

POS 7 : Procédures de nettoyage et de validation des données (version française)

1. Avis de non-responsabilité

Les présentes Procédures Opérationnelles Standardisées (POS) fournissent des conseils étape par étape au personnel directement impliqué dans l'utilisation et l'administration de la plateforme ESPEN Collect. Elles sont destinées à soutenir et à faciliter la mise en œuvre des politiques et procédures de l'OMS relatives à la collecte de données, au suivi et à l'évaluation des programmes de lutte contre les MTN. Ces POS sont fournies à titre indicatif uniquement et n'ont aucun caractère faisant autorité ou juridiquement contraignant. Elles reflètent les pratiques et politiques en vigueur au moment de leur rédaction.

2. Caractéristiques du document

Version	Date de publication	Auteur	Approbateur	Observation
1.0	June 2025	ESPEN Team	ESPEN Collect Project Officer	

3. Introduction

Cette procédure opérationnelle standardisée (POS) est la septième et dernière de la série. Elle fournit des orientations complètes sur les procédures de nettoyage des données pour les enquêtes menées à l'aide de la plateforme ESPEN Collect. Le nettoyage des données est essentiel pour garantir l'exactitude et la fiabilité des données d'enquête, ce qui est indispensable pour évaluer l'impact des maladies tropicales négligées traitées par chimiothérapie préventive (MTN-CP).

Des données de qualité sont essentielles pour prendre des décisions éclairées, comme la reprise ou non des programmes de distribution de médicaments. Pour soutenir ce processus, ESPEN Collect propose des outils de suivi des données via le tableau de bord Metabase, permettant un contrôle en temps réel de la qualité des données pendant les enquêtes en cours. De plus, la plateforme offre un soutien continu pour la gestion des données, tant pendant l'enquête que lors des phases de nettoyage.

Afin d'accélérer le processus de nettoyage et d'améliorer la qualité globale des données, les gestionnaires de données du ministère de la Santé (MS) et des organisations partenaires sont

encouragés à participer activement. Ce guide détaillera comment ils peuvent contribuer efficacement aux efforts de nettoyage, en soulignant les avantages de leur implication.

En suivant cette POS, les utilisateurs seront dotés des procédures et des meilleures pratiques nécessaires pour garantir des données de haute qualité, soutenant ainsi l'élimination réussie des maladies ciblées.

4. Procédures de nettoyage des données

Les procédures de nettoyage des données sont essentielles pour garantir l'exactitude, la cohérence et la fiabilité des données collectées. Elles consistent à identifier et corriger les erreurs, à gérer les données manquantes et à standardiser les formats de données. Dans le cadre de la gestion des enquêtes ESPEN Collect, le nettoyage comprend la correction des erreurs, l'épuration des données et le suivi de la collecte.

Metabase est un outil essentiel pour le processus de nettoyage des données, et son utilisation est détaillée dans la "POS 6 : Guide de Metabase, suivi et supervision des résultats".

Nous pouvons envisager deux scénarios pour le nettoyage des données. Dans le premier scénario, une équipe de gestion des données est présente dans le pays. Cette équipe peut être composée de gestionnaires de données du ministère de la Santé (MS), de partenaires ou de l'OMS. Leur rôle est de se concentrer sur la gestion des données, ce qui inclut la correction des erreurs, le nettoyage et le suivi de la collecte. De cette manière, ils garantissent une bonne qualité des données. Dans le second scénario, il n'y a pas de gestionnaires de données dédiés dans le pays pour travailler sur le nettoyage. Cela peut poser des défis supplémentaires pour assurer la qualité des données. Dans ce cas, les gestionnaires de données d'ESPEN Collect travailleront avec les superviseurs de l'enquête pour effectuer le nettoyage des données.

4.1. Scénario avec une équipe de gestion des données dédiée

Lorsqu'une équipe de gestionnaires de données est présente pour travailler sur le nettoyage des données afin de garantir une bonne qualité, le processus doit être le suivant. Le gestionnaire de données doit posséder au moins une connaissance de base d'Excel. Il doit également comprendre le fonctionnement du tableau de bord Metabase et être capable d'identifier les incohérences à partir de Metabase. Le suivi et le nettoyage des données se déroulent en trois étapes : l'identification des erreurs, le traitement (nettoyage) des erreurs et l'application des corrections dans Metabase.

a. Identification des erreurs :

L'identification des erreurs se fait à partir de Metabase. Le gestionnaire de données se connecte au tableau de bord dans la section « Surveillance et Erreurs » pour voir les détails et les éventuelles incohérences détectées. Chaque fois que des données sont envoyées au serveur, cette section est mise à jour automatiquement. Le gestionnaire de données peut télécharger les détails de toutes les incohérences et les examiner. Il peut également télécharger l'intégralité de la base de données au format XLS ou CSV pour une analyse approfondie de la qualité des données.

b. Nettoyage des erreurs :

Les principaux outils de nettoyage de données dont nous avons besoin sont Metabase et Microsoft Excel. Le gestionnaire de données télécharge les détails des erreurs depuis Metabase. Ensuite, grâce aux registres de données et aux copies papier, il apporte les corrections dans le fichier Excel, puis renvoie ce fichier pour que les corrections puissent être appliquées. Les gestionnaires de données ont la possibilité de télécharger l'ensemble de la base de données, d'effectuer les corrections, puis de renvoyer la base de données afin que les gestionnaires de données d'ESPEN Collect puissent vérifier et appliquer les corrections dans Metabase. Le nettoyage des données doit commencer dès le premier jour pour des raisons d'optimisation. L'avantage est que les agents de saisie peuvent facilement se souvenir des informations saisies et peuvent rapidement reconnaître les erreurs et suggérer des solutions.

c. Application des corrections :

Une fois le nettoyage des données terminé, les gestionnaires de données renvoient les données aux gestionnaires de données d'ESPEN Collect qui effectueront un contre-contrôle. S'il subsiste des incohérences, ils contacteront les gestionnaires de données du pays pour corriger les erreurs restantes. Ce processus de va-et-vient se poursuit jusqu'à ce que les données ne comportent plus aucune erreur. Une fois que les données sont exemptes d'erreurs, les gestionnaires de données d'ESPEN Collect appliqueront ces corrections dans Metabase afin que tout le monde puisse accéder aux données corrigées. Après cette étape, les gestionnaires de données d'ESPEN Collect généreront l'EPIRF, s'il s'agit d'une enquête dont le résultat peut être inclus dans un EPIRF, puis ils partageront cet EPIRF avec le pays pour qu'il puisse vérifier et compléter les informations manquantes. L'EPIRF contient les informations collectées avec ESPEN Collect et les informations soumises par le pays au moment de la demande, mais souvent des informations doivent être complétées, comme la morbidité. L'EPIRF complet est ensuite soumis par le pays via le portail ESPEN en utilisant l'outil de téléchargement JAP.

4.2 Scénario sans équipe de gestion des données dédiée

Dans ce scénario, il n'y a pas de gestionnaires de données dédiés dans le pays pour travailler sur le nettoyage des données. Les gestionnaires de données d'ESPEN Collect surveilleront les données via Metabase et enverront les erreurs identifiées aux superviseurs de terrain. Les superviseurs recevront les détails des erreurs, contacteront les collecteurs de données et vérifieront les copies papier pour obtenir les informations correctes. Ils renverront ensuite les informations corrigées aux gestionnaires de données d'ESPEN Collect. Il est important de noter que les gestionnaires de données d'ESPEN Collect ne peuvent effectuer aucune correction sans l'approbation des gestionnaires de données du pays, car les données appartiennent au pays et non à l'OMS.

Une fois les données nettoyées et validées, les gestionnaires de données d'ESPEN Collect appliqueront les corrections dans Metabase, rendant les données corrigées accessibles à tous. Enfin, si l'enquête le permet, un EPIRF sera généré et partagé avec le pays pour vérification et ajout d'informations supplémentaires avant sa soumission via le portail ESPEN.

5. Outils et logiciels

Pour garantir un nettoyage efficace des données, il est essentiel d'utiliser les bons outils. Les principaux outils recommandés pour les enquêtes ESPEN Collect sont Metabase et Microsoft Excel.

- **Metabase** : Cet outil de visualisation de données open-source permet de surveiller et d'analyser les données en temps réel, en identifiant les erreurs et les incohérences. Pour plus de détails, voir la POS 6 : Guide de Metabase, suivi et supervision des résultats.
- **Microsoft Excel** : Excel est largement utilisé pour le nettoyage et l'analyse des données. Des tutoriels pour améliorer vos compétences sur Excel sont disponibles sur des plateformes telles que YouTube.

En utilisant ces outils, les équipes de gestion des données peuvent garantir la qualité et la fiabilité des résultats d'enquête d'ESPEN Collect.

6. Conclusion

Cette procédure opérationnelle standardisée (POS) sur les procédures de nettoyage et de validation des données est la septième et dernière d'une série de documents destinés à guider les utilisateurs de la plateforme ESPEN Collect. En suivant les étapes décrites dans ce document, les équipes de gestion des données peuvent garantir l'exactitude, la cohérence et la fiabilité des données collectées, ce qui est essentiel pour évaluer l'impact des maladies tropicales négligées traitées par chimiothérapie préventive (MTN-CP).

Les POS suivantes constituent une base solide pour comprendre et utiliser efficacement la plateforme ESPEN Collect :

- POS 1 : Présentation de la plateforme ESPEN Collect
- POS 2 : Procédure de demande d'accès à la plateforme
- POS 3 : Guide étape par étape d'ESPEN Collect
- POS 4 : Manuel de formation pour les instructeurs
- POS 5 : Organisation des enquêtes
- POS 6 : Guide de Metabase, suivi et supervision des résultats

En intégrant les meilleures pratiques de nettoyage et de validation, les utilisateurs peuvent s'assurer que les données sont prêtes pour une analyse approfondie et une prise de décision éclairée. La collaboration entre les gestionnaires de données, les superviseurs et les collecteurs de données est essentielle pour maintenir la qualité des données tout au long du processus.