralisée des données.

Recrutement et formation des membres des équipes de terrain

Le bureau d'études est responsable du recrutement du personnel de terrain. Cependant, l'équipe d'EI a intérêt à veiller que tous les agents recrutés justifient des compétences voulues et répondent à un ensemble d'exigences standard.

- **Nombre de personnes recrutées.** Le bureau d'études devrait recruter et former au moins 15 % de plus de personnes que le nombre jugé nécessaire pour former les équipes de terrain. Il faut en effet prévoir que certains enquêteurs démissionneront même si le plan de travail comprend des mesures visant à limiter les causes de fatigue ou de surmenage — par exemple, mesure de gestion du travail de terrain et conception du questionnaire. Le bureau d'études doit se doter d'un plan d'urgence afin d'éviter toute interruption de la collecte des données.
- **Compétences du personnel.** Le bureau d'études devrait recruter et former ses agents en tenant compte du rôle qu'ils joueront au sein de l'équipe et de la réglementation locale. Par exemple, les spécialistes de l'anthropométrie et les personnes responsables de la collecte des marqueurs biologiques pourraient devoir justifier d'autorisations ou de compétences spéciales —

expérience en soins infirmiers ou dans un autre domaine de la médecine. Le bureau d'études devrait définir ces exigences avant de lancer le processus de recrutement.

Programme de formation et matériaux

La formation des superviseurs, des enquêteurs et des opérateurs de saisie est essentielle pour assurer la qualité des données d'enquête.

- **Programme.** Tous les superviseurs, enquêteurs et opérateurs de saisie devraient suivre une formation générale qui permettra de renforcer la cohésion de l'équipe et de permettre à tous les membres d'acquérir la souplesse requise pour se remplacer mutuellement en cas d'absence temporaire pour cause de maladie ou d'autre situation d'urgence. Les superviseurs devraient recevoir une formation supplémentaire, en sus de la formation générale.
- **Aspects logistiques.** Idéalement, le bureau d'études organisera la formation des superviseurs, des enquêteurs et des opérateurs de saisie en un lieu central, ce qui permettra à tous de recevoir la même formation et d'utiliser à cette fin des présentations Powerpoint standard. Si cela n'est pas possible, le bureau d'études devra veiller à prévoir une période de temps et des ressources financières suffisantes pour offrir une formation standard dans des endroits différents. Pendant les négociations budgétaires avec les bureaux d'études candidats, l'équipe d'EI devrait veiller à ce que ces derniers prévoient les ressources nécessaires pour les frais de transport, d'alimentation (déjeuner et cafés dans tous les cas ; dîner pour les stagiaires de l'extérieur) et de logement pour les superviseurs, les enquêteurs et les opérateurs de saisie pendant la durée de la formation. Il appartiendra à l'enquêteur principal et au bureau d'études de décider de la forme des séances de formation : séances plénières ou, si les stagiaires sont trop nombreux, répartition des participants en sous-groupes. Dans ce dernier cas, le bureau d'études devra veiller à uniformiser la formation en utilisant le même matériel pour tous les sous-groupes.
- **Durée.** La formation devrait s'étendre sur au moins deux semaines.
- **Contenu.** La formation devrait se diviser en quatre grands volets comme suit :
 - **Théorie.** Examen des objectifs de recherche, de la teneur des questionnaires et de chacune des questions qu'ils contiennent pour bien comprendre les objectifs de chaque question. Il convient de promouvoir l'interaction et de prévoir suffisamment de temps pour répondre aux questions.
 - **Exercices pratiques en classe.** Chaque stagiaire devrait pouvoir remplir les questionnaires en classe. Le formateur peut afficher les questions à l'aide d'un projecteur et inviter les stagiaires à répondre à tour de rôle, en encourageant la participation de tous. Le formateur peut également élaborer des scénarios fondés sur des ménages typiques (observés par exemple pendant la formation ou l'apprentissage du superviseur) et demander aux stagiaires de remplir les questionnaires en s'appuyant sur les situations présentées. On peut également filmer une entrevue pilote et demander ensuite aux stagiaires de remplir un questionnaire à partir du cas illustré pour tester la cohérence des résultats obtenus par les divers stagiaires.

- **Exercices sur le terrain.** Après les séances théoriques et pratiques en classe, les stagiaires devraient se rendre sur le terrain pour administrer le questionnaire entier à un petit nombre de ménages (ne faisant pas partie de l'échantillon à l'étude). L'équipe peut se réunir le lendemain pour examiner les résultats, ce qui lui fournira l'occasion de clarifier les concepts, d'apprendre à traiter les cas « difficiles », de maîtriser l'enchaînement des questions (instructions « passez à... ») et de débattre de toute question supplémentaire ayant trait à la présentation du questionnaire et à sa traduction.
- **Évaluation.** À la suite de la formation, il convient d'évaluer les connaissances acquises par les enquêteurs, les superviseurs et les opérateurs de saisie participants, et leur aptitude à recueillir les données correctement en utilisant les scénarios utilisés en classe. Idéalement, les personnes qui ne réussissent pas à satisfaire aux exigences minimales prescrites ne devraient pas être autorisées à participer à l'enquête.

La boîte à outils contient **du matériel de formation ainsi que des programmes de cours et des manuels conçus à l'intention des équipes de terrain.**

Pilotage des questionnaires

Le pilotage des questionnaires poursuit les mêmes objectifs principaux que les tests préalables, à la différence qu'il exige la participation de tous les membres de l'équipe de terrain après la conclusion de la formation. Ses objectifs sont les suivants :

- veiller à ce que les questionnaires soient adaptés au contexte local ;
- réviser la traduction pour veiller à ce qu'elle soit adaptée au contexte local ;
- déterminer les ajustements qu'il convient d'apporter pour réduire le nombre de cas où les personnes questionnées répondent « je ne sais pas » ou refusent de répondre ;
- tester l'organisation et la gestion du travail de terrain — par exemple, le partage des tâches entre les enquêteurs administrant le questionnaire dans les centres de santé et le rôle du superviseur ;
- tester la collecte des mesures biométriques et ajuster les méthodes et l'équipement, le cas échéant ;
- évaluer l'aptitude des enquêteurs à administrer les questionnaires ;
- estimer le temps qu'il faudra pour administrer chaque questionnaire sur le terrain. L'équipe d'EI et le bureau d'études doivent connaître la durée moyenne de chaque composante de l'opération de collecte de données afin de pouvoir planifier correctement le travail de terrain (voir l'outil intitulé **Traçage de la durée des entretiens**) ;
- prendre note des changements et adaptations suggérées par l'équipe.

Il est préférable à cette étape de limiter à un minimum les changements apportés au questionnaire. L'équipe d'EI et le bureau d'études devraient prévoir de consacrer au moins deux jours au pilotage du questionnaire et au peaufinage de tous les aspects logistiques avant le démarrage du travail de terrain.

Gestion du travail sur le terrain

Une fois entamé le travail de collecte des données, c'est au bureau d'études qu'il appartient de veiller au respect du protocole de recherche, du plan d'échantillonnage et du plan de travail sur le terrain. Le bureau d'études devrait prévoir dans son budget l'organisation d'une réunion de deux à trois jours pour l'ensemble des superviseurs, enquêteurs et opérateurs de saisie des équipes de terrain après le début des travaux. Cette réunion donnerait aux membres des équipes l'occasion de se pencher sur les problèmes de supervision, d'organisation du travail de terrain, de présentation et de contenu des questionnaires et de saisie des données, et de les corriger. Nous suggérons d'organiser cette réunion deux ou trois semaines après le début des travaux sur le terrain.

Les superviseurs doivent par ailleurs évaluer la performance des enquêteurs. Un formulaire (**Formulaire d'évaluation de l'enquêteur**) est prévu à cet effet dans la boîte à outils, avec d'autres ressources telles qu'un formulaire de suivi (**Formulaire de traçage pour le superviseur**) et un formulaire de gestion de trésorerie (**Fiche de suivi des décaissements**).

Préparation des rapports

Il est essentiel de maintenir une communication efficace entre l'équipe d'EI et le bureau d'études pendant toutes les étapes de la collecte des données afin de faire en sorte que les problèmes qui surgissent soient réglés le plus rapidement possible. Cependant, l'expérience a démontré que les bureaux d'études ont souvent du mal à préparer des rapports vraiment utiles à l'équipe d'EI. L'équipe globale d'évaluation d'impact des programmes de FBR a donc élaboré un modèle standard de rapport d'avancement de la collecte des données (**Rapport de progrès de la collecte des données**) que les bureaux d'études peuvent utiliser pour préparer les rapports dans lesquels ils font état de la préparation, de la formation de la collecte des données et des activités de gestion des données. Le rapport d'avancement proposé comporte les éléments suivants :

1. Date et phase
2. Résumé des progrès
3. Principaux enjeux
4. Prochaines étapes
5. État d'avancement des travaux de collecte et de saisie des données
6. Résultats
7. État des paiements



Module 6

Stockage des données et accès à ces dernières

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats

www.worldbank.org/health/impacetevaluationtoolkit

Module 6. Stockage des données et accès à ces dernières

Principales recommandations et outils disponibles pour ce module			
Recommandations	Essentiel	Important	Souhaitable
<ul style="list-style-type: none"> Le chef d'équipe doit veiller à la planification et à la coordination d'une documentation complète des activités d'évaluation d'impact. <ul style="list-style-type: none"> ▶ Inclure des notes de synthèse, des protocoles de recherche, des questionnaires, des manuels de formation, etc., à jour. ▶ Déterminer quels sont les éléments d'information qu'il convient de supprimer pour assurer la confidentialité des participants. L'enquêteur principal doit préparer un ou des fichiers d'identification permettant d'établir le lien entre les codes d'identification géographique et les codes d'identification de terrain. L'enquêteur principal doit déterminer qu'elles sont les variables qu'il faudra éviter de publier (par exemple, informations personnelles sensibles). Les fichiers confidentiels (fichiers d'identification et autres données non accessibles au public) doivent être stockés de façon sécurisée, de préférence dans une enclave de données. Les équipes d'évaluation d'impact doivent prévoir suffisamment de temps pour la documentation et le téléchargement des données vers l'amont, afin de garantir la continuité de l'accès aux données au sein de l'équipe, et la facilité du partage futur des données et du processus d'analyse. L'équipe d'évaluation d'impact doit se référer au protocole d'accord (ou à un accord similaire de partage des données) pour ses opérations de documentation, de stockage ou de partage des données. 	<p>✓</p> <p>✓</p> <p>✓</p>	<p>✓</p> <p>✓</p> <p>✓</p>	



Outils
<ul style="list-style-type: none"> 6.01 Data Deposit Form – IE Micro-data Catalog (Formulaire de dépôt des données – Catalogue de micro-données EI) 6.02 Nesstar Data Storage Templates (Modèles de stockage des données Nesstar) 6.03 Login to Micro-data Management Toolkit (Accès à l'outil de la gestion des micro-données) 6.04 How to Access the Data Catalog and Data (Comment accéder au catalogue de données et aux données)

Contenu du module

Pourquoi utiliser un catalogue de données ?	6-4
Types de données	6-5
Étapes du catalogage des données d'enquête	6-6
Catalogue de données FBR : Caractéristiques	6-7
Accès aux données.....	6-8

En moyenne, une évaluation de l'impact des FBR coûte près de 1,5 million de dollars et nous estimons qu'environ 70 à 80 % de ce budget couvre la collecte des données. Les données sont donc chères et fort prisées, mais elles sont fréquemment sous-utilisées, bien souvent parce qu'elles ne sont pas librement accessibles. De plus, il arrive parfois que des données collectées aient été perdues ou mal documentées à cause d'une défaillance informatique ou d'un changement de personnel. Enfin, les données d'évaluation d'impact comprennent souvent des informations confidentielles permettant de localiser et/ou d'identifier les personnes interrogées. Ces données doivent être protégées et la confidentialité des réponses préservée.

C'est dans cet esprit que l'équipe d'EI s'efforce d'aider les équipes pour s'assurer que :

- Les données collectées avec un financement du HRITF sont stockées en toute sécurité et correctement documentées ;
- Les équipes ont accès aux données d'une manière qui réponde à leurs besoins ;
- Les données sont à terme librement accessibles au public conformément à l'initiative Libre accès aux données de la Banque mondiale, tout en respectant leur confidentialité.

Ces trois objectifs peuvent être atteints en utilisant un catalogue de micro-données. Ce module explique i) pourquoi utiliser un catalogue de données pour stocker les données ; ii) les types de données ; iii) comment cataloguer les données ; iv) les caractéristiques du catalogue de données FBR ; et v) les différents types d'accès aux données pouvant être configurés dans le catalogue de données FBR.

Pourquoi utiliser un catalogue de données ?

Les équipes d'EI devront documenter et stocker leurs données soit dans le catalogue de données FBR soit dans le catalogue de données central de la Banque mondiale. Ces deux catalogues ont été spécifiquement créés pour les équipes de FBR et sont conformes aux normes DDI de Dublin sur le stockage et la documentation des données. Le stockage et la documentation des données dans ces catalogues présentent les avantages suivants :

- Les membres de l'équipe, y compris l'enquêteur principal, les chefs d'équipe de projet, l'équipe d'EI et leurs homologues du gouvernement peuvent facilement partager les données des pays et la documentation correspondante. Ils ne dépendent pas de la présence d'un membre spécifique de l'équipe pour préserver les données et la documentation.
- La sécurité des données et de la documentation est améliorée. Les données et la documentation stockées et échangées par courrier électronique, sur les disques durs des ordinateurs et les périphériques de stockage portables sont vulnérables aux défaillances, aux attaques et aux virus. Ces menaces sont minimisées dans un catalogue de données protégé.
- Le catalogue est conforme aux normes internationales de documentation et de stockage des données, ce qui permet de veiller à l'étiquetage, à la fiabilité et à l'anonymat des données.

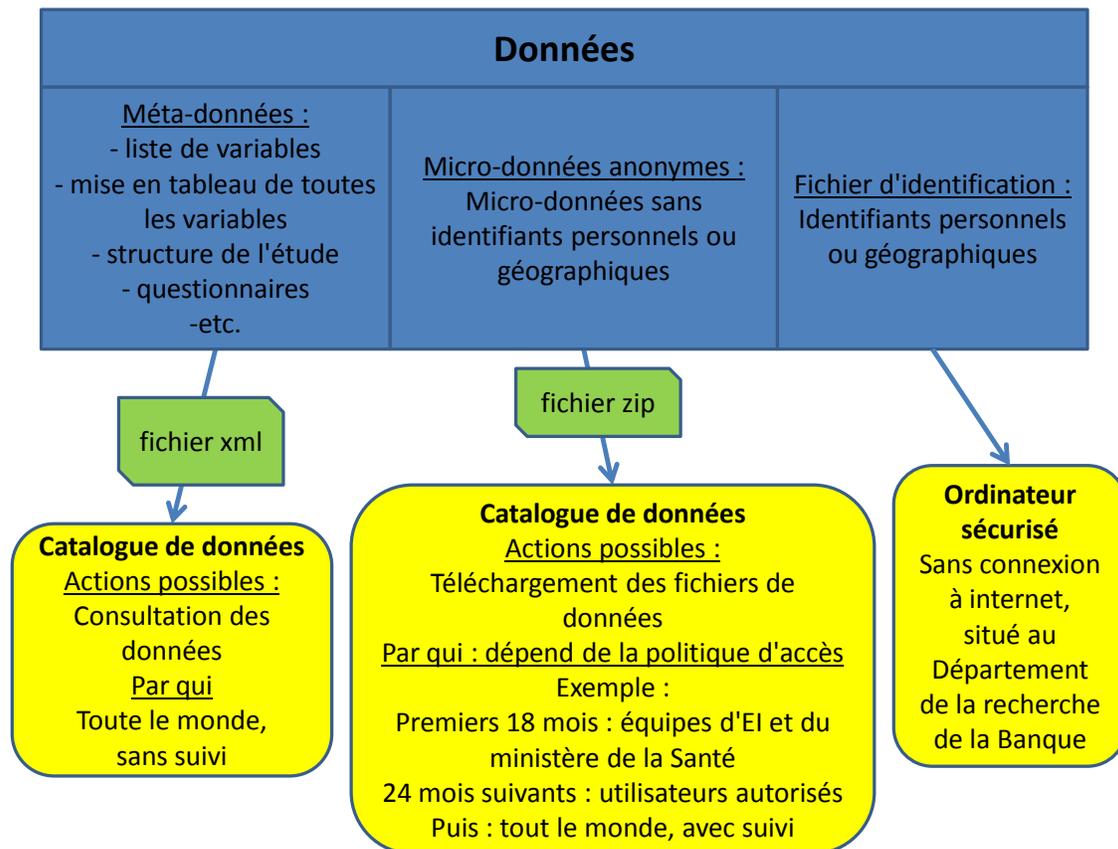
L'accès au catalogue de données FBR est actuellement protégé par un mot de passe. En revanche, le catalogue de données central de la Banque mondiale est accessible au public et consultable sur le web sans mot de passe. Lorsque les données sur les FBR passent dans le domaine public, elles sont migrées du catalogue de données FBR protégé par un mot de passe vers le catalogue de données central de la Banque mondiale.

Types de données

Une enquête d'évaluation d'impact produira généralement trois types de données :

- **Macro-données** : Les macro-données comprennent la liste des variables, la mise en tableau de toutes les variables, la structure de l'étude, les questionnaires, etc. Les macro-données sont en quelque sorte les « données sur les micro-données », ou toutes les informations nécessaires pour interpréter correctement les micro-données.
- **Micro-données anonymes** : Les micro-données correspondent aux données observées, par exemple les revenus des ménages, les mesures de la taille et du poids des enfants, et les données financières des centres de santé. Ces données devraient être anonymes et ne devraient pas contenir les codes géographiques, noms ou autres informations permettant d'identifier les personnes concernées. Elles doivent toutefois contenir les codes d'identification de terrain qui permettent aux chercheurs d'analyser les données sans identifier les unités correspondantes. Dans certains cas, les enquêteurs principaux peuvent décider de ne pas rendre *publiques* certaines micro-données si elles sont de nature confidentielle (résultats de tests sur le VIH, par exemple).
- **Fichier d'identification** : L'enquêteur principal doit protéger le fichier qui permet d'établir le lien entre les codes d'identification géographique et les codes d'identification de terrain, ainsi que les noms collectés dans le cadre de l'enquête. Sans ce fichier, il ne pourra pas recréer l'échantillon lors de la collecte des données en fin de projet si le bureau d'études doit retourner dans les mêmes zones d'échantillonnage. Pour préserver l'anonymat des données, le fichier d'identification ne devrait JAMAIS être inclus dans le catalogue de données. Il devrait être conservé en lieu sûr ; la Banque mondiale dispose d'un serveur sécurisé, non connecté à internet, sur lequel les équipes peuvent sauvegarder et stocker leurs fichiers d'identification. (Voir également le 0, Définition de codes d'identification)

Figure 6 : Gestion des données d'EI



Étapes du catalogage des données d'enquête

L'outil de gestion des micro-données permet de cataloguer les données en deux étapes :

Étape 1 : Organiser les données en utilisant le logiciel Nesstar

La première étape consiste à organiser les données à l'aide du logiciel Nesstar. Ce logiciel gratuit peut être téléchargé à l'adresse suivante : <http://nesstar.com/software/download.html>. Nesstar permet d'extraire, d'organiser et de stocker les macro-données en utilisant des normes internationales de documentation, de protection, de sécurité et de stockage des données¹⁶. Conformément aux normes internationales, toute la documentation finale devrait être en anglais ; si une partie des activités est menée dans une autre langue, l'équipe devrait prévoir un budget pour la traduction des informations en anglais. Le logiciel stocke deux types d'informations : i) la « documentation », stockée dans un fichier DDI (initiative de documentation des données) qui donne des informations de base sur l'enquête, telles

¹⁶ Pour en savoir plus, veuillez consulter le site : <http://www.ihsn.org>

que les membres de l'équipe, les dates de l'enquête, la conception de l'enquête, etc. ; et ii) les « ressources externes » stockées dans un fichier DC (Dublin Core), qui décrit les ressources et le processus de production des données, tel que les questionnaires, notes de synthèse, protocoles de recherche, etc. Il est important de noter que les ressources ne sont pas stockées dans Nesstar.

Les fichiers DDI et DC sont convertis en deux fichiers .xml qui contiennent toutes les macro-données de l'étude, c'est-à-dire une description générale de la méthodologie et des instruments de l'enquête d'évaluation d'impact. Le projet est également intégralement sauvegardé dans un fichier .Nesstar qui peut être rouvert et modifié avec Nesstar.

Étape 2 : Télécharger les données dans le catalogue de données

La deuxième étape consiste à télécharger les deux fichiers de macro-données .xml et les micro-données¹⁷ anonymes dans le catalogue de données. Les autres documents, tels que les questionnaires, la méthodologie de l'enquête et les protocoles de recherche, devraient également être téléchargés car ils sont essentiels pour comprendre comment les données ont été produites et comment elles devraient être interprétées. On pourra pour cela utiliser le Gestionnaire de ressources du catalogue de données.

Les macro-données téléchargées seront visibles de toute personne consultant le catalogue. L'accès aux fichiers de micro-données anonymes peut être restreint en utilisant les options « indisponible » ou « utilisation autorisée » du catalogue (voir ci-dessous). Rappelons que le fichier d'identification ne devrait jamais être téléchargé dans le catalogue de données.

Bien que le traitement et le téléchargement des données puissent être laborieux, l'équipe d'EI sur la santé, la nutrition et la population s'est associée au Département de la recherche de la Banque mondiale pour obtenir de l'aide au besoin. Pour obtenir son adresse électronique, veuillez consulter la page de contacts de la boîte à outils. Les équipes seront priées de remplir un **Formulaire de dépôt des données** pour faciliter le téléchargement.

Catalogue de données FBR : Caractéristiques

Le catalogue de données FBR comprend une liste détaillée et consultable de toutes les données et de la documentation connexe qui ont été collectées dans le contexte du programme de FBR. Veuillez noter que le site web est protégé par un mot de passe et est uniquement accessible aux équipes de FBR. Son adresse est la suivante : <http://www.ihsn.org/apps/hrif/index.php/auth/login/?destination=>

Le catalogue offre les fonctionnalités suivantes :

¹⁷ Dans un fichier au format zip

- **Visualisation et téléchargement de la documentation correspondante** : Les membres de l'équipe peuvent visualiser la note conceptuelle, le protocole de recherche, les questionnaires en anglais (et en langue locale), le plan d'échantillonnage et tous les autres documents sur l'évaluation d'impact nécessaires pour comprendre comment les données ont été produites.
- **Visualisation des macro-données** : Visualisation des macro-données des équipes de FBR, y compris les questionnaires, les réponses récapitulatives, les variables, la conception de la recherche, etc. L'accès aux micro-données dépend de la politique d'accès aux données.
- **Comparaison des macro-données** : Comparaison interpays des fréquences pour les mêmes variables.
- **Suivi des citations** : Suivi des publications qui utilisent les ensembles de données de la base de données.
- **Contrôle de l'accès aux données** : Les équipes de pays peuvent contrôler qui a accès aux micro-données et pendant quelle période. L'accès aux données devrait être clairement défini dans le **Protocole d'accord sur l'accès aux données** de chaque équipe nationale (voir 0).

Accès aux données

Toutes les données et la documentation devraient être stockées et documentées dans le catalogue de données FBR dans un délai de six mois suivant la fin de la collecte des données. Les données stockées dans le catalogue peuvent être étiquetées « non accessible », « utilisation autorisée » ou « utilisation publique ». Les données entrant dans les catégories « utilisation autorisée » ou « utilisation publique » devraient être transférées dans le catalogue de données général de la Banque mondiale car le catalogue FBR est protégé par un mot de passe et n'autorise pas ces types d'accès. Le catalogue de données général de la Banque mondiale autorise les accès de type « non accessible », « utilisation autorisée » et « utilisation publique ».

- **Non accessible.** L'accès aux micro-données peut être limité à l'équipe d'évaluation d'impact du pays et à l'équipe centrale correspondante. En règle générale, l'accès aux micro-données sera limité pendant une période de deux ans à compter de la fin de la collecte des données ou de six mois à compter de la publication du premier rapport, si cette date est antérieure.
- **Utilisation autorisée.** Dans un délai de deux ans à compter de la fin de la collecte des données ou de six mois à compter de la publication du premier rapport, si cette date est antérieure, les données devraient normalement être disponibles dans le cadre d'une utilisation autorisée. Ce type d'utilisation permet aux visiteurs externes de soumettre en ligne une demande d'accès aux données, en précisant le sujet de leurs recherches, les variables qui les intéressent, leur calendrier et leurs projets en matière de diffusion. Cette demande est envoyée au gestionnaire du catalogue de données, qui travaille avec l'équipe de pays pour approuver ou rejeter les demandes. Lorsqu'une demande a été approuvée, le gestionnaire des données envoie un nom d'utilisateur et un mot de passe au demandeur.
- **Utilisation publique.** Les micro-données devraient être librement accessibles au public dans un délai de quatre ans à compter de la fin de la collecte des données. Ce type d'utilisation permet

aux visiteurs externes de soumettre en ligne une demande d'accès aux données, en précisant le sujet de leurs recherches, les variables qui les intéressent, leur calendrier et leurs projets en matière de diffusion. Cette demande est immédiatement approuvée par le système et le demandeur reçoit un nom d'utilisateur et un mot de passe.

2012/02/05

Étude de cas nationale : Politique d'accès aux données pour différents intervenants
Projet de renforcement des actions sanitaires en faveur de la population rurale défavorisée en Afghanistan

L'évaluation d'impact du FBR est réalisée par la Banque mondiale, l'Université Johns Hopkins et le Gouvernement afghan. L'Université Johns Hopkins a collecté les données de référence à la fin de 2010. L'intervention se poursuit et l'enquête de suivi est prévue pour février 2013. Il n'existe pas de protocole d'accord explicite entre les parties concernant l'accès aux données. Des accords contractuels sur l'accès aux données existent par contre entre l'Université Johns Hopkins et le gouvernement.

Après avoir collecté les données de référence, les trois parties prenantes ont mis ces données en commun. L'EI étant constituée de deux groupes représentatifs, il n'était pas nécessaire d'inclure des informations d'identification des ménages ou des établissements de santé pour les retrouver dans les ensembles de données lors de la deuxième phase. L'Université Johns Hopkins, en tant que bureau d'études, a accès aux données nominatives contenues dans les questionnaires papier.

Bien que l'accès aux données ait été simple et facile au sein de l'équipe, on ne sait pas encore comment les données seront accessibles à des tiers. L'initiative Libre accès aux données de la Banque mondiale prévoit que les données collectées et financées par la Banque mondiale soient accessibles au public au bout d'un certain temps. L'équipe d'EI s'est interrogée sur les approches à adopter pour satisfaire cette exigence tout en bénéficiant d'un accès privilégié aux données pendant les premières phases de la diffusion des données de référence. L'équipe a formellement accepté que les données soient accessibles au public. Mais le moment de leur diffusion publique fait toujours débat.

Pour de plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.



Module 7

Analyse des données et diffusion des résultats

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats
Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

www.worldbank.org/health/impacetevaluationtoolkit

- 7.04 STATA ado file for Baseline balance table (Fichier Ado STATA pour tableau d'équilibre de l'enquête de base)
- 7.05 WHO Anthro calculation package (Pack de calcul OMS Anthro)
- 7.06 STATA training (Formation STATA)
- 7.07 STATA Training Design Validation 7.07 Validating IE Design Training(Formation STATA pour la Validation de la conception de l'EI)

Contenu du module

Rapport de référence	7-5
Nettoyage des données	7-6
Création de variables aux fins d'analyse	7-7
Validation de la méthode d'évaluation	7-8
Validité externe.....	7-8
Validité interne	7-8
La puissance statistique est-elle suffisante ?.....	7-9
Les caractéristiques de référence sont-elles équilibrées ?	7-9
Existe-t-il d'autres menaces à la validité interne de l'étude ?.....	7-10
Statistiques descriptives	7-12
Analyse des impacts	7-12
Après l'enquête finale : Déterminer l'impact du programme	7-13
Rapport final	7-13
Diffusion.....	7-14

L'EI a pour principal objectif d'évaluer l'impact d'une intervention de FBR et de partager les résultats obtenus avec les responsables afin d'éclairer les décisions politiques. La diffusion des résultats est une étape essentielle de l'EI, mais les équipes sous-estiment souvent le temps et les ressources nécessaires pour produire une analyse des données et des produits de diffusion de qualité. Les équipes devraient identifier à l'avance des analystes de données qualifiés disponibles après la réalisation des enquêtes de référence, à mi-parcours et en fin de projet (Voir calendrier au 0) afin de produire des résultats en temps opportun.

Les équipes devraient également tenir compte de l'aspect de renforcement des capacités des évaluations d'impact : compte tenu de la technicité des évaluations d'impact, de nombreux pays à faible revenu n'ont pas les capacités nécessaires pour bien comprendre et réaliser des évaluations d'impact contrôlées et randomisées. La collaboration avec des chercheurs et analystes locaux peut contribuer à la promotion des évaluations d'impact auprès des responsables et renforcer les capacités locales à réaliser et analyser ces évaluations.

Plusieurs rapports devraient être produits dans le cadre de l'EI :

- Pour les différentes unités étudiées : ménages, établissements de santé, agents de santé, agents de santé communautaires, etc.
- Pour les différentes étapes du cycle d'EI : enquêtes de référence, à mi-parcours et de fin.

Ce module donne aux équipes d'EI un aperçu des principales recommandations concernant i) l'analyse des données de référence et la rédaction du rapport associé, ii) l'analyse d'impact et iii) la diffusion. Le tableau ci-dessous résume les principaux objectifs des rapports de référence, à mi-parcours et de fin.

Tableau 8 : Les rapports d'évaluation d'impact et leurs fonctions

Rapport de référence	<ul style="list-style-type: none"> • Valider la conception de l'EI : réaliser des tests sur la différence des moyennes entre les groupes de traitement et de référence pour évaluer l'équilibre de l'échantillon. Réaliser des contrôles de validité externe et interne • Produire des statistiques descriptives sur les unités étudiées • Formuler des recommandations concernant la mise en œuvre des enquêtes de suivi
Rapport à mi-parcours	<ul style="list-style-type: none"> • Présenter une première évaluation des impacts et du rapport coût-efficacité (si l'exposition est suffisante) • Formuler des recommandations concernant la mise en œuvre de l'intervention en cas d'impacts imprévus ou de problèmes opérationnels (intrants insuffisants, retard dans le décaissement des primes, etc.)
Rapport de fin	<ul style="list-style-type: none"> • Évaluer l'impact et le rapport coût-efficacité • Formuler des recommandations concernant le développement, la poursuite et les améliorations possibles de l'intervention. Formuler des recommandations de politique générale

Rapport de référence

Objectifs : Le rapport de référence a trois principaux objectifs :

- Contrôler la validité interne et externe de la méthode d'évaluation (voir ci-dessous)
- Présenter des statistiques descriptives sur les unités ou la population étudiées afin de documenter les caractéristiques de l'échantillon
- Formuler des recommandations en vue d'enquêtes futures le cas échéant.

Il est possible que d'autres objectifs doivent être intégrés dans le rapport, en particulier à la demande du chef d'équipe et/ou du gouvernement. Par exemple, le rapport de référence pourrait aborder des questions telles que l'accès aux soins, l'équité, etc.

Plan recommandé : Le plan du rapport dépend des enquêtes et des objectifs poursuivis au-delà de la validation et de la présentation des statistiques descriptives. Les deux plans suivants nous ont semblé utiles : Rapport de référence sur les ménages du projet de FBP axé sur les communautés du Rwanda (2010) et *L'Évaluation d'impact en pratique* (Gertler et al. 2011, p.213).

Figure 7 : Exemple de plan de rapport de référence

Plan 1 (Projet de FBP axé sur les communautés du Rwanda 2010)	Plan 2 (Gertler et al 2011)
1. Présentation	1. Introduction
1.1 Introduction	2. Description de l'intervention (avantages, critères d'admissibilité, etc.)
1.2 Contexte du projet	3. Objectifs de l'évaluation
1.3 Composantes du projet	3.1 Hypothèses, théorie du changement, chaîne de résultats
1.4 Objectifs de l'étude	3.2 Questions stratégiques
2. Méthodologie	3.3 Indicateurs de résultats clés
2.1 Randomisation	4. Méthode d'évaluation
2.2 Conception de l'étude	4.1 Méthode prévue initialement
2.3 Taille de l'échantillon et stratégie	4.2 Participants et non-participants effectifs
2.4 Variables d'analyse des données	5. Échantillonnage et données
2.5 Instruments de collecte des données et d'assurance de la qualité des données	5.1 Stratégie d'échantillonnage
2.6 Stockage et gestion des données et politique d'accès	5.2 Calculs de puissance
3. Représentativité de l'échantillon et validité externe	5.3 Données collectées
3.1 Représentativité géographique	6. Validation de la méthode d'évaluation
3.2 Comparaison entre l'étude de référence et la population du pays	7. Statistiques descriptives détaillées
4. Conclusions (organisées par section du questionnaire ou rubrique pertinente)	8. Conclusion et recommandations pour la mise en œuvre
5. Validité interne de l'étude	
5.1 Calculs de puissance ex-post	
5.2 Menaces à la validité interne	
5.3 Équilibre des échantillons : Résumé des résultats des tests	

Nettoyage des données

Veillez noter tout d'abord que l'enquêteur principal ne devrait pas partager des données non anonymes avec les autres analystes de données. Le fichier d'identification, qui fait le lien entre les données géographiques et autres et les codes d'identification de terrain, devrait être stocké en lieu sûr et les analystes de données devraient avoir uniquement accès aux codes d'identification de terrain (voir 0).

L'équipe d'EI devra nettoyer les fichiers de données brutes avant d'analyser les données et de produire des matériels de diffusion. Le nettoyage des données consiste à identifier les incohérences et les imprécisions figurant dans les données brutes et à réduire ces erreurs en faisant des hypothèses prédéfinies. Lors du processus de nettoyage des données, les analystes devraient documenter toute erreur récurrente ainsi que les solutions adoptées. Ces informations devraient être transmises à l'équipe de mise en œuvre de l'EI afin d'éviter l'utilisation de questionnaires ou de données incorrects dans le

cadre des enquêtes de suivi. Le nettoyage des données prend du temps et les équipes nationales d'EI devraient prévoir un délai suffisant avant le début de l'analyse et de la rédaction du rapport de référence.

Remarque : Le processus de nettoyage des données est directement lié à la qualité de la collecte et de la saisie des données. Les erreurs de données sont dues à une formation, une supervision et une assurance qualité insuffisantes lors de la préparation et de la mise en œuvre de l'évaluation d'impact. La mise en place de solides mécanismes d'assurance de la qualité des données dans le cadre du processus de préparation et de mise en œuvre permet de minimiser le nombre d'erreurs figurant dans les fichiers de données brutes, et donc de réduire les besoins de nettoyage des données.

Création de variables aux fins d'analyse

Les analystes de données peuvent créer deux types de variables aux fins d'analyse :

- **Les indicateurs étudiés sur les intrants, les processus, les extrants et les résultats** : Ces variables sont les plus importantes car elles permettent aux analystes d'évaluer l'efficacité, le rapport coût-efficacité et l'impact de l'intervention de FBR après la ou les enquêtes de suivi. Elles devraient être définies conformément aux définitions internationales et nationales (voir le 0 sur les indicateurs). Les équipes nationales d'EI devraient également consulter les experts nationaux de la santé et/ou des FBR pour identifier d'autres indicateurs spécifiques aux pays devant être inclus. Elles devraient veiller à inclure la population appropriée dans leur calcul, avec des catégories d'âge et une répartition entre les sexes correspondant aux directives et définitions.
- **Covariables** : Ces variables n'évaluent pas l'impact du programme mais documentent les principales caractéristiques et le comportement de la population incluse dans l'enquête.

Les données brutes sont composées des variables continues déjà présentes dans l'ensemble de données (par exemple, l'âge ou les revenus des personnes interrogées) ou de variables dichotomiques 0-1. Pour certaines variables, les analystes de données convertiront les réponses en variables dichotomiques 0-1.

- Par exemple, l'enquête sur les ménages contient la question suivante : « De quoi souffrait NOM ? ». Les réponses sont codées à l'aide d'un chiffre compris entre 01 et 20 et représentant chacun une maladie. Dans ce cas, l'analyste devra créer 20 variables dichotomiques 0-1 pouvant prendre la valeur 1 si la personne souffre de la maladie en question et 0 dans le cas contraire. L'analyste pourra ensuite calculer la moyenne de chaque variable, qui correspondra au pourcentage de personnes interrogées souffrant surtout de cette maladie.

*Il est impératif que les analystes des données documentent le processus de nettoyage des données et de construction des variables, essentiellement à l'aide de **fichiers do Stata®** annotés et ce, pour pouvoir : i) reproduire les résultats et ii) améliorer l'efficacité de toute analyse ultérieure. L'équipe nationale d'EI devrait être en mesure de reproduire les résultats et de fournir les définitions et hypothèses adoptées ainsi que les variables construites. En outre, les instruments d'enquête et les programmes de saisie des données étant semblables dans le cadre des enquêtes de référence et de suivi, les activités de construction et d'analyse des variables de référence assument les coûts fixes associés à l'écriture du code d'analyse, qui peut ensuite être utilisé pour d'autres ensembles de données.*

Validation de la méthode d'évaluation

Le rapport de référence doit évaluer la validité externe et interne de l'EI.

Validité externe

L'évaluation de la validité externe devrait répondre à la question suivante : Dans quelle mesure les résultats de l'enquête peuvent-ils être généralisés à un plus grand groupe de population ? La stratégie d'échantillonnage définit la population ciblée par l'enquête ainsi que ses modalités de sélection mais l'analyse doit indiquer dans quelle mesure elle est comparable à la population nationale ou à des groupes de population infranationaux. L'évaluation de la validité externe est essentiellement fondée sur le plan d'échantillonnage et complétée par les résultats de l'échantillonnage effectué sur le terrain avant le début du programme. Les résultats de l'enquête d'EI peuvent être comparés aux autres données disponibles (par exemple, enquêtes démographiques et sanitaires) pour évaluer la comparabilité des données à d'autres mesures.

Validité interne

L'évaluation de la validité interne est l'un des principaux objectifs du rapport de référence. La majorité des contrôles de validité interne consiste à comparer la conception théorique de l'EI aux résultats de terrain. Les contrôles de validité devraient répondre aux questions suivantes :

- La puissance statistique est-elle suffisante ? La taille de l'échantillon définie pour chaque unité (établissements, agents de santé, ménages, bénéficiaires) est-elle suffisante ?
- Les groupes de traitement et de référence sont-ils équilibrés ? Les moyennes des groupes de traitement et de référence sont-elles équilibrées ?
- Existe-t-il d'autres menaces à la validité interne ? Des caractéristiques préexistantes ou une intervention préalable menacent-elles la validité interne de l'étude ?

La puissance statistique est-elle suffisante ?

Après avoir collecté les données de référence, l'équipe peut utiliser les valeurs de référence des principales variables, la taille de l'échantillon choisie, le niveau de puissance désiré et le niveau de signification statistique souhaité pour déterminer l'effet minimum détectable d'un résultat donné, c'est-à-dire le plus petit effet que l'expérience pourra déceler. Si cet effet minimum détectable est plus petit que l'effet escompté de l'intervention, il est probable que l'expérience détectera cet effet escompté (c'est-à-dire aura la puissance statistique nécessaire) au niveau choisi de signification statistique.

Les calculs de puissance ex-post sur les données de référence sont utiles dans les dialogues sur les politiques nationales menés en cours d'intervention. Par exemple, l'importance de l'effet minimum détectable peut également influencer la durée de l'évaluation : si l'effet minimum détectable est élevé, l'intervention devra avoir un effet significatif pour qu'il soit décelé lors de l'évaluation d'impact. Pour que l'impact soit élevé, il pourra être nécessaire de poursuivre l'intervention pendant plus longtemps dans le groupe de traitement tout en maintenant le groupe de référence, en supposant que l'effet de l'intervention continue d'augmenter avec le temps. Un effet minimum détectable plus élevé nécessitera donc une intervention plus longue.

Les calculs de puissance peuvent être effectués par des logiciels tels que Optimal Design¹⁸, disponible gratuitement auprès de la Fondation William T. Grant (<http://www.wtgrantfoundation.org>). Le logiciel et sa documentation sont disponibles en ligne sur le site de la Fondation à :

http://sitemaker.umich.edu/group-based/optimal_design_software

Le code nécessaire aux calculs de puissance pour les essais randomisés en grappes en R (<http://www.r-project.org/>), un langage et un environnement libres de calcul et de diagrammes statistiques, est également accessible gratuitement auprès des auteurs du logiciel Optimal Design.

Un exemple de **Calculs de puissance** ex-post est disponible dans le 0 de la boîte à outils ; il utilise les résultats des données de référence collectées dans le cadre de l'intervention de FBP axé sur les communautés du Rwanda. Pour plus de détails sur les calculs de puissance, voir également le 0.

Les caractéristiques de référence sont-elles équilibrées ?

Le rapport de référence devrait contenir des tests statistiques sur la différence des caractéristiques entre les groupes de traitement et de référence. Lorsque l'EI a plusieurs branches de traitement, nous recommandons l'utilisation des deux tests statistiques suivants :

¹⁸ Raudenbush, S. W., et al. (2011). Optimal Design Software for Multi-level and Longitudinal Research (Version 3.01) [Logiciel]. Disponible à www.wtgrantfoundation.org ou sitemaker.umich.edu/group-based.

- Le test F général de différence des moyennes des principaux indicateurs et des covariables dans toutes les branches de l'étude.
- Les tests T de différence des moyennes des principaux indicateurs et des covariables entre une branche d'étude et une autre, en mettant l'accent sur la différence entre chaque groupe de traitement et le groupe de référence.

Les tests T sont suffisants si l'EI n'a qu'une branche de traitement.

Le rapport de référence devrait contenir une évaluation générale du pourcentage des indicateurs et covariables non équilibrés, en faisant la distinction entre les résultats des tests F et des tests T. Les résultats des tests T et F obtenus pour chaque section du questionnaire devraient être examinés dans le rapport lorsque les statistiques descriptives sont présentées.

Existe-t-il d'autres menaces à la validité interne de l'étude ?

Le rapport de référence devrait évaluer les autres menaces potentielles à la validité interne de l'étude :

- Des interventions formelles ou informelles existantes dans le pays imitent-elles ou, au contraire, compensent-elles l'intervention ? Pourraient-elles produire des réactions inattendues à l'intervention de FBR dans les groupes de traitement ou de référence ?
- Des personnes interrogées ont-elles mentionné recevoir des récompenses susceptibles d'influencer leur comportement dans la même direction que l'intervention de FBR ou dans la direction opposée ? Dans l'affirmative, l'importance de ces mesures incitatives est-elle suffisante pour menacer celles du FBR et ces mesures sont-elles différentes entre les groupes de traitement et de référence ?
- Respect du plan d'échantillonnage dans le cadre des activités de terrain : les analystes de données devraient comparer les nombres théoriques et effectifs d'unités incluses dans l'enquête (unités géographiques, ménages, établissements de santé, etc.) et rendre compte des taux de non-réponse. Ils devraient vérifier si le non-respect ou la non-réponse sont différents entre les groupes de traitement et de référence. En cas de non-respect ou de taux de non-réponse élevé, les analystes de données devraient formuler des recommandations à l'équipe chargée de l'enquête pour améliorer les enquêtes de suivi.

Étude de cas nationale : Validation de la conception de l'EI et utilisation des calculs de puissance ex-post dans les débats d'orientation

Projet de financement basé sur la performance axé sur les communautés du Rwanda

Dans le cadre de l'évaluation de la conception de l'EI, l'équipe a réalisé des calculs de puissance ex-post sur les données de référence collectées. L'objectif était de veiller à ce que la taille des échantillons et l'échantillonnage de terrain permettent de déceler un effet d'une ampleur raisonnable, pour une puissance et un niveau de confiance donnés, et d'évaluer dans quelle mesure les résultats de l'étude de référence correspondent aux résultats escomptés lors de la phase de conception. L'équipe de recherche a utilisé un ensemble d'indicateurs de résultats calculés à partir des données de référence sur les ménages. Les 14 indicateurs binaires ont été sélectionnés sur la base de la probabilité qu'ils ressentent les effets du programme :

- Couverture des soins prénatals (1 visite ou plus),
- Soins prénatals opportuns (avant le 4^e mois de la grossesse),
- Couverture des soins prénatals (4 visites ou plus),
- Couverture TT2 pendant la grossesse,
- Supplémentation martiale de 90 jours pendant la grossesse,
- Présence de personnel qualifié lors de l'accouchement,
- Accouchement dans un établissement de santé,
- Nouveau-nés de faible poids de naissance,
- Initiation opportune à l'allaitement maternel,
- Allaitement maternel exclusif (0-6 mois),
- Visite de soins postnatals dans un établissement de santé,
- Supplémentation postnatale en vitamine A,
- Prévalence des méthodes contraceptives formelles,
- Besoins non satisfaits en matière de planification familiale.

Le modèle statistique a été défini en se fondant sur la conception de l'étude : un essai randomisé en grappes avec blocs à deux niveaux, où les secteurs étaient définis par niveau de pauvreté et les données regroupées au niveau du secteur. Le traitement a été alloué au niveau du secteur. Le taux d'erreur de type 1 a été défini égal à 0,05 et la puissance désirée à 80 %. Pour l'ensemble des résultats étudiés, l'effet minimum détectable se situait entre 0,06 et 0,12. Les calculs de puissance ex-ante étant fondés sur un effet minimum détectable de 0,2, l'équipe a conclu que l'ampleur des effets minimums détectables ex-post était dans la limite prévue et pouvait être atteinte par l'intervention.

Ces résultats ont suscité deux débats d'orientation importants au sein de l'équipe.

- Pour les indicateurs qui étaient déjà élevés, les résultats des calculs de puissance ont soulevé la question de la capacité à atteindre ces effets minimums détectables. Par exemple, la présence de personnel qualifié lors de l'accouchement devait passer de 89 % (niveau de référence) à 96 % à la fin du programme pour que l'augmentation soit détectée par l'EI. L'équipe a évalué comment les programmes existants non fondés sur les résultats, qui sont axés sur les agents de santé communautaires pour améliorer la santé maternelle, pourraient avoir un effet sur les indicateurs les plus difficiles à atteindre.
- L'ampleur des effets minimums détectables a été mise en contraste avec l'amélioration relativement lente des indicateurs clés indiquée par les données de suivi. L'équipe a donc conclu que la durée de l'expérience devait être étendue pour que ces effets minimums soient produits et donc détectés par l'EI. Le gouvernement a décidé de maintenir les groupes de traitement et de référence jusqu'en janvier 2013 plutôt que jusqu'en juin 2012, comme cela était initialement prévu.

Pour de plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.

Statistiques descriptives

Réalisation de l'analyse et ventilation par catégorie : Le rapport de référence est généralement consacré à la production de statistiques sur les caractéristiques des unités étudiées. Ces statistiques proviennent des principaux indicateurs étudiés sur les intrants, le processus, les extrants et les résultats, ainsi que les covariables. Le rapport devrait présenter des estimations des moyennes de ces variables pour :

- L'intégralité de l'échantillon sondé par questionnaire
- L'échantillon sondé par questionnaire pour chaque branche de l'étude.

Les moyennes peuvent être produites en fonction de critères démographiques ou géographiques : elles peuvent être présentées pour comparer les ménages urbains et ruraux, les établissements de santé ou les agents de santé communautaires, les participants de sexe masculin ou féminin ; les différentes catégories d'âge ; différentes personnes telles que les chefs de ménage, les enfants de moins d'un an, les enfants de moins de 5 ans, etc. ; les établissements de santé publics ou privés, etc. En fonction des objectifs de politique et de la comparaison potentielle des statistiques à d'autres enquêtes, une ventilation des statistiques par catégorie pourra être utile.

Présentation des statistiques descriptives : Les statistiques descriptives obtenues peuvent être présentées dans le corps du rapport ou y être annexées. La boîte à outils contient du **code Stata®** pour produire les tableaux et les diagrammes pouvant être utilisés pour illustrer les principales caractéristiques des personnes interrogées. Les diagrammes sont particulièrement utiles lorsque l'équipe désire diffuser les résultats et les rendre accessibles à un vaste public et aux décideurs.

Analyse des impacts

L'impact et le rapport coût-efficacité des interventions de FBR ne peuvent être évalués que lorsque les enquêtes à mi-parcours et/ou de fin ont été achevées.

Une enquête à mi-parcours peut fournir une analyse précoce des impacts du programme, en ce qu'elle peut évaluer les modifications des intrants, processus, extrants et résultats après un laps de temps limité. Il est toutefois possible que le temps écoulé depuis le début du programme ne soit pas suffisant pour que les résultats aient changé. Les résultats à mi-parcours peuvent être utiles pour formuler des recommandations concernant la modification de l'intervention si des changements inattendus sont observés, tels que des impacts négatifs, une pénurie d'intrants, une répartition incorrecte des incitations, un accroissement des coûts, et une affectation erronée des groupes de traitement et de référence. Dans ce dernier cas, la méthode d'analyse d'impact devrait être redéfinie pour tenir compte du fait que des personnes concernées par l'intervention n'ont pas été traitées ou que des personnes ne remplissant pas les conditions voulues en aient bénéficié.

Après l'enquête finale : Déterminer l'impact du programme

L'analyse d'impact devrait être réalisée conformément à la stratégie d'identification définie dans le document de conception de l'EI. *L'évaluation d'impact en pratique* (Gertler *et al.* 2011) examine en détail comment évaluer les impacts en fonction de l'affectation des groupes de traitement et de référence. Comme pour les données de référence, l'analyse des impacts en fin de programme devrait être réalisée sur des données propres une fois que les indicateurs étudiés et les covariables ont été créés. Si l'intervention comprend plusieurs groupes de traitement, l'impact de l'intervention devrait être évalué dans chaque groupe de traitement en le comparant au groupe de référence. Les groupes de traitement devraient également être comparés entre eux pour identifier la meilleure stratégie de FBR.

D'autres aspects peuvent être pris en compte lors de l'analyse des impacts. En particulier, les impacts de répartition du FBR et son effet sur les plus démunis sont d'importants aspects dont il convient de rendre compte. L'analyse de l'équité peut éclairer les décisions de politique nationale et contribuer aux connaissances mondiales sur les FBR et ainsi permettre de déterminer s'il s'agit d'un instrument stratégique pour améliorer les systèmes de santé tout en préservant ou améliorant l'accès aux soins des populations démunies.

Nous recommandons que les équipes d'EI utilisent des instruments d'enquête et des programmes de saisie des données semblables entre les enquêtes de référence et de fin de programme : le nettoyage et l'analyse des données en fin de programme seront semblables au travail réalisé sur les données de référence. Cela permettra de minimiser le temps et les efforts nécessaires au moment de consigner les résultats.

Rapport final

Les principaux résultats de l'évaluation d'impact feront vraisemblablement l'objet d'un article rédigé par l'équipe et son enquêteur principal. Toutefois, un rapport plus détaillé présentant tous les résultats de l'évaluation, y compris ceux qui ne seraient peut-être pas retenus aux fins de publication, pourra souvent être plus utile au gouvernement concerné. Nous recommandons donc que l'équipe d'EI prépare également un rapport final détaillé en plus d'articles destinés à être publiés. En outre, l'équipe devrait réfléchir à la meilleure façon de présenter et diffuser les résultats aux décideurs : la présentation des résultats et la formulation de recommandations en matière de politique ne devraient pas être les uniques buts du rapport. Le tableau ci-dessous présente un exemple de plan de rapport final.

Figure 8 : Exemple de plan de rapport final

Plan de rapport d'impact	
1. Présentation	
1.1 Introduction	
1.2 Contexte du projet	
1.3 Composantes du projet	
1.4 Objectifs de l'étude	
2. Méthodologie	
2.1 Randomisation	
2.2 Conception de l'étude	
2.3 Taille de l'échantillon et stratégie	
2.4 Variables d'analyse des données	
2.5 Instruments de collecte des données et d'assurance de la qualité des données	
2.6 Stockage et gestion des données et politique d'accès	
3. Validité interne et validité externe	
3.1 Validité externe	
3.2 Validité interne	
4. Limites de l'étude	
5. Principales statistiques descriptives par branche d'étude	
6. Conclusions	
6.1 Analyse des impacts (par intrant, processus, extrant, résultat ou indicateurs d'impact et comparaison des groupes de traitement)	
6.2 Rapport coût-efficacité (par intrant, processus, extrant, résultat ou indicateurs d'impact et comparaison des groupes de traitement)	
7. Conclusions intéressantes sur les FBR – À définir, par exemple impacts de répartition des FBR	
8. Recommandations et possibilités d'action	

Diffusion

La diffusion opportune et appropriée est essentielle pour veiller à ce que l'évaluation d'impact influence les politiques. Nous recommandons les méthodes de diffusion suivantes :

- **Document écrit :** Les rapports de référence et d'impact devraient être diffusés au gouvernement, à la direction de la Banque et à toutes les autres parties prenantes avant d'être publiés. La série des Documents de travail de la Banque mondiale publie de longue date les résultats préliminaires des évaluations d'impact. Il est toutefois important de garder à l'esprit que plusieurs revues médicales (telles que *The Lancet* et le *New England Journal of Medicine*) peuvent rejeter un article si les résultats sont contenus dans un autre rapport (antérieur) disponible en ligne, y compris des documents de travail. Il est donc recommandé de se renseigner sur les politiques des revues avant de publier des résultats en ligne.
- **Présentation dans les pays et ateliers de diffusion :** Il peut être utile de diffuser les résultats aux parties prenantes nationales et locales dans le cadre d'un atelier. Les parties prenantes locales peuvent contribuer à l'interprétation de résultats inattendus tandis qu'un public plus vaste peut être atteint à l'échelon national.

- **Notes d'orientation** : Des résumés analytiques sur la conception du programme, la méthode d'évaluation et les résultats peuvent être utiles dans le cadre des dialogues sur les politiques nationales menés dans d'autres pays, à condition qu'ils soient facilement disponibles. Les séries « En Breve » et « HD Brief » peuvent être utiles pour ce type de diffusion.
- **Ateliers et événements internationaux** : Le recueil de données nationales permettra à terme d'évaluer si le FBR est un mécanisme efficace et d'un bon rapport coût-efficacité pour améliorer les systèmes de santé et la prestation des services. La diffusion des résultats entre les pays est également fort utile aux pays moins avancés en termes d'évaluation d'impact et d'exécution de projets.

Nous recommandons vivement de faire participer les chercheurs et analystes locaux au processus de diffusion pour faciliter des dialogues sur les politiques nationales reposant sur des bases factuelles.

Remarque : L'équipe d'EI devrait prévoir à l'avance le type d'activités de diffusion requises dans le pays et au siège pour veiller à leur inclusion dans le budget (voir 0).

2012/04/03

Étude de cas nationale : Diffusion des résultats en début et fin de programme et dialogue sur les politiques nationales

Projets de financement basé sur la performance axé sur les établissements de santé et les communautés du Rwanda

En septembre 2011, le Gouvernement rwandais a organisé un atelier de trois jours sur les programmes de FBP axés sur les établissements de santé et les communautés du pays. Cet atelier avait pour objectif de diffuser les résultats disponibles sur ces programmes de sorte à susciter un dialogue sur les politiques aux niveaux communautaire, local et central. À terme, les décideurs du gouvernement central voulaient recenser les problèmes de mise en œuvre et les goulots d'étranglement et élaborer des solutions en collaboration avec les représentants infranationaux. De plus, le ministère de la Santé voulait diffuser les résultats de ces deux programmes afin d'utiliser les enseignements tirés du premier programme de FBP lors de la mise en œuvre du second programme de FBP visant les communautés. L'atelier a également encouragé la participation d'étudiants de l'École nationale de santé publique, partenaire local de recherche : les étudiants ont ainsi présenté les résultats de l'analyse des données des interventions menées dans le pays dans le secteur sanitaire.

Les résultats diffusés sur le programme de FBP visant les établissements de santé reposaient sur la publication dans une revue à comité de lecture des résultats finaux concernant l'impact du programme sur la santé infantile (Basinga et al. Lancet 2011). Ces résultats ont indiqué un effet positif du programme sur plusieurs résultats en matière de santé infantile et ont appuyé la promotion du FBP par le gouvernement auprès des parties prenantes locales, des donateurs et du grand public en présence des médias.

L'atelier était essentiellement consacré au plus récent programme de FBP axé sur les communautés. Les premiers résultats présentés reposaient sur l'analyse des données de référence collectées dans le cadre de l'EI. L'équipe a utilisé les moyens les plus conviviaux pour partager ses résultats avec le public, notamment de simples tableaux de statistiques descriptives et de nombreux diagrammes. Les participants ont été invités à s'exprimer et à poser des questions sur les résultats le plus souvent possible. Des sessions de groupe ont été organisées, au cours desquelles les participants pouvaient utiliser les résultats de référence (y compris des tableaux de statistiques synthétiques sur des sujets pertinents) pour examiner un aspect spécifique de la mise en œuvre du programme (par exemple,

participation financière des usagers, suivi et vérification). Ils ont recensé des problèmes de mise en œuvre reflétés au niveau des données ou, au contraire, des anomalies entre les données et leur propre expérience de terrain et ont proposé des solutions susceptibles d'être prises en compte et intégrées dans le programme. Les participants ont ensuite présenté leur analyse et proposé leurs solutions à toute l'assemblée.

Pour de plus amples détails, voir la section « Pleins feux sur le pays » de la boîte à outils.



Module 8

Suivi et documentation des programmes FBR

Boîte à outils pour l'évaluation d'impact
Mesure de l'impact sur la santé maternelle et infantile du financement basé sur les résultats
Christel Vermeersch, Elisa Rothenbühler, Jennifer Renee Sturdy

www.worldbank.org/health/impactevaluationtoolkit

Module 8. Suivi et documentation des programmes FBR

Principales recommandations et outils disponibles pour ce module			
Recommandations	Critique	Important	Utile
<ul style="list-style-type: none"> Le suivi et la documentation des activités du projet constituent un complément crucial de l'évaluation d'impact parce qu'ils fournissent des informations sur les interventions effectives sur le terrain et par conséquent sur l'intervention que l'on est en train d'évaluer. Les équipes d'évaluation d'impact souhaiteront que le programme identifie deux risques majeurs pour l'évaluation d'impact : 1) une compensation offerte au groupe de comparaison sous la forme d'une intervention ou d'un programme alternatifs ; et 2) l'imitation du traitement par le groupe de comparaison. 		<p style="text-align: center;">✓</p> <p style="text-align: center;">✓</p>	
			
Outils			
<ul style="list-style-type: none"> 8.01 Monitoring Indicators Rwanda Example (Modèle d'indicateurs de suivi pour le Rwanda) 8.02 Field Supervision Visit Rwanda Templates (Modèles de visites de supervision sur le terrain) 			

Contenu du module

Suivi des interventions FBR.....	8-3
Suivi et documentation dans le contexte de l'évaluation d'impact	8-6
Risque #1 : Compensation accordée au groupe de comparaison	8-7
Risque #2 : Imitation du traitement par le groupe de comparaison	8-8
Sources de données complémentaires.....	8-9

L'évaluation d'impact fait partie d'une vaste gamme de méthodes complémentaires à l'appui d'une politique fondée sur des éléments probants, qui inclut le suivi et l'évaluation. On peut trouver chez Görgens et Kusek (2009) une description générale des systèmes de suivi. Dans ce module, nous nous concentrerons sur les aspects du suivi qui ont un rapport direct avec l'évaluation des effets des projets FBR, à savoir : i) le suivi des interventions FBR, ii) l'adhésion au modèle d'évaluation et à la façon de réduire les risques pour la validité interne durant les opérations, et iii) l'utilisation de sources de données complémentaires de l'EI.

Suivi des interventions FBR

Dans ce module, nous utiliserons la définition suivante du suivi : « Collecte systématique d'informations sur les intrants, les activités, et les produits d'un programme, ainsi que sur son contexte et ses autres caractéristiques clés » (Centres de lutte et de prévention des maladies, 2008). Cette description de la mécanique et de la fonction d'un programme dans le temps est également appelée communément « évaluation du processus ». Le suivi ou l'évaluation du processus devraient faire partie intégrante de l'évaluation globale de tout plan FBR.

Un système de suivi efficace comporte de nombreux éléments :

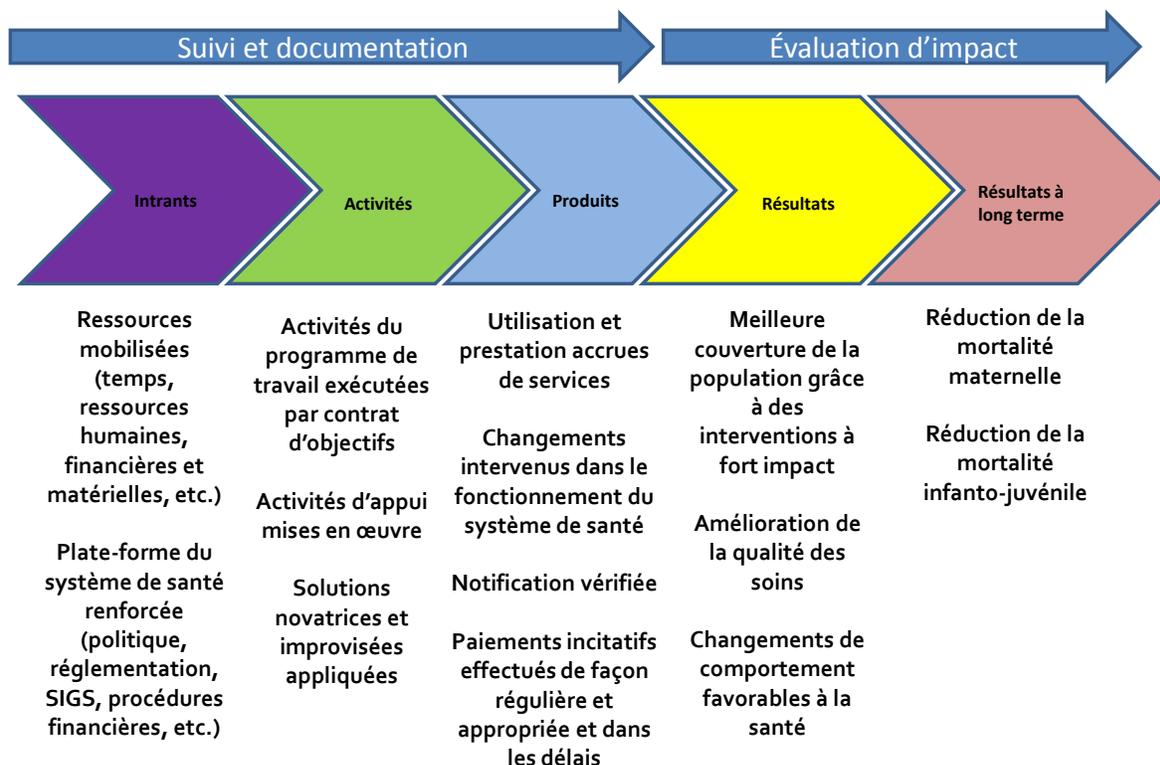
- Établissement d'**indicateurs des intrants, des activités et des produits du programme**
- Mise en place de systèmes de collecte d'informations relatives à ces indicateurs
- Collecte et enregistrement de l'information
- Analyse de l'information
- Utilisation de l'information pour faciliter la gestion courante.

Objet du suivi et de la documentation. Le suivi et la documentation continus de l'exécution d'un plan FBR répondent à plusieurs fins. Tout d'abord, le suivi permet d'identifier les problèmes opérationnels et autres susceptibles de donner lieu à des mesures correctives à un stade précoce de l'exécution. Par ailleurs, le suivi permet de tenir les responsables du gouvernement et les partenaires de développement régulièrement informés des progrès réalisés par rapport aux objectifs des interventions. De plus, les informations tirées du suivi et de la documentation aideront les équipes EI à mieux comprendre et interpréter les conclusions de leur évaluation et en particulier à déterminer si les effets escomptés ont été ou non obtenus. Enfin, en fournissant une riche description de la mécanique de l'exécution du FBR — intrants, activités et produits — et le contexte dans lequel elle s'inscrit, le suivi aide à répondre à la question de savoir *pourquoi* ces interventions réussissent ou échouent. Ces informations sont cruciales pour déterminer dans quelle mesure elles peuvent être reproduites dans différents contextes et répondre aux demandes d'informations sur la façon de les mener à bien.

Questions fondamentales de suivi et de documentation. Une grande quantité d'informations pourrait être obtenue au moyen du suivi. Il est donc important d'avoir une vision stratégique de la nature des informations à recueillir. Les concepteurs des plans FBR devraient s'efforcer de distinguer entre les informations essentielles et complémentaires qui aident à choisir des indicateurs et les méthodes à

utiliser, et à prendre des décisions en matière de budget. Nous proposons une stratégie pratique destinée à répondre à trois questions fondamentales. On trouvera ci-après une description des domaines d'investigation complémentaires qui font généralement partie du champ couvert par le suivi, ou qui se trouvent eux-mêmes au point de convergence du suivi et de l'évaluation du FBR.

Figure 1. Cadre courant de suivi et d'évaluation du FBR



Q1. [Intrants] Quels intrants a-t-on utilisés pour assurer l'exécution du FBR, notamment quelles ressources humaines et financières, les politiques et les procédures ?

La stratégie de suivi et de documentation devrait être suffisamment solide pour capter la nature, le volume et l'opportunité des diverses ressources qui sont mobilisées par l'intervention FBR, ainsi que les types de décisions et de mesures prises pour faciliter l'exécution, notamment, si possible, leurs coûts connexes. Il convient par exemple de suivre et documenter les intrants suivants :

- Investissements initiaux sous la forme d'intrants : bien que le concept de « paiement pour des résultats » soit au cœur des interventions FBR, les centres nécessiteront quelques investissements initiaux sous la forme d'intrants.

- Modification des lois et règlements existants et/ou politiques et procédures en matière de santé.
- Signature d'accords de performance, de quasi-contrats et de contrats (Loevinsohn, 2008).
- Amélioration du système d'information de gestion sur la santé (SIGS).
- Révisions et améliorations des procédures de gestion financière.
- Ouverture de comptes bancaires par les centres de santé.

Q2. [Activités] Quelles activités ont été entreprises dans le cadre de l'intervention FBR? Quels ont été les facteurs positifs et négatifs qui ont influé sur l'exécution de ces activités?

La mise en place d'un FBR nécessitera diverses activités menées au niveau aussi bien central que local. La stratégie de suivi devrait collecter des informations permettant de déterminer si et dans quelle mesure ces activités souhaitées, qui se résument souvent à des plans de travail assortis de délais, ont été menées à bien comme prévu. Il conviendrait également de documenter tout facteur positif ou négatifs. Voici quelques exemples d'activités que les équipes devraient suivre :

- Formation des agents de santé pour qu'ils sachent ce qu'ils peuvent attendre de leur participation au plan FBR, notamment les récompenses et sanctions éventuelles ;
- Supervision additionnelle dans les centres de santé ;
- Activités d'éducation et de communication, ainsi qu'efforts de vulgarisation au sein des communautés et des ménages afin d'accroître la demande de services ;
- Octroi de subventions de transport aux femmes pour les aider à surmonter les obstacles à l'accès aux services ;
- Assistance technique provenant de sources aussi bien internes qu'externes, fournie au niveau national et infranational.
- De plus, les plans FBR peuvent fournir des solutions novatrices, souvent improvisées pour surmonter les obstacles qui entravent la prestation et l'utilisation des services. Il peut s'agir de mesures simples et peu coûteuses prises au point de prestation des services ou au sein des ménages et des communautés. Du fait que ces obstacles sont imprévisibles et qu'il est impossible de préciser les solutions à y apporter dans des accords ou contrats de performance, les équipes d'EI et du projet devront s'appuyer sur le suivi de la documentation pour comprendre exactement comment le plan FBR a éliminé les obstacles au niveau local.

Q3. [Produits] Dans quelle mesure les services liés à des objectifs de performance ont-ils été fournis et utilisés ? Dans quelle mesure l'exactitude des rapports a-t-elle été vérifiée ? Les incitations, financières ou non, ont-elles été offertes et reçues comme prévu ?

Les produits découlent directement des activités FBR. La crédibilité des interventions FBR dépend de façon cruciale de i) la disponibilité d'informations fiables et valides sur les services fournis et utilisés (c'est-à-dire sur les produits), comme indiqué dans des accords ou contrats de performance ; et ii) de la question de savoir si les paiements ont été effectués en temps voulu au profit des bons prestataires et/ou bénéficiaires, pour la bonne raison et au bon moment. Pour pouvoir expliquer pourquoi un plan FBR a fonctionné ou a échoué, l'équipe d'EI aura besoin de ces informations cruciales. Les facteurs qui ont pu faciliter ou entraver la fourniture ou l'utilisation des services devraient également être documentés. On trouvera ci-après une liste de produits que les équipes souhaiteront éventuellement suivre :

- Fourniture de services ciblés ou réalisation d'objectifs pour des indicateurs de produits préétablis, comme indiqué dans des accords ou contrats de performance. Parmi ces produits peuvent figurer divers indicateurs de services concernant la nutrition, la santé infantile et les services aux mères et aux nouveaux nés, en particulier dans le cadre des plans axés sur les OMD 4 et 5.
- Changements explicites attendus dans le fonctionnement du système de santé.
- Fréquence, ponctualité et caractère approprié des paiements aux bénéficiaires, qu'il s'agisse des prestataires de services ou des ménages. Lorsque les paiements ne sont pas effectués en temps voulu ou que leur montant n'est pas celui qui était prévu, la crédibilité du plan FBR, qui repose sur le lien étroit entre performance et incitations, risque d'être compromise.
- Toute sanction, comme la retenue d'une partie des paiements, lorsque les résultats sont inférieurs à ce qu'ils auraient dû être.
- Toute sanction en cas de « trucage » ou de fraude, comme le retrait de prestataires du plan, ou bien la suspension de bénéficiaires.

On trouvera à la figure 1 un résumé de chaque étape du processus de suivi comme il ressort des trois questions fondamentales, et le rapport entre suivi et évaluation,.

Suivi et documentation dans le contexte de l'évaluation d'impact

Dans une évaluation d'impact prospective, les évaluateurs conçoivent l'évaluation avant le début de l'exécution du programme, quand le programme proprement dit en est encore au stade de la conception. À mesure que le programme se déploie, les exécutants peuvent apporter des changements à sa conception initiale. Un suivi et une documentation appropriés de ces changements informeront les

évaluateurs de la nature des changements intervenus et de leur date, et leur permettront de décider de les prendre en compte dans l'analyse. Si les modifications apportées à la conception du programme sont importantes, comme au niveau de la tranche d'âge des bénéficiaires prévus ou un système d'incitations modifié, les évaluateurs peuvent être contraints de modifier profondément la méthode d'évaluation.

Comme on l'a vu au Module 3, la validité interne de l'évaluation d'impact repose sur la capacité de comparer les groupes de traitement et de comparaison sans aucune différence en dehors de l'intervention FBR. Nous examinons au Module 5 comment la validité interne de la conception de l'évaluation d'impact pourrait être menacée par certaines pratiques de collecte des données. La validité interne de la conception de l'évaluation risque également d'être menacée durant l'exécution de l'intervention FBR, et l'atténuation de ces menaces exige un suivi et une documentation suffisants, et une collaboration intense sur place avec le gouvernement et d'autres partenaires. Nous examinons ci-après deux exemples courants :

Risque #1 : Compensation accordée au groupe de comparaison

Bien que l'équipe du projet et les homologues du gouvernement puissent être totalement d'accord sur la conception de l'évaluation, il est possible que le gouvernement ou d'autres partenaires souhaitent accorder une compensation au groupe de comparaison (au niveau des districts ou des centres de santé) pour le fait d'être exclu de l'expérience FBR. Pour cette raison, d'autres programmes axés sur la santé maternelle et infantile peuvent être introduits spécifiquement dans le groupe de comparaison. L'exemple suivant montre comment la validité interne s'en trouve affectée :

- Supposons qu'il y ait un accord entre l'équipe du projet et le gouvernement sur la conception de l'évaluation ; les groupes de traitement et de comparaison sont clairement définis et le groupe de comparaison attendra 24 mois avant de commencer à bénéficier de l'intervention FBR. La collecte des données de base est effectuée avec succès et les données de base montrent que les groupes ont des moyennes similaires pour les indicateurs de résultat clés. Au 12^e mois, les représentants des districts du groupe de comparaison font pression sur le gouvernement pour qu'il leur accorde une compensation du fait qu'ils sont exclus de l'expérience FBR. Un autre partenaire de développement envisage, hors FBR, un projet à grande échelle de santé maternelle et infantile et le gouvernement décide d'autoriser le partenaire à augmenter immédiatement l'échelle du projet dans le groupe de comparaison de façon à atténuer quelque peu la pression politique. Cela n'est pas discuté avec les équipes du projet FBR ou EI car on estime que cela n'est pas contraire aux accords conclus au sujet du projet FBR ni à la conception de l'évaluation. Les données finales sont collectées au 24^e mois comme convenu et montre des améliorations sensibles des résultats de SMI dans l'ensemble de l'échantillon d'évaluation. Toutefois, l'analyse d'impact des résultats ne révèle pas de différence entre le groupe de traitement et le groupe de comparaison, ce qui laisse à penser que l'intervention FBR n'influe pas sur les résultats de la santé maternelle et infantile. En réalité, l'intervention FBR et l'autre programme de SMI ont probablement amélioré l'un comme l'autre les résultats de SMI, mais il est impossible d'estimer le véritable impact du programme FBR puisque le groupe de

comparaison ne constitue plus une estimation précise de la comparaison contrefactuelle. L'évaluation d'impact finit donc par sous-estimer l'impact de l'intervention FBR. Sans résultats positifs à présenter, le gouvernement doit maintenant se battre pour défendre son investissement dans le programme et a des difficultés à trouver des bailleurs de fonds disposés à lui apporter leur appui.

Cela signifie-t-il donc que le gouvernement ne peut lancer AUCUN autre programme de SMI durant l'évaluation d'impact ? Pas vraiment. Des programmes additionnels peuvent être lancés tant que le groupe de traitement et le groupe de comparaison de l'évaluation en bénéficient autant l'un que l'autre.

- Dans l'exemple cité plus haut, le gouvernement aurait pu autoriser l'autre bailleur de fonds à lancer l'autre programme SMI tant qu'il le faisait dans tous les centres. Si cela n'était pas possible, il devait veiller à ce que le même nombre de centres de traitement et de comparaison bénéficient du programme. En langage technique, le gouvernement aurait dû veiller à ce que de nouveaux programmes soient déployés « orthogonalement à l'affectation aux groupes de traitement et de comparaison aux fins de l'évaluation FBR ».

Risque #2 : Imitation du traitement par le groupe de comparaison

Dans certains cas, il est possible que les districts/centres de santé affectés au groupe de comparaison n'entendent jamais parler de l'intervention FBR et continuent à fonctionner comme si de rien n'était. Toutefois, compte tenu de l'accessibilité de l'information et dans certains cas, de la promotion très publique de programmes innovants comme le FBR, des personnes travaillant dans le groupe de comparaison peuvent être assez entreprenantes et motivées pour mettre en place leur propre forme de FBR. Cela peut être difficile pour certaines réformes complexes du système de santé comme le PBF, mais c'est possible pour une stratégie d'incitation au niveau de la demande. On trouvera ci-après un exemple de la façon dont cela influe sur la validité interne :

- Supposons que l'équipe du projet et le gouvernement soient d'accord sur la conception de l'évaluation ; les groupes de traitement et de comparaison sont clairement définis et le groupe de comparaison attendra 24 mois avant de commencer à bénéficier de l'intervention FBR. Le recueil des données de base s'est bien passé et les données de base montrent que les moyennes des groupes sont similaires pour des indicateurs de résultats clés. Toutefois, au 12^e mois, il y a une conférence nationale sur la santé maternelle et infantile à laquelle participent les directeurs des centres de santé du pays tout entier. À cette conférence, les directeurs de centres de santé des districts de comparaison apprennent une nouvelle stratégie visant à accroître le recours aux services de soins prénatals, d'accouchement et post-natals qui prévoit des mesures d'incitation en nature en faveur des femmes pour les amener à recourir à ces services dans les centres. Les directeurs de centres de santé du groupe de comparaison reconnaissent que cela peut constituer une méthode très utile pour leur faire utiliser davantage les services de leurs propres centres. Utilisant leur budget basé sur les moyens du ministère des Finances, certains centres de santé situés dans les districts de comparaison commencent à offrir aux femmes ces mêmes ensembles d'incitations. Les données finales sont recueillies au 24^e mois comme convenu. Les

résultats de l'analyse d'impact montrent qu'il y a eu des progrès dans les résultats de SMI dans les groupes de traitement et de comparaison, et que l'amélioration est légèrement plus marquée dans le premier que dans le second, ce qui laisse à penser que le FBR n'a eu qu'un effet limité sur les résultats de SMI. Malheureusement, du fait que l'intervention FBR (dans ce cas, une stratégie d'incitation au niveau de la demande) a été également menée (partiellement) dans un sous-échantillon de districts de comparaison, une comparaison des résultats entre les centres de traitement et de comparaison nous amènera finalement à sous-estimer le véritable impact de l'intervention FBR.

Sources de données complémentaires

Nous avons estimé que le coût de la collecte des données d'enquête représentait environ 80 % du coût total d'une évaluation d'impact, et que par conséquent, les sources de données complémentaires peuvent donc constituer un important moyen d'économiser des ressources financières et de fournir des résultats en temps opportun¹⁹. Des sources de données complémentaires susceptibles d'être utiles dans le contexte des plans FBR sont par exemple :

- Des enquêtes régulières (recensement, enquête démographique et sanitaire (EDS))
- Un suivi régulier (tests de performance annuels)
- Dossiers administratifs (carnets de santé, effectifs scolaires).

¹⁹ Avant d'utiliser des sources de données complémentaires pour l'évaluation d'impact, nous recommandons que les équipes d'EI évaluent la qualité de ces sources.

**Pleins feux sur le pays : Utilisation des données administratives pour suivre l'impact du FBR
et justifier le renforcement des capacités locales
Plan Nacer d'Argentine**

L'EI nationale [du Plan Nacer] s'est déroulée en deux phases, dans neuf provinces durant la Phase I et dans les provinces restantes durant la Phase II. En 2007, le programme a été étendu au reste du pays, ce qui a en fait contaminé le groupe de comparaison pour les provinces de la phase I, tandis qu'aucune donnée complémentaire n'avait encore été collectée du fait du délai nécessaire pour faire appel à un bureau d'études. (...) L'équipe de la Banque mondiale a recherché des stratégies alternatives pour obtenir des résultats provisoires à l'appui du dialogue de politique générale : (...) des données administrative. Dans deux provinces, Misiones et Tucuman, on a trouvé des sources de données parmi lesquelles figuraient tous les services de santé fournis dans les centres de soins de santé primaires du secteur public et dans les maternités (...) durant la période 2005-2008 dans la province de Tucumán et durant la période 2007-2009 dans celle des Misiones. Une caractéristique importante de ces données de base est le fait que les registres couvrent absolument tous les types de soins assurés, tant pour les participants au Plan Nacer que pour les autres. Au total, la base de données couvre plus de 2 750 000 services fournis durant la période d'analyse. La grande quantité de données et leur haute qualité offrent un excellent potentiel statistique en ce qui concerne l'identification des impacts, même dans le cas d'indicateurs à faible incidence comme la mortalité néonatale.

Les données des deux provinces pilotes ont été analysées en six mois, ce qui a permis à l'équipe de produire des résultats intermédiaires et d'organiser un événement de diffusion au début de 2010. Les résultats positifs n'ont fait que confirmer l'importance du programme et incité les autorités nationales et provinciales à réfléchir à l'utilisation de systèmes de suivi pour un traçage plus poussé de l'intervention et l'obtention de résultats supplémentaires.

On trouvera de plus amples détails dans la section « Pleins feux sur le pays » de la boîte à outils.

References

- Basinga P., P.J. Gertler, A. Binagwaho, A. Soucat, J. Sturdy and C. Vermeersch. 2011. "Effect on maternal and child health services in Rwanda of payment to primary health care providers for performance: an impact evaluation". *The Lancet*, Vol. 377(9775), pp. 1421-28.
- Bessinger, R. E. and J. T. Bertrand. 2001. "Monitoring quality of care in family planning programs: A comparison of observations and client exit interviews". *International Family Planning Perspectives*, Vol. 27(2), pp. 63-70.
- Bloom, E., I. Bhushan, D. Clingingsmith, R. Hong, E. King, M. Kremer, B. Loevinsohn and J.B. Schwartz. 2006. Contracting for Health: Evidence from Cambodia. Unpublished manuscript.
- Das, J. and J. Hammer. 2007. "Location, location, location: residence, wealth, and the quality of medical care in Delhi, India." *Health Affairs*, Vol. 26(3), pp. w338-w351.
- Das, J. and J. Hammer. 2004. "Strained Mercy: The Quality of Medical Care in Delhi". *Policy Research Working Paper*, No. 3228, The World Bank, Washington, D.C.
- Donabedian, A. 1980. "The definitions of quality and approaches to its assessment". *Health Administration Press, School of Public Health, The University of Michigan, Ann Arbor*.
- Donabedian, A. 1988. "The quality of care. How can it be assessed?" *JAMA*, Vol. 260, pp. 1743-8.
- Doran T., C. Fullwood, H. Gravelle, D. Reeves, E. Kontopantelis, U. Hiroeh, and M. Roland. 2006. "Pay-for-performance programs in family practices in the United Kingdom." *New England Journal of Medicine*, Vol. 355, pp. 375-384.
- Fiszbein, A. and N. Schady. 2009. "Conditional Cash Transfers: Reducing Present and Future Poverty". The World Bank, Washington, DC.
- Fleetcroft R., N. Steel, R. Cookson, S. Walker and A. Howe. 2012. "Incentive payments are not related to expected health gain in the pay for performance scheme for UK primary care: cross-sectional analysis.", *BMC Health Serv Res.*, Vol. 16, pp. 12-94.
- Franco, L. M., C. Daly, D. Chilongozi and G. Dallabetta. 1997. "Quality of case management of sexually transmitted diseases: Comparison of the methods for assessing the performance of providers." *Bulletin of World Health Organization*, Vol. 75(6), pp.523-32.
- Franco, L. M., C. Franco, M. Kumwenda and W. Nkhoma, W. 2002. "Methods for assessing quality of provider performance in developing countries." *International Journal of Quality in Health Care*. Vol. 14, pp. 17-24.

Gertler P.J., S. Martinez, P. Premand, L. Rawlings, and C. Vermeersch. 2011. "Impact Evaluation in Practice." The World Bank, Washington, D.C.

Gertler P.J. and C. Vermeersch. 2012. "Using Performance Incentives to Improve Health Outcomes," *Policy Research Working Paper Series*, Forthcoming, The World Bank, Washington, D.C.

Glickman, S., F. Ou, E. DeLong, M. Roe, B. Lytle, J. Mulgund, J. Rumsfeld, B. Gibler, M. Ohman, K. Schulman, E. Peterson. 2007. "Pay for Performance, Quality of Care, and Outcomes in Acute Myocardial Infarction." *JAMA*, Vol. 297(21), pp. 2373-2380.

Görgens, M. and J.Z.Kusek. 2009. "Making Monitoring and Evaluation Systems Work." *The World Bank, Washington, DC*.

Hermida, J., D. D. Nicholas and S.N. Blumenfeld. 1999. "Comparative validity of three methods for assessment of the quality of primary health care." *International Journal of Quality in Health Care*, Vol. 11(5), pp. 429-33.

Huntington, D., H. Zaky, S. Shawky, F. Fattah, and E. El-Hadary. 2010. "Impact of a Service Provider Incentive Payment Scheme on Quality of Reproductive and Child-health Services in Egypt," *Journal of Health, Population and Nutrition*, Vol. 28(3), pp. 273-280.

Jha, A., K. Joynt, J. Orav and A. Epstein. 2012. "The Long-Term Effect of Premier Pay for Performance on Patient Outcomes," *New England Journal of Medicine*, Vol. 366(17), pp. 1606-1615.

Leonard, K. L. and N. C. Masatu. 2005. "The use of direct clinician information and vignettes for health service quality evaluation in developing countries." *Social Sciences and Medicine*, Vol. 61(9), pp. 1944-51.

Leonard, K. L. and N. C. Masatu. 2006. "Outpatient process quality evaluation and the Hawthorne effect." *Social Sciences and Medicine*, Vol. 63(9), pp. 2330-40.

Levine, R. and R. Eichler. 2009. "Performance Incentives for Global Health". *Center for Global Development, Washington, DC*.

Lindenauer P., D. Remus, S. Roman et al. 2007. "Pay for performance in hospital quality improvement," *New England Journal of Medicine*, Vol. 356, pp. 486-496.

Martinez, A., S.W. Raudenbush and Jessaca Spybrook. 2009. "The Design of Blocked Cluster-Randomized Trials," Unpublished Manuscript.

Meyer, C., N. Bellows, M. Campbell, M. Potts. 2011. *The Impact of Vouchers on the Use and Quality of Health Goods and Services in Developing Countries: A systematic review*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Musgrove, P. 2010. "Rewards for Good Performance or Results: A Short Glossary of RBF". www.rbfhealth.org

Olken, B.A., J. Onishi and S. Wong. 2011. "Should Aid Reward Performance? Evidence from a Field Experiment on Health and Education in Indonesia", unpublished working paper.

Peabody, J. W., J. Luck, P. Glassman, T.R. Dresselhaus and M. Lee. 2000. "Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality." *JAMA*, Vol. 283(13), pp. 1715-22.

Petersen L. A., D. L. Woodard, T. Urech, C. Daw and S. Sookanan. 2006. "Does pay-for-performance improve the quality of health care?" *Annals of Internal Medicine*, Vol. 145, pp. 265-272.

Raudenbush, S. W., H. Bloom, R. Congdon, C. Hill, A. Martinez and J. Spybrook. 2011. Optimal Design Software for Multi-level and Longitudinal Research (Version 3.01) [Software]. Available from www.wtgrantfoundation.org or from sitemaker.umich.edu/group-based.

Raudenbush, S. W.. 1997. "Statistical Analysis and Optimal Design for Cluster Randomized Trials," *Psychological Methods*, Vol. 2(2), pp. 173-185.

Spybrook, J., H. Bloom, R. Congdon, C. Hill, A. Martinez and S. Raudenbush. 2011. Optimal Design Plus Empirical Evidence: Documentation for the "Optimal Design" Software."

Witter, S., A. Fretheim, F.L. Kessy, and A.K. Lindahl. 2012. Paying for performance to improve the delivery of health interventions in low- and middle-income countries (Review). *Cochrane Database of Systematic Reviews 2012*, Issue 2. Art. No.: CD007899.

World Health Organization. 2011. World Health Statistics 2011: Indicator Compendium. Geneva, Switzerland.