

Vérification de la performance dans le  
cadre du Financement Basé sur les  
Résultats : le cas du Burundi

*Adrien Renaud, consultant*

*Juillet 2013*

## TABLE DES MATIERES

Remerciements .....	3
Liste des abréviations.....	4
Résumé exécutif.....	5
Introduction.....	7
1. Méthodologie.....	8
2. Contexte : le FBP au Burundi .....	9
3. Principales caractéristiques de la méthode de vérification.....	11
3.1. Considérations générales sur la méthode de vérification.....	12
3.2. Comment la quantité de services fournis est-elle vérifiée ?.....	13
3.3. Comment la qualité technique est-elle évaluée ?.....	15
3.4. Comment les enquêtes communautaires sont-elles effectuées ?.....	16
3.5. Comment les données de la vérification sont-elles contre-vérifiées ?.....	19
4. Résultats de la vérification.....	21
4.1. Quels sont les résultats de la vérification de la quantité ? .....	21
4.2. Quels sont les résultats de l'évaluation de la qualité technique ?.....	27
4.3. Quels sont les résultats des enquêtes communautaires ?.....	29
4.4. Quels sont les résultats de la contre-vérification ?.....	32
4.5. A quoi les résultats sont-ils utilisés ?.....	36
5. Les coûts de la vérification.....	40
6. Leçons .....	42
Annexes .....	46
Annexe 1 : Bibliographie .....	46
Annexe 2 : Formulaire de déclaration de la performance dans les CS et les hôpitaux.....	46
Annexe 3 : Factures mensuelles pour les indicateurs des CS et des hôpitaux .....	46
Annexe 4 : Check-lists d'évaluation de la qualité technique des CS et des hôpitaux.....	46
Annexe 5 : Questionnaire pour les enquêtes communautaires .....	47
Annexe 6 : Cadre d'analyse pour les études de cas .....	47

## REMERCIEMENTS

*Je voudrais saluer la mémoire du Dr. Etienne Sekaganda, décédé en décembre 2010 durant une mission de contre-vérification, quelques jours après avoir répondu à mes questions.*

*Je remercie la CT-FBP, et tout particulièrement son président le Dr. Jean Kamana qui m'a permis d'avoir accès aux données, son conseiller technique le Dr. Olivier Basenya qui a apporté de précieux conseils et contributions sur différentes versions de ce document, et le Dr. Emmanuel Seheye qui a fourni la base de données Excel. Les CPVV ont pleinement collaboré avec moi dans le cadre de cette étude, et celle-ci leur doit beaucoup. J'exprime enfin ma gratitude envers toutes les personnes qui ont accepté de répondre à mes questions au Burundi comme à l'étranger.*

*Enfin, je voudrais préciser que cette étude n'aurait pas pu être possible sans l'aide, les conseils, le soutien et les contributions du personnel de la Banque Mondiale. Je remercie particulièrement Petra Vergeer, Anna Heard et Moulay Driss Zine-Eddine El-Idrissi.*

## LISTE DES ABREVIATIONS

ASLO	Association Locale
BDS	Bureau du District de Santé
BPS	Bureau Provincial de la Santé
CPVV	Comité Provincial de Vérification et de Validation
CS	Centre de Santé
CT-FBP	Cellule Technique du FBP
FBP	Financement Basé sur la Performance
FBR	Financement Basé sur les Résultats
HDP	Health, Development & Performance
MSPLS	Ministère de la Santé Publique et de la Lutte contre le SIDA
ONG	Organisation Non-Gouvernementale
SIS	Système d'Information Sanitaire

*NB : Le taux de change utilisé dans ce document est le taux de change en novembre 2012 :  
1 USD = 1464 BIF*

## RESUME EXECUTIF

Le Burundi est un pays à faible revenu situé dans la région des Grands Lacs, en Afrique. Il compte environ 8,6 millions d'habitants et ses indicateurs économiques et de développement sont parmi les plus bas au monde. Les indicateurs de santé sont, en conséquence, très faibles. Le Burundi a été l'un des premiers pays africains à mettre en œuvre le Financement Basé sur les Résultats (FBR). Son programme de Financement Basé sur les Performances (FBP) est mené par le MSPLS, et est appliqué dans l'ensemble du pays pour tous les CS et hôpitaux à but non-lucratif, publics ou privés. La mise en œuvre du programme aux niveaux intermédiaire et opérationnel est menée par des comités provinciaux appelés Comités Provinciaux de Vérification et de Validation (CPVV) et impliquant les antennes locales du Ministère de la Santé Publique et de la Lutte contre le SIDA (MSPLS), des ONG, l'administration locale et la société civile. Le paiement des structures de soins est fondé sur un paiement à l'acte mensuel, et sur une composante portant sur la qualité des soins, le respect des protocoles cliniques, la recherche des patients dans la communauté (vérification de l'existence des patients et de leur confirmation des soins), et sur la satisfaction des patients.

Cette étude vise à décrire les méthodes utilisées dans ce programme pour vérifier la performance, à présenter leurs résultats et à discuter les difficultés auxquelles elles ont été confrontées, la manière dont elles ont été résolues et les défis qui demeurent.

Le programme inclut quatre types de vérification :

- Vérification de la quantité de services fournis, effectuée mensuellement par les CPVV pour chaque structure de santé directement contractualisée et pour chaque. Cette vérification consiste à recompter les services dans les registres des fournisseurs.
- Evaluation de la qualité technique des services fournis, effectuée trimestriellement en utilisant une grille prédéfinie, dans toutes les structures de santé par les antennes locales du MSPLS pour les Centres de Santé (CS) et par des pairs pour les hôpitaux.
- Evaluation de la mesure dans laquelle les patients ont effectivement reçu ces services et de leur satisfaction, effectuée semestriellement par des ONG locales pour chaque structure de santé et pour des indicateurs choisis. 80 patients sont sélectionnés aléatoirement et interrogés pour chaque structure de soins chaque semestre.
- Contre-vérification des informations fournies par ces trois mécanismes, effectuée trimestriellement par une organisation indépendante (Health, Development and Performance, HDP). Quatre provinces sont sélectionnées chaque trimestre, et les mêmes outils que pour la vérification sont utilisés pour la contre-vérification.

D'après cette étude, les résultats de la vérification sont les suivants :

- La vérification de la quantité dans les CS par les CPVV entre janvier et août 2012 a trouvé que la performance quantitative était correctement estimée par les structures de soins dans 69% des cas. Elle était surestimée dans 22% des cas (surestimation moyenne : 22%), et sous-estimée dans 9% des cas (sous-estimation moyenne : 4%). La vérification de la quantité dans les hôpitaux par les CPVV entre janvier et juillet 2012 a trouvé que la performance quantitative était correctement estimée par les structures de soins dans 62% des cas. Elle était surestimée dans 26% des cas (surestimation moyenne : 13%), et sous-estimée dans 7% des cas (sous-estimation moyenne : 7%).
- L'évaluation de la qualité technique par les antennes locales du MSPLS et par les pairs a trouvé que durant la période 2010-2012, le score de qualité technique était de 70% pour les CS, 77% pour les hôpitaux de district et 74% pour les hôpitaux nationaux. Cependant, ces scores peuvent fortement varier en fonction des changements qui peuvent être effectués dans les grilles d'évaluation de la qualité technique.
- Dans les CS de 13 provinces pour lesquelles des données désagrégées sur les enquêtes communautaires ont pu être collectées, la proportion de patients qui n'ont pas pu être retrouvés dans la communauté variait de 8,6% à 6,9% entre le premier trimestre 2011 et le

premier semestre 2012. Dans les hôpitaux des mêmes 13 provinces et durant la même période, la proportion de patients qui n'ont pas pu être retrouvés dans la communauté variait de 9,8% à 15,4%.

- Pour les CS, aucune différence significative n'a été trouvée durant les huit rounds de contre-vérification entre la performance quantitative telle que vérifiée par les CPVV et la performance quantitative telle que contre-vérifiée par HDP. Cependant, les données désagrégées n'étaient pas disponibles et les sous-estimations annulaient les surestimations, ce qui conduit à sous-estimer les différences. Pour les hôpitaux au contraire, la différence entre les données de vérification et de contre-vérification étaient importantes et systématiques : la moyenne des valeurs absolues des différences entre la performance quantitative telle que vérifiée par les CPVV et la performance quantitative telle que contre-vérifiée par HDP est de 31%. Cette étude a également trouvé d'importantes et systématiques différences entre la qualité technique telle qu'évaluée par les antennes locales du MSPLS et par les pairs et la qualité technique telle qu'évaluée par HDP : en comparaison des résultats de HDP, l'évaluation de la qualité technique dans les CS par le MSPLS était surestimée dans 79% des cas (surestimation moyenne de 20%) et sous-estimée dans 21% des cas (sous-estimation moyenne de 24%). En comparaison des résultats de HDP, l'évaluation de la qualité technique dans les hôpitaux par les paires était surestimée dans 84% des cas (surestimation moyenne de 24%) et sous-estimée dans 16% des cas (sous-estimation moyenne de 20%).

La principale utilisation des résultats de la vérification est le paiement des prestataires : les résultats de la vérification de la quantité servent au paiement mensuel de subsides aux prestataires, et les résultats de l'évaluation de la qualité technique, combinés avec ceux des enquêtes communautaires, se traduisent dans le bonus ou un malus de qualité que les structures peuvent recevoir trimestriellement. Les résultats de la contre-vérification sont utilisés pour évaluer la mesure dans laquelle les CPVV et les antennes locales du MSPLS ont rempli leurs tâches correctement. Toutes les sanctions qui sont définies dans le manuel des procédures du FBP ne sont pas encore mises en œuvre, car le système de vérification n'est pas qu'un instrument de contrôle : il est également considéré comme un instrument pédagogique à l'intention des structures de soins.

De l'avis de l'ensemble des acteurs interrogés, le système de vérification de la performance dans le FBP au Burundi permet un niveau satisfaisant de détection de la fraude. Il existe un consensus pour considérer que la plupart des erreurs sont non-intentionnelles plutôt que des tentatives de tricherie de la part des structures de soins qui viseraient à augmenter leurs revenus. Ceci est confirmé par le fait que les sous-estimations de la performance sont fréquentes (bien que moins fréquentes que les surestimations).

Au Burundi, une forte intégration au sein des organes de régulation du MSPLS d'une part, et l'existence d'un système de contre-vérification indépendant d'autre part, produit un degré de confiance qui permet que les données soient jugées satisfaisantes par plupart des acteurs, et ce à un coût acceptable (environ 16% des dépenses totales du FBP). Cette étude a montré que le niveau de discordance entre les résultats avant et après la vérification était encore élevé, mais ceci peut être résolu par la modification des incitations d'une part (par exemple par la mise en œuvre de sanctions plus strictes), et par la formation continue des CPVV, des BPS, des BDS et des pairs hospitaliers d'autre part.

## INTRODUCTION

Le Burundi est un pays à faible revenu situé dans la région des Grands Lacs, en Afrique. Il compte environ 8,6 millions d'habitants et ses indicateurs économiques et de développement sont parmi les plus bas au monde : en 2011, son PIB par habitant était d'environ 270\$<sup>1</sup>, 67% de la population vivait sous le seuil de pauvreté. Avec un Indice de Développement Humain de 0,316, il se classait 185<sup>e</sup> sur 187 pays dans le Rapport sur le développement humain 2011 établi par le Programme des Nations Unies pour le Développement (PNUD, 2011). Sans surprise, les indicateurs de santé sont aussi très bas : la dépense totale de santé est environ de 20\$ par habitant, l'espérance de vie à la naissance ne dépasse pas 50 ans, la mortalité des moins de 5 ans est au-delà de 140 pour 1 000 naissances vivantes et la mortalité maternelle est estimée à 800 pour 100 000 naissances vivantes.

Éléments de contexte pour le Burundi	
Population	8,6 millions
PNB par habitants (\$ courants)	271\$
Taux de pauvreté	67%
Dépense totale de santé par habitant (\$ courants)	21\$
Espérance de vie à la naissance	50 ans
Mortalité des moins de 5 ans (pour 1 000 naissances vivantes)	142
Ratio de mortalité maternelle (pour 100 000 naissances vivantes)	800

Tableau 1 : Éléments de contexte pour le Burundi (source : Banque Mondiale, 2011)

Le Burundi a été l'un des premiers pays africains à introduire le Financement Basé sur les Résultats (FBR, appelé au Burundi Financement Basé sur la Performance, FBP). En 2010, il a été le second sur le continent (après le Rwanda) à faire de cette approche une politique nationale, et à la mettre en œuvre sur l'ensemble de son territoire. Le FBP burundais a un double objectif :

- Il est utilisé comme moyen d'améliorer les performances du système de santé ;
- Il constitue une méthode de mise en œuvre de la politique nationale de gratuité ciblée des soins pour les femmes enceintes et les enfants de moins de cinq ans.

Il est financé pour moitié par le gouvernement burundais (MSPLS, 2012), ce qui constitue une caractéristique très rare pour les programmes de FBR en Afrique. L'autre moitié du budget est fournie par différents partenaires internationaux (Banque Mondiale, Union Européenne, coopération belge, USAID...), ce qui nécessite un important et louable effort de coordination de l'aide. Ce contexte singulier confère au FBP burundais un intérêt particulier pour l'étude des différentes méthodes de vérification utilisées dans les programmes de FBR.

Après deux ans de mise en œuvre du FBP au Burundi, cette étude vise à décrire le système de vérification que ce pays a mis en place, mais aussi à identifier les résultats qu'il a obtenus, les écueils auxquels il a fait face, la manière dont il les a résolus et les défis qu'il lui reste à relever.

<sup>1</sup> En dollars US courants. Ces chiffres, comme ceux qui sont présentés dans cette section, est tiré (sauf mention contraire) de *Banque Mondiale, 2011*.



Illustration 1 : Le Burundi en Afrique (source : Banque Mondiale)

## 1. METHODOLOGIE

La collecte des données nécessaires à la réalisation de cette étude s'est déroulée en deux temps distincts. Tout d'abord, à la fin de l'année 2010, la documentation pertinente (annexe 1) a été étudiée, les principaux acteurs du FBP ont été interrogés par téléphone ou par e-mail, et en accord avec le Ministère de la Santé Publique et de la Lutte contre le SIDA (MSPLS), quatre provinces<sup>2</sup> ont été sélectionnées pour des entretiens téléphoniques avec les personnes qui mettent en œuvre la vérification sur le terrain, ainsi qu'avec les Bureaux Provinciaux de la Santé (BPS). Cette première phase a permis d'obtenir une bonne description de la manière dont la vérification se déroulait, mais elle a été conduite moins d'un an après la généralisation du FBP à tout le pays et ne pouvait donc pas donner une image satisfaisante des résultats de la vérification.

Deux ans plus tard, en 2012, une deuxième phase de collecte de données a été entreprise. Une mission d'évaluation du FBP a fourni une opportunité pour documenter les changements intervenus dans le processus de vérification depuis 2010, d'interviewer davantage d'acteurs de terrain et de récolter des données quantitatives détaillées sur les résultats des différents mécanismes de vérification.

Il y a 4 mécanismes de vérification dans le système de FBP au Burundi (voir parties 2 et 3), et les périodes pour lesquelles les données quantitatives ont pu être collectées diffèrent d'un mécanisme à l'autre :

- Pour la vérification quantitative, les données sur les services déclarés et vérifiés étaient disponibles de janvier à août 2012 dans les Centres de Santé (CS), et de janvier à juillet 2012 dans les hôpitaux ;
- Pour l'évaluation de la qualité technique, les résultats étaient disponibles du début de la mise à l'échelle du FBP (deuxième trimestre 2010) au troisième trimestre 2012 ;
- Pour les enquêtes communautaires, les résultats étaient disponibles du premier trimestre 2011 au premier semestre 2012, et seulement pour 13 provinces sur 17 ;
- Pour la contre-vérification, les résultats étaient disponibles du début de la mise à l'échelle du FBP au troisième trimestre 2012.

De plus, des données ont également été collectées sur l'estimation des coûts de mise en œuvre de la vérification, à travers l'étude des rapports d'activité élaborés par le MSPLS, qui met en œuvre le FBP.

<sup>2</sup> Bubanza, Muramvya, Muyinga et Ruyigi.

Toutes ces données ont été rendues disponibles par le MSPLS, et elles concernent l'ensemble du pays (sauf pour les enquêtes communautaires). Elles ont notamment été utilisées pour calculer le niveau d'erreur, défini comme la différence entre les données avant vérification et les données après vérification. Les formules de calcul pour chaque mécanisme de vérification ont été présentées dans les sections appropriées de cette étude.

L'analyse des données quantitatives et qualitatives a été effectuée en fonction d'un cadre qui a été utilisé dans d'autres études de cas sur les systèmes de vérification dans différents pays (annexe 6). Ce cadre a été conçu de manière à permettre la comparabilité des expériences. Il comprend 5 éléments principaux qui ont déterminé les chapitres principaux de cette étude : présentation du système de FBP (partie 2), description des systèmes de vérification à l'œuvre (partie 3), résultats des méthodes de vérification (partie 4), coût de la vérification (partie 5), et enfin leçons qui peuvent être tirées de l'application des méthodes de vérification (partie 6).

## 2. CONTEXTE : LE FBP AU BURUNDI

Le FBP a été d'abord introduit au Burundi comme projet pilote dans trois provinces en 2006, avec le soutien des Organisations Non Gouvernementales (ONG) hollandaises Cordaid et HealthNet TPO. Ces expériences avaient pour objectif d'améliorer les faibles performances du système de santé en augmentant et en remodelant les incitations pour les prestataires de soins, et en passant d'un système basé sur le financement d'intrants à un système basé sur le financement de résultats. Les résultats de ces expériences ont été jugés encourageants (comme le montre une étude publiée plus tard : Cordaid et MSPLS, 2011), et elles ont été progressivement étendues à d'autres districts et à d'autres provinces. De plus en plus de partenaires sont devenus intéressés par le FBP au Burundi. En 2010, le gouvernement a décidé de généraliser le FBP à tout le pays, et lui a donné un objectif supplémentaire : financer la gratuité des soins pour les femmes enceintes et les enfants de moins de cinq ans (une priorité nationale depuis 2006). L'approche bénéficie d'un fort soutien des partenaires techniques et financiers, qui, en plus de financer environ la moitié du budget, apportent une assistance technique tant au niveau central qu'au niveau provincial.

Le fonctionnement du FBP au Burundi est décrit en détail ailleurs (MSPLS, 2011b), et nous ne donnerons ici qu'un bref aperçu. La participation au programme est obligatoire pour tous les prestataires publics du pays. Les prestataires privés peuvent également participer. En pratique, tous les prestataires privés à but non-lucratif participent, et quelques prestataires privés<sup>3</sup> ont des sous-contrats avec des prestataires publics ou privé à but non-lucratif. Le programme compte en tout 565 CS et 50 hôpitaux (45 hôpitaux de district et 5 hôpitaux nationaux), plus 151 prestataires privés qui bénéficient d'un sous-contrat.

Le FBP paie les formations sanitaires en deux étapes. Le premier paiement est un paiement à l'acte qui a pour objectif d'inciter les structures sanitaires à augmenter le volume de prestations fournies (le présumé est celui d'une sous-utilisation des services) : un certain nombre d'indicateurs de résultat<sup>4</sup> a été défini, et à chaque indicateur correspond un subside unitaire. Les paiements ne sont pas plafonnés : plus une formation sanitaire soigne, plus elle reçoit d'argent. Les structures de soins sont payées chaque mois après vérification de leurs registres. La seconde étape est un paiement trimestriel lié à une évaluation de la qualité des soins. Il a un double objectif :

- Inciter les prestataires à respecter les normes (« qualité technique ») ;

<sup>3</sup> A but lucratif ou à but non-lucratif : il s'agit en général de CS de petite taille ou situés dans des zones reculées.

<sup>4</sup> Les listes d'indicateurs ont été établies pendant un atelier national au cours duquel différents acteurs du système de santé ont listé tous les indicateurs pertinents, et ont sélectionné les plus importants en utilisant la méthode de priorisation DELPHI.

- Inciter à l'amélioration de la relation des prestataires avec les patients en détectant d'éventuels « patients-fantômes » d'une part (vérification de la mesure dans laquelle les services ont effectivement été fournis), et évaluant la satisfaction des patients d'autre part (« qualité subjective »).

Cette évaluation est trimestrielle pour ce qui est de la « qualité technique » et semestrielle pour ce qui est de la « qualité subjective », les résultats de cette dernière étant utilisés pour le paiement de deux trimestres. Elle se traduit par un score, puis par un bonus de qualité (incitation financière) si l'évaluation est jugée satisfaisante (score de qualité supérieur ou égal à 70%). Si au contraire cette évaluation est jugée insuffisante (score de qualité inférieur à 50%), une partie des subsides prévus pour le premier paiement (incitation quantitative) peut être déduite.

Les subsides du FBP ne sont pas payés directement au personnel de santé, mais aux formations sanitaires. Ces dernières disposent d'une autonomie de gestion partielle, et utilisent les fonds reçus pour financer leurs activités habituelles. Ces subsides constituent un revenu complémentaire pour les structures de soins, qui reçoivent par ailleurs les paiements directs des patients, des subventions de l'Etat (salaires, etc.), des primes et des intrants de partenaires et de programmes verticaux, etc... Ils représentent environ 30 à 35 % des recettes générées par l'activité des structures pour les hôpitaux, et environ 80 à 85% pour les CS. Les prestataires peuvent en allouer une partie à la motivation du personnel si elles sont capables de générer des bénéfices (ce qui est généralement le cas) et si leurs comptes bancaires sont suffisamment approvisionnés (solde équivalant à au moins 2 mois de fonctionnement). Les incitations financières sont en effet supposées motiver le personnel de santé à améliorer sa performance, et elles peuvent représenter jusqu'au tiers de leur salaire de base, comme le montre le tableau 2.

	<b>Salaire mensuel en début de carrière</b>	<b>Motivation financière moyenne par mois du FBP en 2010</b>
<b>Infirmiers</b>	200 US\$	72 US\$
<b>Médecins</b>	400 US\$	100 US\$

Tableau 2 : Motivation financière mensuelle moyenne du FBP pour le personnel de santé en comparaison avec le salaire mensuel moyen en début de carrière (source : MSPLS, 2011a)

Le programme est dirigé par le MSPLS, où une Cellule Technique pour le FBP (CT-FBP), qui dépend de la Direction Générale de la Santé Publique et de la lutte contre le SIDA, coordonne toutes les activités qui y sont liées. Aux niveaux intermédiaire et périphérique, des Comités Provinciaux de Vérification et de Validation (CPVV), avec les démembrements du MSPLS que sont les BPS et les Bureaux de District de Santé (BDS), sont chargés de la mise en œuvre. L'illustration 2 résume les interactions entre les principaux acteurs du programme FBP.

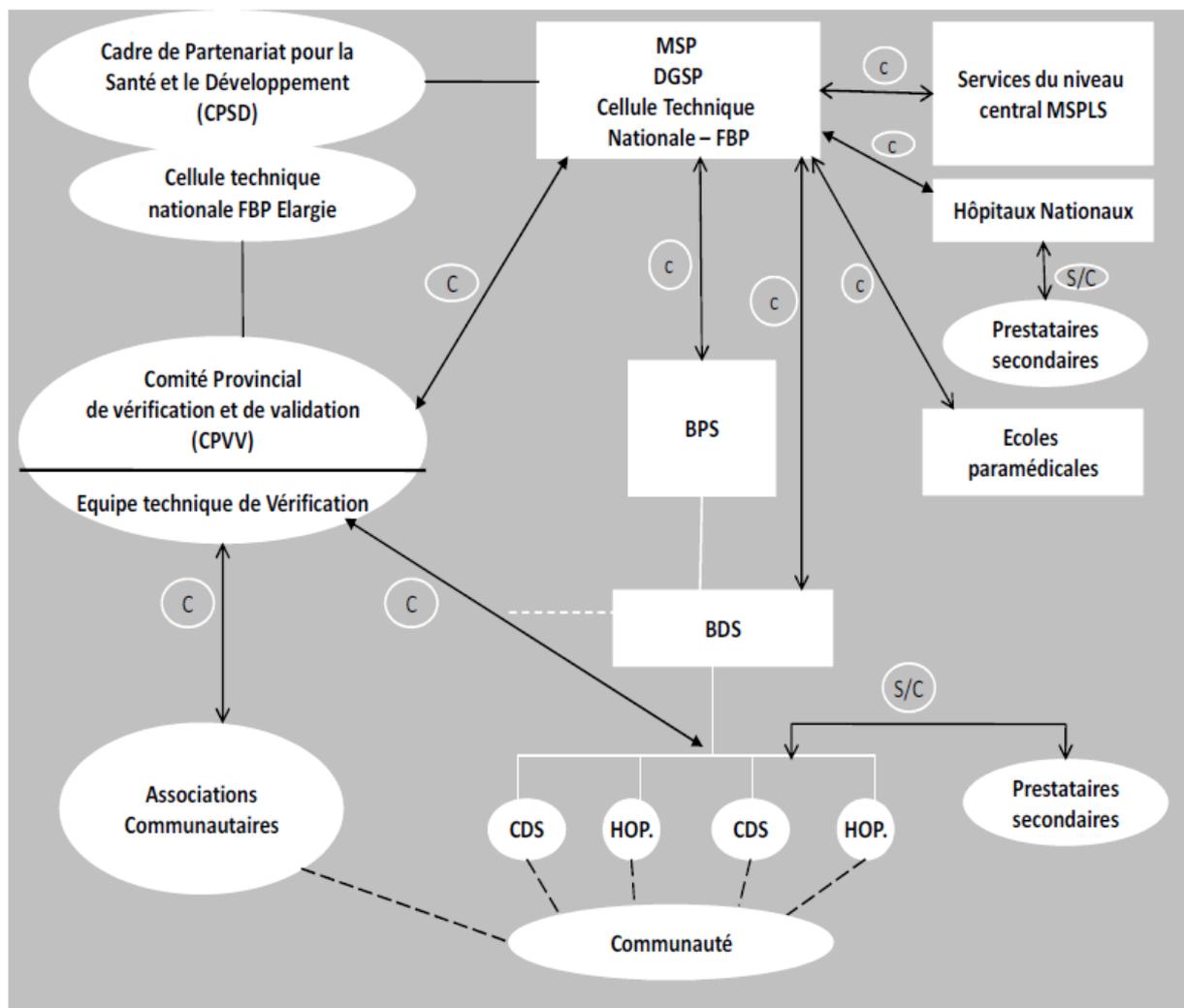


Illustration 2 : Montage institutionnel du FBP (source: MSPLS, 2011b. « C » signifie « contrat », et « S/C » signifie « sous-contrat »)

### 3. PRINCIPALES CARACTERISTIQUES DE LA METHODE DE VERIFICATION

Le système de vérification du FBP au Burundi a été élaboré durant les projets pilotes menés entre 2006 et 2010. Lors de sa généralisation à l'ensemble du pays, il a été adapté par consensus entre les différents acteurs : le MSPLS, les partenaires techniques et financiers, les programmes verticaux, les ONG qui avaient mis en œuvre les projets pilote, etc. Les procédures à suivre ont été synthétisées dans un manuel (MSPLS, 2010), et qui a été révisé par la suite (MSPLS, 2011b). Ce document, ainsi que les entretiens qui ont été menés avec les personnes qui mettent en œuvre la vérification, constituent la principale source utilisée pour cette section. Il convient de noter que la vérification est en un processus évolutif, et que cette étude n'a pas pu prendre en compte les changements intervenus après la fin de l'année 2012.

### 3.1. Considérations générales sur la méthode de vérification

Le système comprend quatre types de vérification<sup>5</sup> :

- Vérification de la quantité de services fournis (appelée « vérification de la quantité » dans le reste de ce document). Cette vérification vise à déterminer si l'on peut retrouver les services déclarés dans les registres du prestataire, et elle donne lieu au paiement mensuel à l'acte des services des prestataires de santé ;
- Evaluation de la qualité technique des services fournis (appelée « évaluation de la qualité technique » dans la suite de ce document). Cette évaluation vise à déterminer si les prestataires sont capables de fournir des services répondant aux normes du MSPLS, et constitue une partie du bonus ou malus qualitatif trimestriel que peuvent recevoir les formations sanitaires ;
- Evaluation de la mesure dans laquelle les patients ont véritablement reçu ces services et de leur satisfaction (appelée « enquête communautaire » dans le reste de ce document) : cette vérification a un double objectif. Tout d'abord, elle vise à déterminer si les patients déclarés existent véritablement et s'ils confirment avoir été traités dans la structure, et constitue donc un outil de détection de la fraude (identification d'éventuels « patients-fantômes »). D'autre part, elle vise à déterminer le niveau de satisfaction des patients, et constitue donc un outil d'évaluation de la qualité perçue des soins. Avec l'évaluation de la qualité technique, l'enquête communautaire constitue la seconde partie du bonus ou malus qualitatif que les formations peuvent recevoir ;
- Contre-vérification des informations fournies par ces trois mécanismes (appelée « contre-vérification » dans le reste de ce document), effectuée par une organisation indépendante (appelée « contre-vérification » dans le reste de ce document). Cette contre-vérification vise à évaluer la qualité des données produites par le système FBP, mais aussi à assurer un suivi de la mise en œuvre de ce système.

Afin de mettre en œuvre le FBP au niveau local, le gouvernement du Burundi a mis en place des CPVV, déjà cités. Ceux-ci constituent la cheville ouvrière de l'approche au niveau provincial. Ils sont responsables de la contractualisation des prestataires au nom du MSPLS, de la mise en œuvre des vérifications, de la décision d'autoriser les paiements FBP, de l'analyse des données, de l'élaboration et du suivi des recommandations qui en découlent, et de la mise en œuvre de sanctions si nécessaire. Au début de la généralisation du FBP, un débat a eu lieu entre le MSPLS, les partenaires financiers et les ONG pour savoir si le vérificateur devait être une agence gouvernementale plutôt qu'une institution indépendante, comme c'était le cas lors des projets pilote. Une solution de compromis a été trouvée entre ces deux options : les CPVV sont des institutions mixtes où la société civile, les partenaires techniques et financiers et le gouvernement sont représentés. Le secrétariat du CPVV est assuré par le médecin-chef du BPS, c'est-à-dire par le représentant du MSPLS au niveau provincial. Le CPVV est composé de deux sous-unités :

- Une « sous-unité vérification » chargée de mener les activités de vérification dans les formations sanitaires, qui est composée de fonctionnaires et de contractuels ;
- Une « sous-unité validation<sup>6</sup> » chargée de la validation des données de vérification au niveau provincial, de l'élaboration d'une facture compilée provinciale, du suivi des contrats et plus généralement de la mise en œuvre du FBP au niveau local, et qui est composée de :

<sup>5</sup> Un cinquième type de vérification existe : l'évaluation de la performance des organes de régulation et de mise en œuvre aux niveaux central, intermédiaire et périphérique : la CT-FBP, les BPS, les BDS et les CPVV. Il n'a pas été inclus dans cette étude, qui se limite à la vérification de la performance des formations sanitaires.

<sup>6</sup> Le terme de « validation » peut ici prêter à confusion. Il est également utilisé pour désigner les données issues du travail de la sous-unité vérification (prestations vérifiées qui correspondent exactement aux définitions des indicateurs, voir partie 3.2).

- Un représentant de l'administration territoriale;
- Des représentants des BDS et BPS ;
- Des représentants des partenaires techniques et financiers ;
- Des représentants de la société civile.

Lors de la généralisation du FBP à tout le pays, le CPVV n'était pas subdivisé de la sorte : il avait tout simplement en son sein une équipe technique chargée de mener les tâches actuellement dévolues à la sous-unité vérification. Il a été décidé de séparer les deux fonctions de validation et de vérification car cela permettait de mieux inciter chacune des deux sous-unités à une meilleure performance. Cette décision avait pour but d'éviter les conflits internes au comité qui étaient apparus dans certaines provinces et qui étaient liés à la charge de travail des membres. L'objectif de la réforme était que chacun des membres du CPVV soit rétribué en fonction de sa contribution à la performance collective du comité.

Le CPVV bénéficie de l'assistance technique d'agences locales d'accompagnement, qui sont souvent dirigées par les ONG qui ont mis en œuvre les projets pilote de FBP au Burundi. L'utilité de ces agences d'accompagnement est cependant débattue au Burundi : leur contribution à la qualité du système de vérification dépend des capacités de leur personnel, qui ne correspondent pas toujours aux capacités requises. Leur coût est relativement important (6% du budget total du FBP, voir partie 5), et des discussions sont en cours pour savoir si leur mission sera maintenue dans le futur.

Les membres du CPVV élisent un président et un vice-président, mais les décisions au sein du comité sont prises par consensus. On peut remarquer les formations sanitaires publiques sont hiérarchiquement responsables devant certains membres du comité. Aucun exemple de conflit d'intérêt dû à cette situation n'a été mentionné à l'auteur de cette étude, mais il y a des raisons de penser qu'elle peut mener à une certaine forme de complaisance, notamment en ce qui concerne l'évaluation de la qualité (voir partie 4.2). Ce montage institutionnel était cependant la seule manière de trouver un compromis acceptable entre les tenants d'une vérification mise en œuvre par le gouvernement et ceux qui préféreraient qu'elle soit mise en œuvre par une équipe indépendante.

### **3.2. Comment la quantité de services fournis est-elle vérifiée ?**

La quantité de services est déclarée par les prestataires via le Système d'Information Sanitaire (SIS) de routine : un formulaire de déclaration (annexe 2) rempli manuellement est annexé au formulaire du SIS et reprend, pour les indicateurs contractualisés, les données qui sont pertinentes pour le FBP. Cette déclaration est ensuite vérifiée par la sous-unité vérification du CPVV, et les informations sont utilisées pour créer une facture. Chaque structure de soins est vérifiée chaque mois, de manière à correspondre au rythme de paiement mensuel. Les vérifications aléatoires ne sont pour l'instant pas considérées comme une option envisageable au Burundi par les acteurs interrogés dans le cadre de cette étude, pour au moins deux raisons :

- La vérification est liée au paiement, et le résultat de la visite de vérification est une facture ;
- La vérification quantitative est vue comme l'un des instruments pédagogiques qui permettent aux prestataires d'apprendre les mécanismes et les principes du FBP ;
- La facture mensuelle sur laquelle le paiement est basé est élaborée au cours des visites de vérification, ce qui les rend indispensables au paiement des prestataires.

Au cours de la visite de vérification, le(s) vérificateur(s) du CPVV et le responsable de la structure rassemblent les registres et recomptent tous les services fournis par la structure pendant le mois pour les indicateurs concernés : 22 indicateurs pour les CS, et 24 indicateurs pour les hôpitaux. Si la formation sanitaire vérifiée a des contrats avec un ou plusieurs autres prestataires, les registres de ces dernières sont acheminés vers la formation sanitaire principale et les prestations de ce(s) dernier(s) sont recomptées sur place. L'équipe établit alors une facture (annexe 3) qui mentionne, pour chaque indicateur, trois quantités :

- La quantité de services déclarée, qui correspond à la quantité telle qu'évaluée par la formation sanitaire et reportée dans le formulaire de déclaration avant l'arrivée de l'équipe de vérification ;
- La quantité de services vérifiée, qui correspond à la quantité de services telle que comptée par la formation sanitaire et l'équipe de vérification dans les registres pour le mois en écoulé, sans prendre en compte les définitions exactes des indicateurs : n'importe quelle prestation effectuée pendant le mois et ayant été reportée dans le registre sera comptabilisée ;
- La quantité de services validée, qui est celle qui est utilisée pour le paiement. Elle correspond à la quantité de services qui se trouvent dans les registres pour le mois écoulé et qui respectent la définition exacte de l'indicateur. Par exemple, si un accouchement a été effectué au CS et est enregistré sans que la date d'accouchement soit mentionnée, ou si un enfant est complètement vacciné mais a achevé le processus de vaccination après son premier anniversaire, ces services seront comptés comme prestations vérifiées mais pas comme prestations validées. Cette validation ne doit pas être confondue avec la validation effectuée par la sous-unité validation du CPVV (voir ci-dessous).

La facture indique également, en cas de différence trop importante entre la quantité déclarée et la quantité vérifiée, les pénalités éventuelles qui doivent être appliquées et qui sont décrites dans le tableau 3. Ces sanctions n'existaient pas dans le modèle initial du système de vérification. Elles ont été introduites au moment de la révision du manuel de procédures (2011), afin d'augmenter la rigueur du processus de vérification envers les prestataires.

Différence entre quantité déclarée et quantité vérifiée <sup>7</sup>	Pénalité
<5%	Pas de pénalité
5-10%	-5% du montant de l'indicateur concerné
10-20%	-10% du montant de l'indicateur concerné
>20%	Pas de subside pour l'indicateur concerné

Tableau 3 : Critères pour les sanctions en cas de différences entre déclaration et vérification (source : MSPLS, 2011b)

On remarque que ces sanctions s'appliquent à la différence entre quantité déclarée et quantité vérifiée, et non, comme on aurait pu s'y attendre, à la différence entre quantité déclarée et quantité validée, qui correspond à l'ensemble de la définition de l'indicateur. Le MSPLS a préféré se fonder sur les quantités vérifiées pour les sanctions de manière à renforcer son SIS : les quantités vérifiées sont celles qui sont rapportées dans le SIS (une femme qui a accouché mais dont on a oublié de mentionner l'adresse exacte doit tout de même apparaître dans le SIS). Les quantités validées constituent pourtant une information plus pertinente du point de vue du FBP, puisqu'elles déterminent le paiement des prestataires. Il semblerait donc logique de faire évoluer le modèle de sanctions de manière à prendre aussi en compte les différences qu'elles présentent avec les quantités déclarées.

Sur la base de la facture, l'équipe de vérification peut donner dans un rapport des recommandations pour améliorer la qualité des données, si nécessaire. La structure et l'équipe conservent chacun une copie de la facture, et une copie de l'éventuel rapport.

La sous-unité vérification du CPVV est généralement composée d'infirmiers, avec diverses qualifications. La moitié de ses membres sont des fonctionnaires, et la moitié sont des contractuels

<sup>7</sup> Les sanctions s'appliquent aux écarts positifs (sous-estimation) comme aux écarts négatifs (surestimation).

recrutés par les partenaires techniques et financiers<sup>8</sup>. Le coordinateur de l'équipe est toujours un fonctionnaire. La taille de l'équipe varie de quatre à cinq personnes.

Vérifier la quantité dans un CS mobilise généralement entre un et trois membres de l'équipe, et prend deux à trois heures. Vérifier la quantité dans un hôpital mobilise habituellement de quatre à cinq membres, et prend de quatre heures à une journée. La vérification de la quantité pour une province prend habituellement d'une à deux semaines. Les résultats du mois sont saisis au niveau provincial dans une base de données nationale intégrée en ligne, et l'analyse est effectuée par la sous-unité validation du CPVV au cours d'une réunion mensuelle. Au cours de cette réunion, la performance de chaque structure de soins est analysée de façon à pouvoir identifier les sous-performances possibles, à analyser leurs causes et à effectuer des recommandations. La réunion peut aussi détecter les possibles incohérences dans les données (erreurs de compilation). Elle se termine par l'élaboration de la facture provinciale compilée.

### 3.3. Comment la qualité technique est-elle évaluée ?

La qualité technique des services est évaluée trimestriellement. Cette évaluation est effectuée par le personnel du BPS accompagné du BDS pour les CS. Elle est effectuée par des pairs (c'est-à-dire par des professionnels hospitaliers venant d'autres provinces) pour les hôpitaux, et des délégués du niveau central du MSPLS y participent depuis 2012. En théorie, les formations sanitaires qui bénéficient d'un sous-contrat devraient être évaluées par la structure dont elles dépendent. Dans la pratique, la qualité technique des sous-contractants n'est jamais évaluée. Les sous-contractants représentent plus de 20% des CS qui bénéficient du FBP. Il s'agit la plupart du temps de petites formations sanitaires privées situées dans des endroits reculés et moins fréquentées que les formations sanitaires bénéficiant d'un contrat principal, mais ce n'est pas toujours le cas. En tout état de cause, le fait qu'ils ne soient jamais évalués sur la qualité est un sujet de préoccupation majeur qui devrait faire l'objet d'une attention particulière du MSPLS.

Une partie de la prime de performance (évaluée trimestriellement par le niveau central) que les BPS reçoivent du programme FBP est fonction de la complétude de leur travail d'évaluation de la qualité technique des CS<sup>9</sup>. Les BPS ont été choisis pour cette activité car la qualité technique relève des normes : faire appliquer les normes est précisément le rôle du BPS. Les pairs ont été choisis pour les hôpitaux afin de leur permettre de se comparer entre eux. Le choix des évaluateurs, pour les CS comme pour les hôpitaux, a été guidé par la volonté de faire de l'évaluation de la qualité technique un instrument pédagogique permettant aux formations sanitaires de s'améliorer, autant qu'un instrument de vérification.

Au début de la mise en œuvre du FBP, l'évaluation de la qualité technique des hôpitaux était effectuée par des pairs provenant de la même province que ceux qu'ils évaluaient. Il arrivait donc fréquemment que les évaluateurs d'un jour aient déjà été évalués par des pairs d'hôpitaux qu'ils venaient juger. Ce système pouvait donner lieu à des collusions ou au contraire à des situations conflictuelles. Par exemple, en plusieurs occasions, certains évaluateurs ont donné de mauvaises notes en représailles à de mauvaises notes qu'ils avaient reçues dans le passé de la part de l'hôpital qu'ils étaient en train d'évaluer. A partir de 2012, il a donc été décidé que les pairs évaluateurs pourraient venir de n'importe quel hôpital du pays faisant partie du programme. Ceci n'a pas supprimé le problème (seulement 50 hôpitaux font partie du programme), mais l'a atténué. Il a également été décidé d'ajouter un facilitateur à l'équipe des évaluateurs. Il s'agit d'un tiers, désigné

<sup>8</sup> Le coordinateur et les autres fonctionnaires sont choisis par le MSPLS. Les contractuels sont des membres des organisations partenaires (comme Cordaid, HealthNet TPO, Pathfinder, etc.). Ils reçoivent tous un salaire. Les fonctionnaires peuvent recevoir en plus une prime basée sur la performance du CPVV (évaluée trimestriellement par le niveau central).

<sup>9</sup> Le reste de cette prime de performance est basée sur les tâches normales du BPS.

par la CT-FBP, qui fait généralement partie de l'assistance technique au niveau provincial ou central, et qui est chargé de représenter la CT-FBP lors de l'évaluation par les pairs et de prévenir et de résoudre les éventuels conflits qui pourraient survenir lors de cette activité. Au cours d'une visite d'évaluation de la qualité à laquelle l'auteur de cette étude a pu assister, il a été observé que le rôle de ce facilitateur était crucial, à la fois pour répondre aux questions concernant l'interprétation de la grille, et pour apaiser les tensions entre les évaluateurs et le personnel de l'hôpital.

Il n'y a pas d'échantillonnage : chaque structure est évaluée tous les trois mois, ce qui correspond au rythme de paiement de la qualité. L'évaluation des hôpitaux fait l'objet d'une planification par le niveau central, et les structures connaissent donc à l'avance la date à laquelle elles seront évaluées. Pour les CS, la situation peut varier en fonction des provinces, et la décision de les prévenir ou non revient au BPS. Certaines provinces préviennent les structures longtemps à l'avance, d'autres la veille, et d'autres encore ne préviennent pas du tout les structures. Dans tous les cas, les CS connaissent la période à laquelle l'évaluation va se dérouler : au cours de la première quinzaine du premier mois de chaque trimestre. Il aurait été intéressant de comparer les scores de structures qui sont averties en avance avec ceux de celles qui ne le sont pas, particulièrement dans des domaines comme le stock de médicament, l'hygiène, etc., pour lesquels il est facile de se préparer. Malheureusement, les données disponibles dans le cadre de cette étude ne permettaient pas de telles comparaisons.

Les superviseurs utilisent une check-list (annexe 4) basée sur les normes définies par le MSPLS. Cette check-list comprend 13 domaines pour les hôpitaux, et 14 pour les CS. Un certain nombre de points est défini pour chaque item de la check-list, ce qui donne un score de qualité technique (en pourcentage). La liste peut être modifiée, mais ceci arrive rarement : la grille d'évaluation des hôpitaux a été révisée seulement une fois de manière substantielle. En 2011, davantage d'aspects cliniques, ainsi que des éléments relatifs au dossier médical, ont été intégrés dans la check-list des hôpitaux. En 2012, une révision de moindre ampleur de la check-list pour les CS a été mise en œuvre.

L'évaluation de la qualité technique dans les CS mobilise presque tout le personnel des BPS, qui est composé d'infirmiers avec différents niveaux de qualification et d'un médecin (le directeur). Evaluer un CS mobilise entre trois et cinq personnes, et prend de trois à quatre heures. L'évaluation de la qualité technique dans les hôpitaux mobilise habituellement quatre personnes : le médecin-chef d'un BPS, le directeur d'un hôpital (un médecin), le responsable administratif d'un hôpital (un gestionnaire), et le chef nursing d'un hôpital (un infirmier). Chaque membre de l'équipe se concentre sur son domaine de compétence : les administratifs évaluent la gestion, et les cliniciens évaluent les aspects techniques. La visite peut prendre entre deux heures et toute une journée : cette durée dépend de la taille de l'hôpital, des résultats de l'évaluation et du degré de méticulosité des évaluateurs.

Avant de quitter la structure, l'équipe d'évaluation donne au responsable une copie de la check-list, avec des recommandations pour améliorer la qualité. De plus, une réunion trimestrielle est organisée au niveau provincial, où les résultats de l'évaluation de la qualité technique sont présentés publiquement. Au cours de cette réunion, les structures peuvent comparer leurs stratégies et échanger sur la meilleure manière d'améliorer la qualité.

Evaluer la qualité pour une province prend habituellement une à deux semaines. Comme pour la vérification de la quantité, les résultats des évaluations de la qualité sont saisis au niveau provincial dans la base de données nationale intégrée en ligne.

### **3.4. Comment les enquêtes communautaires sont-elles effectuées ?**

En plus de son travail sur la vérification de la quantité, le CPVV est chargé de contractualiser des Associations Locales (ASLO : par exemple des associations de développement local, des associations

de femmes, etc.) au niveau des structures pour effectuer semestriellement des enquêtes communautaires. Ces enquêtes ont deux principaux objectifs :

- Déterminer si les services déclarés dans les registres des structures ont effectivement eu lieu (détection des « patients-fantômes ») ;
- Déterminer le niveau de satisfaction des patients.

Chaque semestre, un ou plusieurs membres de la sous-unité vérification du CPVV se rendent dans les structures et choisissent aléatoirement dans les registres un échantillon de 80 patients qui ont visité la structure. On peut noter que pour des raisons de confidentialité, trois registres sont exclus du tirage : le registre du VIH, de la tuberculose, ainsi que ceux de la planification familiale. Les indicateurs qui se rapportent à ces activités ne sont donc jamais vérifiés dans le cadre de l'enquête communautaire. Cela représente 9 indicateurs sur 22 pour les centres de santé, et 8 indicateurs sur 24 pour les hôpitaux de district. Une part substantielle de l'activité des formations sanitaire échappe donc à l'enquête communautaire, ce qui est un sujet de préoccupation auquel le MSPLS devrait apporter des réponses.

Le tirage aléatoire est effectué en présence d'un membre du Comité de Santé, afin de renforcer la transparence. Les ASLO se rendent ensuite sur le terrain pour les entretiens. Les patients ne sont pas avertis que l'organisation va venir pour l'enquête. Ils n'ont pas la possibilité de refuser de figurer dans l'échantillon tiré aléatoirement, mais ils peuvent refuser de répondre aux questions. Le questionnaire (annexe 5) permet de déterminer si :

- Le patient existe réellement ;
- Il/elle a réellement fréquenté la structure pour le service approprié ;
- Il/elle était satisfait des soins prodigués ;
- Il/elle a trouvé les prix raisonnables ;
- Il/elle était satisfait/satisfaite de l'accueil ;
- Il/elle était satisfait/satisfaite de la capacité du personnel à le/la traiter ;
- Il/elle a des suggestions pour améliorer les soins dans la structure.

Seuls les trois premiers éléments sont pris en compte dans le calcul du score obtenu par les prestataires lors de l'enquête communautaire. Les 6 autres apportent aux structures de soins des informations qui les aident à s'améliorer en fonction des besoins exprimés par les patients. Ces trois éléments « payés » renvoient à deux objectifs disparates (contrôle de la fraude et satisfaction des patients), qui ne sont associés dans le système d'incitation que parce qu'ils le sont dans le système de vérification.

En combinant le score obtenu lors de l'enquête communautaire avec celui obtenu lors de l'évaluation de la qualité technique, on obtient un « score de qualité globale », dans lequel la qualité technique représente 60% et l'enquête communautaire 40%. Ces 40% se répartissent comme suit:

- Si plus de 5% des patients sélectionnés aléatoirement ne sont pas retrouvés lors de l'enquête communautaire, la formation sanitaire perd 10%. Si moins de 5% ne sont pas retrouvés, elle obtient 10% ;
- Si plus de 5% des patients qui ont été retrouvés ne confirment pas avoir été traités dans la structure, la formation sanitaire perd de nouveau 10%. Si moins de 5% ne confirment pas, elle obtient de nouveau 10% ;
- Si plus de 20% des patients qui confirment avoir été soignés déclarent ne pas être satisfaits des soins qu'ils ont reçus, la formation sanitaire perd 20%. Si moins de 20% se déclarent insatisfaits, elle obtient 20%.

Le score de qualité globale peut se traduire par un bonus ou par un malus, comme défini dans le tableau 4.

Score de qualité globale	Bonus ou malus de qualité
>90%	Subsides du trimestre précédent $\times$ 30% $\times$ score de qualité
70% - 89,9%	Subsides du trimestre précédent $\times$ 25% $\times$ score de qualité
50% - 69,9%	Pas de bonus de qualité
40% - 49,9%	-10% $\times$ subsides du trimestre précédent
30% - 39,9%	-20% $\times$ subsides du trimestre précédent
<30%	-25% $\times$ subsides du trimestre précédent

Tableau 4 : Critères d'attribution du bonus ou du malus de qualité (source : MSPLS, 2011b)

On peut remarquer que le rythme des enquêtes communautaires est semestriel, alors que le rythme de l'évaluation de la qualité technique est trimestriel. Pour les trimestres où l'enquête communautaire n'est pas réalisée, le score du trimestre précédent est utilisé pour calculer le score de qualité globale.

Ce système de bonus et de malus est venu remplacer en 2011 un ancien système, où il n'y avait pas de sanctions, mais seulement un bonus si le score de qualité globale. Le nouveau système, plus progressif, constitue une incitation plus forte à l'amélioration de la qualité. En revanche, ce mélange des résultats des enquêtes communautaires avec ceux de l'évaluation de la qualité technique a un double désavantage :

- Les prestataires reçoivent une incitation pour limiter le nombre de patients inscrits de manière erronée (et/ou frauduleuse) dans leurs registres. Il semblerait plus logique d'utiliser les résultats de l'enquête communautaire (du moins la partie qui porte sur l'« existence » et la « confirmation ») uniquement à des fins de sanction, et non d'incitation ;
- Les résultats des enquêtes communautaires sont rarement analysés au niveau national, parce qu'ils ne sont pas désagrégés des résultats de l'évaluation de la qualité technique : cela signifie qu'il y a peu d'analyse de la proportion de « patients-fantômes », au moins au niveau national.

Il y a une ASLO contractualisée par structure sanitaire. Chacune dispose de six membres formés à l'enquête, dont quatre participent chaque semestre (les deux autres sont l'équipe « de réserve »). Les qualités requises pour effectuer l'enquête sont :

- De savoir lire et écrire ;
- D'être considéré comme étant « de bonne moralité ».

Les capacités des enquêteurs sont faibles, mais les ASLO ont été choisies pour cette activité car une profonde connaissance de la communauté est nécessaire pour effectuer les enquêtes. Les organisations locales sont sélectionnées par un appel d'offre local, auquel toutes les associations implantées dans l'aire d'attraction de la structure peuvent participer. Le choix est habituellement le fruit d'une rude compétition. Les organisations locales sont payées à la fin de chaque enquête, sur la base du nombre de patients tirés au sort qu'elles ont réussi à interroger. Il s'agit d'un paiement basé sur la performance : les organisations locales ne sont pas payées pour les patients qu'elles ne retrouvent pas de manière à ce qu'elles fassent de leur mieux pour les retrouver. De plus, si elles s'avèrent ne pas remplir leurs obligations (mauvaise qualité systématique du remplissage des fiches, par exemple), leur contrat peut être résilié.

Chaque membre des associations locales a environ 20 patients à interroger chaque semestre. Les patients sont généralement répartis entre les membres en fonction de leur lieu de résidence, afin de réduire les coûts de transport. Les entretiens ne durent pas plus de 15 minutes, mais les membres peuvent passer beaucoup de temps à trouver le patient. L'enquête communautaire dans une province entière prend une à deux semaines. Le rythme des enquêtes communautaires était auparavant trimestriel, mais cette activité était considérée comme trop onéreuse et lourde, ce qui a conduit à passer à un rythme semestriel.

Les résultats des enquêtes communautaires sont compilés par les CPVV et non par les ASLO, qui ne disposent pas des compétences nécessaires. Ils sont présentés aux structures en même temps que les résultats de l'évaluation de la qualité technique. Ils sont saisis au niveau provincial dans la base de données nationale intégrée en ligne, mais seuls les scores agrégés pour l'ensemble de l'enquête sont disponibles dans cette base. Il n'est donc actuellement pas possible de connaître au niveau national les résultats désagrégés pour les trois éléments qui entrent dans le système d'incitation (existence des patients, confirmation des soins et satisfaction). L'absence de connaissance de la proportion de « patients-fantômes » au niveau national était une faiblesse majeure du système d'information du FBP burundais au moment de la collecte des données pour cette étude (l'information a dû être collectée et compilée au niveau provincial). Depuis, la base de données a été modifiée pour prendre en compte cette préoccupation.

### 3.5. Comment les données de la vérification sont-elles contre-vérifiées ?

La qualité des données fournies par les trois mécanismes de vérification décrits ci-dessus est assurée par différents mécanismes. Tout d'abord, tous les acteurs ont été formés sur les outils qu'ils doivent utiliser. Lors de la généralisation du FBP à tout le pays, des « formations en cascade » ont été organisées dans tout le pays : les médecins-chefs des BPS et des BDS ont été formés au niveau national, et ont ensuite organisé des ateliers locaux de formation au niveau du district. Toutes les ASLO ont effectué un pré-test du questionnaire avant leur première enquête communautaire. Enfin, les CPVV sont encadrés par des équipes d'accompagnateurs, expérimentés dans le FBP et financés par les partenaires techniques et financiers (voir partie 3.1).

Cependant, une assurance-qualité supplémentaire est jugée nécessaire, car de nombreux facteurs peuvent provoquer des erreurs (tous les mécanismes de collecte des données sont manuels, par exemple), et parce que le système de vérification repose principalement sur des acteurs internes au système de santé : la contre-vérification fournit une assurance supplémentaire de transparence et de validité. C'est pourquoi les trois mécanismes de vérification/évaluation sont contre-vérifiés par une organisation sélectionnée sur appel d'offre international pour une période de deux ans : Health, Development and Performance (HDP).

La contre-vérification a été ajoutée au processus en tant que contrôle : c'était un pré-requis des partenaires techniques et financiers afin d'éviter de payer pour des résultats à la fiabilité inconnue. Chaque trimestre, quatre provinces sont donc sélectionnées de manière aléatoire (à la fin de l'année, chaque province a donc été contre-vérifiée une fois). HDP se rend dans ces provinces, sélectionne aléatoirement un district dans lequel il refait la vérification de la quantité, l'évaluation de la qualité technique et l'enquête communautaire<sup>10</sup> dans l'hôpital et dans 25% des CS (échantillon de CS tiré aléatoirement). La contre vérification suit exactement les mêmes procédures que la vérification : les seules différences sont l'identité des vérificateurs et le degré de rigueur. Dans le cadre du processus de contre-vérification, les critères utilisés sont appliqués, sans aucune tolérance pour des différences mineures.

En théorie, seulement 5% de discordance entre les résultats obtenus par la vérification et ceux de la contre-vérification sont tolérés, au-delà de laquelle des sanctions financières peuvent être appliquées contre le CPVV. Dans la pratique, ces sanctions n'ont jamais encore été appliquées. La suite de cette étude montrera qu'il s'agit d'une faiblesse du FBP burundais.

En plus de rendre compte et d'analyser les discordances entre vérification et contre-vérification, les rapports d'HDP présentent de nombreux développements sur la performance des structures elle-

<sup>10</sup> Les organisations qui effectuent l'enquête communautaire dans le cadre de la contre-vérification doivent être différentes de celles qui l'effectuent habituellement dans le cadre de la vérification.

même, et sur l'évolution de celle-ci. Ils analysent les raisons d'éventuelles sous-performances, et effectuent des recommandations afin d'améliorer la quantité et la qualité des services fournis par les structures de santé au lieu de se concentrer sur la qualité des données. Cette tâche a été insérée dans les termes de référence de la contre-vérification à la demande du MSPLS, de manière à diversifier les canaux permettant de l'aider à améliorer le système. On peut cependant se questionner sur l'opportunité de telles analyses, qui sont redondantes avec celles opérées par la CT-FBP et par l'assistance technique nombreuse dont bénéficie le FBP au Burundi (rapports d'évaluation ou d'activité).

L'équipe de contre-vérification de HDP est composée de deux médecins, un statisticien, deux infirmiers et un comptable (plus quatre membres du personnel pour la logistique). Le niveau de qualification des contre-vérificateurs est en moyenne supérieur à celui des vérificateurs, et leur expérience du FBR en général, et de la vérification en particulier, est plus importante. La contre-vérification d'une province prend une semaine, et le temps passé dans chaque structure par les équipes de contre-vérification est supérieur à celui que passent les vérificateurs, car les premières sont plus rigoureuses. Chaque round de contre-vérification est conclu par un rapport détaillé, discuté au niveau national. HDP fournit de plus chaque année un rapport regroupant les quatre rounds effectués et comprenant l'ensemble des provinces du pays.

Le tableau 5 résume les quatre types de vérification à l'œuvre dans le cadre du FBP au Burundi.

	Vérification de la quantité	Evaluation de la qualité technique	Enquête communautaire	Contre-vérification
<b>Responsable</b>	CPVV	BPS et BDS	ASLO	HDP
<b>Objectif</b>	Vérification de la véracité des déclarations des prestataires	Evaluation de la conformité des soins subventionnés aux normes nationales	Evaluation de la satisfaction des patients et détection de la fraude	Audit qualité des données et suivi du processus FBP dans son ensemble
<b>Fréquence</b>	Mensuelle	Trimestrielle	Semestrielle	Trimestrielle
<b>Taille de l'échantillon</b>	Pas d'échantillonnage (chaque structure est vérifiée chaque mois)	Pas d'échantillonnage (chaque structure est évaluée chaque trimestre)	80 patients par structure	L'hôpital et 25% des CS dans 1 district de chacune des 4 provinces sélectionnées aléatoirement chaque trimestre
<b>Méthode</b>	Recomptage des cas dans les registres	Evaluation au moyen d'une check-list	Enquête auprès des patients	Toutes les méthodes utilisées par la vérification
<b>Nombre de personnes nécessaires pour vérifier une structure</b>	1 à 3 pour un CS, 4 à 5 pour un hôpital	3 à 5 pour un CS, 4 pour un hôpital	6	10 pour tout le pays
<b>Temps nécessaire pour vérifier une structure</b>	2 à 3 heures pour un CS. 4 heures à une journée pour un hôpital.	3 à 4 heures pour un CS, de 2h à une journée pour un hôpital	15 mn par patient	2 à 3 fois plus que pour la vérification
<b>Temps nécessaire pour vérifier une province</b>	1 à 2 semaines	1 à 2 semaines	1 à 2 semaines	1 semaine

Tableau 5 : Résumé des quatre méthodes de vérification utilisées dans le cadre du FBP au Burundi

## 4. RESULTATS DE LA VERIFICATION

Cette section présente les résultats produits par les quatre mécanismes de vérification que nous venons de décrire. Sauf mention contraire, les données utilisées proviennent de la base de données nationale intégrée en ligne (pour des raisons pratiques, une version Excel de cette base de données été utilisée au lieu de la version officielle en ligne).

### 4.1. Quels sont les résultats de la vérification de la quantité ?

La vérification de la quantité est effectuée par les CPVV dans les hôpitaux et les CS pour des indicateurs différents, et nous présenterons donc les résultats séparément pour ces deux types de structures. La formule utilisée pour calculer l'écart pour chaque structure est :

$$(Quantité vérifiée^{11} - Quantité déclarée) / (Quantité déclarée)$$

Un écart négatif signifie donc que la formation sanitaire concernée a surestimé sa performance. Nous avons décidé d'écarter du calcul les écarts supérieurs à 100%, car ils sont probablement majoritairement dus à des erreurs de saisie. La moyenne arithmétique des écarts nous a permis d'avoir une idée de l'orientation générale de ces derniers, mais elle ne permettait pas de donner un aperçu satisfaisant de leur taille car elle agrège des écarts positifs et des écarts négatifs. Nous avons donc également calculé la moyenne des valeurs absolues des écarts : les écarts négatifs ont été comptés comme des écarts positifs, de manière à ce qu'ils n'annulent pas les véritables écarts positifs.

Dans les CS, il a été possible d'obtenir les données sur les quantités déclarées et les quantités vérifiées pour toutes les structures du pays et pour tous les indicateurs sur la période janvier-août 2012. Avant cette date, le mode de compilation des données rend malheureusement ce calcul impossible. Il y a 565 CS, ce qui pour 8 mois et 19 indicateurs faisait un nombre d'observations possibles de 85 580. Toutes les prestations n'étant pas fournies pour tous les indicateurs tous les mois dans tous les CS et certains résultats aberrants ayant été écartés<sup>12</sup>, le nombre total d'observations est de 50 790. Les résultats sont présentés dans le tableau 6.

---

<sup>11</sup> Comme mentionné plus haut, la terminologie utilisée au Burundi peut prêter des confusions entre la vérification et la validation. Nous avons choisi, dans la suite de ce document, d'abandonner le terme de « quantités validées » (à moins qu'il n'en soit fait mention), et de nommer « quantités vérifiées » les données sur la base desquelles les paiements sont effectués (données vérifiées par la sous-unité vérification et validées par la sous-unité validation du CPVV).

<sup>12</sup> La plupart des observations manquantes le sont parce que les services ne sont pas fournis par les structures de santé, plutôt que parce que les données étaient incohérentes. Par exemple, peu de CS au Burundi fournissent le traitement ARV, ce qui explique le faible nombre d'observations pour cet indicateur.

	Moyenne des écarts	Moyenne des val. abs. des écarts	n <sup>13</sup>
Consultation curative ≥ 5 ans	-2%	4%	4 085
Consultation curative < 5 ans	-1%	2%	4 085
Journée d'observation ≥ 5 ans	-5%	8%	3 802
Journée d'observation < 5 ans	-4%	8%	2 896
Petite chirurgie	-6%	11%	3 987
Référence	-12%	14%	3 424
Enfant complètement vacciné	-3%	4%	4 065
Vaccination antitétanique 2-5	0%	3%	4 055
Femme enceinte sous ARV	-5%	6%	375
Prise en charge nouveau né ARV	-6%	7%	288
Dépistage volontaire du VIH	-2%	2%	3 737
Nouveau cas sous ARV	-7%	8%	253
Patient sous ARV depuis 6 mois	-12%	12%	245
IST	-2%	6%	3 950
Cas de tuberculose détecté	-1%	1%	599
Cas de tuberculose guéri	-5%	5%	918
Consultation curative femme enceinte	-5%	7%	4 062
Accouchement eutocique	-1%	2%	3 822
Planification familiale (pilule ou inj.)	-1%	2%	3 458
Implant ou DIU	-1%	2%	2 798
Consultation post-natale	-5%	6%	3 784
Consultation prénatale	-2%	3%	4 074
<b>Total</b>	<b>-4%</b>	<b>5%</b>	<b>50 790</b>

Tableau 6 : Ecart entre prestations déclarées par les CS dans leurs formulaires mensuels et prestations payées par le FBP après le processus de vérification par les CPVV dans tous les CS contractualisés au Burundi entre janvier et août 2012

L'écart moyen entre données déclarées et données vérifiées dans les CS était entre janvier et août 2012 de -4%, ce qui veut dire que les CS ont en moyenne surestimé leurs prestations. La moyenne des valeurs absolues des écarts est de 5%. Notons que les raisons de ces écarts sont multiples (erreurs de calcul, mauvaise identification des patients dans les registres, cas ne correspondant pas exactement aux définitions, etc.) et ne sont donc pas forcément liées à la fraude. Les écarts ne sont pas homogènes pour tous les indicateurs et pour toutes les structures de soins. Le tableau 7 décrit la répartition des erreurs entre déclaration exactes, surestimations et sous-estimations.

<sup>13</sup> Notons que *n* ne se rapporte pas au nombre de services déclarés (par exemple le nombre de consultations) mas au nombre de déclarations effectuées (chaque centre de santé doit déclarer le nombre de consultations chaque mois).

	% décl. correcte	Surestimation		Sous-estimation		n
		% décl. surest.	Surest. moy.	% décl. sous-est.	Sous-est. moy.	
Consultation curative ≥ 5 ans	22%	54%	-5%	25%	21%	4 085
Consultation curative < 5 ans	30%	49%	-2%	21%	19%	4 085
Journée d'observation ≥ 5 ans	44%	40%	-16%	16%	7%	3 802
Journée d'observation < 5 ans	69%	23%	-25%	8%	3%	2 896
Petite chirurgie	53%	31%	-28%	16%	5%	3 987
Référence	70%	26%	-48%	3%	1%	3 424
Enfant complètement vacciné	67%	23%	-13%	10%	5%	4 065
Vaccination antitétanique 2-5	68%	16%	-8%	16%	8%	4 055
Femme enceinte sous ARV	91%	8%	-74%	1%	0%	375
Prise en charge nouveau-né ARV	91%	8%	-81%	1%	1%	288
Dépistage volontaire du VIH	71%	21%	-9%	8%	5%	3 737
Nouveau cas sous ARV	89%	10%	-71%	1%	1%	253
Patient sous ARV depuis 6 mois	82%	17%	-73%	0%	0%	245
IST	72%	18%	-21%	10%	4%	3 950
Cas de tuberculose détecté	97%	2%	-53%	1%	0%	599
Cas de tuberculose guéri	93%	6%	-74%	0%	0%	918
Consultation curative femme enceinte	44%	40%	-14%	17%	8%	4 062
Accouchement eutocique	85%	11%	-12%	4%	2%	3 822
Planification familiale (pilule ou inj.)	58%	30%	-4%	12%	8%	3 458
Implant ou DIU	90%	6%	-24%	4%	2%	2 798
Consultation post-natale	78%	18%	-28%	4%	2%	3 784
Consultation prénatale	67%	23%	-10%	9%	8%	4 074
<b>Total</b>	<b>69%</b>	<b>22%</b>	<b>-19%</b>	<b>9%</b>	<b>4%</b>	<b>50 790</b>

Tableau 7 : Distribution des déclarations entre déclarations correctes, surestimées et sous-estimées par rapport prestations payées par le FBP après le processus de vérification par les CPVV dans tous les CS contractualisés au Burundi entre janvier et août 2012

Entre janvier et août 2012, dans 69% des déclarations des centres de santé, l'erreur était nulle, ce qui signifie que le montant déclaré par le CS au début du processus de vérification était le même que celui qui a été payé à la fin de celui-ci. Les 5% d'erreur moyenne constatés dans le tableau 6 sont donc dus à seulement 31% des déclarations. Parmi ces 31% de déclarations incorrectes, plus des deux tiers sont des surestimations : au cours de la période d'observation, 22% des déclarations étaient surestimées (la surestimation moyenne était de 19%), alors que 9% étaient sous-estimées (la sous-estimation moyenne était de 4%).

Dans les hôpitaux, il a été possible d'obtenir des données sur les quantités déclarées et vérifiées sur la période janvier-juillet 2012, ce qui faisait pour 50 hôpitaux et 24 indicateurs un nombre d'observations possibles de 8 400. Tous les hôpitaux ne délivrant pas tous les services concernés tous les mois et certains résultats aberrants ayant, comme pour les CS, été écartés, le nombre d'observations a été de 6 763. Les résultats sont présentés dans le tableau 8.

	Moyenne des écarts	Moyenne des val. abs. des écarts	n
Consultation curative ≥ 5 ans	-3%	4%	355
Consultation curative < 5 ans	-2%	3%	355
Consultation curative femme enceinte	-5%	8%	346
Contre-référence	-8%	9%	321
Chirurgie majeure	-2%	3%	275
Petite chirurgie	-6%	8%	343
Accouchement eutocique	-1%	2%	348
Césarienne	0%	0%	338
Accouchement dystocique	-1%	2%	341
Journée d'hospitalisation ≥ 5 ans	0%	3%	348
Journée d'hospitalisation < 5 ans	0%	3%	348
Femme enceinte sous ARV	-2%	3%	251
Prise en charge nouveau-né ARV	-2%	3%	244
Dépistage volontaire du VIH	-1%	2%	346
Nouveau cas sous ARV	-2%	2%	261
Patient sous ARV depuis 6 mois	-5%	6%	259
IST	-2%	9%	274
Cas de tuberculose détecté	-1%	2%	231
Circoncision	-2%	4%	251
Implant ou DIU	1%	1%	125
Planification familiale (pilule ou inj.)	-2%	4%	209
Planification familiale (méthode définitive)	-1%	1%	144
Consultation post-natale	-7%	9%	207
Consultation prénatale	-4%	7%	243
<b>Total</b>	<b>-2%</b>	<b>4%</b>	<b>6 763</b>

Tableau 8 : Ecart entre prestations déclarées par les hôpitaux et mensuels et prestations payées par le FBP après le processus de vérification par les CPVV dans tous les hôpitaux contractualisés au Burundi entre janvier et juillet 2012

L'écart moyen entre déclaration par les hôpitaux et vérification par les CPVV était entre janvier et juillet 2012 de -2%, ce qui veut dire que les hôpitaux ont en moyenne surestimé leurs prestations. La moyenne des valeurs absolues des écarts est de 4%. Comme pour le cas des CS, les raisons des erreurs sont multiples et ne sont pas liées à la fraude (définition des indicateurs, identification des patients, etc.). La taille moyenne des erreurs cache d'importantes différences entre les hôpitaux. Le tableau 9 donne un aperçu de la répartition des erreurs entre déclaration exacte, surestimation et sous-estimation.

	% décl. correcte	Surestimation		Sous-estimation		n
		% décl. surest.	Surest. moy.	% décl. sous-est.	Sous-est. moy.	
Consultation curative ≥ 5 ans	11%	60%	-6%	29%	3%	355
Consultation curative < 5 ans	18%	59%	-4%	23%	4%	355
Consultation curative f. enceinte	29%	47%	-14%	24%	6%	346
Contre-référence	60%	36%	-23%	5%	10%	321
Chirurgie majeure	78%	14%	-19%	8%	6%	275
Petite chirurgie	47%	40%	-17%	13%	9%	343
Accouchement eutocique	69%	22%	-6%	8%	4%	348
Césarienne	92%	6%	-7%	1%	3%	338
Accouchement dystocique	78%	14%	-14%	8%	7%	341
Journée d'hospitalisation ≥ 5 ans	14%	43%	-4%	43%	4%	348
Journée d'hospitalisation < 5 ans	26%	39%	-4%	35%	4%	348
Femme enceinte sous ARV	93%	6%	-43%	1%	43%	251
Prise en charge nouveau-né ARV	94%	4%	-57%	2%	19%	244
Dépistage volontaire du VIH	57%	31%	-6%	13%	3%	346
Nouveau cas sous ARV	93%	5%	-43%	2%	16%	261
Patient sous ARV depuis 6 mois	86%	13%	-40%	1%	56%	259
IST	66%	20%	-25%	14%	25%	274
Cas de tuberculose détecté	96%	3%	-51%	1%	75%	231
Circoncision	89%	8%	-39%	3%	33%	251
Implant ou DIU	95%	3%	-12%	2%	55%	125
Planification familiale (pilule ou inj.)	70%	24%	-11%	6%	19%	209
Planification familiale (méthode déf)	99%	1%	-38%	0%	n/a	144
Consultation post-natale	73%	23%	-33%	4%	25%	207
Consultation prénatale	57%	32%	-18%	12%	14%	243
<b>Total</b>	<b>62%</b>	<b>26%</b>	<b>-13%</b>	<b>12%</b>	<b>7%</b>	<b>6 763</b>

Tableau 9 : Distribution des déclarations entre déclarations correctes, surestimées et sous-estimées par rapport prestations payées par le FBP après le processus dans tous les hôpitaux de district contractualisés au Burundi entre janvier et juillet 2012

Entre janvier et juillet 2012, dans 62% des déclarations des hôpitaux, l'erreur était nulle, ce qui signifie que le montant déclaré par le CS au début du processus de vérification était celui qui a été payé à la fin de celui-ci. Les 4% d'erreur moyenne constatés dans le tableau 8 sont donc dus à seulement 38% des déclarations. Parmi ces 38% de déclaration, plus des deux tiers sont des surestimations : au cours de la période d'observation, 26% des déclarations étaient surestimées (la surestimation moyenne était de 13%), et 12% étaient sous-estimées (la sous-estimation moyenne était de 7%).

Il existe un consensus parmi les acteurs interrogés pour estimer que la fraude délibérée est rare, dans les CS comme dans les hôpitaux. Ceci est confirmé par les différents rapports disponibles (MSPLS 2011a, MSPLS 2012, rapports de contre-vérification de HDP, etc.). Certains des membres des CPVV interviewés ont rapporté des pratiques malhonnêtes (ils ont par exemple découvert durant les visites de vérification que certains prestataires inscrivent par exemple les anciens cas comme nouveaux cas), mais celles-ci restent d'après eux extrêmement minoritaires. La plupart des erreurs constatées dans les tableaux présentés-ci-dessus sont d'après eux non-intentionnelles : elles sont dues à une mauvaise compréhension des indicateurs, à des erreurs de remplissage, à un mauvais comptage, etc.

L'observation des données présentées dans les tableaux 6 à 9 semble confirmer les déclarations faites par les personnes interrogées dans le cadre de cette étude. Le plus haut niveau d'erreur peut être trouvé pour les indicateurs où taux d'utilisation est élevé (par exemple, les consultations curatives au CS), c'est-à-dire pour les indicateurs où le risque d'erreur lors du comptage est le plus important. Un autre facteur qui semble favoriser les erreurs est la définition des indicateurs : quand

celle-ci est compliquée, les structures ont plus de chance de compter un service qui ne correspond pas exactement à la définition, et qui ne sera pas validé par l'équipe de vérification (on peut ici prendre l'exemple des consultations pour les femmes enceintes dans les CS : seules les consultations liées à la grossesse doivent être comptées, mais beaucoup de CS continuent de déclarer des consultations données à des femmes enceintes pour d'autres raisons).

Bien que cette situation ne soit pas liée à la fraude, on peut s'étonner qu'après plus de deux ans de mise en œuvre du FBP dans le pays, le niveau d'erreur non-intentionnelle soit encore aussi élevé. Il aurait été intéressant d'étudier les données pour les années précédentes afin de voir quelle était l'évolution de l'erreur moyenne, mais cela n'a malheureusement pas été possible. Comme expliqué plus haut, le mode de compilation des données qui a permis les calculs présentés dans cette section n'a été instauré qu'à partir de l'année 2012.

## **4.2. Quels sont les résultats de l'évaluation de la qualité technique ?**

Sur la période 2010-2012, la moyenne des scores de qualité technique était de 70% dans les CS. Elle était de 77% sur la même période dans les hôpitaux de districts, et de 74% dans les hôpitaux nationaux. L'évaluation des hôpitaux est supérieure à celle des CS, mais d'après les personnes interrogées, les conséquences que l'on peut tirer de cette observation doivent davantage porter sur les modes d'évaluation que sur la performance des structures elle-même. L'illustration 3 montre l'évolution de ces scores depuis la généralisation du FBP pour les trois catégories de structures sanitaires.

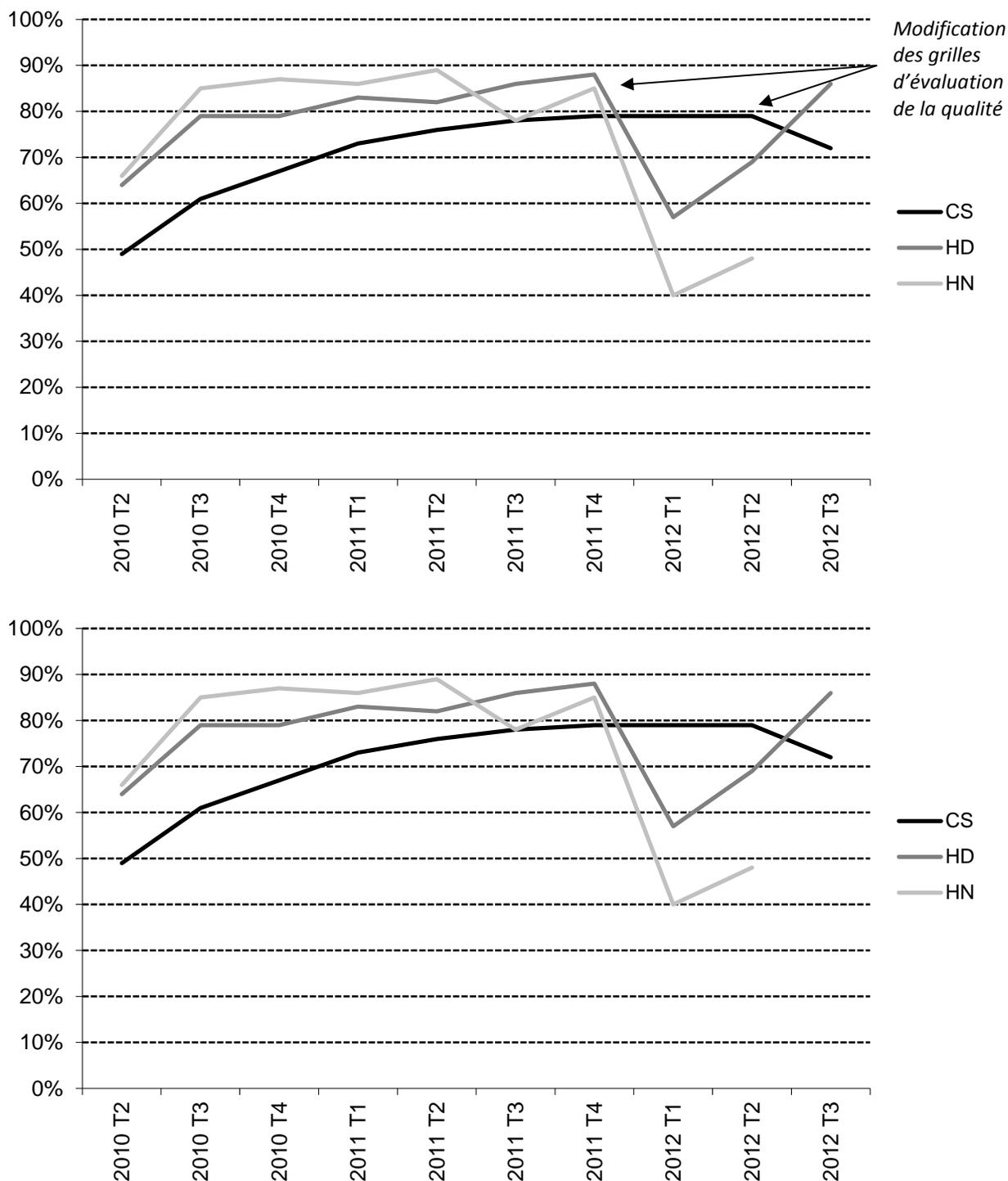


Illustration 3 : Evolution des scores moyens de qualité technique dans les CS, hôpitaux de districts et hôpitaux nationaux contractualisés au Burundi entre janvier 2010 et novembre 2012

Ce graphique montre que les évaluations de la qualité technique des structures ont donné des résultats de plus en plus satisfaisants pour les CS jusqu'au deuxième trimestre 2012, et pour les hôpitaux jusqu'au quatrième trimestre 2011. Les baisses observées à ces dates sont dues à des changements dans la grille d'évaluation (changements majeurs pour les hôpitaux, mineurs pour les CS), auxquels les prestataires n'étaient pas préparés (par exemple, introduction de l'analyse des dossiers patients dans l'évaluation de la qualité des hôpitaux). Ils ont rapidement retrouvé leurs

scores des évaluations précédentes, ce qui montre qu'il y a une attention plus grande des prestataires pour les éléments qui sont mesurés que pour ceux qui ne le sont pas.

Certaines des personnes interrogées ont pointé des limites de l'évaluation de la qualité technique. Elles ont notamment observé qu'il était difficile de maintenir le niveau de qualité entre deux évaluations : il est très probable que les formations sanitaires, connaissant la date ou la période à laquelle ils seront évalués, se préparent en conséquence. D'autre part, le fait que l'évaluation des hôpitaux soit effectuée par des pairs est probablement un facteur de complaisance. Enfin, notons que les grilles d'évaluation de la qualité, bien qu'elles aient été modifiées, sont encore focalisées sur les conditions dans lesquelles les soins sont délivrés, et moins sur la manière dont ils sont délivrés.

En résumé, il est probable que l'évaluation de la qualité technique ait incité, surtout au début de la mise en œuvre du FBP, les prestataires à améliorer leurs pratiques. Mais avec le temps, les possibilités de *gaming*, c'est-à-dire les opportunités pour les prestataires de jouer avec les règles pour augmenter leur score sans pour autant modifier réellement leur qualité réelle, peuvent voir le jour. Des mécanismes permettant de lutter contre cette tendance doivent être recherchés par le MSPLS. Certains de ces changements passent par les modifications régulières de la grille d'évaluation, la systématisation des évaluations non-annoncées, etc., et ont été mis en œuvre dès le début 2013, après que la collecte des données pour cette étude ait été achevée.

#### 4.3. Quels sont les résultats des enquêtes communautaires ?

Comme cela a été précisé dans la partie 3.3, les résultats de l'enquête communautaire comprennent trois types de données liées au FBP :

- Existence : le patient sélectionné aléatoirement dans les registres de la formation sanitaire a-t-il pu être retrouvé dans la communauté ?
- Confirmation : le patient confirme-t-il qu'il a fréquenté la formation sanitaire pour le service concerné ?
- Satisfaction : était-il satisfait des soins reçus ?

Le score pour ces trois indicateurs représente ce qu'on appelle le « score de qualité subjective ». Entre 2010 et 2012, ce score est passé de moins de 50% à plus de 75% pour les centres de santé, et de moins de 40% à 60% dans les hôpitaux. Mais cette progression du score agrégé ne renseigne pas sur l'évolution de la proportion de possibles « patients-fantômes » qui pourraient être ajoutés par les formations sanitaires dans leurs registres pour maximiser leur revenu. Comme cela a déjà été évoqué, les données désagrégées sur les enquêtes communautaires ne sont pas encore disponibles au niveau central : seuls les scores combinés pour les trois éléments (« score de qualité subjective ») sont enregistrés dans la base de données nationale en ligne. Pour les besoins de cette étude, il a donc été demandé aux CPVV de saisir les résultats désagrégés dans une feuille de calcul Excel. 13 CPVV sur 17 ont envoyé les données pour 2011 et pour le premier semestre 2012.

Les données collectées, bien que ne provenant pas d'un échantillon aléatoire, sont informatives. L'illustration 4 montre l'évolution de la proportion de patients qui n'ont pas pu être retrouvés dans la communauté (existence) et qui n'ont pas confirmé avoir reçu les soins (confirmation) dans 389 CS des 13 provinces de l'échantillon (soit 75% des CS du pays).

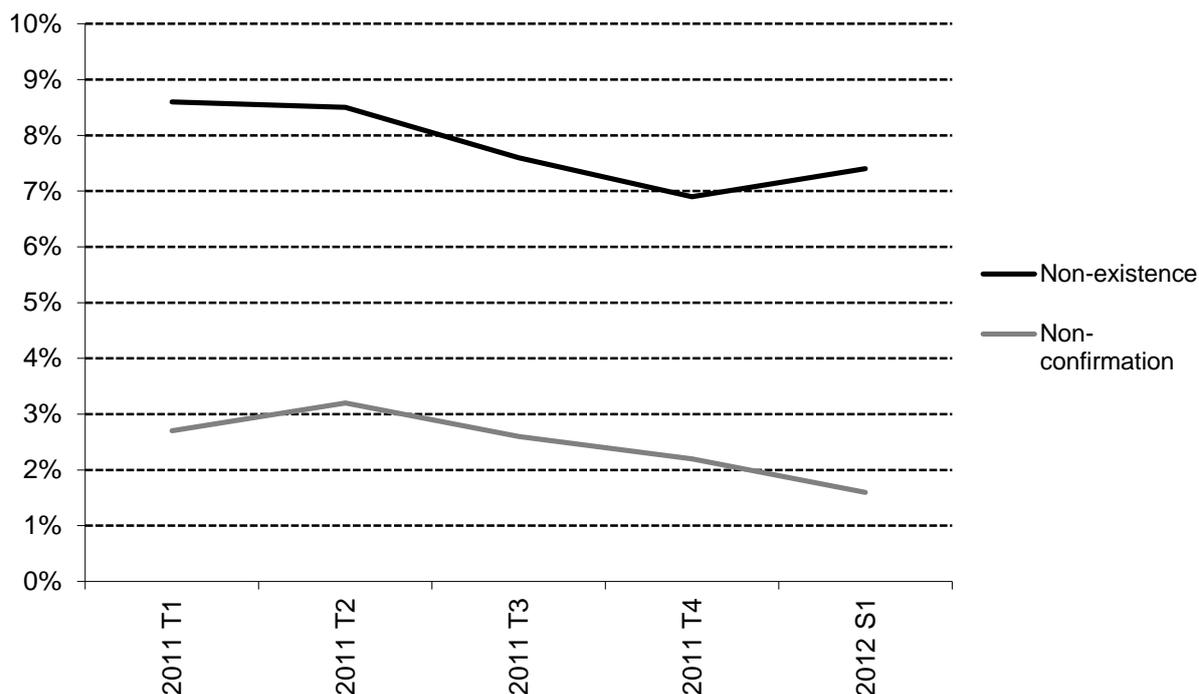


Illustration 4: Evolution de la proportion de patients sélectionnés pour l'enquête communautaire qui n'ont pas pu être retrouvés dans la communauté (non-existence) et qui n'ont pas confirmé avoir reçu les soins (non-confirmation) dans les centres de santé des 13 provinces de l'échantillon entre le premier trimestre 2011 et le premier semestre 2012<sup>14</sup>

Dans les centres de santé des 13 provinces de l'échantillon, la proportion de patients qui n'ont pas pu être retrouvés dans la communauté au cours des enquêtes communautaires a varié entre 8,6% et 6,9% entre le premier trimestre 2011 et le premier semestre 2012, avec une légère tendance à la baisse. La moyenne reste supérieure à la limite de 5% au-dessus de laquelle les centres de santé perdent 10% du score de qualité global. La proportion de centres de santé avec un pourcentage supérieur à 5% est passée de plus de 50% au premier trimestre 2011 à 40% au premier semestre 2012. Parmi les patients qui ont pu être retrouvés, la proportion de ceux qui n'ont pas confirmé avoir été traité au centre de santé a varié entre 3,2% et 1,6%, avec la même tendance à la baisse. La moyenne est cette fois-ci inférieure à la limite de 5% au-delà de laquelle les centres de santé perdent 10% du score de qualité global.

L'illustration 4 montre l'évolution de la proportion de patients qui n'ont pas pu être retrouvés dans la communauté (existence) et qui n'ont pas confirmé avoir reçu les soins (confirmation) dans les 38 hôpitaux des 13 provinces de l'échantillon (soit 84% des hôpitaux de district du pays).

<sup>14</sup> Comme expliqué plus haut, le rythme des enquêtes communautaires était trimestriel, mais il est devenu semestriel à partir de 2012.

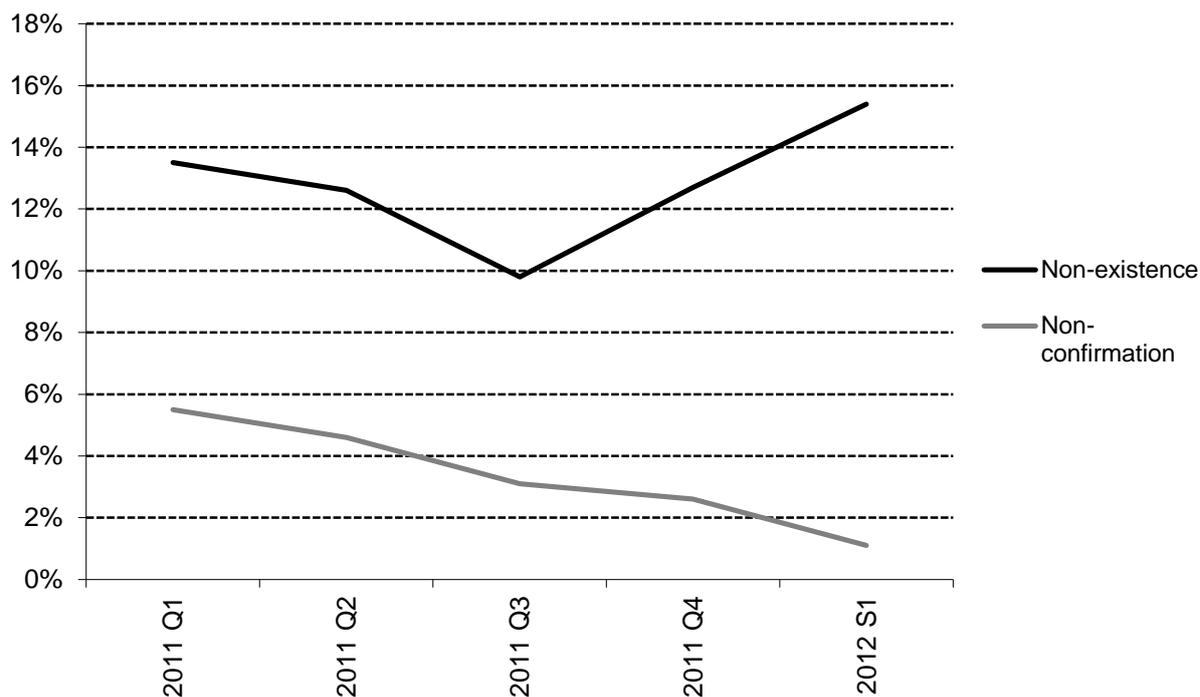


Illustration 5: Evolution de la proportion de patients sélectionnés pour l'enquête communautaire qui n'ont pas pu être retrouvés dans la communauté (non-existence) et qui n'ont pas confirmé avoir reçu les soins (non-confirmation) dans les hôpitaux des 13 provinces de l'échantillon entre le premier trimestre 2011 et le premier semestre 2012

Dans les hôpitaux des 13 provinces de l'échantillon, la proportion de patients qui n'ont pas pu être retrouvés dans la communauté au cours des enquêtes communautaires a varié entre 9,8% et 15,4% entre le premier trimestre 2011 et le premier semestre 2012, et contrairement à ce qui avait été constaté pour les CS, aucune tendance à la baisse n'a pu être remarquée. La moyenne reste largement supérieure à la limite de 5% au-dessus de laquelle les hôpitaux perdent 10% du score de qualité global. La proportion d'hôpitaux avec un pourcentage supérieur à 5% a varié entre 60% et 80% entre le premier trimestre 2011 et le premier semestre 2012. Parmi les patients qui ont pu être retrouvés, la proportion de ceux qui n'ont pas confirmé avoir été traité à l'hôpital a varié entre 5,5% et 1,4%, avec une forte tendance à la baisse. Comme dans le cas des centres de santé, la moyenne est cette fois-ci inférieure à la limite de 5% au-delà de laquelle les hôpitaux perdent 10% supplémentaires de leur score de qualité global.

Le fait que la proportion de patients non-retrouvés soit plus élevée dans les hôpitaux que dans les CS peut être expliqué par la nature des patients qui se rendent à l'hôpital : ils proviennent de zone plus éloignées, sont davantage susceptibles d'avoir voyagé pour se soigner que les patients des CS et sont par conséquent plus difficiles à retrouver. La proportion relativement élevée de patients non-retrouvés au cours des enquêtes communautaires est un sujet de préoccupation et devrait faire l'objet d'une attention sérieuse, mais elle ne signifie pas qu'une même proportion de patients sont des « patients-fantômes ». De nombreuses raisons peuvent expliquer pourquoi un patient n'est pas retrouvé :

- Il peut avoir été insuffisamment identifié dans la formation sanitaire (par exemple, adresse inexacte ou incomplète. Il arrive même que des patients donnent de fausses adresses de manière intentionnelle) ;
- Il peut avoir temporairement ou définitivement quitté la zone (ceci est particulièrement le cas dans les zones où les réfugiés ou les travailleurs saisonniers sont nombreux) ;

- Il peut refuser de répondre à l'enquête.

Cependant, comme cela a été expliqué plus haut, le principal sujet de préoccupation en ce qui concerne les enquêtes communautaires n'est pas leurs résultats mais leur mode de mise en œuvre : ils omettent un grand nombre d'indicateurs pour des raisons de confidentialité, et leurs résultats ne pouvaient pas être analysés de manière désagrégée au niveau central au moment de la collecte des données pour cette étude, ce qui empêchait d'analyser les tendances au niveau des structures sanitaires.

#### 4.4. Quels sont les résultats de la contre-vérification ?

L'organisation indépendante qui effectue la contre-vérification, HDP, rend compte de ses activités dans des rapports trimestriels qui ont été exploités et qui constituent la source des données présentées dans cette section. Les résultats de la contre-vérification de la quantité dans les CS y sont malheureusement présentés de manière agrégée (au niveau de la province). Cette manière de présenter les résultats a une double conséquence :

- Elle sous-estime le niveau d'erreur moyen en agrégeant les résultats de 2 ou 3 CS sélectionnés aléatoirement au niveau provincial, les rapports et en additionnant des différences positives et négatives (sous et surestimations) ;
- Elle rend impossible l'analyse de l'évolution des résultats de la contre-vérification dans le temps pour les CS qui seraient sélectionnés aléatoirement plus d'une fois.

Ces deux faiblesses peuvent facilement être corrigées dans les prochains rapports de contre-vérification. Le tableau 10 donne les résultats de la contre-vérification dans les CS, agrégés au niveau des provinces pour les deux années de contre-vérification. Il y a au maximum, pour chaque indicateur, 32 observations : chaque province a été contre-vérifiée deux fois. Pour chaque province, la formule de calcul utilisée est la suivante :

$$\frac{(\text{Quantité provinciale contre-vérifiée par HDP} - \text{Quantité provinciale vérifiée par le CPVV})}{(\text{Quantité provinciale vérifiée par le CPVV})}$$

	Moyenne des écarts provinciaux	n
Consultation curative ≥ 5 ans	0%	32
Consultation curative < 5 ans	0%	32
Consultation curative femme enceinte	0%	32
Journée d'observation ≥ 5 ans	2%	32
Journée d'observation < 5 ans	-1%	29
Petite chirurgie	0%	32
Référence	3%	32
Enfant complètement vacciné	1%	32
Vaccination antitétanique 2-5	-2%	32
Distribution de moustiquaires	1%	32
Nouvelles latrines construites	0%	32
Femme enceinte sous ARV	0%	12
Prise en charge nouveau né ARV	0%	14
Dépistage volontaire du VIH	0%	31
Nouveau cas sous ARV	0%	14
Patient sous ARV depuis 6 mois	0%	14
IST	-1%	32
Cas de tuberculose détecté	6%	18
Cas de tuberculose guéri	1%	27
Accouchement eutocique	0%	32
Planification familiale (pilule ou inj.)	0%	30
Implant ou DIU	0%	28
Consultation post-natale	1%	32
Consultation prénatale	0%	32
<b>Total</b>	<b>0%</b>	<b>665</b>

Tableau 10 : Ecarts entre total provincial vérifié par les CPVV et payé par le programme, et total provincial contre-vérifié par HDP pour les indicateurs des CS dans toutes les provinces contre-vérifiées au Burundi au cours des huit rounds de contre-vérification

Les résultats présentés dans le tableau 10 sont sous-estimés car ils agrègent des écarts positifs et des écarts négatifs. Cependant, ils permettent de montrer que les écarts entre contre-vérification et vérification sont en moyenne moins importants que les écarts entre déclaration et vérification. A part pour un indicateur (cas de tuberculose détecté), les marges d'erreurs de 5% prévues par le manuel de procédure ne sont en moyenne pas dépassées au niveau provincial. La vérification par les CPVV dans les CS montre peu de différence entre la vérification et la contre-vérification, même s'il serait souhaitable, pour avoir une idée plus précise, de disposer des résultats au niveau désagrégé pour avoir une lecture plus fine des résultats et une analyse de leur évolution dans le temps.

Pour les hôpitaux en revanche, les rapports de contre-vérification donnent les résultats désagrégés par structure. Nous avons donc pu calculer les écarts moyens avec exactement la même méthodologie que pour les résultats de la vérification présentés dans la section 4.1, qui ne conduit pas à une sous-estimation du niveau d'erreur. Le tableau 11 donne les résultats de la contre-vérification dans les hôpitaux. Il y a au maximum, pour chaque indicateur, 32 observations : chaque province a été contre-vérifiée deux fois. La formule de calcul pour chaque hôpital est la suivante :

$$(Quantité\ contre-vérifiée\ par\ HDP - Quantité\ vérifiée\ par\ le\ CPVV) / (Quantité\ vérifiée\ par\ le\ CPVV)$$

	Moyenne des écarts	Moyenne des val. abs. des écarts	n
Consultation curative ≥ 5 ans	-10%	21%	32
Consultation curative < 5 ans	-12%	18%	32
Consultation curative femme enceinte	0%	27%	32
Contre-référence	-65%	77%	31
Chirurgie majeure	-18%	26%	29
Petite chirurgie	-2%	27%	32
Accouchement eutocique	-4%	12%	32
Césarienne	-2%	3%	31
Accouchement dystocique	-28%	37%	30
Journée d'hospitalisation ≥ 5 ans	-13%	20%	32
Journée d'hospitalisation < 5 ans	-12%	18%	32
IST	50%	74%	30
Cas de tuberculose détecté	22%	32%	29
Cas de tuberculose guéri	-46%	63%	7
Implant ou DIU	4%	21%	17
Planification familiale (pilule ou inj.)	-9%	16%	18
Planification familiale (méthode définitive)	-9%	41%	22
Consultation prénatale	-19%	27%	24
Consultation post-natale	-35%	42%	23
Circoncision	8%	66%	7
<b>Moyenne</b>	<b>-9%</b>	<b>31%</b>	<b>522</b>

Tableau 11 : Ecart entre vérification par les CPVV et contre-vérification par HDP pour les indicateurs des hôpitaux dans tous les hôpitaux contre-vérifiés au Burundi au cours des huit rounds de contre-vérification

Dans les hôpitaux, la différence entre les données de vérification et les données de contre-vérification sont importantes et systématiques. La performance quantitative est en moyenne surévaluée par les CPVV de 9%, la moyenne des valeurs absolues des écarts est de 31%, et la différence pour tous les indicateurs sauf un (césariennes) est en moyenne supérieure à la marge de 5% tolérée par le manuel de procédures. Si les sanctions prévues par ce dernier étaient appliquées, le montant des subsides aux CPVV aurait été réduit.

7 hôpitaux ont été contre-vérifiés deux fois entre 2010 et 2012 : ils ont été aléatoirement sélectionnés pour les deux rounds de contre-vérification. Ceci nous permet d'analyser l'évolution des résultats de la contre-vérification pour ce sous-échantillon, afin de détecter de possibles effets d'apprentissage (on s'attend à ce que les CPVV s'améliorent et que le niveau d'erreur diminue). Le tableau 12 présente les résultats de cette analyse. 20 indicateurs sont contre-vérifiés, ce qui pour 7 hôpitaux et deux rounds de contre-vérification donne un nombre total d'observations de 280. Tous les hôpitaux ne délivrent pas tous les services tous les mois, ce qui explique pourquoi notre nombre total d'observations est seulement de 103.

	n	%	Erreur moy. round 1	Erreur moy. round 2	Différence
Indicateurs pour lesquels le niveau d'erreur a augmenté entre les rounds 1 et 2	58	56%	-6%	-21%	14%
Indicateurs pour lesquels le niveau d'erreur a diminué entre les rounds 1 et 2	45	44%	-26%	-12%	-13%
<b>Total</b>	<b>103</b>	<b>100%</b>	<b>-15%</b>	<b>-17%</b>	<b>2%</b>

Tableau 12: Evolution du niveau d'erreur effectuées par les CPVV tel que détecté par la contre-vérification dans les hôpitaux qui ont été aléatoirement sélectionnés dans les 2 rounds de contre-vérification (7 hôpitaux)

Entre les rounds 1 et 2 de la contre-vérification, pour les indicateurs qui ont pu être contre-vérifiés deux fois dans le même hôpital, le niveau moyen d'erreur a augmenté d'une surestimation moyenne de 15% à une surestimation moyenne de 17% (on peut noter que le niveau moyen d'erreur dans l'échantillon d'hôpitaux qui a été contre-vérifié deux fois est inférieur aux résultats pour l'ensemble du pays présentés dans le tableau 11). Pour 56% de ces indicateurs, le niveau d'erreur était plus élevé au cours du second round qu'au cours du premier (l'augmentation moyenne était de 14 points). Pour les 44% restants, le niveau d'erreur était moins élevé au cours du second round que du premier (la diminution moyenne était de 13 points). Le tableau 12 montre donc qu'il n'y a pas eu d'effet d'apprentissage, et que la performance des CPVV dans la vérification quantitative de la performance des hôpitaux ne s'est pas améliorée entre les 2 rounds de contre-vérification.

Ces résultats tendent à montrer que la performance des CPVV dans la vérification de la quantité des hôpitaux est insuffisamment rigoureuse, mais les différences mises en avant peuvent, d'après les personnes interrogées dans le cadre de cette étude, largement être expliquée par l'absence de registres standardisés pour l'échelon hospitalier. Les hôpitaux construisent eux-mêmes leurs registres, ce qui conduit à une grande variabilité dans la manière dont leurs résultats sont rapportés. D'autre part, la contre-vérification applique la plus grande rigueur, et elle est souvent conduite à annuler la totalité d'un indicateur. Une autre explication est que la contre-vérification intervient longtemps après la période où les soins ont été délivrés (six mois après en moyenne) : les registres dans lesquels les activités ont été consignées ne sont parfois plus accessibles (archivés dans une pièce dont une personne absente a la clé, par exemple). Une mission conduite par le MSPLS après la contre-vérification d'un hôpital a par exemple trouvé que les la raison des importantes discordances dans cet hôpital s'expliquaient par le fait que trois registres n'avaient pas été retrouvés par HDP : la contre-vérification n'a donc pas pu inclure les services consignés dans ces registres, ce qui rendait le niveau d'erreur artificiellement élevé.

Ces différentes raisons expliquent les résultats, la plupart du temps négatifs (surestimation de la performance par les CPVV), présentés dans le tableau 11. Elles expliquent aussi pourquoi le MSPLS a préféré ne pas appliquer contre les CPVV les sanctions prévues par le manuel de procédures du FBP en cas de discordance relevée lors de la contre-vérification. On peut cependant penser que ces sanctions pourraient inciter les CPVV à exiger davantage de rigueur et une meilleure gestion de la part des hôpitaux, de manière à améliorer leurs propres résultats et à faire en sorte que les registres soient disponibles. Il semble donc souhaitable d'envisager leur mise en œuvre, et de clarifier ce que les CPVV doivent exiger de l'hôpital.

La contre-vérification ne se cantonne pas à la vérification de la quantité. Elle effectue aussi à nouveau les évaluations de la qualité technique effectuées par les BPS et les BDS dans les CS, et par les pairs dans les hôpitaux. Dans les rapports de la contre-vérification, les résultats sont comparés avec les résultats pour le trimestre précédent. La contre-vérification intervenant généralement longtemps après la vérification, la qualité dans les structures contre-vérifiées a largement pu varier dans les structures concernées. C'est pourquoi nous avons préféré dans le cadre de cette étude comparer les résultats de l'évaluation de la qualité effectuée par les équipes de contre-vérification avec ceux de l'évaluation la plus proche dans le temps effectuée par les équipes de vérification (soit avant, soit après la contre-vérification). Cette comparaison n'est pas exempte de problèmes, car bien que l'intervalle de temps entre les deux évaluations soit réduit, la qualité peut avoir changé dans les formations sanitaires. Cependant, nous estimons que cette comparaison est plus pertinente que celle effectuée dans les rapports de contre-vérification. Les résultats sont présentés dans le tableau 13<sup>15</sup>. Pour chaque structure de santé, la formule de calcul utilisée pour établir l'écart est la suivante :

<sup>15</sup> Contrairement à ce qui se passe pour la contre-vérification de la quantité, les rapports de contre-vérification présentent les résultats de l'évaluation de la qualité technique pour les CS de manière désagrégée. Nous pouvons donc effectuer les mêmes calculs pour les CS et pour les hôpitaux.

(Score établi par HDP - Score établi par les BDS, BPS ou les pairs) / (Score établi par les BDS, BPS ou les pairs)

	Moyenne des écarts	% avec sur-évaluation <sup>16</sup>	Suréval. moyenne	% avec sous-évaluation <sup>17</sup>	Sous-éval. moy.	n
CS	-11%	79%	-20%	21%	24%	101
Hôpitaux	-17%	84%	-24%	16%	20%	32

Tableau 13 : Ecarts entre évaluation de la qualité technique par les BPS, BDS ou les pairs et contre-vérification par HDP dans tous les CS et hôpitaux contre-vérifiés au Burundi au cours des huit rounds de contre-vérification

Si on la compare aux résultats de HDP, la qualité technique telle qu'évaluée par les BPS et les BDS est surévaluée de 11% dans les CS et de 17% dans les hôpitaux. Cette conclusion doit être nuancée car, comme cela a été mentionné plus haut, les deux évaluations ont lieu à des moments différents. Cependant, cette surévaluation a un caractère systématique (elle intervient dans 79% des CS contre-vérifiés et dans 85% des hôpitaux contre-vérifiés), ce qui tend à montrer que les résultats présentés dans le tableau 13 ne peuvent pas uniquement être expliqués par la distance temporelle entre la vérification et la contre-vérification.

Cette surestimation peut s'expliquer par la grande rigueur appliquée par les contre-vérificateurs, mais les personnes interrogées ont aussi avancé des facteurs institutionnels : les BPS et les BDS sont les supérieurs hiérarchiques des structures qu'ils évaluent, et ils n'ont donc pas intérêt à mettre en évidence leurs lacunes aux yeux de tous. Quant aux pairs, ils ne peuvent manquer de penser que leurs évaluations peuvent influencer la manière dont ils seront évalués par leurs collègues par la suite. En ce qui concerne les CS, ces problèmes institutionnels peuvent avoir une solution partielle dans l'extension (et la mise en œuvre) aux BPS et BDS des sanctions prévues contre les CPVV dans le manuel de procédures du FBP jusqu'ici : si le revenu des BPS et BDS était lié aux résultats de la contre-vérification, ils pourraient augmenter le niveau de rigueur de leur évaluation de la qualité.

En ce qui concerne les enquêtes communautaires, les rapports de contre-vérification ne comparent pas leurs résultats avec ceux des enquêtes communautaires effectuées par les ASLO sous la supervision des CPVV. Ceci est justifié par le fait que les échantillons de patients interrogés ne sont pas les mêmes lors des deux enquêtes : les résultats sont différents, par construction. Dans ces conditions, on peut se demander pourquoi la contre-vérification effectue ces enquêtes communautaires de toute façon : il s'agit d'une activité coûteuse qui nécessite le recrutement d'une autre association locale pour effectuer un travail qui a déjà été effectué et qui n'apporte pas d'élément nouveau. Il serait préférable soit d'exclure l'enquête communautaire de la contre-vérification, soit de l'effectuer avec le même échantillon utilisé lors de la vérification.

#### 4.5. A quoi les résultats sont-ils utilisés ?

La principale utilisation des résultats de la vérification est le paiement des prestataires : les résultats de la vérification de la quantité servent au paiement mensuel de subsides aux prestataires, et les résultats de l'évaluation de la qualité technique, combinés avec ceux des enquêtes communautaires, se traduisent dans le bonus ou un malus de qualité que les structures peuvent recevoir trimestriellement.

En ce qui concerne la vérification de la quantité, nous avons tenté de reconstituer les montants économisés grâce à la vérification en comparant les montants qui auraient été payés en l'absence de

<sup>16</sup> Pourcentage de contre-vérifications ayant conclu à une surévaluation de la qualité technique des formations sanitaires par les BPS, BDS ou les pairs.

<sup>17</sup> Pourcentage de contre-vérifications ayant conclu à une surévaluation de la qualité technique des formations sanitaires par les BPS, BDS ou les pairs.

vérification (montants déclarés) avec ceux qui ont effectivement été payés. Cette comparaison est théorique pour deux raisons :

- La vérification, surtout quand elle est fréquente et intégrale comme c'est le cas pour la quantité au Burundi, a un effet dissuasif sur la fraude : il est probable que les montants déclarés seraient beaucoup plus importants en l'absence de vérification ;
- Pour arriver à ce résultat, nous avons dû reconstituer un montant déclaré théorique : le montant déclaré ne figure en effet pas dans les bases de données. Seules les quantités déclarées sont enregistrées : il n'a pas été possible de calculer la valeur monétaire exacte des services déclarés en utilisant ces dernières, car la valeur monétaire de chaque indicateur varie d'une province à l'autre pour inciter le personnel de santé à travailler dans les zones reculées. D'autre part, les tarifs unitaires peuvent varier dans le temps. Nous avons donc dû considérer que les indicateurs avaient la même valeur monétaire dans toutes les provinces : cette hypothèse est fautive et conduit à légèrement sous-estimer la valeur monétaire totale des quantités déclarées et vérifiées, mais elle a peu d'impact sur le pourcentage du budget qui a été économisé car elle affecte à la fois le montant déclaré et le montant vérifié.

Les résultats de ces calculs sont présentés dans le tableau 14.

	CS	Hôpitaux	Total
<b>Estimation du montant déclaré</b>	9 695 185	7 179 405	<b>16 874 590</b>
<b>Montant vérifié</b>	9 480 727	7 098 941	<b>16 579 668</b>
<b>Estimation du montant économisé</b>	214 458	80 464	<b>294 922</b>
<b>% économisé</b>	2%	1%	<b>2%</b>

Tableau 14 : Estimation des montants économisés grâce à la vérification quantitative (En US\$). Extrapolations pour l'année 2012 à partir des données disponibles : janvier-août dans les CS et janvier-juillet dans les hôpitaux

D'après nos calculs, la vérification quantitative a donc économisé 2% du montant attribué aux CS, et 1% du montant attribué aux hôpitaux. Ces pourcentages sont différents des pourcentages d'erreurs présentés dans les tableaux 6 et 7. Ceci se comprend aisément car dans ces deux tableaux, chaque déclaration avait le même poids, indépendamment du nombre de prestations qu'elle concernait et donc de sa valeur monétaire. Dans le tableau 14 en revanche, les résultats sont pondérés par la quantité et par le montant du subsidie accordé à chaque indicateur. Par exemple, si un centre de santé déclare avoir effectué 11 accouchements et que seulement 10 sont vérifiés par le CPVV, la surestimation est de 10%, exactement comme dans un centre de santé qui aurait déclaré 22 accouchements et pour lequel seulement 20 auraient été vérifiés. Mais les économies dues à la vérification quantitative dans le premier centre seront deux fois moins importantes que dans le second.

Les résultats des vérifications du FBP sont également utilisés pour sanctionner les prestataires. Ceci était peu le cas pendant les premiers mois de mise en œuvre du FBP au Burundi, mais les pénalités financières sont de plus en plus mises en application. Comme cela a été mentionné plus haut (tableau 3), toute différence de plus de 5% entre les quantités déclarées et les quantités vérifiées entraîne le retrait d'une partie des subsides que la formation sanitaire aurait dû recevoir pour l'indicateur concerné. Dans le cadre de ces sanctions, la différence qui est prise en compte n'est pas la différence entre la quantité telle qu'elle est auto-évaluée par les structures (appelée quantité « déclarée » dans les factures des structures) et la quantité telle qu'évaluée après la fin du processus de vérification (appelée quantité « validée » dans les factures), mais la différence entre la quantité « déclarée » et la quantité qui est rapportée dans les registres des structures, indépendamment des définitions exactes des indicateurs (appelée quantité « vérifiée » dans les factures)<sup>18</sup>. L'illustration 6

<sup>18</sup> Voir partie 3.2 pour plus de détails sur les quantités « déclarées », « vérifiées » et « validées ».

présente la proportion de déclarations qui, entre janvier et août 2012<sup>19</sup>, ont fait l'objet de sanctions (fondées sur la différence entre les quantités « déclarées » et « vérifiées »), mais aussi la proportion de déclarations qui auraient fait l'objet de sanctions si la différence entre les quantités « déclarées » et « validées » avaient été prises en compte.

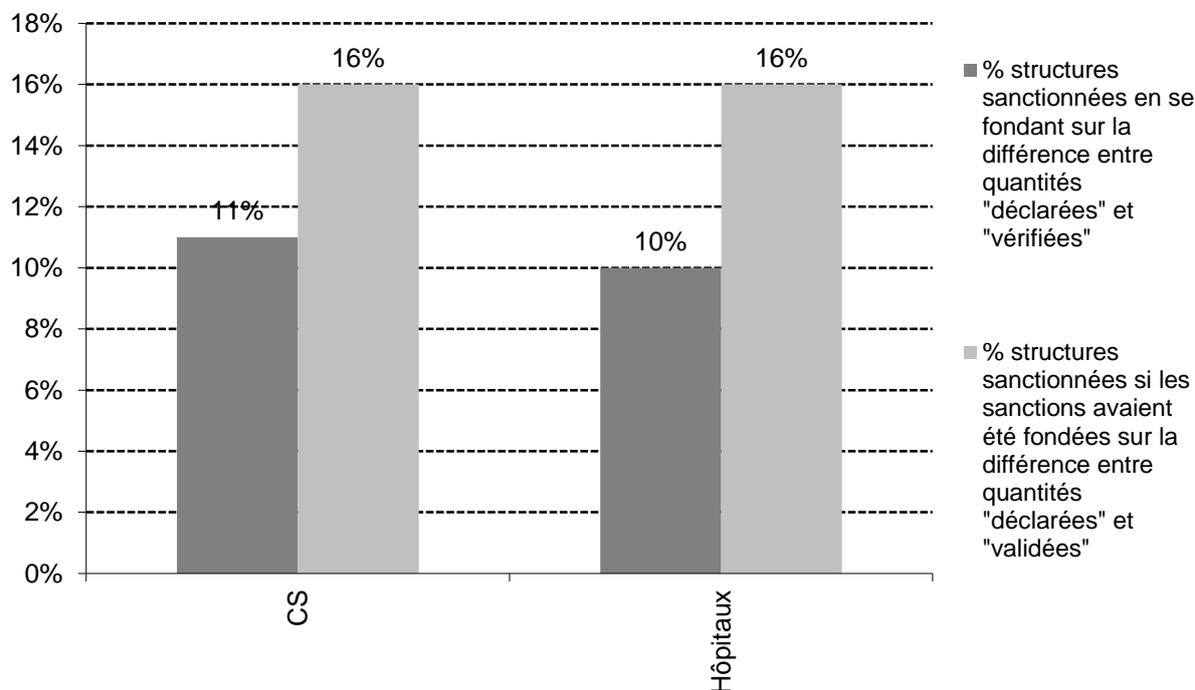


Illustration 6 : Proportion de déclarations qui ont fait l'objet de sanctions entre janvier et août 2012 (janvier et juillet 2012 pour les hôpitaux) car la différence entre les données « déclarées » et « vérifiées » était supérieure à 5%, comparée à la proportion de déclarations qui auraient fait l'objet de sanctions si la base des sanctions avait été la différence entre données « déclarées » et « validées »

Entre janvier et août 2012, 11% des déclarations effectuées par les CS et 10% des déclarations effectuées par les hôpitaux ont fait l'objet de sanctions car la différence entre « déclaration » et « vérification » était supérieure à 5%. Ces proportions auraient été largement supérieures si les données « validées » avaient été prises comme base de calcul pour ces sanctions en lieu et place des données « vérifiées » (16% des déclarations effectuées par les CS et les hôpitaux auraient alors été sanctionnées). Cette option aurait été plus pertinente du point de vue du FBP car les données « validées » sont celles qui déterminent le paiement. Il est donc souhaitable que le système de sanctions soit structuré à l'avenir de manière à prendre en compte les écarts entre données « déclarées » et « validées » au lieu des écarts entre données « déclarées » et « vérifiées ». Malheureusement, il n'a pas été possible de calculer les montants économisés par le programme FBP grâce à ces sanctions, ni celles qui auraient été réalisées si les données validées avaient été prises comme base de paiement : les bases de données disponibles ne disposaient pas de cette information.

Depuis le premier trimestre 2011, un malus est également instauré pour les structures dont le score de qualité globale (score de qualité technique et score pour l'enquête communautaire) est inférieur à 50%. Le tableau 15 présente le nombre de structures qui se sont trouvées dans ce cas de figure.

<sup>19</sup> Entre janvier et juillet 2012 pour les hôpitaux. De plus, pour les CS, les quantités vérifiées du mois de juin étaient manquantes, et ce mois n'a pas pu être pris en compte. Le nombre total de déclarations sur lesquelles porte l'illustration 6 est de 44 291 pour les CS et 6 762 pour les hôpitaux.

	CS		Hôpitaux	
	Nb.	Part	Nb.	Part
<b>2011 T1</b>	33	6%	1	2%
<b>2011 T2</b>	12	2%	1	2%
<b>2011 T3</b>	33	6%	1	2%
<b>2011 T4</b>	6	1%	0	0%
<b>2012 T1</b>	6	1%	9	18%
<b>2012 T2</b>	3	1%	2	4%
<b>Moyenne</b>	<b>16</b>	<b>3%</b>	<b>2</b>	<b>5%</b>

Tableau 15 : Nombre de formations sanitaires ayant été sanctionnées par un malus de qualité depuis le premier trimestre 2011 parmi l'ensemble des formations sanitaires contractualisées au Burundi

Avant de d'analyser ces sanctions, on peut noter qu'elles ne sont pas entièrement motivées par des informations erronées transmises par les formations sanitaires : 60% du score sur lequel elles sont basées évalue la qualité technique, et 20% évalue la satisfaction des utilisateurs. Seulement 20% évalue le fait que les patients reçoivent effectivement les services : la principale raison de ces sanctions est donc la sous-performance. Mais en l'absence d'information désagrégée sur les sanctions liées aux « patients-fantômes », on ne peut que fonder l'analyse sur les scores de qualité globale. En moyenne, chaque trimestre, 3% des structures ont vu une partie de leurs subsides retranchés car leur score de qualité globale était inférieur à 50%. Leur nombre est en constante diminution depuis le premier trimestre 2011, date de la mise en place de ce malus (sauf pour le premier trimestre 2012 pour les hôpitaux, date de l'introduction de la nouvelle grille de la qualité). En moyenne, le malus de qualité a représenté 2% du montant total de l'incitation à la qualité pour les CS, et 4% pour les hôpitaux (l'incitation à la qualité représente entre 10 et 15% de l'incitation totale). Ces montants pourraient varier si les sanctions liées aux « patients-fantômes » (tels que détectés par les enquêtes communautaires : environ 7-9% de patients non retrouvés dans les CS, 10-15% dans les hôpitaux) étaient désagrégées des sanctions liées à la qualité technique et à la satisfaction des patients.

De plus, comme cela a été mentionné plus haut, des sanctions devraient être prises contre les CPVV quand les résultats de la contre-vérification effectuée par HDP et les résultats de CPVV diffèrent de plus de 5%. Les modalités et les montants de ces sanctions sont précisément définies dans le manuel de procédures, mais il n'a pas été possible d'évaluer les montants qui auraient été économisés si elles avaient été mises en œuvre : comme expliqué plus haut (voir partie 4.4), les rapports de contre-vérification ne fournissent pas les différences entre la vérification et la contre-vérification au niveau des CS, ce qui rend impossible le calcul de sanctions hypothétiques.

En plus de ces sanctions qui touchent aux incitations financières des formations sanitaires, le système de vérification peut permettre de détecter des problèmes qui peuvent donner lieu à des sanctions administratives. Le tableau 16 montre le nombre de ruptures de contrats recensées en 2010 et 2011 pour les formations sanitaires bénéficiant d'un contrat principal, les formations sanitaires bénéficiant d'un contrat secondaire et les ASLO.

	2010	2011
<b>Ruptures de contrat avec les FOSA bénéficiant d'un contrat principal</b>	0	0
<b>Ruptures de contrat avec les FOSA bénéficiant d'un contrat secondaire</b>	2	1
<b>Ruptures de contrat avec les ASLO</b>	11	6

Tableau 16 : Nombre de ruptures de contrats recensées en 2010 et 2011 pour les formations sanitaires bénéficiant d'un contrat principal, les formations sanitaires bénéficiant d'un contrat secondaire et les ASLO

On remarque que les ruptures de contrat avec les formations sanitaires sont rares, et qu'elles sont même inexistantes en ce qui concerne les formations sanitaires bénéficiant d'un contrat principal. Au lieu de mettre en œuvre des sanctions coercitives, le système FBP du Burundi préfère accompagner les structures, de manière à ce qu'elles puissent améliorer leurs résultats. Le système peut notamment profiter du fait que les organes de régulation au niveau intermédiaire et opérationnel (BPS et BDS), en tant que membres du CPVV, sont fortement impliqués dans la validation des résultats, et peuvent agir pour résoudre les incohérences dans les données. Si des erreurs répétées sont remarquées, des recommandations sont effectuées par le CPVV, des investigations supplémentaires sont effectuées, et le BPS ou le BDS essaient de trouver un moyen, en collaboration avec la structure, d'améliorer la qualité des données. Par exemple, les BDS peuvent vérifier les registres durant leurs supervisions mensuelles si la sous-unité vérification du CPVV déclare avoir du mal à identifier les patients. Il faut souligner que la rupture des contrats peut se révéler problématique pour le système de santé dans son ensemble en termes d'accessibilité financière : le système FBP finance en effet également la gratuité des soins pour les femmes enceintes et les enfants. Si un contrat est rompu avec un prestataire, ce dernier devra facturer les soins à ces populations.

Notons que la politique envers les ASLO qui mettent en œuvre les enquêtes communautaires est différente, et que les ruptures de contrats sont plus fréquentes avec ces associations, généralement motivées par le comportement inadéquat des membres. Ceci s'explique par les plus faibles capacités de leurs membres, mais aussi par le fait que celles-ci ne sont pas impliquées dans la délivrance de soins de santé vitaux pour les populations : la résiliation de leur contrat est moins problématique, et il est plus facile de leur trouver un remplaçant. Les sanctions à leur encontre sont toutefois devenues moins fréquentes au cours de la seconde année de mise en œuvre, ce qui peut être dû à un effet d'apprentissage, ou à l'embauche de meilleures organisations.

En résumé, les sanctions ont encore un impact marginal dans le système FBP au Burundi :

- Les sanctions quantitatives ne concernent que 16% des déclarations, et dans la grande majorité des cas (92% des cas sanctionnés dans les CS, 94% dans les hôpitaux) il s'agit d'erreurs inférieures à 20%, donc ne donnant pas lieu à la suppression de l'ensemble des subsides pour l'indicateur concerné (comme expliqué dans le tableau 3) ;
- Seulement 3% des CS et 5% des hôpitaux subissent un malus de qualité chaque trimestre ;
- Les sanctions contre les CPVV ne sont pas appliquées ;
- Les sanctions administratives sont rares.

Ceci illustre une nouvelle fois le caractère interne plutôt qu'indépendant du système de vérification burundais, qui privilégie l'aspect pédagogique du processus par rapport à son aspect coercitif.

## 5. LES COÛTS DE LA VERIFICATION

Dans son rapport 2011 (MSPLS, 2012), la CT-FBP calcule que les paiements aux prestataires représentent 88% du coût total du FBP, mais son estimation de ce dernier omet un nombre de coûts importants : les salaires des fonctionnaires travaillant dans les CPVV et à la CT-FBP, l'assistance technique à la CT-FBP, les contributions des partenaires au fonctionnement des CPVV, l'amortissement des investissements, etc. Dans le tableau 17, nous avons tenté de reconstituer la plupart de ces coûts pour l'année 2011.

	US\$	%
<b>Subsides centres de santé</b>	10 300 743	45%
<b>Subsides hôpitaux</b>	7 675 419	33%
<b>Subsides BPS et BDS</b>	1 357 278	6%
<b>Total paiements aux prestataires</b>	<b>19 333 440</b>	<b>84%</b>
<b>CT-FBP (y.c. assistance technique)</b>	581 787	3%
<b>CPVV</b>	978 601	4%
<b>Contre-vérification</b>	281 108	1%
<b>Accompagnement<sup>20</sup></b>	1 393 443	6%
<b>Associations locales</b>	344 970	1%
<b>Renforcement des capacités</b>	113 697	0%
<b>Coûts de vérification</b>	<b>3 693 605</b>	<b>16%</b>
<b>Coût total</b>	<b>23 027 045</b>	<b>100%</b>

Tableau 17 : Estimation des coûts de vérification du FBP pour l'année 2011 (hors amortissement et une partie de l'assistance technique au niveau provincial)

Les coûts présentés ici montrent que les coûts de la vérification sont d'environ 16% des dépenses totales du programme. Il s'agit encore d'une approximation, car :

- Nous n'avons pas pu estimer l'amortissement des équipements ;
- Nous n'avons pas pu estimer le coût de l'assistance technique apporté par certains partenaires au niveau provincial ;
- Certains des coûts inclus dans les coûts de vérification ne sont pas entièrement consacrés à la vérification, mais n'ont pas pu être désagrégés (la CT-FBP est aussi en charge de la gestion du programme, les CPVV sont aussi en charge de la signature des contrats, etc.).

Malgré tout, ce chiffre est à comparer avec les expériences pilote du FBP, où les coûts de mise en œuvre représentaient entre 25 et 30% du budget total d'après les estimations de Cordaid. La généralisation du FBP semble donc avoir permis d'effectuer des économies d'échelle substantielles, mais les coûts peuvent encore être rationalisés. 6% du budget total (soit 38% du budget de la vérification) sont notamment actuellement consacré à l'accompagnement, c'est à dire au travail d'encadrement des CPVV au niveau provincial par les ONG qui ont mis en œuvre les projets pilotes du FBP. Etant donné les débats actuels sur les services effectivement rendus par ces ONG, leur rapport coût-efficacité et les utilisations alternatives de ce budget pourraient être étudiées.

On peut remarquer que le montant du budget de la vérification est largement supérieur aux montants qui ont été économisés grâce à la vérification quantitative (de l'ordre de 300 000 US\$ en 2012, cf. tableau 14) ou de la qualité. Cependant, la vérification a une vertu dissuasive, et l'on ne peut donc ni estimer correctement ni comparer directement les montants économisés aux montants dépensés pour effectuer la vérification. De plus, comme cela a été décrit plus haut (voir partie 3.1), la valeur de certaines sanctions n'a pas pu être calculée, et certaines d'entre elles ne sont pas encore appliquées, ce qui a pour effet de réduire les montants économisés. Enfin, la vérification contribue à l'exactitude du SIS et à ses utilisations ultérieures pour une prise de décision fondée sur les faits, ainsi qu'à une attention plus poussée aux résultats. La contribution de la vérification ne devrait donc pas seulement être trouvée dans les montants économisés.

<sup>20</sup> Hors assistance technique au niveau provincial effectuée par des partenaires tels que la coopération belge et l'Union Européenne, pour laquelle il n'a pas été possible de séparer ce qui relevait du FBP de ce qui relevait du renforcement du système de santé en général.

## 6. LEÇONS

La généralisation du FBP à l'ensemble du pays a été (et est toujours) un défi important pour le Burundi. En passant d'un système d'incitations basé sur les intrants à un système d'incitations basé sur les résultats, le FBP est en train de modifier progressivement l'ensemble du système de santé en accroissant l'autonomie des prestataires, la séparation des fonctions et la transparence.

L'existence même du mécanisme de vérification est une amélioration majeure de la gouvernance du système de santé. *La vérification a obligé les prestataires à s'ouvrir à des regards extérieurs* (bien que cela se passe souvent dans le cadre hiérarchique), ce qui est à la fois dans l'intérêt des prestataires et des patients. On aurait pu craindre que ce système soit ressenti comme un contrôle et mal accepté par les prestataires, mais les témoignages des équipes de vérification et des prestataires eux-mêmes montrent qu'il n'y a pas ou peu de tensions entre les prestataires et les vérificateurs : la vérification est généralement vue comme une incitation pour les prestataires à améliorer la quantité et la qualité des soins qu'ils fournissent.

Les équipes de vérification témoignent du fait que leur travail a provoqué des changements dans les pratiques de prestataires. Par exemple, *une plus grande attention est portée à la bonne gestion des registres* (bien que d'importants progrès restent à faire, comme cela a été décrit plus haut) : les prestataires savent que tout patient mal identifié dans les registres peut se traduire en une baisse de ses revenus (à travers les subsides de quantité ou le bonus de qualité). De plus, de nombreux acteurs du FBP estiment qu'au moins certains aspects de la qualité des soins se sont améliorés au Burundi depuis que le FBP a été généralisé (BM, Cordaid, OMS, UE, 2012), grâce à l'évaluation de la qualité technique et aux enquêtes communautaires (où la satisfaction des patients représente 50% du score).

Nous avons vu au cours de cette étude que le modèle de vérification original choisi par le Burundi, avec une grande participation des acteurs du MSPLS, a permis de *faire entrer dans le processus une grande part d'éducation des prestataires* : la régulation prend une part active au système de vérification (le BPS et les BDS sont des membres prééminents de la sous-unité validation du CPVV, et ils sont en charge de l'évaluation de la qualité technique), elle en tire un avantage non négligeable pour l'amélioration de ses propres performances. Le renforcement des capacités est privilégié par rapport à la sanction dans les modes de résolution des problèmes. Ceci permet une grande appropriation du processus de la part des organes de régulation comme de la part des prestataires.

Tout système de vérification dans le cadre du FBR nécessite un arbitrage entre les priorités de transparence, d'indépendance, d'intégration et de rigueur que nécessite (Ergo et Paina, 2012) : le Burundi a choisi une méthode originale qui repose sur une forte implication du MSPLS, mais qui donne également un certain rôle à la société civile. *La méthode de vérification est évolutive* : elle a d'abord mis l'accent sur l'accompagnement des formations sanitaires contractualisées plutôt que sur leur punition, puis elle s'est graduellement durcie et a introduit des sanctions plus fermes. Elle reste cependant fondée sur le dialogue davantage que sur la coercition.

Cela étant, de nombreux aspects de la vérification peuvent être améliorés au Burundi. L'analyse des résultats des vérifications montre que le choix en faveur d'une vérification intégrée au système de santé (par opposition à une vérification indépendante) s'accompagne d'une certaine tolérance : *le niveau d'erreur reste important*. En ce qui concerne la vérification quantitative par exemple, 69% des déclarations dans les CS et 62% dans les hôpitaux sont exactes si on les compare aux montants qui ont été finalement payés à l'issue du processus de vérification, mais cela laisse un tiers de déclarations inexactes. Dans les cas de surestimation, celle-ci est en moyenne supérieure à 10%, tandis que les sous-estimations sont supérieures à 5%. Les facteurs qui expliquent ces erreurs sont des problèmes de définition des indicateurs, d'identification des patients et de standardisation des outils de collecte : on pourrait s'attendre à ce qu'ils soient réglés après plus de deux ans de mise en œuvre du FBP dans le pays. Peu d'attention est portée aux données déclarées par les prestataires et à leurs différences avec les données vérifiées. Le système de sanctions mis en place à l'encontre des

déclarations qui différaient de plus de 5% des données vérifiées ne semble pas être assez sévère pour prévenir ces problèmes.

Il est urgent que le MSPLS (et notamment le SIS, en collaboration avec la CT-FBP), pour chacun des indicateurs concernés, se penche sur les causes précises des erreurs systématiques et propose des solutions (par exemple, effectuer une nomenclature des cas à prendre en compte pour certains indicateurs problématiques comme la petite chirurgie, modifier les registres dans lesquels les prestations sont consignées et dans certains cas créer des registres standards, etc.). *Un durcissement des sanctions financières contre les prestataires en cas de déclaration erronée pourrait également être envisagé*, notamment en les faisant porter sur les données « validées » et non sur les données « vérifiées » comme c'est le cas actuellement.

En ce qui concerne l'évaluation de la qualité, l'une des faiblesses majeures du montage burundais réside dans le fait que *l'évaluation de la qualité technique n'est pas mise en œuvre dans la totalité des formations sanitaires, et la recherche des patients dans la communauté ne comprend pas tous les indicateurs*. Tout d'abord, la qualité technique des sous-contractants, qui représentent plus de 20% des CS, n'est jamais évaluée. De plus, certains registres sont exclus des enquêtes communautaires pour des raisons de confidentialité : en conséquence, l'existence de « patients-fantômes » pour presque la moitié des indicateurs des CS, et un tiers des indicateurs des hôpitaux, n'est jamais vérifiée. Effectuer l'évaluation de la qualité dans toutes les structures et pour tous les indicateurs a un coût et pose des problèmes de confidentialité qui devront être pris en compte. Afin de réduire le coût de l'extension de l'évaluation de la qualité technique et de l'enquête communautaire à tous les indicateurs et à toutes les formations sanitaires, un tirage aléatoire des indicateurs et des structures pourrait être envisagé : toutes les structures et indicateurs auraient une chance d'être vérifiés, tous ne le seraient pas mais aucun ne serait systématiquement exclu.

Cette étude a également remarqué que *l'évaluation de la qualité mêle dans un seul système d'incitation des éléments de qualité technique, de satisfaction des patients et de détection de la fraude*. Ceci est justifié par le fait que les éléments liés à la satisfaction des patients et ceux liés à la détection de la fraude sont évalués par les mêmes acteurs et avec les mêmes moyens au cours du processus d'enquête communautaire. Mais le résultat est inapproprié, au moins pour deux raisons. Tout d'abord, il mélange la détection des « patients-fantômes » avec d'autres éléments, ce qui les rend moins visibles comme en témoigne le fait que la proportion de « patients-fantômes » n'est jamais analysée au niveau national. D'autre part, *ce système récompense les formations sanitaires pour ne pas frauder* : dans le montage actuel, si les patients sont retrouvés et confirment qu'ils ont bien reçu les soins, la formation sanitaire gagne des points et donc de l'argent. Il serait plus cohérent d'intégrer de lier la détection des « patients-fantômes » aux sanctions prévues lorsque la quantité déclarée est différente de la quantité vérifiée. Cependant, le système est actuellement en train de prendre une direction différente : il est prochainement prévu de faire passer la proportion relative de la « qualité technique » et de la « qualité perçue » de 60%/40% à 70%/30%. Cela réduira l'incitation financière positive à ne pas tricher, mais réduira aussi l'importance de la recherche des « patients-fantômes » dans l'ensemble du système de vérification.

Notre étude a aussi montré que le processus d'évaluation de la qualité technique est encore fortement vulnérable. *Il semble que les prestataires soient tentés de concentrer leurs efforts sur les aspects de la qualité qui sont mesurés, sans pour autant améliorer la qualité globale des soins qu'ils dispensent*. Deux éléments permettent d'illustrer ce constat. D'une part, le score de qualité technique des hôpitaux a fortement chuté au début de l'année 2012 à la suite de la modification de la grille d'évaluation de la qualité technique. Ceci semble montrer que les hôpitaux s'étaient habitués à un mode d'évaluation plutôt qu'à effectuer un effort réel sur la qualité. De telles modifications sont donc encore trop rares : il est important que les prestataires ne s'habituent pas à un mode d'évaluation routinier. Le niveau d'exigence doit au contraire augmenter avec les compétences des prestataires. Les modifications de la grille devraient aussi intervenir plus régulièrement pour les CS. Le principal défi en la matière consistera à faire entrer le patient dans l'évaluation, et notamment en

renforçant les éléments impliquant l'observation des soins. En effet, la vérification se concentre toujours beaucoup sur les conditions dans lesquelles les soins sont fournis, plutôt que sur les soins eux-mêmes : il n'y a pas de patient-mystère, par exemple. Dans tout le processus de vérification, la seule phase au cours de laquelle le patient est présent est l'enquête communautaire. Des progrès sont déjà en cours dans ce domaine, avec l'inclusion de plus d'éléments cliniques liés au dossier médical des patients dans la grille d'évaluation de la qualité, et ils sont à poursuivre. Rappelons que la mesure de la qualité technique est un sujet complexe, et qu'il ne peut y avoir de solutions toutes faites : de constantes expérimentations sont au contraire nécessaires.

Un autre point faible de l'évaluation de la qualité technique est *la difficulté, soulignée par plusieurs des personnes interrogées, à maintenir le niveau de qualité entre deux évaluations*. Cette difficulté tiendrait principalement au fait que les prestataires connaissent la période, et parfois la date précise, à laquelle ils vont être évalués, et sont donc en mesure de se préparer en conséquence. Une évolution possible dans ce domaine serait de ne plus annoncer en avance les visites d'évaluation de la qualité technique aux CS, et que celles-ci puissent se dérouler n'importe quand dans le trimestre plutôt que d'être concentrées sur une quinzaine donnée. Ceci modifierait le climat dans lequel l'évaluation de la qualité technique se déroule et elle y perdrait de son aspect éducatif, mais elle obligerait ainsi les prestataires à accroître leurs efforts pour le maintien de la qualité. Pour les hôpitaux, il semble difficile d'envisager le passage à des évaluations de la qualité inopinées sans renoncer au système d'évaluation par les pairs, car celui-ci nécessite intrinsèquement une planification. De plus, la taille du secteur hospitalier au Burundi est relativement réduite, le nombre d'acteur est limité, ce qui rend la conservation d'un secret difficile. Cependant, la crédibilité de l'évaluation de la qualité technique dans les hôpitaux doit être améliorée, comme le montrent les résultats de la contre-vérification dans ce domaine. Le système de vérification au Burundi doit donc trouver un compromis entre la rigueur du processus d'une part (évaluation par un tiers indépendant), et l'intégration de celui-ci au système de santé d'autre part (évaluation par les pairs).

La contre-vérification a apporté de nombreux éléments très positifs au système. En dépit de certains problèmes méthodologiques, elle a montré que la vérification de la quantité par les CPVV dans les CS était fiable dans son ensemble. Il faut cependant rappeler que la manière dont les données concernant les CS sont présentées dans les rapports de contre-vérification est problématique : en agrégeant les CS par province, la contre-vérification sous-estime le niveau d'erreur parce qu'elle fait la moyenne des sous et des surestimations. De plus, cette agrégation rend impossible l'analyse de l'évolution du niveau d'erreur à travers le temps, pour les CS qui auraient été aléatoirement sélectionnés pour plusieurs rounds de contre-vérification. *Des données désagrégées, avec séparation des sur et des sous-estimations sont nécessaires de manière à évaluer la fiabilité de la vérification des CPVV dans les CS*. De telles données permettraient également d'identifier les problèmes, de détecter les tendances dans les erreurs, et d'aider à l'amélioration de la performance.

*Les résultats de la contre-vérification sont encore sous-utilisés*. Ils montrent d'importants niveaux de discordance dans la vérification de la quantité pour les hôpitaux, mais cela n'a pas véritablement donné lieu à une action contre les CPVV ou les hôpitaux à la mesure de ce qui serait nécessaire. Ils ont aussi montré d'importants niveaux de discordance entre les évaluations de la qualité technique effectuées par les BPS, les BPS et les pairs et celles effectuées par HDP, mais là aussi, cela n'a été suivi d'aucune sanction. La mise en œuvre de sanctions liées au processus de contre-vérification contre les CPVV, les BPS et les BDS pourrait grandement améliorer la qualité des données fournies par le système FBP.

D'autre part, il semblerait pertinent de clarifier le rôle de la contre-vérification. Celle-ci assume aujourd'hui un double rôle d'évaluation de la qualité et données et de suivi du processus de mise en œuvre du FBP, et effectue des recommandations visant à la fois à améliorer la qualité des données et le système FBP lui-même. Le fait que la contre-vérification effectue des enquêtes communautaires, mais ne compare pas ses résultats avec l'enquête communautaire effectuée sous la supervision du CPVV, est une illustration éloquent de cette ambiguïté : ne pas comparer les résultats est justifié

pour des raisons méthodologiques (les échantillons ne sont pas les mêmes), mais pourquoi dans ce cas la contre-vérification effectue-t-elle des enquêtes communautaires ? Il serait préférable que la contre-vérification se concentre sur la qualité des données, et non sur la qualité du système : le MSPLS bénéficie en effet d'une forte assistance technique au niveau central qui lui permet d'assurer un suivi du processus satisfaisant. En revanche, une attention poussée à la qualité des données est primordiale, et l'œil externe que constitue la contre-vérification peut aider dans ce sens.

Au-delà des pistes qui viennent d'être suggérées pour améliorer le système de vérification à court terme, il existe de nombreuses possibilités de développement à moyen terme.

*La prochaine étape pour la vérification quantitative serait par exemple de procéder par vérification aléatoire sur un échantillon de formations sanitaires au lieu de procéder à la vérification systématique mensuelle de toutes les formations sanitaires.* En effet, cette dernière est très coûteuse en temps et en ressources qui pourraient être employées à meilleur escient. Les formations sanitaires seraient ainsi payées sur la base de leur déclaration, et non sur la base de la vérification systématique, comme c'est le cas dans de nombreux systèmes de vérification dans les pays à niveau de revenu plus élevé. Etant donné le haut niveau de discordance entre les données déclarées par les formations sanitaires et les données vérifiées qui a été constaté dans cette étude, cela pourrait ne pas être possible dans l'immédiat. Une fois que les discordances auront été réduites, il conviendra naturellement d'être prudent dans l'expérimentation de cette vérification par échantillonnage, et de la mettre en œuvre d'abord dans une zone pilote avant de la généraliser à tout le pays : le Burundi n'a probablement pas encore atteint un niveau de confiance dans ses institutions pour qu'une telle réforme puisse être effectuée rapidement.

De l'avis de l'ensemble des acteurs interrogés, le système de vérification de la performance dans le FBP au Burundi permet un niveau satisfaisant de détection de la fraude. Ce système, notamment grâce au niveau de confiance que permettent une forte intégration au sein des organes de régulation du MSPLS d'une part, et l'existence d'un système de contre-vérification indépendant d'autre part, produit des données que la plupart des acteurs jugent satisfaisantes, et ce à un coût acceptable (environ 16% des dépenses totales du FBP). Cette étude a montré que le niveau de discordance entre les résultats avant et après la vérification était encore élevé, mais ceci peut être résolu par la modification des incitations d'une part (par exemple par la mise en œuvre de sanctions plus strictes), et par la formation continue des CPVV, des BPS, des BDS et des pairs hospitaliers d'autre part. Le défi majeur consistera donc dans le futur à continuer à améliorer graduellement ce système de vérification pour éviter qu'il ne se transforme en un exercice de routine, et pour qu'il continue à bénéficier de la confiance des différents acteurs amenés à utiliser les données qu'il produit.

## ANNEXES

### Annexe 1 : Bibliographie

Banque Mondiale, Cordaid, Organisation Mondiale de la Santé, Union Européenne, *Rapport de la 2<sup>e</sup> revue externe du Financement Basé sur la Performance*, 2012

Banque Mondiale, *World Development Indicators Database*, 2011

Cordaid et Ministère de la Santé Publique et de la Lutte contre le Sida, *Résultats de l'enquête ménage et de l'enquête qualité de base, 2006-208-2010*, Bujumbura, 2011

Ergo, Alex et Ligia Paina, *Verification in Performance-Based Incentive Schemes*, Bethesda, MD: Health Systems 20/20, Abt Associates Inc., August 2012

Ministère de la Santé Publique et de la Lutte contre le Sida, *Manuel des procédures pour la mise en œuvre du Financement Basé sur la Performance au Burundi, version révisée*, Bujumbura, 2011b

Ministère de la Santé Publique et de la Lutte contre le Sida, *Manuel des procédures pour la mise en œuvre du Financement Basé sur la Performance au Burundi, version originale*, Bujumbura, 2010

Ministère de la Santé Publique et de la Lutte contre le Sida, *Rapport de mise en œuvre du FBP et de la gratuité des soins pour l'année 2011*, Bujumbura, 2012

Ministère de la Santé Publique et de la Lutte contre le Sida, *Rapport de la mise en œuvre du Financement Basé sur la Performance au Burundi, avril 2010 à mars 2011*, Bujumbura, 2011a

Programme des Nations Unies pour le Développement, *World development Report 2011 – Sustainability and equity: a better future for all*, 2011

### Annexe 2 : Formulaires de déclaration de la performance dans les CS et les hôpitaux



CS



Hôpitaux

Source : MSPLS, 2011b

### Annexe 3 : Factures mensuelles pour les indicateurs des CS et des hôpitaux



CS



Hôpitaux

Source : MSPLS, 2011b

### Annexe 4 : Check-lists d'évaluation de la qualité technique des CS et des hôpitaux



CS ancienne version



Hôpitaux ancienne version

Source : MSPLS, 2010



CS nouvelle version



Hôpitaux nouvelle  
version

Source : MSPLS, 2011b

## Annexe 5 : Questionnaire pour les enquêtes communautaires

En kirundi



Enquête  
communautaire

Source : MSPLS, 2011b

## Annexe 6 : Cadre d'analyse pour les études de cas

En anglais



Cadre d'analyse