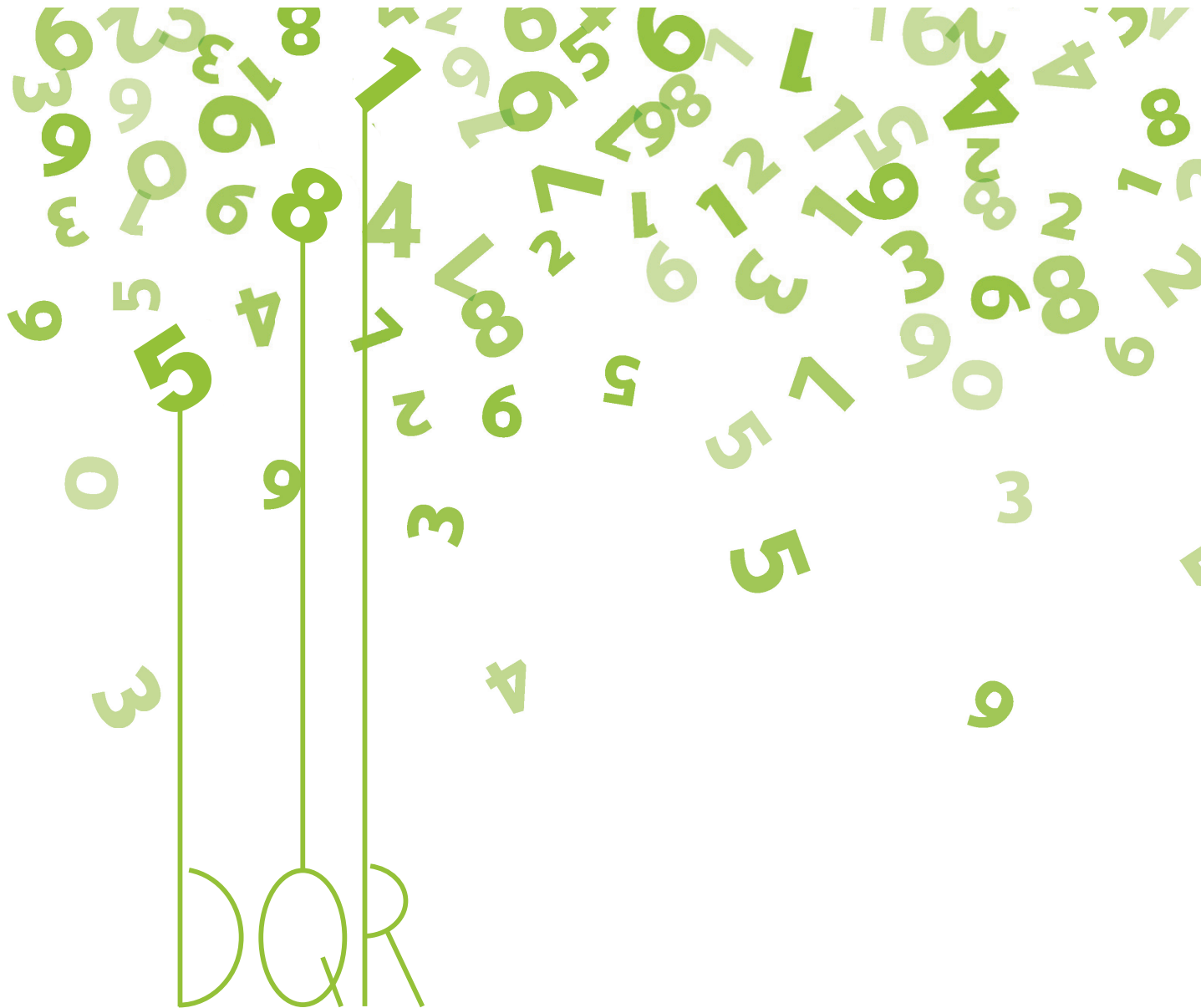




DATA QUALITY REVIEW (EXAMEN DE LA QUALITÉ DES DONNÉES)

Module 3 Vérification des données et évaluation des systèmes

DATA QUALITY REVIEW (EXAMEN DE LA QUALITÉ DES DONNÉES)

Module 3 Vérification des données et évaluation des systèmes

Data quality review (examen de la qualité des données) : une boîte à outils pour l'évaluation de la qualité des données des établissements de santé. Module 3. Vérification des données et évaluation des systèmes [Data quality review: a toolkit for facility data quality assessment. Module 3. Data verification and system assessment]

ISBN 978-92-4-251274-8

© **Organisation mondiale de la Santé 2019**

Certains droits réservés. La présente publication est disponible sous la licence Creative Commons Attribution – Pas d'utilisation commerciale – Partage dans les mêmes conditions 3.0 IGO (CC BY-NC-SA 3.0 IGO ; <https://creativecommons.org/licenses/by-nc-sa/3.0/igo>).

Aux termes de cette licence, vous pouvez copier, distribuer et adapter l'oeuvre à des fins non commerciales, pour autant que l'oeuvre soit citée de manière appropriée, comme il est indiqué ci-dessous. Dans l'utilisation qui sera faite de l'oeuvre, quelle qu'elle soit, il ne devra pas être suggéré que l'OMS approuve une organisation, des produits ou des services particuliers. L'utilisation de l'emblème de l'OMS est interdite. Si vous adaptez cette oeuvre, vous êtes tenu de diffuser toute nouvelle oeuvre sous la même licence Creative Commons ou sous une licence équivalente. Si vous traduisez cette oeuvre, il vous est demandé d'ajouter la clause de non responsabilité suivante à la citation suggérée : « La présente traduction n'a pas été établie par l'Organisation mondiale de la Santé (OMS). L'OMS ne saurait être tenue pour responsable du contenu ou de l'exactitude de la présente traduction. L'édition originale anglaise est l'édition authentique qui fait foi ».

Toute médiation relative à un différend survenu dans le cadre de la licence sera menée conformément au Règlement de médiation de l'Organisation mondiale de la propriété intellectuelle.

Citation suggérée. Data quality review (examen de la qualité des données) : une boîte à outils pour l'évaluation de la qualité des données des établissements de santé. Module 3. Vérification des données et évaluation des systèmes [Data quality review: a toolkit for facility data quality assessment. Module 3. Data verification and system assessment]. Genève : Organisation mondiale de la Santé ; 2019. Licence : CC BY-NC-SA 3.0 IGO.

Catalogage à la source. Disponible à l'adresse <http://apps.who.int/iris>.

Ventes, droits et licences. Pour acheter les publications de l'OMS, voir <http://apps.who.int/bookorders>. Pour soumettre une demande en vue d'un usage commercial ou une demande concernant les droits et licences, voir <http://www.who.int/about/licensing>.

Matériel attribué à des tiers. Si vous souhaitez réutiliser du matériel figurant dans la présente oeuvre qui est attribué à un tiers, tel que des tableaux, figures ou images, il vous appartient de déterminer si une permission doit être obtenue pour un tel usage et d'obtenir cette permission du titulaire du droit d'auteur. L'utilisateur s'expose seul au risque de plaintes résultant d'une infraction au droit d'auteur dont est titulaire un tiers sur un élément de la présente oeuvre.

Clause générale de non responsabilité. Les appellations employées dans la présente publication et la présentation des données qui y figurent n'impliquent de la part de l'OMS aucune prise de position quant au statut juridique des pays, territoires, villes ou zones, ou de leurs autorités, ni quant au tracé de leurs frontières ou limites. Les traits discontinus formés d'une succession de points ou de tirets sur les cartes représentent des frontières approximatives dont le tracé peut ne pas avoir fait l'objet d'un accord définitif.

La mention de firmes et de produits commerciaux ne signifie pas que ces firmes et ces produits commerciaux sont agréés ou recommandés par l'OMS, de préférence à d'autres de nature analogue. Sauf erreur ou omission, une majuscule initiale indique qu'il s'agit d'un nom déposé.

L'Organisation mondiale de la Santé a pris toutes les précautions raisonnables pour vérifier les informations contenues dans la présente publication. Toutefois, le matériel publié est diffusé sans aucune garantie, expresse ou implicite. La responsabilité de l'interprétation et de l'utilisation dudit matériel incombe au lecteur. En aucun cas, l'OMS ne saurait être tenue responsable des préjudices subis du fait de son utilisation.

Conception et mise en page graphique : L'IV Com Sàrl, Villars-sous-Yens, Suisse.

Imprimé par le Service de production des documents de l'OMS, Genève, Suisse.

Table des matières

Remerciements	2
Abréviations	4
3.1 Aperçu	5
Enquête en établissement	5
Indicateurs clés	5
Dimensions de la qualité des données	6
3.2 Mise en œuvre de l'enquête en établissement dans le cadre de la procédure d'examen de la qualité des données	9
Préparation et mise en œuvre de l'enquête en établissement	9
3.3 Analyse et interprétation	12
Paramètres de mesure de la qualité des données recueillis lors de l'enquête en établissement	12
Analyse des paramètres de qualité des données et d'autres mesures recueillies dans le cadre des enquêtes en établissement	13
Évaluation des systèmes	24
Annexe 1: Indicateurs recommandés	27
Annexe 2: Calcul des paramètres de qualité des données dans le cadre de l'enquête en établissement	29
Annexe 3: Calcul des indicateurs « traceurs » du système qui sont susceptibles d'influer sur la qualité des données	33
Annexe 4: Documents sources recommandés et vérification des données par recoupement ou contrôle ponctuel	37
Annexe 5: Méthodes et problèmes d'échantillonnage	39
Annexe 6: Instruments de collecte des données et outils d'analyse	47



Remerciements

Cette boîte à outils, qui est le fruit d'une collaboration entre l'Organisation mondiale de la Santé, le Fonds mondial, l'Alliance GAVI et l'USAID/MEASURE Evaluation, propose une approche unifiée de la qualité des données. Elle intègre et complète des outils et méthodes passés et actuels d'évaluation de la qualité des données dans les établissements, en s'appuyant sur les meilleures pratiques et les leçons tirées de l'expérience de nombreux pays.

Kavitha Viswanathan a supervisé les aspects techniques liés à la mise au point de ce manuel, sous la direction générale de Kathryn O'Neill. Ties Boerma, David Boone, Robert Pond, Olav Poppe, Claire Preaud, Ashley Sheffel, Amani Siyam et Marina Takane ont apporté des contributions techniques.

Tous les modules de cette boîte à outils s'appuient sur les contributions techniques de différents départements et programmes de l'OMS et d'autres organismes. Nous tenons en particulier à citer les personnes suivantes : Thomas Cherian, Marta Gacic-Dobo, Jan Van Grevendonk ; Richard Cibulski, Michael Lynch ; Katherine Floyd, Philippe Glaziou, Hazim Timimi ; Isabelle Bergeri, Txema Callejas, Chika Hayashi, Serge Xueref ; Ryuichi Komatsu, John Puvimanasinghe, Annie Schwartz, Alka Singh, Nathalie Zorzi ; Peter Hansen, Chung-won Lee.

Nous tenons à remercier tout particulièrement les pays partenaires qui ont testé différents éléments de cette approche et nous ont fait part de leurs précieux commentaires (Bénin, Burkina Faso, Cambodge, Kenya, Mauritanie, Ouganda, République démocratique du Congo, Sierra Leone, Togo, Zambie et Zanzibar, en République-Unie de Tanzanie).

Cette boîte à outils a été mise au point grâce à des subventions de l'initiative Data for Health de Bloomberg Philanthropies ; de l'Alliance GAVI ; du Fonds mondial de lutte contre le sida, la tuberculose et le paludisme ; et de l'Agence norvégienne de coopération pour le développement (Norad).

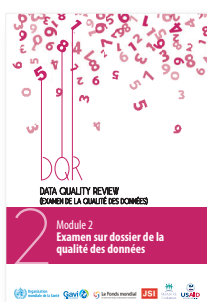


Contenu de la boîte à outils

La boîte à outils pour l'examen de la qualité des données contient des lignes directrices et d'autres ressources. Les lignes directrices sont présentées dans les trois modules suivants. Des ressources supplémentaires relatives à la collecte et à l'analyse des données seront publiées en ligne à des fins de téléchargement. Des informations plus détaillées sur ces ressources supplémentaires sont fournies dans le Module 1, Cadre et paramètres.



Module 1
Cadre et
paramètres



Module 2
Examen sur dossier
de la qualité des
données

DOCUMENT ACTUEL



Module 3
Vérification des
données et évaluation
des systèmes

Abréviations

ANC	soins prénatals
ANC1	première consultation prénatale
AT	vaccin à base d'anatoxine tétanique
CSPPro	Census and survey processing system
DHIS 2	logiciel open source en ligne que les pays utilisent principalement comme système d'information sanitaire pour la gestion des données et le suivi des programmes de santé
DTC	vaccin antidiphtérique-antitétanique-anticoquelucheux
DTC3	trois doses de vaccin antidiphtérique-antitétanique-anticoquelucheux
Fonds mondial	Fonds mondial de lutte contre le sida, la tuberculose et le paludisme
FV	facteur de vérification
HMIS	système d'information pour la gestion sanitaire (health management information system)
MCV	vaccin à valence rougeole
OMS	Organisation mondiale de la Santé
Penta	vaccin pentavalent
PTME	prévention de la transmission mère-enfant
SARA	évaluation de la disponibilité et de la capacité opérationnelle des services
TAR	traitement antirétroviral
TDR	test de diagnostic rapide
TPI3	troisième dose de traitement préventif intermittent
TPIp	traitement préventif intermittent pendant la grossesse
USAID	United States Agency for International Development
VIH	virus de l'immunodéficience humaine
VPC	vaccin antipneumococcique conjugué



3.1 Aperçu

Enquête en établissement

La mesure de la qualité des données au moyen d'une enquête en établissement offre une occasion unique de vérifier la qualité des données dans un échantillon d'établissements de santé choisis de manière aléatoire. Les résultats obtenus peuvent être comparés à ceux de l'examen sur dossier réalisé dans la cadre de la procédure d'examen de la qualité des données. On trouvera ci-dessous une description de l'analyse à effectuer et des résultats recommandés pour les indicateurs de qualité des données lors de l'enquête en établissement. L'enquête étant fondée sur un échantillon représentatif d'établissements de santé, un facteur de pondération approprié devra être appliqué pour obtenir des estimations correctes. L'annexe 5 contient des informations détaillées sur cette pondération.

En choisissant un échantillon d'établissements et en pondérant les résultats observés lors de l'enquête, on peut calculer une valeur moyenne nationale des paramètres de qualité des données (pour les indicateurs programmatiques sélectionnés) qui soit représentative de tous les établissements de santé du pays. Il ne faut cependant pas oublier que ces valeurs moyennes peuvent masquer des variations dues aux caractéristiques de chaque établissement de santé, comme l'autorité de gestion (établissement public ou privé à but lucratif), le type d'établissement (par exemple, hôpital, centre de santé ou dispensaire) et la région géographique. C'est pourquoi il peut être nécessaire de mener une analyse stratifiée (c'est-à-dire ventilée) afin d'établir une estimation pour chaque grande catégorie de la caractéristique prise en compte (strate). Les strates proposées comprennent le type d'établissement, l'autorité de gestion et la région géographique, bien qu'elles ne soient pas toutes nécessairement pertinentes pour chaque enquête. La stratification de l'échantillon a également pour effet d'accroître la taille de l'échantillon.

Indicateurs clés

Les indicateurs clés proposés pour l'enquête en établissement sont les mêmes que ceux de l'examen sur dossier. Dans l'idéal, les paramètres calculés dans l'enquête en établissement et dans l'examen sur dossier dresseront un portrait global de la qualité des données et des systèmes et permettront la mise en place de mécanismes solides d'amélioration. Ces indicateurs clés sont présentés dans le Tableau 3.1.

Il est recommandé aux pays d'adopter des indicateurs figurant sur cette liste de base, mais ils peuvent également en choisir d'autres ou élargir la sélection selon leurs besoins et les ressources dont ils disposent. Une liste complète d'indicateurs clés et d'indicateurs supplémentaires est fournie en annexe 1.

Tableau 3.1 Indicateurs clés recommandés pour la procédure d'examen de la qualité des données.

Indicateurs recommandés pour l'examen de la qualité des données		
Domaine programmatique	Nom de l'indicateur	Définition complète de l'indicateur
Santé maternelle	Couverture de la 1 ^{ère} consultation prénatale (ANC1)	Nombre et pourcentage de femmes enceintes ayant bénéficié d'au moins une consultation pendant la grossesse
Vaccination	Couverture de la vaccination DTC3/ Penta3	Nombre et pourcentage d'enfants de <1 an ayant reçu trois doses de vaccin DTC/pentavalent
VIH	Actuellement sous TAR	Nombre et pourcentage de personnes vivant avec le VIH qui sont actuellement sous TAR
Tuberculose	Taux de notification de la tuberculose	Nombre de nouveaux cas et de cas de rechute de la tuberculose qui sont notifiés pour 100 000 habitants
Paludisme	Cas confirmés de paludisme ¹	Nombre de cas de paludisme confirmés (par examen microscopique ou TDR) pour 1000 personnes par an

Remarque : ANC = soins prénatals ; DTC3 = 3 doses de vaccin antidiphthérique-antitétanique-anticoquelucheux ; Penta = vaccin pentavalent ; TAR = traitement antirétroviral ; TDR = test de diagnostic rapide.

Dimensions de la qualité des données

Le cadre d'examen de la qualité des données étudie chaque indicateur choisi selon quatre perspectives, ou dimensions :

- Dimension 1 : Exhaustivité et ponctualité des données ;
- Dimension 2 : Cohérence interne des données notifiées ;
- Dimension 3 : Cohérence externe (c'est-à-dire concordance avec d'autres sources de données, comme les enquêtes) ;
- Dimension 4 : Comparaison externe des données sur la population (examen des dénominateurs utilisés pour calculer les taux relatifs aux indicateurs de performance).

Exhaustivité et ponctualité

L'exhaustivité des données est évaluée en observant si toutes les entités censées notifier des données le font effectivement. Cela vaut pour les rapports transmis par les établissements de santé aux districts, ainsi que pour ceux communiqués par les districts aux niveaux régional et provincial. Pour évaluer la ponctualité des données, on détermine si les entités qui ont soumis des données l'ont fait avant une date butoir donnée. Les paramètres relatifs à l'exhaustivité et à la ponctualité utilisés dans le cadre de la procédure d'examen de la qualité des données sont les suivants :

- **Exhaustivité et ponctualité de la notification au niveau des districts** – Ces paramètres mesurent les performances des districts en matière d'exhaustivité et de ponctualité de la notification.

¹ Si le nombre de cas confirmés de paludisme n'est pas recueilli, on peut lui substituer le nombre total de cas de paludisme.

- **Exhaustivité et ponctualité de la notification au niveau des établissements** – Ces paramètres mesurent les performances des établissements en matière d'exhaustivité et de ponctualité de la notification.
- **Exhaustivité des données relatives aux indicateurs (données individuelles)** – Cet indicateur évalue dans quelle mesure les établissements supposés notifier des données sur les indicateurs clés choisissent le font effectivement. Cet aspect est à distinguer de l'exhaustivité globale de la notification en cela qu'il s'agit d'examiner l'exhaustivité de données individuelles spécifiques, et pas seulement de vérifier que le formulaire mensuel de notification a bien été reçu.
- **Stabilité de l'exhaustivité de la notification** – Cet indicateur examine les tendances en matière d'exhaustivité de la notification.

Cohérence interne des données notifiées

Dans cette dimension, on s'intéresse à la cohérence interne des données évaluées. Les paramètres correspondants portent sur les aspects suivants : 1) cohérence des mêmes données à différents moments ; 2) cohérence entre des données apparentées ; et 3) comparaison des données entre les documents sources et les bases de données nationales.

La procédure d'examen de la qualité des données comprend quatre paramètres de cohérence interne :

- **Présence de données atypiques (outliers)** : Il s'agit de voir si une valeur est extrême par rapport aux autres valeurs de la même série de données.
- **Cohérence dans le temps** : La plausibilité des résultats notifiés pour les indicateurs programmatisés sélectionnés est examinée en se fondant sur l'historique de la notification de ces indicateurs. Les tendances sont évaluées en vue de déterminer si les valeurs signalées sont extrêmes par rapport aux autres valeurs notifiées pendant l'année écoulée ou sur plusieurs années.
- **Cohérence entre les indicateurs** : Les indicateurs programmatisés liés par une relation prévisible sont examinés pour déterminer si la relation attendue entre ces deux indicateurs est bien observée. En d'autres termes, il s'agit d'examiner si le lien observé entre les indicateurs, tel qu'il ressort des données notifiées, correspond à la relation attendue.
- **Cohérence entre les données notifiées et les documents d'origine** : Il s'agit d'évaluer l'exactitude de la notification pour les indicateurs concernés en passant en revue les documents sources dans les établissements de santé. Cet élément de la cohérence interne est mesuré au moyen d'un exercice de vérification des données qui impose de conduire un examen des registres dans un échantillon d'établissements de santé. C'est la seule dimension de la qualité des données qui suppose de recueillir des données primaires supplémentaires.

Cohérence externe avec d'autres sources de données

On évalue ici le niveau de concordance entre deux sources de données mesurant le même indicateur de santé. Les deux sources généralement comparées sont, d'une part, les données qui transitent par le système d'information pour la gestion sanitaire (HMIS) ou le système d'information propre à chaque programme et, de l'autre, les données issues des enquêtes en population périodiques. Les données du système HMIS peuvent également être comparées aux données des registres de pharmacie ou à d'autres types de données pour vérifier qu'elles se situent dans une fourchette comparable.

Comparaison externe des données sur la population

L'objectif est de déterminer l'adéquation des données sur la population employées pour évaluer la performance des indicateurs sanitaires. Les données sur la population servent de dénominateur dans le calcul des taux ou des proportions et fournissent des informations importantes sur la couverture. Cette mesure de la qualité des données consiste à comparer deux sources différentes d'estimations de la population (dont les valeurs sont calculées de manière différente) afin de vérifier le degré de concordance de ces deux sources. Si les deux estimations de la population divergent, les estimations de la couverture, pour un indicateur donné, pourront avoir des valeurs très différentes, même si le résultat programmatique (c'est-à-dire le nombre d'événements) est identique. Plus le niveau de convergence est élevé entre les dénominateurs de différentes sources, plus il est probable que les valeurs obtenues correspondent à la valeur réelle de la population.



3.2 Mise en œuvre de l'enquête en établissement dans le cadre de la procédure d'examen de la qualité des données

Préparation et mise en œuvre de l'enquête en établissement

Exigences relatives à la vérification des données et à l'évaluation des systèmes

Les documents sources recommandés et les contrôles de vérification par recoupement des données sont répertoriés dans les annexes 4 et 5.

Échantillonnage des établissements de santé

Un échantillon représentatif d'établissements de santé doit être choisi pour la vérification des données et l'application du module d'évaluation des systèmes. Il est indispensable de disposer d'une « liste centrale des établissements », ou d'une liste des établissements de santé comprenant des informations sur leurs caractéristiques (par exemple, autorité de gestion, type d'établissement, emplacement en termes de région et de district), pour procéder à la vérification des données et à l'évaluation des systèmes dans le cadre de la procédure d'examen de la qualité des données. Une fois que les objectifs de l'examen de la qualité des données ont été définis, la méthode d'échantillonnage peut être élaborée. Par exemple, certaines évaluations en établissement, comme l'évaluation de la disponibilité et de la capacité opérationnelle des services (SARA), reposent habituellement sur un échantillonnage par liste et/ou par zone, tandis que d'autres évaluations de la qualité des données utilisent une méthode modifiée d'échantillonnage en grappes à deux degrés. S'il est nécessaire d'établir des estimations régionales de l'exactitude des données ou des estimations spécifiques pour certains types d'établissements (par exemple, selon l'autorité de gestion ou le type d'établissement), la méthode d'échantillonnage devra tenir compte de ces impératifs. Certains services spécialisés (comme le diagnostic et le traitement de la tuberculose ou le dépistage et le traitement de l'infection à VIH) ne sont pas offerts dans tous les établissements. Il pourra donc être nécessaire d'ajuster l'échantillonnage si les indicateurs relatifs à ces domaines programmatiques doivent être évalués. Les exigences techniques liées à la sélection de l'échantillon et au calcul des estimations à partir des données obtenues ne sont pas négligeables. Lors de l'élaboration de la méthode d'échantillonnage, il convient d'être attentif aux exigences propres à chaque pays. Il est recommandé de consulter un statisticien pour veiller à ce que l'échantillon soit convenablement sélectionné. L'annexe 5 fournit des informations complémentaires sur l'échantillonnage des établissements de santé aux fins de la vérification des données et de l'évaluation des systèmes dans le cadre de l'examen de la qualité des données.

Identification, adaptation et reproduction des outils d'enquête (sur support papier et/ou électronique)

Des outils standardisés de vérification des données et d'évaluation des systèmes ont été mis au point pour aider les pays à mener l'examen de la qualité des données au niveau des établissements de santé et des districts. Ils ont été élaborés à titre de modules de la boîte à outils SARA, mais peuvent être utilisés comme outils autonomes lorsque le principal objectif est d'évaluer la qualité des données.

Il convient que ces outils soient adaptés au contexte national avant leur mise en application (par exemple, selon les domaines programmatiques spécifiques, les indicateurs et les documents sources). S'il est prévu que les données soient saisies par voie électronique (par exemple sur un ordinateur tablette), une base de données doit être établie pour faciliter la saisie. Il convient que cette base soit prérenseignée avec la liste des établissements de santé choisis lors de l'échantillonnage et que son contenu soit disponible sur les tablettes sur le terrain. Les modules de vérification des données et d'évaluation des systèmes ont été mis au point sous CSPro 6.2 et peuvent être obtenus auprès de l'OMS. À l'instar des outils d'enquête sur support papier, ces modules devront être adaptés au contexte national avant la mise en œuvre de l'examen de la qualité des données.

Formation des agents sur le terrain (enquêteurs)

Les agents de terrain chargés de réaliser l'enquête en établissement doivent être formés aux méthodes de vérification des données et d'évaluation des systèmes. La vérification des données dans plusieurs domaines programmatiques exige une bonne connaissance des différents outils de collecte des données employés (registres, dossiers de patients, feuilles de décompte, etc.) selon l'indicateur et le domaine programmatique. Dans l'idéal, les enquêteurs auront de l'expérience en matière d'enregistrement des données de santé publique, ainsi qu'une bonne connaissance des outils de collecte des données utilisés sur le terrain. La formation des enquêteurs doit comprendre des exercices leur permettant de s'entraîner à compiler les indicateurs relatifs à chaque domaine programmatique au moyen des outils qu'ils sont susceptibles de rencontrer sur le terrain.

Aviser les sites et les autorités infranationales

Plusieurs semaines avant la mise en œuvre, les établissements échantillonnés pour l'examen de la qualité des données devront être avisés de la date de la visite des équipes d'évaluation. Les personnels compétents chargés de la gestion des données, ainsi que leurs superviseurs, devront être présents dans l'établissement le jour de la visite afin de faciliter l'accès aux registres pertinents, de répondre aux questions relatives à l'évaluation des systèmes et de contribuer à la réalisation de l'enquête au sein de l'établissement. De même, les autorités infranationales de



gestion du système HMIS, notamment les gestionnaires HMIS travaillant au niveau du district et/ou de la région, devront également être informés, à la fois pour satisfaire aux protocoles administratifs éventuels et pour solliciter leur soutien ou leur coopération dans la réalisation de l'enquête.

Réalisation de l'enquête dans l'établissement de santé

Il convient que les équipes d'évaluation travaillent en binôme pour une efficacité optimale et un contrôle adéquat de la qualité lors des visites dans les établissements. Il est recommandé de faire porter la vérification des données sur un maximum de cinq indicateurs (un par domaine programmatique). Les équipes devront prévoir de passer une journée complète dans chaque établissement si les activités de vérification des données et d'évaluation des systèmes sont menées conjointement à une autre évaluation existante de l'établissement, comme l'évaluation de la disponibilité et de la capacité opérationnelle des services (SARA). Si seuls les modules de vérification des données et d'évaluation des systèmes sont mis en œuvre, il faudra consacrer au moins une demi-journée à la collecte des données ; cependant, la réalisation de l'enquête pourrait prendre plus de temps, en particulier dans les sites caractérisés par une clientèle nombreuse (avec un nombre important de registres à examiner) et une qualité et une organisation insuffisantes des données (difficultés d'extraction et de contrôle des données). L'évaluation des systèmes ne devrait pas prendre plus d'une heure dans l'établissement de santé. Dans l'idéal, la personne interrogée pour l'évaluation des systèmes sera le gestionnaire des données de l'établissement (ou la personne responsable de la compilation et de la notification des données).

Réalisation de l'enquête au niveau des districts

La procédure d'examen de la qualité des données est également mise en œuvre au niveau des districts, dans les unités de gestion du système HMIS impliquées dans la circulation des données provenant des établissements de santé échantillonnés. Au niveau des districts, l'équipe d'évaluation agrégera de nouveau les valeurs obtenues dans le district pour les indicateurs choisis en s'appuyant sur les données contenues dans les formulaires mensuels de notification soumis par tous les établissements du district (pas seulement ceux qui appartiennent à l'échantillon). L'équipe évaluera également l'exhaustivité et la ponctualité de la notification à ce niveau. Au niveau du district, l'évaluation des systèmes sera menée sous forme d'un entretien avec le gestionnaire des données ou l'administrateur de programme. Les équipes d'évaluation devraient prévoir de passer environ une demi-journée dans l'unité de gestion HMIS du district.

Supervision et contrôle de la qualité de l'enquête

Sur le terrain, les équipes d'évaluation doivent être supervisées par du personnel désigné à cette fin. Chaque superviseur sera responsable d'une zone géographique prédéfinie et d'un

nombre déterminé d'équipes d'évaluation. Le rôle du superviseur consiste à appuyer les équipes dans la réalisation des enquêtes (si nécessaire), à recueillir et examiner les questionnaires remplis et à résoudre les problèmes éventuels. Il incombe aux superviseurs de revisiter un petit échantillon d'établissements (par exemple 10 %) et de vérifier les résultats de l'enquête dans ces établissements pour s'assurer que les résultats sont correctement consignés. Si possible, des moniteurs indépendants provenant d'organismes partenaires nationaux (par exemple, bailleurs de fonds) peuvent également contribuer à la surveillance de l'exécution de l'enquête.

Compilation des résultats

Les superviseurs des équipes d'évaluation transmettront les questionnaires remplis au personnel désigné chargé de la gestion des données de l'examen de la qualité des données au niveau national. Une petite équipe sera constituée, parmi les personnels du ministère de la santé et/ou des organismes partenaires, pour examiner les formulaires soumis, corriger les erreurs et saisir les données dans le système informatique (tel que CSPro 6.2) afin d'en faciliter l'analyse. Selon le nombre d'établissements inclus dans l'échantillon et le nombre d'indicateurs vérifiés, il faudra compter jusqu'à une semaine pour qu'une équipe de 4-5 gestionnaires de données parvienne à nettoyer et à saisir toutes les données.



3.3 Analyse et interprétation

Paramètres de mesure de la qualité des données recueillis lors de l'enquête en établissement

Bien que l'examen de la qualité des données comporte quatre dimensions de la qualité des données (voir section 3.1), seuls certains paramètres, relevant des dimensions suivantes, peuvent être examinés dans le cadre des enquêtes en établissement :

Dimension 1 : exhaustivité et ponctualité des données ;

Dimension 2 : cohérence interne des données notifiées.

Exhaustivité et ponctualité

L'évaluation de l'exhaustivité des données consiste à déterminer si toutes les entités censées notifier des données soumettent effectivement des rapports, et si elles le font dans les délais prescrits. Les mesures de l'exhaustivité et de la ponctualité comprises dans l'enquête en établissement lors de l'examen de la qualité des données sont notamment :

- **Exhaustivité et ponctualité de la notification au niveau des établissements** – Ce paramètre examine si les établissements de santé appartenant à l'échantillon représentatif de l'enquête ont transmis leurs formulaires mensuels de notification et si ces rapports ont été soumis dans les délais impartis.
- **Exhaustivité des données relatives aux indicateurs** – Ce paramètre évalue dans quelle mesure les établissements de santé appartenant à l'échantillon représentatif de l'enquête ont inclus des informations sur chacun des indicateurs choisis dans leurs formulaires mensuels de notification, s'ils proposent ce service.
- **Exhaustivité des données sur la tuberculose dans les documents sources** – Selon les normes et critères B1.4 relatifs à la tuberculose,¹ on doit disposer de données sur un ensemble minimal de variables pour ≥95 % des cas de tuberculose notifiés dans l'unité de gestion de base (UGB). Des conclusions erronées risquant d'être établies si les données de l'UGB sont inexactes ou incomplètes, on évaluera la proportion de cas de tuberculose pour lesquels au moins l'une des six variables est manquante dans le registre sur la tuberculose (année d'enregistrement, sexe, âge, classification de la maladie, type de patient, résultats bactériologiques).

¹ Standards and benchmarks for tuberculosis surveillance and vital registration systems: checklist and user guide. Geneva: World Health Organization; 2014 (WHO/HTM/TB/2014.02; http://apps.who.int/iris/bitstream/10665/112673/1/9789241506724_eng.pdf?ua=1, consulté le 11 juin 2015).



Cohérence interne des données notifiées

La cohérence interne des données peut être mesurée en examinant la cohérence de mêmes données à différents moments, la cohérence entre des données apparentées, et la cohérence des données entre les documents sources et les bases de données nationales.

Dans le cadre des enquêtes en établissement, cette mesure est effectuée en comparant les données contenues dans les documents sources et dans la base de données nationale, comme suit :

- **Vérification de la cohérence de la notification** : cette étape consiste à examiner les documents sources des établissements de santé afin d'évaluer l'exactitude de la notification pour les indicateurs choisis. Cet élément de la cohérence interne est mesuré au travers d'un exercice de vérification des données qui impose d'examiner les registres dans un échantillon d'établissements de santé. La vérification des données repose sur la comparaison du nombre total de services dispensés selon qu'il provient des documents sources de l'établissement de santé ou du système de notification (HMIS ou système de notification propre à chaque programme) pour les indicateurs sélectionnés. Pour une période de notification donnée, les valeurs des indicateurs sélectionnés sont recalculées à partir des sources primaires de données sur ces indicateurs. La valeur recalculée est alors comparée à la valeur initialement notifiée par le système pour la période de notification concernée. Le rapport entre la valeur recomptée et la valeur notifiée, appelé « facteur de vérification », est une mesure de l'exactitude de l'indicateur. Il convient que cet exercice soit effectué dans les établissements de santé, puis répété au niveau des districts et des provinces, avec un calcul du rapport de vérification à chaque niveau.

Analyse des paramètres de qualité des données et d'autres mesures recueillies dans le cadre des enquêtes en établissement

Les sections suivantes contiennent des tableaux recommandés pour la présentation et l'interprétation des indicateurs de la qualité des données recueillis lors de l'enquête en établissement dans le cadre de l'examen de la qualité des données.

Informations générales sur l'établissement de santé

Cette section présente des tableaux destinés à décrire l'échantillon et à fournir un contexte pour l'interprétation des paramètres de qualité des données.

Disponibilité des services et notification des données

Les résultats présentés doivent comprendre des informations sur la proportion d'établissements de l'échantillon qui fournissent les services de santé concernés, ainsi que sur les établissements notifiant des données via le système HMIS ou un autre système de notification relevant du



ministère de la santé. Cela permet de déterminer le nombre d'établissements sur lesquels se fondent les résultats ultérieurs de vérification des données. Les Tableaux 3.2 et 3.3 donnent un exemple de la manière dont les données peuvent être présentées.

Tableau 3.2 Pourcentage d'établissements de l'échantillon fournissant chacun des services de santé concernés, par strate et par indicateur

	Global	Strate 1 (par exemple, type d'établissement)			Strate 2 (par exemple, autorité de gestion)	
		Type 1	Type 2	Type 3	Type 1	Type 2
ANC1 (n =)						
DTC3/Penta3 (n =)						
Cas de paludisme (n =)						
Cas notifiés de tuberculose (n =)						
Actuellement sous TAR (n =)						

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Tableau 3.3 Pourcentage d'établissements fournissant les services concernés qui notifient des données via un système de notification du ministère de la santé, par strate et par indicateur

	Global	Strate 1 (par exemple, type d'établissement)			Strate 2 (par exemple, autorité de gestion)	
		Type 1	Type 2	Type 3	Type 1	Type 2
ANC1 (n =)						
DTC3/Penta3 (n =)						
Cas de paludisme (n =)						
Cas notifiés de tuberculose (n =)						
Actuellement sous TAR (n =)						

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Disponibilité des documents sources et des rapports mensuels

Si un établissement offre un service particulier, il faut également qu'il puisse fournir, le jour de l'enquête de vérification des données, les documents sources nécessaires (registres, feuilles de décompte, etc.), ainsi que des rapports mensuels pour toute la période de vérification de trois mois. Il convient que les registres, feuilles de décompte et autres documents employés par les établissements de santé pour consigner les activités quotidiennes relatives aux indicateurs programmatiques sélectionnés (et aux services correspondants) soient des documents standard du ministère de la santé. Il est possible que les établissements de santé aient recours à plusieurs documents pour enregistrer les services dispensés, mais il est important d'identifier s'il existe un document source principal dont sont extraites les données aux fins du rapport mensuel. Le Tableau 3.4 indique le pourcentage de disponibilité de ces documents pour toute la période de trois mois. L'équation suivante donne le pourcentage de disponibilité des documents sources et des rapports mensuels.

% de disponibilité des documents sources et des rapports mensuels pour chaque établissement =

$$\frac{\sum_{i=1}^n \text{Disponible mois } 1_i + \text{Disponible mois } 2_i + \text{Disponible mois } 3_i}{3n} \times 100$$

où n est le nombre total d'établissements fournissant un service particulier.

Tableau 3.4 Pourcentage de mois-établissements (pour les établissements fournissant un service donné) pour lesquels l'équipe d'évaluation a pu obtenir tous les documents sources requis et le rapport mensuel, par strate

	Global	Strate 1 (par exemple, type d'établissement)			Strate 2 (par exemple, autorité de gestion)	
		Type 1	Type 2	Type 3	Type 1	Type 2
ANC1 (n =)						
DTC3/Penta3 (n =)						
Cas de paludisme (n =)						
Cas notifiés de tuberculose (n =)						
Actuellement sous TAR (n =)						

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphthérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Concordance entre les documents sources et les rapports mensuels

Le nombre d'événements recomptés à partir du document source principal doit être identique au nombre signalé dans le formulaire mensuel de notification. Le Tableau 3.5 indique le degré de concordance entre les résultats de prestation des services qui sont notifiés dans les rapports mensuels et qui sont recomptés à partir des documents sources pendant toute la période de trois mois.

% de concordance entre les résultats notifiés et recomptés =

$$\frac{\sum_{i=1}^n (\# \text{ d'établissements à concordance exacte mois}_1 + \# \text{ d'établissements à concordance exacte mois}_2 + \# \text{ d'établissements à concordance exacte mois}_3)}{3n} \times 100$$

où n est le nombre total d'établissements fournissant un service particulier.



Tableau 3.5 Pourcentage de mois-établissements (pour les établissements fournissant un service donné) pour lesquels la somme des données sources est exactement égale aux données notifiées, par strate

	Strate 1 (type d'établissement)				Strate 2 (autorité de gestion)	
	Global	Hôpitaux	Centres de santé	Dispensaires	Gouvernement	Privé, à but lucratif
ANC1 (n =)						
DTC3/Penta3 (n =)						
Cas de paludisme (n =)	%	%	%	%	%	%
Cas notifiés de tuberculose (n =)	%	%	%	%	%	%
Actuellement sous TAR (n =)						

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphthérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Indicateurs de la qualité des données

Exhaustivité de la notification des établissements

Cet indicateur mesure le pourcentage de rapports mensuels reçus par le bureau de district pour les établissements inclus dans l'échantillon de l'enquête en établissement. Le nombre de formulaires mensuels de notification varie selon le pays. Dans certains pays, un formulaire unique regroupe tous les indicateurs clés, tandis que dans d'autres, différents formulaires sont employés pour différents services/domaines programmatiques. Dans la procédure d'examen de la qualité des données, l'exhaustivité de la notification peut être mesurée pour plusieurs formulaires mensuels de notification.

Dans l'idéal, l'exhaustivité de la notification des établissements est mesurée sur la base des rapports mensuels reçus par le bureau de district. Que le système d'information sanitaire du pays soit électronique ou sur support papier, il est recommandé de mesurer l'exhaustivité de la notification des établissements au niveau du district en demandant des informations sur les rapports mensuels reçus de la part des établissements inclus dans l'enquête. Exceptionnellement, si une visite dans le bureau de district n'est pas possible, une variable de mesure indirecte de l'exhaustivité peut être calculée en se fondant sur la disponibilité des rapports mensuels au sein des établissements de santé. Le Tableau 3.6 donne un exemple de présentation des données.



Tableau 3.6 Pourcentage de mois-établissements (pour les mois étudiés et pour les établissements visités qui fournissent le service concerné) pour lesquels des rapports mensuels comprenant les indicateurs suivants ont été reçus par le bureau de district, par strate

	Global	Strate 1 (par exemple, type d'établissement)			Strate 2 (par exemple, autorité de gestion)	
		Type 1	Type 2	Type 3	Type 1	Type 2
ANC1 (n =)						
DTC3/Penta3 (n =)						
Cas de paludisme (n =)	%	%	%	%	%	%
Cas notifiés de tuberculose (n =)	%	%	%	%	%	%
Actuellement sous TAR (n =)						

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphthérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Ponctualité de la notification des établissements

Il est essentiel pour les administrateurs d'avoir rapidement accès aux informations. Cet indicateur, recueilli au niveau des districts, évalue dans quelle mesure les établissements inclus dans l'enquête soumettent leurs rapports au bureau de district dans les délais (c'est-à-dire avant la date de réception définie dans les modes opératoires normalisés de gestion des données). Le Tableau 3.7 illustre comment présenter les données.

Tableau 3.7 Pourcentage de mois-établissements (pour les mois étudiés et pour les établissements visités qui fournissent le service concerné) pour lesquels les rapports mensuels ont été reçus par le bureau de district dans les délais de notification, par strate

	Global	Strate 1 (type d'établissement)			Strate 2 (autorité de gestion)	
		Type 1	Type 2	Type 3	Type 1	Type 2
ANC1 (n =)						
DTC3/Penta3 (n =)						
Cas de paludisme (n =)	%	%	%	%	%	%
Cas notifiés de tuberculose (n =)	%	%	%	%	%	%
Actuellement sous TAR (n =)						

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphthérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Ponctualité de la notification par les districts

La ponctualité de la notification par les districts est mesurée au point de destination des rapports transmis par les districts, c'est-à-dire généralement au niveau national. Des problèmes de ponctualité peuvent survenir aussi bien au niveau de la notification par les districts qu'à des niveaux supérieurs d'agrégation. Un effet de chaîne peut se produire lorsqu'une notification incomplète ou tardive de la part des établissements a des répercussions sur la notification par les districts et par les autres niveaux d'agrégation. Le Tableau 3.8 présente la ponctualité de la notification par une unité supérieure d'agrégation (par exemple, le bureau de district). Cet indicateur n'est pas calculé dans les pays où les données sont exclusivement transmises sous forme électronique entre le district et le niveau national.

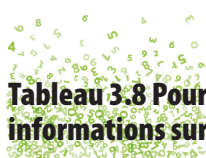


Tableau 3.8 Pourcentage de rapports mensuels de district (pour les trois mois définis, contenant des informations sur les indicateurs suivants) soumis dans les délais par le bureau de district

	Strate (région)			
	Global	Région 1	Région 2	Région 3
ANC1 (n =)				
DTC3/Penta3 (n =)				
Cas de paludisme (n =)				
Cas notifiés de tuberculose (n =)				
Actuellement sous TAR (n =)				

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Exhaustivité des données individuelles

S'il est essentiel d'atteindre un niveau élevé d'exhaustivité de la notification par les établissements, il est également important de veiller à ce que les établissements censés transmettre des données sur un indicateur donné fournissent bien les informations correspondantes dans leurs rapports mensuels. Cet indicateur mesure le degré d'exhaustivité des données individuelles transmises par les établissements de l'échantillon. Le Tableau 3.9 donne un exemple de présentation des données.



Tableau 3.9 Pourcentage de mois-établissements (pour les établissements visités qui fournissent le service concerné et notifient des données) pour lesquels des données sur les indicateurs suivants sont incluses dans les rapports mensuels, par strate

	Global	Strate 1 (type d'établissement)			Strate 2 (autorité de gestion)	
		Type 1	Type 2	Type 3	Type 1	Type 2
ANC1 (n =)						
DTC3/Penta3 (n =)						
Cas de paludisme (n =)						
Cas notifiés de tuberculose (n =)						
Actuellement sous TAR (n =)						

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphthérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Exhaustivité des informations sur un ensemble minimal de variables pour la tuberculose

Pour permettre aux systèmes de surveillance de la tuberculose de bien évaluer l'incidence et les tendances de la maladie, il faut que des données soient notifiées sur un ensemble minimal de variables. Cet ensemble minimal de variables devra notamment comporter, pour tous les cas, des données sur : l'âge, le sexe, l'année, les résultats bactériologiques (confirmation en laboratoire ou confirmation clinique), les antécédents de traitement (nouveaux cas ou cas précédemment traités) et le site anatomique de la maladie (tuberculose pulmonaire ou extrapulmonaire). L'exhaustivité des données sur cet ensemble minimal de variables est évaluée pour déterminer si les normes B1.5, B1.6 et B1.7 sont respectées,² comme indiqué dans le Tableau 3.10.

Tableau 3.10 Fréquence des données manquantes concernant des variables particulières dans les registres de la tuberculose

	n	%
Nombre total d'établissements comptant des cas pour lesquels des données sont manquantes		
Cas pour lesquels il manque des données concernant des variables particulières		
Année d'enregistrement		
Sexe		
Âge		
Classification de la maladie (pulmonaire ou extrapulmonaire)		
Type de patient (nouveau cas ou cas précédemment traité)		
Résultats d'examen bactériologique		
Nombre de cas pour lesquels il manque des données pour au moins l'une des six variables suivantes : année d'enregistrement, sexe, âge, classification de la maladie, type de patient ou résultats bactériologiques		

2 Standards and benchmarks for tuberculosis surveillance and vital registration systems: checklist and user guide. Geneva: World Health Organization; 2014 (WHO/HTM/TB/2014.02; http://apps.who.int/iris/bitstream/10665/112673/1/9789241506724_eng.pdf?ua=1, consulté le 11 juin 2015).

Facteur de vérification (FV)

Même si les valeurs notifiées ne correspondent pas exactement aux valeurs recomptées, il est utile d'évaluer leur degré de disparité.

Pour un indicateur donné, le facteur de vérification (FV) dans un établissement est calculé en divisant le nombre d'événements recomptés à partir des documents sources par le nombre d'événements notifiés dans le système HMIS.

$$\text{Facteur de vérification} = \frac{\text{Nombre d'événements recomptés à partir des documents sources}}{\text{Nombre d'événements notifiés dans le HMIS}}$$

Un FV supérieur à 1 est signe d'une sous-notification des événements dans le système HMIS pendant la période d'analyse, tandis qu'un FV inférieur à 1 indique une surnotification des événements dans le HMIS pendant cette même période. Pour un indicateur traceur donné, les données provenant d'établissements ne fournissant pas le service en question seront naturellement exclues du calcul du FV. Il convient également de noter que les valeurs recomptées peuvent être supérieures aux valeurs notifiées en cas de rapports manquants, l'inverse étant vrai en cas de documents sources manquants. C'est pourquoi le FV est calculé uniquement pour les établissements de santé dotés à la fois des documents sources et des rapports mensuels nécessaires ; il n'est pas calculé pour les établissements dans lesquels il manque soit les données sources, soit un ou plusieurs rapports mensuels. En cela, l'évaluation de l'exactitude de la notification se distingue de l'évaluation de l'exhaustivité de la tenue des registres et de l'archivage.

Les Tableaux 3.11 et 3.12 présentent les facteurs FV nationaux généraux calculés au niveau des établissements, les FV par strate et le pourcentage d'établissements dans lesquels une surnotification ou une sous-notification est observée. Le FV est une moyenne pondérée. Comme toute moyenne, sa valeur peut masquer une variation des facteurs FV entre différents établissements, certains pouvant enregistrer une valeur de FV beaucoup plus faible (surnotification plus importante que la moyenne ne le laisse supposer) et d'autres un facteur FV beaucoup plus élevé (sous notification plus importante que ne l'indique la moyenne). Il arrive qu'on observe une surnotification dans certaines catégories d'établissements (par exemple, établissements publics) et une sous-notification dans d'autres (par exemple, établissements privés à but lucratif). Il est également utile d'examiner la distribution des facteurs FV parmi les établissements individuels : le pourcentage d'établissements présentant une surnotification de plus de 10 % (FV < 0,90), le pourcentage d'établissements présentant une sous-notification de plus de 10 % (FV > 1,10) et le pourcentage d'établissements dont les données sources correspondent exactement aux données notifiées. Si la taille de l'échantillon le permet, on comparera aussi les unités infranationales (régions) pour déterminer où il convient de concentrer les ressources à des fins de renforcement des systèmes.

Les estimations pondérées des facteurs FV pour les indicateurs analysés devront être comparées aux résultats des évaluations précédentes de la qualité des données afin de dégager des tendances en matière d'exactitude des données.



Tableau 3.11 Facteur de vérification au niveau des établissements pour les indicateurs sélectionnés, par strate

	Facteur de vérification national	Strate 1 (type d'établissement)			Strate 2 (autorité de gestion)	
		Type 1	Type 2	Type 3	Type 1	Type 2
ANC1 (n =)						
DTC3/Penta3 (n =)						
Cas de paludisme (n =)						
Cas notifiés de tuberculose (n =)						
Actuellement sous TAR (n =)						

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Tableau 3.12 Paramètres pertinents pour la vérification des données au niveau des établissements

	Données				
	ANC1 (n =)	DTP3/ PENTA3 (n =)	Cas de paludisme (n =)	Cas notifiés de tuberculose (n =)	Actuellement sous TAR (n =)
Pourcentage d'établissements fournissant le service concerné et notifiant des données qui disposent de tous les documents sources et rapports nécessaires					
Pourcentage d'établissements dont les données sources correspondent exactement aux données notifiées					
Pourcentage d'établissements présentant une surnotification de plus de 10 % (FV<0,90)					
Pourcentage d'établissements présentant une sous notification de plus de 10 % (FV>1,10)					

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Facteur de vérification pour les unités supérieures d'agrégation

L'exercice de vérification des données devrait être effectué à tous les niveaux où des informations sanitaires sont physiquement agrégées (par exemple, établissement → district → province → niveau national). Dans les pays dotés d'un système électronique d'information sanitaire dans lequel les districts saisissent toutes les données relatives aux établissements de santé, la vérification des données sera effectuée au niveau des établissements et des districts. Dans les autres pays, où il existe plusieurs niveaux d'agrégation, la vérification doit être réalisée à tous les niveaux. L'exemple fourni dans le Tableau 3.13 ci-dessous présente une analyse tabulaire des informations de vérification au niveau des districts. Une opération semblable doit être effectuée pour les autres niveaux d'agrégation, le cas échéant.

Le Tableau 3.13 ci-dessous indique qu'au niveau des districts, le FV est calculé en procédant à une nouvelle agrégation de la valeur des indicateurs sélectionnés transmise par les établissements qui soumettent des rapports mensuels récapitulatifs au district. La valeur réagrégée est divisée par la valeur notifiée par le district pour la période de notification concernée afin d'obtenir le facteur FV du district. Le facteur FV du district est une mesure indépendante de l'exactitude de la notification pour le système HMIS du district ou le bureau de gestion du programme. Le FV du district n'est pas pris en compte dans le facteur FV composite dérivé de l'échantillon complet d'établissements de santé.



Tableau 3.13 Paramètres pertinents pour la vérification des données au niveau des districts

	Données				
	ANC1 (n =)	DTP3/ PENTA3 (n =)	Cas de paludisme (n =)	Cas notifiés de tuberculose (n =)	Actuellement sous TAR (n =)
Pourcentage d'établissements fournissant le service concerné et notifiant des données qui disposent de tous les documents sources et rapports nécessaires					
Facteur FV national au niveau des districts					
Nombre et liste des districts avec un FV<0,90					
Nombre et liste des districts avec un FV>1,10					
Pourcentage de districts présentant une surnotification (FV<0,90)					
Pourcentage de districts présentant une sous notification (FV>1,10)					

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Causes de divergence entre les données des rapports mensuels et des documents sources

Au niveau des établissements

Le Tableau 3.14 passe en revue les causes de divergence entre les données recomptées à partir des documents sources et les données notifiées dans les rapports mensuels. Le Tableau 3.15 examine les causes d'indisponibilité des rapports mensuels, lorsqu'il manque un ou plusieurs rapports mensuels. Il est utile d'étudier chaque programme séparément car les résultats obtenus peuvent indiquer si un problème donné est systémique ou propre à un programme particulier. Des analyses complémentaires peuvent être réalisées selon le type d'établissement ou l'entité qui en est propriétaire.



Tableau 3.14 Causes de divergence entre les données sources et les données notifiées au niveau des établissements, par domaine programmatique

	Données				
	ANC1 (n =)	DTP3/ PENTA3 (n =)	Cas de paludisme (n =)	Cas notifiés de tuberculose (n =)	Actuellement sous TAR (n =)
Pourcentage d'établissements sans divergence					
Pourcentage d'établissements présentant des erreurs arithmétiques					
Pourcentage d'établissements présentant des erreurs de transcription					
Pourcentage d'établissements où il manquait certains documents lors de la préparation des rapports					
Pourcentage d'établissements où il manquait certains documents lors de la réalisation de l'enquête					
Autres causes					

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Tableau 3.15 Causes d'indisponibilité des rapports mensuels, par domaine programmatique

	Données				
	ANC1 (n =)	DTP3/ PENTA3 (n =)	Cas de paludisme (n =)	Cas notifiés de tuberculose (n =)	Actuellement sous TAR (n =)
Pourcentage d'établissements où les trois rapports mensuels sont disponibles					
Pourcentage d'établissements ayant soumis un rapport qui ne peut actuellement être retrouvé					
Pourcentage d'établissements ne disposant pas du personnel qualifié nécessaire pour la notification					
Pourcentage d'établissements ne disposant pas de formulaire de notification					
Pourcentage d'établissements ayant connu une interruption des services durant au moins l'un des mois de la période concernée					
Autres causes					

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

District ou niveaux supérieurs d'agrégation

Le Tableau 3.16 indique si le bureau de district chargé de traiter les rapports mensuels contenant des informations sur les indicateurs programmatiques sélectionnés dispose d'un système de suivi de l'exhaustivité et de la ponctualité des rapports mensuels reçus de la part des établissements de santé. Il est possible que plusieurs bureaux de district soient impliqués, surtout lorsque les programmes utilisent des systèmes parallèles de notification. Dans ce cas, la question sera posée au niveau des programmes. Cependant, si un seul bureau de district contrôle le flux d'information (par exemple, le bureau HMIS), le suivi de l'exhaustivité et de la ponctualité ne sera demandé qu'une seule fois.

Remarque : Selon la stratégie d'échantillonnage adoptée pour l'enquête en établissement, si le district n'est pas l'unité principale d'échantillonnage, il ne sera pas possible, avec l'information recueillie, de tirer des conclusions applicables à tous les districts du pays. On peut toutefois espérer que l'information obtenue demeure utile à titre illustratif, permettant de guider les discussions au niveau national sur les problèmes rencontrés en matière de gestion des données dans les districts. Cette restriction s'applique à toutes les analyses réalisées au niveau des districts.

Tableau 3.16 Disponibilité d'un système de suivi de l'exhaustivité et de la ponctualité, au niveau des districts

	Global	Données				
		ANC1 (n =)	DTP3/ PENTA3 (n =)	Cas de paludisme (n =)	Cas notifiés de tuberculose (n =)	Actuellement sous TAR (n =)
Pourcentage de districts disposant d'un système de suivi de la ponctualité						
Pourcentage de districts disposant d'un système de suivi de l'exhaustivité						

Remarque : ANC = soins prénatals ; DTP3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Le Tableau 3.17 identifie les causes de divergence entre les données agrégées provenant des rapports mensuels de tous les établissements de santé et les rapports transmis par le bureau de district vers le niveau supérieur de notification. L'information présentée dans ce tableau est ventilée par domaine programmatique. Si plusieurs bureaux de district participent au processus de vérification des données, l'analyse de district peut révéler une variation de l'exactitude des données relatives à différents programmes. Même si les données sont compilées par un seul agent au niveau du district, des informations pertinentes propres à chaque programme peuvent être obtenues. Les Tableaux 3.18 et 3.19 examinent, du point de vue d'un administrateur de district, les raisons de la non-soumission ou de la soumission tardive des rapports par les établissements du district. Il est utile d'étudier chaque programme séparément car les résultats obtenus peuvent indiquer si les divergences observées sont systémiques ou propres à des programmes particuliers. Des analyses complémentaires peuvent être réalisées selon le type d'établissement ou l'entité qui en est propriétaire.

Tableau 3.17 Causes signalées de la divergence entre les données sources et les données notifiées au niveau des districts, par domaine programmatique

	Données				
	ANC1 (n =)	DTP3/ PENTA3 (n =)	Cas de paludisme (n =)	Cas notifiés de tuberculose (n =)	Actuellement sous TAR (n =)
Pas de divergence					
Erreurs arithmétiques ou erreurs de saisie des données					
Rapports supplémentaires reçus de la part des établissements après la notification par le district	%	%	%	%	%
Il manque certains rapports d'établissements après la notification par le district	%	%	%	%	%
Autres causes					

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Tableau 3.18 Point de vue des districts sur les causes de non-exhaustivité de la notification, par domaine programmatique

	Données				
	ANC1 (n =)	DTP3/ PENTA3 (n =)	Cas de paludisme (n =)	Cas notifiés de tuberculose (n =)	Actuellement sous TAR (n =)
Exhaustivité de 100 % de la notification					
Manque de personnel qualifié dans les établissements					
Manque de formulaires de notification dans les établissements	%	%	%	%	%
Difficultés de transport/communication	%	%	%	%	%
Certains établissements ne fournissent plus le service en question					
Certains établissements ne suivent pas les lignes directrices					
Le système de suivi de l'exhaustivité dont dispose le district est inadéquat					
Autres causes					

Remarque : ANC = soins prénatals ; DTC3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Tableau 3.19 Point de vue des districts sur les causes de retard de la notification, par domaine programmatique

	Données				
	ANC1 (n =)	DTP3/ PENTA3 (n =)	Cas de paludisme (n =)	Cas notifiés de tuberculose (n =)	Actuellement sous TAR (n =)
Ponctualité de la notification de 100 %					
Difficultés de transport/communication					
Certains établissements complètent tardivement leurs rapports	%	%	%	%	%
Le système de suivi de la ponctualité dont dispose le district est inadéquat					
Autres causes					

Remarque : ANC = soins prénatals ; DTP3 = vaccin antidiphtérique-antitétanique-anticoquelucheux ; TAR = traitement antirétroviral.

Évaluation des systèmes

L'évaluation des systèmes a pour objet de mesurer la capacité du système à produire des données de bonne qualité. Pour ce faire, on étudie dans quelle mesure les éléments critiques du système de notification se conforment à un ensemble minimal de normes acceptables. L'évaluation porte sur différents domaines du système, examinant l'existence de lignes directrices, la présence de personnels qualifiés et la disponibilité d'outils de collecte des données, ainsi que la supervision et la rétro-information sur la qualité des données. L'annexe 3 fournit des informations détaillées sur la définition de chacun de ces domaines et le calcul du score correspondant. Le Tableau 3.20 propose une méthode de présentation des résultats relatifs à ces domaines du système. Il est recommandé d'adopter une méthode de présentation analogue au niveau des districts. Le codage couleur conditionnel indique une variation de la performance pour chaque rubrique des différentes strates. À noter que les chiffres et les estimations figurant dans ce tableau sont fournis à titre purement illustratif.

Tableau 3.20 Pourcentage d'établissements ayant transmis des données sanitaires via un système de notification relevant du ministère de la santé et ayant obtenu les scores suivants pour les différents domaines du système de gestion des données, par strate

	Global	Type d'établissement			Entité propriétaire		Emplacement	
		Hôpital	Centre de santé	Poste sanitaire	Public	Privé	Urbain	Rural
	n=231	n=85	n=86	n=60	n=173	n=58	n=88	n=143
Établissements notifiant des statistiques sur les services au ministère de la santé (%)								
Parmi les établissements qui offrent des soins prénatals, pourcentage transmettant des données via un système de notification du ministère de la santé	89	40	92	94	92	81	81	93
Parmi les établissements qui offrent des services de vaccination, pourcentage transmettant des données via un système de notification du ministère de la santé	89	39	90	97	94	75	78	95
Parmi les établissements qui dispensent des soins anti-VIH, pourcentage transmettant des données via un système de notification du ministère de la santé	59	91	68	22	55	71	73	52
Parmi les établissements qui dispensent des soins antituberculeux, pourcentage transmettant des données via un système de notification du ministère de la santé	57	86	62	35	55	64	61	55
Parmi les établissements qui proposent un traitement antipaludique, pourcentage transmettant des données via un système de notification du ministère de la santé	100	96	100	100	100	99	99	100
Scores relatifs aux différents domaines du système de gestion des données (%)								
	n=231	n=82	n=86	n=60	n=171	n=58	n=88	n=143
Disponibilité de lignes directrices	57	66	63	36	54	61	69	54
Disponibilité de personnels qualifiés	42	49	47	26	41	45	56	39
Aucune rupture de stock des feuilles de décompte, registres et formulaires de notification au cours des 6 derniers mois	73	88	77	57	73	73	73	73
Bénéficie d'une supervision et d'une rétro-information sous forme écrite, y compris sur la qualité des données	16	48	20	2	14	20	22	15
Analyse et utilisation des données	37	45	43	20	38	47	56	33
Respect de tous les critères	2	17	3	0	2	4	4	2
Moyenne des différentes rubriques	46	65	49	33	44	50	54	44
Score général	35	30	44	17	37	32	38	34

Le score général est obtenu en multipliant le pourcentage d'établissements transmettant des données via un quelconque système de notification du ministère de la santé par le score moyen.

D'autres analyses simples (par exemple, test d'indépendance du chi carré) peuvent être menées pour examiner un lien significatif éventuel entre ces différentes rubriques (aussi bien sur le plan individuel qu'en tant qu'indices) et la qualité des données (FV). Le Tableau 3.21 donne un exemple de mise en tableau de la relation entre la disponibilité d'une rubrique particulière – formation reçue par un agent chargé de la saisie/compilation des données – et le facteur FV de vérification des données. Des tableaux comparables peuvent être établis pour d'autres rubriques. Bien que ne permettant pas d'établir une relation de causalité, de telles analyses présentent un intérêt certain pour définir les mesures prioritaires à mettre en œuvre en vue d'améliorer certaines de ces caractéristiques physiques. D'autres analyses, comme les régressions, peuvent être effectuées pour étudier la relation entre la disponibilité des indicateurs d'évaluation du système et la qualité des données (FV).



Tableau 3.21 Variation du facteur moyen de vérification des données selon qu'une formation a été dispensée ou non au personnel chargé de la saisie/compilation des données, globalement et par strate

	Moyenne globale du facteur de vérification	Strate 1 (par exemple, type d'établissement)			Strate 2 (par exemple, autorité de gestion)	
		Type 1	Type 2	Type 3	Type 1	Type 2
Oui – ruptures de stock						
Non – ruptures de stock						



Annexe 1 : Indicateurs recommandés

Indicateurs clés

Indicateurs recommandés pour l'examen de la qualité des données		
Domaine programmatique	Nom de l'indicateur	Définition complète de l'indicateur
Santé maternelle	Couverture de la première consultation prénatale (ANC1)	Nombre et pourcentage de femmes enceintes ayant bénéficié d'au moins une consultation pendant la grossesse
Vaccination	Couverture de la vaccination DTC3/ Penta3	Nombre et pourcentage d'enfants de <1 an ayant reçu trois doses de vaccin DTC/pentavalent
VIH	Actuellement sous TAR	Nombre et pourcentage de personnes vivant avec le VIH qui sont actuellement sous TAR
Tuberculose	Taux de notification de la tuberculose	Nombre de nouveaux cas et de cas de rechute de la tuberculose qui sont notifiés pour 100 000 habitants
Paludisme	Nombre total de cas confirmés de paludisme ¹	Nombre de cas de paludisme confirmés (par examen microscopique ou TDR) pour 1000 personnes par an

Remarque : ANC = soins prénatals ; DTC3 = trois doses de vaccin antidiphtérique-antitétanique-anticoquelucheux ; Penta = vaccin pentavalent ; TAR = traitement antirétroviral ; TDR = test de diagnostic rapide.

Indicateurs supplémentaires

Indicateurs supplémentaires de la procédure d'examen de la qualité des données		
Domaine programmatique	Nom de l'indicateur	Définition complète de l'indicateur
Général	Recours aux services	Nombre de consultations ambulatoires par personne par an
Santé maternelle	Quatrième consultation prénatale (ANC4)	Nombre (pourcentage) de femmes âgées de 15 à 49 ans ayant eu un enfant né vivant durant une période donnée qui ont bénéficié d'au moins quatre consultations de soins prénatals
	Accouchements pratiqués en établissement	Nombre (pourcentage) de femmes ayant accouché dans un établissement de santé
	Couverture des soins post-partum	Nombre (pourcentage) de mères et d'enfants ayant bénéficié de soins post-partum dans les deux jours suivant la naissance (quel que soit le lieu d'accouchement)
	Première dose d'anatoxine tétanique	Nombre (pourcentage) de femmes enceintes ayant reçu la première dose de vaccin à base d'anatoxine tétanique
Vaccination	DTC1-3/Penta1-3	Nombre (pourcentage) d'enfants de <1 an ayant reçu 1 dose, 2 doses ou 3 doses de vaccin DTC/pentavalent
	MCV1	Nombre (pourcentage) de nourrissons ayant reçu au moins une dose de vaccin à valence rougeole (MCV) avant l'âge de 1 an
	VPC 1-3 ²	Nombre (pourcentage) d'enfants de <1 an ayant reçu 1 dose, 2 doses ou 3 doses de vaccin antipneumococcique

¹ Si le nombre de cas confirmés de paludisme n'est pas recueilli, on peut lui substituer le nombre total de cas de paludisme.

² Si ce vaccin n'est pas utilisé dans le pays, le remplacer par un autre vaccin inclus dans le programme national.

Indicateurs supplémentaires, suite

Indicateurs supplémentaires de la procédure d'examen de la qualité des données		
Domaine programmatique	Nom de l'indicateur	Définition complète de l'indicateur
VIH	Personnes vivant avec le VIH qui ont été diagnostiquées	Nombre (pourcentage) de personnes vivant avec le VIH qui ont été diagnostiquées
	Couverture des soins anti-VIH	Nombre (pourcentage) de personnes vivant avec le VIH qui bénéficient de soins anti-VIH (dont le TAR)
	Couverture du TAR pour la PTME	Nombre (pourcentage) de femmes enceintes séropositives pour le VIH qui ont bénéficié d'un TAR durant la grossesse
	Rétention sous TAR	Nombre (pourcentage) de personnes vivant avec le VIH et bénéficiant d'un TAR qui demeurent sous TAR 12 mois après le début du traitement (ainsi qu'à 24, 36, 48 et 60 mois)
	Suppression de la charge virale	Nombre (pourcentage) de personnes sous TAR qui présentent une suppression de la charge virale
Tuberculose	Cas notifiés de tuberculose, toutes formes confondues	Nombre de nouveaux cas et de cas de rechute de la tuberculose qui sont notifiés pour 100 000 habitants — Évaluer si les tableaux 1 et 2 ¹ du rapport trimestriel de notification des cas sont corrects selon les normes et critères (B1.4) applicables aux systèmes sur support papier ²
	Taux de réussite du traitement antituberculeux	Nombre (pourcentage) de cas de tuberculose traités avec succès (guérison, et traitement mené à terme) parmi les cas notifiés aux autorités sanitaires nationales au cours d'une période donnée — Évaluer si le Tableau 1 du rapport trimestriel sur les issues thérapeutiques est correct selon les normes et critères (B.14) applicables aux systèmes sur support papier.
	Taux de réussite du traitement antituberculeux de deuxième intention	Nombre (pourcentage) de cas de tuberculose traités avec succès (guérison, et traitement mené à terme) parmi tous les cas confirmés de tuberculose résistante à la rifampicine ou de tuberculose multirésistante ayant commencé un traitement de deuxième intention au cours de la période d'évaluation
Tuberculose/VIH	Proportion de patients enregistrés comme tuberculeux (cas nouveaux et rechutes) dont le statut par rapport au VIH est documenté	Nombre de patients tuberculeux (cas nouveaux et rechutes) pour lesquels un résultat de dépistage du VIH est consigné dans le registre sur la tuberculose, exprimé sous forme de pourcentage du nombre de patients enregistrés au cours de la période de notification
	Proportion de patients tuberculeux (cas nouveaux et rechutes) séropositifs pour le VIH qui sont sous TAR pendant le traitement antituberculeux	Nombre de patients tuberculeux (cas nouveaux et rechutes) séropositifs pour le VIH qui ont bénéficié d'un TAR pendant le traitement antituberculeux, exprimé sous forme de pourcentage des patients enregistrés au cours de la période de notification
Paludisme	Taux de dépistage diagnostique du paludisme	Nombre (pourcentage) de sujets ayant fait l'objet d'un test parasitologique parmi tous les cas suspects de paludisme [= nombre de cas testés / (nombre de cas testés + nombre de cas présumés)]
	Cas confirmés de paludisme bénéficiant d'un traitement	Nombre (pourcentage) de cas confirmés de paludisme bénéficiant d'un traitement antipaludique de première intention conformément à la politique nationale dans des établissements publics
	Cas de paludisme (présumés et confirmés) bénéficiant d'un traitement	Nombre (pourcentage) de cas de paludisme (présumés et confirmés) bénéficiant d'un traitement antipaludique de première intention
	TPIp3	Nombre (pourcentage) de femmes enceintes en soins prénatals qui ont reçu au moins trois doses de traitement préventif intermittent contre le paludisme

Remarque : ANC = soins prénatals ; DTC = vaccin antidiphtérique-antitétanique-anticoquelucheux ; MCV = vaccin à valence rougeole ; PTME = prévention de la transmission mère-enfant ; TAR = traitement antirétroviral ; TPIp = traitement préventif intermittent pendant la grossesse ; VPC = vaccin antipneumococcique conjugué

¹ Définitions et cadre de notification pour la tuberculose – Révision 2013. Genève, Organisation mondiale de la Santé, 2013 (WHO/HTM/TB/2013.2 ; http://apps.who.int/iris/bitstream/10665/110724/1/9789242505344_fre.pdf, consulté le 11 juin 2015).

² Standards and benchmarks for tuberculosis surveillance and vital registration systems: checklist and user guide. Geneva: World Health Organization; 2014 (WHO/HTM/TB/2014.02; http://apps.who.int/iris/bitstream/10665/112673/1/9789241506724_eng.pdf?ua=1, consulté le 11 juin 2015).



Annexe 2 : Calcul des paramètres de qualité des données dans le cadre de l'enquête en établissement de santé

Tableau A2.1 Paramètres de qualité des données recueillis dans le cadre des enquêtes en établissement de santé

Paramètre de qualité des données	Description de l'analyse au niveau de l'établissement	Description de l'analyse au niveau du district
a) Exhaustivité de la notification des établissements : Pourcentage de rapports escomptés qui ont été archivés (pour les trois mois d'analyse) pour les établissements de l'échantillon d'étude DVD_123a=1 – Rapport observé pour le mois 1 pour les soins prénatals DVD_124a=1 – Rapport observé pour le mois 2 pour les soins prénatals DVD_125a=1 – Rapport observé pour le mois 3 pour les soins prénatals	Exemple pour les soins prénatals <i>Score global pour tous les mois-établissements :</i> $\frac{\sum_{i=1}^n DVD_123a_i + DVD_124a_i + DVD_125a_i}{3n} \times 100^1$ <i>Remarque 1 : n est le nombre total d'établissements de l'échantillon censés soumettre des rapports sur les soins prénatals (DVD_121=1 and DVD_122=1)</i> La même logique s'applique à la mesure de l'exhaustivité de la notification pour d'autres indicateurs. Si le système d'information du pays regroupe tous les indicateurs sur un même formulaire de notification, l'exhaustivité sera la même pour tous les indicateurs. Cependant, si les informations relatives aux indicateurs sont consignées sur différents formulaires de notification, l'exhaustivité variera selon l'indicateur.	
b) Ponctualité de la notification : pourcentage des rapports archivés qui ont été reçus dans les délais (pour les trois mois d'analyse) pour les établissements de l'échantillon d'étude DVD_123b=1, DVD_124b=1, DVD_125b=1 – Rapports reçus dans les délais pour les mois 1,2 et 3 respectivement, pour les soins prénatals DVD_132 = Nombre de rapports soumis dans les délais par le district	Exemple pour les soins prénatals $\frac{\sum_{i=1}^n DVD_123b_i + DVD_124b_i + DVD_125b_i}{3n} \times 100^2$ <i>où n est le nombre total d'établissements de l'échantillon censés soumettre des rapports sur les soins prénatals (DVD_121=1 and DVD_122=1)</i> La même logique s'applique à la mesure de la ponctualité de la notification pour d'autres indicateurs programmatiques.	Exemple pour les soins prénatals $\frac{\sum_{i=1}^n DVD_132_i}{n * 12} \times 100$ <i>où n est le nombre total de districts</i>

¹ En supposant que ces variables sont égales à 1 si le rapport archivé est observé par l'équipe d'évaluation et de 0 sinon.

² En supposant que a) les variables du dénominateur sont égales à 1 si le rapport archivé est observé par l'équipe d'évaluation et de 0 sinon ; et les variables du numérateur sont égales à 1 si le rapport a été soumis dans les délais et de 0 sinon.

Tableau A2.1, suite

Paramètre de qualité des données	Description de l'analyse au niveau de l'établissement	Description de l'analyse au niveau du district
c) Exhaustivité des données individuelles : pourcentage de rapports mensuels escomptés archivés qui contiennent des informations sur les indicateurs programmatiques examinés (pour les trois mois d'analyse) pour les établissements de l'échantillon d'étude DVD_123c = résultats des services de soins prénatals notifiés pour le mois 1 DVD_124c = résultats des services de soins prénatals notifiés pour le mois 2 DVD_125c = résultats des services de soins prénatals notifiés pour le mois 3	Exemple pour les soins prénatals $[\text{Count}(\text{DVD_123c} \neq \text{manquant} + \text{DVD124c} \neq \text{manquant} + \text{DVD125c} \neq \text{manquant}) / 3n] \times 100$ <i>où n est le nombre total d'établissements de l'échantillon censés soumettre des rapports sur les soins prénatals (DVD_121=1 and DVD_122=1)</i> La même logique s'applique à la mesure de l'exhaustivité des données individuelles notifiées pour d'autres indicateurs programmatiques.	

¹ Variables = 1. âge ou tranche d'âge ; 2. sexe ; 3. année d'enregistrement ; 4. résultats bactériologiques ; 5. antécédents de traitement ; 6. site anatomique de la maladie

³ En supposant que ces variables sont égales à 1 si le rapport archivé est observé par l'équipe d'évaluation et de 0 sinon



Tableau A2.1, suite

Paramètre de qualité des données	Description de l'analyse au niveau de l'établissement	Description de l'analyse au niveau du district
d) Exhaustivité des informations sur un ensemble minimal de variables pour la tuberculose : - Pourcentage d'établissements pour lesquels il manque des informations sur l'une quelconque des variables de l'ensemble minimal de variables¹ pour le trimestre d'étude DV_406_07 = Nombre de cas pour lesquels il manque des données sur l'une quelconque des variables de l'ensemble minimal de variables DV_405 = Nombre total de cas de tuberculose consignés dans le document source, déduction faite des cas transférés. - Pourcentage de cas pour lesquels il manque des informations sur un élément de données spécifique requis pour le trimestre d'étude DV_406_01 = Nombre de cas pour lesquels il manque des informations sur l'année d'enregistrement - Pourcentage de cas pour lesquels il manque des informations sur au moins un des éléments de données requis pour le trimestre d'étude	où n est le nombre d'établissements censés notifier des données sur la tuberculose (DV_400 = 1 and DV_401 = 1) $[Count(DV_406_07 \neq 0) / n] \times 100$ Pourcentage de cas pour lesquels il manque des informations sur un élément de données spécifique requis (par exemple, année d'enregistrement) $\frac{\sum_{i=1}^n DV_406_01_i}{\sum_{i=1}^n DV_405_i} \times 100$ Remarque : La même logique s'applique à la mesure de l'exhaustivité de la notification pour les autres éléments de données requis (sexe, âge, classification de la maladie, antécédents de tuberculose, résultats bactériologiques). Pourcentage de cas pour lesquels il manque des informations sur au moins un des éléments de données : $\frac{\sum_{i=1}^n DV_406_07_i}{\sum_{i=1}^n DV_405_i} \times 100$	

¹ Variables = 1. âge ou tranche d'âge ; 2. sexe ; 3. année d'enregistrement ; 4. résultats bactériologiques ; 5. antécédents de traitement ; 6. site anatomique de la maladie

² En supposant que ces variables sont égales à 1 si le rapport archivé est observé par l'équipe d'évaluation et de 0 sinon



Tableau A2.1, suite

Paramètre de qualité des données	Description de l'analyse au niveau de l'établissement	Description de l'analyse au niveau du district
d) Vérification des données : pourcentage de concordance entre les données figurant dans les registres des établissements échantillonnés et les registres nationaux pour ces mêmes établissements DV_103_01_B, DV_103_02_B, DV_103_03_B = Recomptage des données relatives aux soins prénatals à partir du document source pour les mois 1, 2 et 3 respectivement. DV_104_01_B, DV_104_02_B, DV_104_03_B = Données relatives aux soins prénatals notifiées dans les rapports mensuels pour les mois 1, 2 et 3 respectivement DVD_126_a, DVD_126_b, DVD_126_c = Somme des consultations prénatales notifiées au bureau de district pour les mois 1, 2 et 3 respectivement DVD_127_a, DVD_127_b, DVD_127_c = Consultations prénatales notifiées par le bureau de district au niveau supérieur	Exemple pour les soins prénatals $\frac{\sum_{i=1}^n DV_{103_01B_i} + DV_{103_02B_i} + DV_{103_03B_i}}{\sum_{i=1}^n DV_{104_01B_i} + DV_{104_02B_i} + DV_{104_03B_i}} \times 100^1$ où n est le nombre total d'établissements de l'échantillon qui disposent de tous les documents sources et tous les rapports requis (DV_103_01_A = 1 et DV_103_02_A = 1 et DV_103_03_A = 1 et DV_104_01_A = 1 et DV_104_02_A = 1 et DV_104_03_A = 1)	Exemple pour les soins prénatals $((DVD_{126_a} + DVD_{126_b} + DVD_{126_c}) / (DVD_{127_a} + DVD_{127_b} + DVD_{127_c}))$



Annexe 3 : Calcul des indicateurs « traceurs » du système qui sont susceptibles d'influer sur la qualité des données

Tableau A3.1 Calcul des scores relatifs aux différents domaines du système de gestion des données^{1,2}

Domaine et rubriques indicatrices	Description de l'analyse au niveau de l'établissement	Description de l'analyse au niveau du district
Disponibilité de personnels qualifiés		
Disponibilité de personnels désignés pour la saisie/ compilation des données DV_600=1 – établissement DVD_103=1 – district	Score d'établissement relatif à la disponibilité de personnels qualifiés = moyenne des scores des différentes rubriques, en pourcentage Score global pour tous les établissements :	Score de district relatif à la disponibilité de personnels qualifiés = moyenne des scores des différentes rubriques, en pourcentage Score global pour tous les districts :
Disponibilité de personnels désignés pour l'examen de la qualité des données avant la soumission : DV_601=1 – établissement DVD_104=1 – district	$\frac{\sum_{i=1}^n DV_600_i + 601_i + 602_i + 603_i}{4n} \times 100$ où n est le nombre total d'établissements de l'échantillon qui notifient des données sanitaires (DV_599 = 1)	$\frac{\sum_{i=1}^n DVD_103_i + 104_i + 105_i + 106_i}{4n} \times 100$ où n est le nombre total de districts dans l'échantillon
Formation dispensée au personnel sur la saisie/ compilation des données : DV_602=1 – établissement DVD_105=1 – district		Si plusieurs bureaux de district font l'objet d'une visite, le calcul devra être effectué pour chaque bureau et les numéros des questions devront être ajustés en conséquence.
Formation dispensée au personnel sur l'examen et le contrôle des données : DV_603=1 – établissement DVD_106=1 – district		
Disponibilité de lignes directrices		
Disponibilité de lignes directrices au niveau des établissements : DV_604=1	Score d'établissement relatif à la disponibilité de lignes directrices = score en pourcentage Score global pour tous les établissements :	
	$\frac{\sum_{i=1}^n DV_604_i}{n} \times 100$ où n est le nombre total d'établissements de l'échantillon qui notifient des données sanitaires (DV_599=1)	

¹ Les scores relatifs aux domaines doivent être calculés pour chaque strate (type d'établissement, autorité de gestion, +/- région géographique).

² Le calcul se fonde sur l'hypothèse que la variable aura un score égal à 1 si le résultat est « Oui, observé » et égal à zéro sinon.

Tableau A3.1, suite

Domaine et rubriques indicatrices	Description de l'analyse au niveau de l'établissement	Description de l'analyse au niveau du district
Disponibilité de lignes directrices relatives à la saisie/ compilation des données au niveau des districts : DVD_107		Score de district relatif à la disponibilité de personnels qualifiés = moyenne des scores des différentes rubriques, en pourcentage Score global pour tous les établissements : $\frac{\sum_{i=1}^n DVD_107_i + DVD_108_i + DVD_109_i}{3n} \times 100$ où n est le nombre total de districts dans l'échantillon Si plusieurs bureaux de district font l'objet d'une visite, le calcul devra être effectué pour chaque bureau et les numéros des questions devront être ajustés en conséquence.
Disponibilité de lignes directrices relatives à l'examen et au contrôle des données au niveau des districts : DVD_108=1		
d'information sanitaire habituel au niveau des districts : DVD_109=1		

Ruptures de stock

Aucune rupture de stock des feuilles de décompte, registres et formulaires de notification au cours des six derniers mois : DV_605=2 (établissement) DVD_111=2 (district)	$\frac{\sum_{i=1}^n DV_605_i}{n} \times 100$ où n est le nombre total d'établissements de l'échantillon qui notifient des données sanitaires (DV_599=1) Pour calculer le score relatif à ce domaine, les valeurs de DV_605 sont remplacées par 1 s'il n'y a pas eu de rupture de stock et par 0 s'il y a eu une rupture de stock.	$\frac{\sum_{i=1}^n DVD_111_i}{n} \times 100$ où n est le nombre total de districts de l'échantillon fournissant des feuilles de décompte, registres et formulaires aux établissements de santé (DVD_110=1) Pour calculer le score relatif à ce domaine, les valeurs de DVD_111 sont remplacées par 1 s'il n'y a pas eu de rupture de stock et par 0 s'il y a eu une rupture de stock. Si plusieurs bureaux de district font l'objet d'une visite, le calcul devra être effectué pour chaque bureau et les numéros des questions devront être ajustés en conséquence.
---	--	---

Supervision et rétro-information

Visite de supervision dans les trois derniers mois : DV_606≠6	$\frac{\sum_{i=1}^n DV_606_i + DV_607_i}{2n} \times 100$ où n est le nombre total d'établissements de l'échantillon qui notifient des données sanitaires (DV_599=1) Calcul reposant sur l'hypothèse que a) DV_606=0 si DV_606=6 et DV_6=1 dans le cas contraire ; et que DV_607=1 si DV_607 et DV_607=0 sinon.	
Rétro-information sur la qualité des données reçue sous forme écrite : DV_607=1 (établissement)		



Tableau A3.1, suite

Domaine et rubriques indicatrices	Description de l'analyse au niveau de l'établissement	Description de l'analyse au niveau du district
Rétro-information sur la qualité des données fournie sous forme écrite DVD_113=1 (district) Rétro-information sur la performance des services fournie sous forme écrite DVD_114=1 (district)		$\frac{\sum_{i=1}^n DVD_113_i + DVD_114_i}{2n} \times 100$ où n est le nombre total de districts dans l'échantillon Pour calculer le score relatif à ce domaine, les valeurs de DVD_113 et DVD_114 sont remplacées par 1 si une rétro-information écrite pertinente a été observée et par 0 dans le cas contraire. Si plusieurs bureaux de district font l'objet d'une visite, le calcul devra être effectué pour chaque bureau et les numéros des questions devront être ajustés en conséquence.

Analyse et utilisation des données

Présence de supports visuels (format papier ou électronique) dans l'établissement : DV_608=1	Score d'établissement relatif à l'utilisation des données = moyenne des scores des différentes rubriques, en pourcentage $\frac{\sum_{i=1}^n DV_608_i + DV_609_i + DV_610_i + DV_611_i}{4n} \times 100$ où n est le nombre total d'établissements de l'échantillon qui notifient des données sanitaires (DV_599=1) Pour calculer le score relatif à ce domaine, les valeurs de DV_608, DV_609, DV_110 et DV_111 sont remplacées par 1 si la preuve correspondante d'analyse des données a été observée et par 0 dans le cas contraire.	
Visualisation des données en sus de la vaccination : DV_609 = 1 si (DV_609_03=1 et (DV_609_01=1 ou DV_609_02=1 ou DV_609_04=1 ou DV_609_05=1))		
Utilisation des données pour l'examen de la performance : DV_610=1		
Utilisation des données pour la planification : DV_611=1		
Présence de supports visuels (format papier ou électronique) dans l'établissement : DVD_115=1		Score de district relatif à l'utilisation des données = moyenne des scores des différentes rubriques, en pourcentage $\frac{\sum_{i=1}^n DVD_115_i + 116_i + 117_i + 118_i + 119_i}{5n} \times 100$ où n est le nombre total de districts dans l'échantillon Si plusieurs bureaux de district font l'objet d'une visite, le calcul devra être effectué pour chaque bureau et les numéros des questions devront être ajustés en conséquence.
Production d'un rapport/ bulletin sur la base des données du système d'information sanitaire habituel : DVD_116=1		
Exemple attesté d'une mesure de suivi : DVD_117		
Utilisation des données pour l'examen de la performance : DVD_118=1		
Utilisation des données pour la planification : DVD_119=1		

Tableau A3.1, suite

Domaine et rubriques indicatrices	Description de l'analyse au niveau de l'établissement	Description de l'analyse au niveau du district
Autres rubriques présentant un intérêt		
Système de suivi de la ponctualité de la notification DVD_102=1		$\frac{\sum_{i=1}^n DVD_102_i}{n} \times 100$ où n est le nombre total de districts dans l'échantillon Si plusieurs bureaux de district font l'objet d'une visite, le calcul devra être effectué pour chaque bureau.
Scores récapitulatifs		
Pourcentage avec toutes les rubriques indicatrices	Countif (DV_600=1 et _601=1 et _602=1 et _603=1 et _604=1 et _605=1 et _606=1 et _607=1 et _608=1 et _609=1 et _610=1 et _611=1)*100/n où n est le nombre total d'établissements de l'échantillon qui notifient des données sanitaires (DV_599=1) DV_609 = 1 si (DV_609_03=1 et (DV_609_01=1 ou DV_609_02=1 ou DV_609_04=1 ou DV609_05=1))	Countif (DVD_102=1 et _103=1 et _104=1 et _105=1 et _106=1 et _107=1 et _108=1 et _109=1 et _110=1 et _111=1 et _112=1 et _113=1 et _114=1 et _115=1 et _116=1 et _117=1 et _118=1 et _119=1)*100/n où n est le nombre total de districts dans l'échantillon
Moyenne des rubriques indicatrices	Average (DV_600, _601, _602, _603, _604, _605, _606, _607, _608, _609, _610, _611)*100 où la valeur de chaque indicateur = 1 s'il est présent et observé et = 0 dans le cas contraire.	Average (DVD_102, _103, _104, _105, _106, _107, _108, _109, _110, _111, _112, _113, _114, _115, _116, _117, _118, _119)*100 où la valeur de chaque indicateur = 1 s'il est présent et observé et = 0 dans le cas contraire.
Score général	$\frac{\sum_{i=1}^n DV_599_i}{n} \times \text{moyenne des rubriques indicatrices}$	



Annexe 4 : Documents sources recommandés et vérification des données par recoupement ou contrôle ponctuel

Le Tableau A4.1 ci-dessous présente, pour les indicateurs clés et les indicateurs supplémentaires, les sources de données appropriées et les vérifications par recoupement des données pouvant être effectuées. Il est toutefois recommandé de procéder à ces vérifications par recoupement dans le cadre des procédures approfondies d'examen de la qualité des données.

Tableau A4.1 Vérification des données par recoupement et par contrôle ponctuel

Programme	Indicateur	Source de données	Vérifications par recoupement et contrôles ponctuels
Statistiques des services généraux	Recours aux services	Registre des consultations ambulatoires	
Santé maternelle	- Première consultation prénatale - Quatrième consultation prénatale - Accouchements pratiqués en établissement - Consultation postnatale 1 - AT1	- Registre de l'unité de travail et d'accouchement - Registre des consultations prénatales - Registre des consultations postnatales	- Si les fiches de santé des patientes sont conservées dans l'établissement, elles peuvent être comparées aux registres de consultations prénatales et postnatales. - Parler aux patientes présentes dans l'établissement au moment de la vérification des données et leur poser des questions sur les services reçus. Comparer les informations obtenues avec les registres correspondants pour voir si les services et traitements ont été correctement consignés.
Vaccination	- DTC1—3 /Penta 1—3 - MCV1 - VPC 1—3¹	Feuilles de décompte	- Les registres de vaccination peuvent être comparés au nombre de doses de vaccin utilisées (en n'oubliant pas que certains vaccins sont fournis par lots de 10 doses et qu'un lot peut être utilisé pour moins de 10 enfants). - Les informations tirées d'un échantillon de carnets de vaccination appartenant à des enfants peuvent être comparées au registre de vaccination des enfants de l'établissement le jour de la visite de vérification.
VIH²	- Actuellement sous TAR - Couverture des soins anti-VIH - Couverture du TAR pour la PTME - Rétention sous TAR - Suppression de la charge virale	- Registres des programmes (registre TAR, fiches TAR des patients) - Registres TAR des établissements - Données relatives aux établissements, agrégées à partir du système de suivi des patients	- Les registres TAR peuvent être comparés aux registres de pharmacie. - Les dossiers des patients peuvent être comparés aux informations contenues dans la base de données sur les patients (si l'établissement en possède une). - Contrôles ponctuels : demander aux patients présents dans l'établissement au moment de la vérification de parler des services qu'ils ont reçus. La confidentialité est une considération primordiale ; si la confidentialité du patient ne peut être garantie, on ne procédera pas au contrôle ponctuel.

Table A4.1, suite

Programme	Indicateur	Source de données	Vérifications par recoupement et contrôles ponctuels
Tuberculose³	- Cas notifiés de tuberculose toutes formes confondues - Taux de réussite du traitement antituberculeux - Taux de réussite du traitement antituberculeux de deuxième intention - Proportion de patients enregistrés comme tuberculeux (cas nouveaux et rechutes) dont le statut par rapport au VIH est documenté - Proportion de patients tuberculeux (cas nouveaux et rechutes) séropositifs pour le VIH qui sont sous TAR pendant le traitement antituberculeux	Registres de l'unité de soins antituberculeux	Vérifications par recoupement : comparaison entre les cas de tuberculose détectés (dans les registres de laboratoire) et les cas de tuberculose notifiés (sujets initialement perdus de vue) - Le registre de l'unité de soins antituberculeux peut être comparé aux fiches de traitement de la tuberculose. - Le registre de l'unité de soins antituberculeux peut être comparé au registre de laboratoire pour vérifier que les sujets chez lesquels la tuberculose a été diagnostiquée ont bien été consignés (si le diagnostic est effectué dans l'établissement). - Le registre de l'unité de soins antituberculeux peut être comparé aux registres de pharmacie.
Paludisme	- Nombre total de cas confirmés de paludisme - Taux de dépistage diagnostique du paludisme - Cas confirmés de paludisme bénéficiant d'un traitement - Cas de paludisme (présumés et confirmés) bénéficiant d'un traitement - TPIp³	- Registre de l'établissement - Registre de laboratoire de l'établissement	- Le registre de l'établissement peut être comparé au registre de laboratoire (examen microscopique et TDR) pour les cas suspects ayant fait l'objet d'un test parasitologique. - Le registre de l'établissement peut être comparé aux registres de pharmacie indiquant les traitements administrés. - Le registre des consultations prénatales peut être comparé aux fiches de santé des patientes pour vérifier les données relatives au TPI, sous réserve que ces fiches soient conservées dans l'établissement. - Le rapport HMIS peut être comparé au rapport du programme de lutte antipaludique si des données sont notifiées dans ces rapports distincts.

Remarque : ANC = soins prénataux ; AT = vaccin à base d'anatoxine tétanique ; DTC = vaccin antidiphthérique-antitétanique-anticoquelucheux ; MCV = vaccin à valence rougeole ; PTME = prévention de la transmission mère-enfant ; TAR = traitement antirétroviral ; TDR = test de diagnostic rapide ; TPIp = traitement préventif intermittent pendant la grossesse ; VPC = vaccin antipneumococcique conjugué

¹ Si ce vaccin n'est pas utilisé dans le pays, le remplacer par un autre vaccin inclus dans le programme national.

² L'échantillonnage des établissements de santé exige une stratification par type d'établissement pour garantir qu'un nombre adéquat d'établissements fournissent des services anti-VIH.

³ L'échantillonnage des établissements de santé exige une stratification par type d'établissement pour garantir qu'un nombre adéquat d'établissements fournissent des services de lutte contre la tuberculose.



Annexe 5 : Méthodes et problèmes d'échantillonnage

Calcul de la taille d'échantillon

La taille de l'échantillon dépendra de la précision souhaitée pour les principales estimations de l'enquête en établissement (y compris l'exactitude des données), ainsi que de la marge d'erreur acceptable. D'autres facteurs sont à prendre en compte, notamment la disponibilité des ressources requises et le niveau d'application souhaité des estimations (à noter que les estimations au niveau provincial exigent un échantillon de plus grande taille que les estimations au niveau national). L'équipe de coordination de l'examen de la qualité des données devra travailler en collaboration avec un statisticien d'enquête et avec les organisateurs de l'enquête en établissement pour définir la taille adéquate de l'échantillon en fonction des priorités du pays concernant le niveau d'application des estimations, les ressources disponibles et la précision souhaitée des estimations.

On trouvera ci-après de brèves orientations sur les facteurs essentiels à prendre en compte pour le calcul de la taille d'échantillon selon qu'il est prévu de mener l'exercice de vérification des données de façon isolée ou conjointement à une autre enquête en établissement. L'objectif est de déterminer la taille que doit avoir l'échantillon pour parvenir à une puissance statistique ou une précision adéquate de l'estimation. Il s'agit donc de décider du nombre minimum d'établissements à inclure dans l'échantillon pour obtenir un résultat statistiquement significatif ou un intervalle de confiance suffisamment étroit pour pouvoir juger du degré de concordance.

La plupart des estimations décrites dans le présent document reposent sur la « concordance » entre les valeurs recomptées à partir des documents sources et les valeurs figurant dans les rapports mensuels. Cette concordance est ici comprise comme le produit de i) une prévalence marginale (c'est-à-dire la probabilité de retrouver aussi bien le document source que le rapport mensuel), et ii) le pourcentage attendu de concordance des résultats pour les principaux services vérifiés (par exemple, Penta3, ANC1, cas confirmés de paludisme, etc.) entre les documents sources et les rapports mensuels. Il est donc impératif que la taille de l'échantillon soit suffisante pour permettre une mesure fiable de la concordance (appelée « kappa »), au-delà de ce que donnerait le seul hasard. Kappa est une mesure (comprise entre 0 et 1) de la concordance corrigée du hasard, calculée à partir du pourcentage global de concordance et de la concordance qui serait attendue du fait du seul hasard.¹ Le Tableau A5.1 présente différentes tailles d'échantillon calculées selon trois scénarios fondés sur la prévalence marginale et la plage admissible définie par les deux niveaux nécessaires de concordance (concordance minimale acceptable (P_0) et concordance attendue par l'étude (P_A)), ainsi que leurs valeurs kappa ajustées correspondantes.

¹ Hyunsook Hong, Yunhee Choi, Seokyoung Hahn, Sue Kyung Park Byung-Joo Park (2014). Nomogram for sample size calculation on a straightforward basis for the kappa statistic. *Annals of Epidemiology* 24 (2014) 673e680



Dans le scénario A, on suppose que l'équipe de coordination de l'examen de la qualité des données ne dispose pas de suffisamment d'information sur la disponibilité des documents sources et des rapports mensuels, auquel cas elle pourra envisager une valeur de 0,30 de la prévalence marginale (c'est-à-dire une probabilité de 30 % de trouver les deux documents). De même, l'équipe doit avoir des indications sur le niveau minimum acceptable de concordance entre les deux valeurs, qu'il est recommandé de fixer à au moins 70 %. Ainsi, avec une concordance minimale de 70 % (soit $P_0=0,70$) et un niveau de concordance supérieur aux attentes fixé par prudence à 80 % (soit $P_A=0,80$), la taille minimale de l'échantillon national est de $n=144$, représentant le nombre d'établissements requis pour obtenir une puissance de 80 % et un intervalle de confiance de 95 % pour toutes les estimations clés établies à partir de l'échantillon. En outre, l'échantillon fournit une fiabilité interobservateurs (valeurs fournies recomptées à partir des documents sources, comparées aux valeurs notifiées dans les rapports mensuels) et une mesure satisfaisante de la concordance (valeur de kappa comprise entre 0,29 et 0,52), au-delà de ce que donnerait le seul hasard.

Dans le scénario B, on suppose que l'équipe de coordination de l'examen de la qualité des données a une idée assez claire des chances de disponibilité des documents sources et des rapports mensuels, auquel cas elle pourra envisager une valeur de 0,50 de la prévalence marginale (c'est-à-dire une probabilité de 50 % de trouver les deux documents). L'équipe doit alors se concerter pour choisir un niveau minimum acceptable de concordance entre les deux valeurs présentées dans les documents – par exemple 80 % (soit $P_0=0,80$) – et un degré de concordance supérieur au niveau attendu de 90 % (soit $P_A=0,90$). Dans ces conditions, la taille minimale de l'échantillon national est de $n=126$ établissements, fournissant également une fiabilité interobservateurs et une mesure solide de la concordance (valeur de kappa comprise entre 0,60 et 0,80) au-delà de ce que donnerait le seul hasard. Si l'équipe de coordination de l'examen de la qualité des données ne dispose pas d'informations suffisantes pour décider d'un niveau minimal acceptable de concordance, il est recommandé de lui attribuer une valeur d'au moins 70 % (soit $P_0=0,70$), comme indiqué dans le Tableau A5.1, et d'établir le niveau prudent de concordance « supérieur aux attentes » à 80 % (soit $P_A=0,80$), ce qui donne une taille minimale de $n=165$ établissements pour l'échantillon national, garantissant également une valeur de kappa modérée, comprise entre 0,40 et 0,60.

Dans le scénario C, on suppose que l'équipe de coordination de l'examen de la qualité des données a une idée très claire des chances de disponibilité des documents sources et des rapports mensuels, auquel cas elle pourra envisager une valeur de 0,80 de la prévalence marginale. De même, si l'équipe de coordination de l'examen de la qualité des données escompte une forte concordance entre les documents sources et les rapports mensuels, elle pourra fixer le niveau minimum acceptable de concordance à 80 % (soit $P_0=0,80$) et le niveau de concordance supérieur aux attentes à 90 % ($P_A=0,90$). Dans ces conditions, une taille minimale de $n=100$ établissements dans l'échantillon national est suffisante (avec une valeur de kappa comprise entre 0,38 et 0,53, ce qui correspond plus ou moins à la plage de valeurs définie comme modérée).

Enfin, l'examen du Tableau A5.1 fait ressortir deux autres points qui méritent d'être mentionnés :

- la taille de l'échantillon augmente lorsque l'écart entre le niveau minimum acceptable et le niveau attendu de concordance diminue (par exemple, avec une prévalence marginale

de 50 % (soit 0,5) et une valeur de $P_0=0,80$, si $P_A=0,85$, la différence est de 5 % et la taille d'échantillon requise est $n=502$, tandis que si $P_A=0,90$, la différence est de 10 % et la taille d'échantillon requise est $n=126$;

- le calcul de la taille d'échantillon peut aussi être appliqué à des situations où une représentation infranationale de l'échantillon d'examen de la qualité des données est nécessaire. Par exemple, dans un pays où la disponibilité attendue des documents sources et des rapports mensuels est susceptible de varier fortement d'une région à l'autre, l'équipe de coordination de l'examen de la qualité des données peut choisir une prévalence marginale prudente de 30 % et un niveau minimum acceptable de concordance de 75 % ($P_0=0,75$) avec un niveau attendu de concordance plus élevé ($P_A=0,95$), auquel cas une taille d'échantillon minimale de $n=37$ établissements par région est adéquate.

Tableau A5.1 Exemples de calcul de la taille d'échantillon pour différentes valeurs de la prévalence marginale, du pourcentage de concordance et de la variable kappa correspondante

Scénario	Prévalence marginale*	Pourcentage de concordance		Kappa**		N***
		P_0	P_A	Selon la concordance minimale P_0	Selon la concordance attendue P_A	
A	0,3	0,90	0,95	0,76	0,88	276
	0,3	0,85	0,95	0,64	0,88	96
	0,3	0,80	0,90	0,52	0,76	118
	0,3	0,80	0,85	0,52	0,64	471
	0,3	0,75	0,95	0,40	0,88	33
	0,3	0,75	0,85	0,40	0,64	134
	0,3	0,75	0,80	0,40	0,52	535
	0,3	0,70	0,80	0,29	0,52	144
B	0,5	0,90	0,95	0,80	0,90	283
	0,5	0,80	0,90	0,60	0,80	126
	0,5	0,80	0,85	0,60	0,70	502
	0,5	0,75	0,95	0,50	0,90	37
	0,5	0,75	0,85	0,50	0,70	147
	0,5	0,75	0,80	0,50	0,60	589
	0,5	0,70	0,80	0,40	0,60	165
C	0,8	0,90	0,95	0,69	0,84	262
	0,8	0,85	0,90	0,53	0,69	348
	0,8	0,80	0,90	0,38	0,69	100
	0,8	0,80	0,85	0,38	0,53	400
	0,8	0,75	0,80	0,22	0,38	408
	0,8	0,75	0,85	0,22	0,53	102
	0,8	0,75	0,95	0,22	0,84	25

* La prévalence marginale est la probabilité de trouver à la fois le document source et le rapport mensuel

** Plages de kappa : 0,21-0,40 : satisfaisante ; 0,41-0,60 : modérée ; 0,61-0,80 : solide

*** Taille d'échantillon calculée pour des valeurs de kappa positives (erreur de type 1=5 %, puissance=80 %)

Pondération des estimations de vérification des données

Les estimations de vérification des données établies à partir de l'échantillon d'établissements de santé doivent être pondérées pour tenir compte de la divergence entre l'échantillon et la base d'échantillonnage en termes de distribution du nombre d'interventions sanitaires sujettes à analyse (par exemple, accouchements en présence de personnel qualifié). Si l'échantillon est stratifié, il faudra pondérer les estimations relatives à l'exactitude des données dans chaque strate. En général, le facteur de pondération d'un indicateur donné pour une strate particulière est calculé en divisant le nombre d'événements de cette strate dans la population par le nombre d'événements de cette strate dans l'échantillon. Étant donné que le nombre d'événements mesuré dans l'échantillon et dans la population (c'est-à-dire dans le système HMIS) sera différent pour chaque indicateur analysé, les estimations devront être pondérées séparément pour chaque indicateur.

Cette approche constitue une forme de pondération poststratification. Prenons comme exemple une situation dans laquelle les établissements de l'échantillon ne fournissent pas tous des services de vaccination et, parmi ceux qui le font, certains ne notifient pas de données ou ne transmettent pas de rapport mensuel au HMIS. Dans ce cas, deux corrections sont nécessaires pour tenir compte i) de l'absence de couverture, et ii) de l'absence de réponse, qui a une incidence sur l'estimation nationale générale de chaque indicateur étudié.

Le Tableau A5.2a donne l'exemple d'un pays A hypothétique, où le nombre total d'établissements est de $N=900$, répartis dans 4 strates (types d'établissements). Dans chaque strate, un échantillon d'environ 35 % a été constitué pour la représentation nationale. La colonne C donne le nombre variable d'établissements fournissant un service de vaccination sur l'ensemble des strates. Parmi ces établissements, ceux pour lesquels on dispose à la fois des documents sources et des rapports pour le « mois X » sont indiqués dans la colonne D. La colonne F récapitule le facteur de pondération de l'échantillonnage pour chaque établissement par strate, et les colonnes G et H présentent les facteurs de correction nécessaires pour tenir compte de l'absence de couverture et de l'absence de réponse, respectivement, par strate. Par exemple, pour la strate « hôpitaux généraux », le facteur de correction est de 1,12 (soit $65/48$) pour l'absence de couverture et de 1,208 (soit $58/48$) pour l'absence de réponse. Il est important de noter que pour l'absence de couverture comme pour l'absence de réponse, on présume que l'information manquante ou non mesurable relève d'une indisponibilité aléatoire ou de données non informatives.

Les chiffres étudiés sont ceux qui portent sur le nombre de vaccinations durant le « mois X », figurant dans la colonne I (recomptés lors de la vérification des données) et la colonne J (notifiés), totalisés par strate. La colonne FV donne le facteur de vérification national brut, calculé en divisant les vaccinations recomptées par les vaccinations notifiées (colonne I/ colonne J), soit $10\,150/11\,750=0,864$. En appliquant une pondération stratifiée de l'échantillonnage, pour tenir compte de l'absence de couverture et de l'absence de réponse, on obtient dans les colonnes K et L les valeurs ajustées de la vaccination, ce qui donne un facteur de vérification national ajusté de $47\,438/55\,567=0,859$.



Dans certaines situations, on peut obtenir une meilleure représentation en ajustant les estimations nationales selon les résultats de prestation des services (lorsque les résultats sont généralement supérieurs dans certains types de strates que dans d'autres, par exemple dans les hôpitaux par rapport aux centres sanitaires). Cette approche constitue une forme de pondération analytique ; le Tableau A.5.2b reprend l'exemple fourni ci-dessus pour en étendre la représentation. Dans ce tableau, la colonne A donne le nombre de vaccinations pendant le « mois X » dans tous les établissements du pays ayant notifié des données au système HMIS par strate. Le facteur de pondération analytique (colonne E) est le nombre total de vaccinations notifiées durant le « mois X », divisé par les vaccinations notifiées par les établissements inclus dans l'échantillon d'étude (colonne A/colonne C). Le facteur de pondération analytique est multiplié par le facteur de vérification ajusté de chaque strate (colonne D x colonne E) et le résultat est indiqué dans la colonne F. Ensuite, l'estimation nationale du facteur de vérification ajusté (pondéré par le nombre notifié dans le système HMIS) est obtenue en divisant la somme de la colonne F par la somme de la colonne E, soit $3,227/3,510 = 0,908$.

En résumé, dans le pays A, le facteur de vérification brut du nombre de vaccinations durant le « mois X » s'élève à 0,864, une valeur qui baisse légèrement, pour se chiffrer à 0,859, une fois que l'on tient compte de la pondération d'échantillonnage et de poststratification. En outre, cette estimation augmente pour atteindre 0,908 lorsque l'on tient compte des résultats de prestation des services par strate.

REMARQUE :

Lors de la vérification des données, l'équipe de coordination de l'examen de la qualité des données peut se trouver dans une situation où, pour certains paramètres et indicateurs, le service en question (par exemple, les soins antituberculeux) n'est disponible que dans un sous-ensemble d'établissements de l'échantillon. Dans ce cas, la couverture attendue du service sera inférieure à 80 % (c'est-à-dire que le facteur d'ajustement de la colonne F du Tableau A5.2a sera supérieur à 1,20). Il peut aussi arriver que parmi les établissements fournissant un service donné, un nombre plus faible qu'escompté réponde aux exigences de notification dans le HMIS pour le mois « X », faisant passer le taux de réponse des établissements en dessous de 80 % (c'est-à-dire que le facteur d'ajustement de la colonne G du Tableau A5.2a sera supérieur à 1,20). Si l'une ou l'autre de ces situations survient, il est recommandé à l'équipe de l'examen de la qualité des données :

- d'utiliser le facteur de vérification brut (colonne FV du Tableau A5.2a), calculé à partir du nombre effectif de vaccinations recomptées et notifiées (valeurs des colonnes I et J dans le Tableau A5.2a) ;
- si nécessaire, de procéder à un ajustement supplémentaire du facteur de vérification brut par le biais d'une pondération analytique en s'appuyant sur les résultats de prestation des services notifiés à l'échelle nationale dans le système HMIS. Ainsi, en faisant les mêmes calculs que ci-dessus pour le Tableau A5.2b, le facteur de vérification brut de 0,864 sera ajusté pour se chiffrer à 0,897.



Selon le type d'échantillonnage effectué pour choisir les établissements à inclure dans l'enquête d'examen de la qualité des données, les valeurs des districts peuvent ou non faire l'objet d'une pondération d'échantillonnage. Actuellement, l'approche la plus courante consiste à mener l'enquête en établissement de l'examen de la qualité des données conjointement avec une autre enquête en établissement, comme l'évaluation SARA. Dans l'évaluation SARA, les établissements de santé sont le plus souvent sélectionnés selon une méthode d'échantillonnage stratifié dans laquelle l'unité primaire d'échantillonnage est l'établissement, et non le district. Par conséquent, les estimations de district présentées sont exemptes de pondération.

Si une méthode d'échantillonnage en grappes à deux degrés est utilisée pour choisir les établissements de santé, le facteur de vérification propre à chaque grappe (généralement le district) est pondéré sur la base du volume des services dispensés dans la grappe. On applique à chaque grappe un facteur d'ajustement, calculé comme étant le rapport entre la valeur de district provenant du bureau de district et la valeur de district provenant du niveau national. Une moyenne pondérée des facteurs de vérification ajustés propres à chaque grappe est alors calculée pour obtenir l'estimation de l'exactitude au niveau national sur la base de l'échantillon étudié



Tableau A5.2a Tableau récapitulatif d'un échantillon représentatif d'établissements d'étude (n=310)

Strate	Établissements dans le pays (A)	Établissements de l'échantillon d'étude (B)	Établissements de l'échantillon d'étude qui fournissent des services de vaccination (C)	Établissements de l'échantillon qui fournissent des services de vaccination et répondent au HMIS (disponibilité à la fois des documents sources et du rapport mensuel pour le « mois X ») (D)	Probabilité d'échantillonnage de chaque établissement selon son type	Facteur de pondération d'échantillon
Hôpitaux généraux	185	65	58	48	0,351	2,849
Centres sanitaires de référence	175	65	56	52	0,371	2,695
Centres sanitaires	400	130	120	100	0,325	3,077
Postes sanitaires	140	50	50	45	0,357	2,801
TOTAL	900	310	284	245		

N=900; n=310.

* Par exemple, le nombre pondéré de 10 214 = le nombre moyen recompté par établissement dans la strate (2650/48), multiplié par le facteur de pondération d'échantillonnage pour chaque établissement dans la strate (2,849), multiplié par le nombre d'établissements répondant dans la strate (n=48), multiplié par le facteur d'ajustement tenant compte de l'absence de couverture (1,121), multiplié par le facteur d'ajustement tenant compte de l'absence de réponse (1,208), soit (2650/48)x2,849x48x1,121x1,208;

Facteur d'ajustement pour l'absence de couverture des services de vaccination (G=B/C)	Facteur d'ajustement pour l'absence de réponse (H=C/D)	Nombre de vaccinations pendant le « mois X » (en excluant les vaccinations effectuées les mois pour lesquels le document source ou le rapport est indisponible)		Nombre pondéré de vaccinations pendant le « mois X » (ajusté pour tenir compte de l'absence de couverture et de réponse)		Facteur de vérification brut (FV=I/J)	Facteur de vérification pondéré (ajusté pour tenir compte de l'absence de couverture et de réponse) (FVaj=K/L)
		Recompté dans l'échantillon (I)	Notifié (J)	Recompté dans l'échantillon (K)	Notifié (L)		
1,121	1,208	2650	3300	10 214	12 719	0,803	0,803
1,161	1,077	2650	3300	8918	11 106	0,803	0,803
1,083	1,200	3250	3800	13 000	15 200	0,855	0,855
1,000	1,111	1600	1350	4978	4200	1,185	1,185
		10 150 Facteur de vérification brut = 10 150/11 750 = 0,864	11 750	37 110	43 225	0,864	0,859

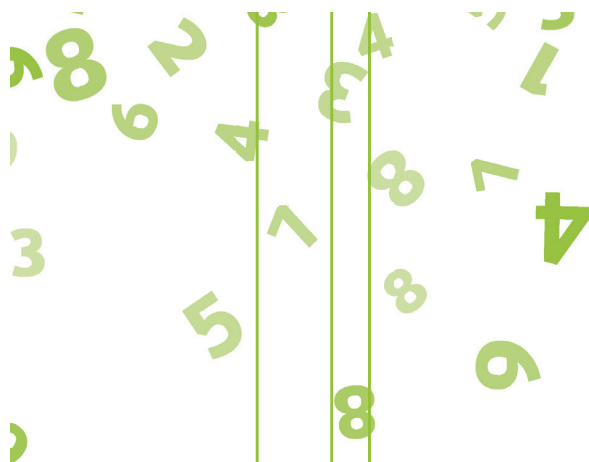


Tableau A5.2b Calcul du facteur de vérification des données et du facteur de pondération selon les résultats de prestation des services notifiés dans le HMIS

Strate	Nombre total de vaccinations notifiées dans le système HMIS pendant le mois « X » (A)	Nombre pondéré de vaccinations pendant le mois « X » (ajusté pour tenir compte de l'absence de couverture et de réponse)		Facteur de vérification ajusté (FVaj=D=B/C)	Pondération analytique (E=A/C)	Facteur de pondération selon les résultats notifiés dans le HMIS (F=DxE)
		Recompté (B)	Notifié (C)			
Hôpitaux généraux	16 170	10 214	12 719	0,803	1,270	1,020
Centres sanitaires de référence	13 860	8918	11 106	0,803	1,247	1,001
Centres sanitaires	11 550	13 000	15 200	0,855	0,760	0,650
Postes sanitaires	4620	4978	4200	1,185	1,100	1,304
TOTAL	46 200	37 132	43 225	0,859	4,377	3,975
		Facteur de vérification ajusté	0,859	Facteur de vérification ajusté selon les résultats notifiés dans le HMIS [F/E]	0,908	

Annexe 6 : Instruments de collecte des données et outils d'analyse

Les instruments de collecte des données incluent la composante de vérification des données et l'outil d'évaluation des systèmes au niveau des établissements et des districts. Des travaux sont en cours pour incorporer l'examen de la qualité des données dans le logiciel DHIS2, ce qui sera particulièrement utile pour les pays où ce logiciel est employé. Un outil Microsoft Excel a été élaboré pour faciliter l'analyse annuelle de la qualité des données dans les pays utilisant un autre logiciel ou un système sur support papier. En outre, un outil d'analyse Microsoft Excel destiné à la vérification des données et à l'évaluation des systèmes est en cours de développement. Les instruments de collecte des données et les outils Excel concernés ne sont pas inclus dans le présent document ; ils font partie de la boîte à outils proposée et seront fournis en pièces jointes à ce document d'orientation.



DQR

DATA QUALITY REVIEW
(EXAMEN DE LA QUALITÉ DES DONNÉES)

ISBN 978 92 4 251274 8



9 789242 512748