

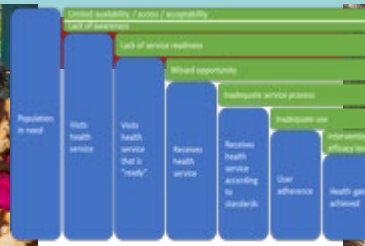
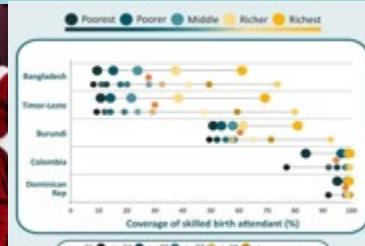
Countdown to 2030

Renforcement des capacités d'analyse au niveau mondial, régional et national sur la santé des femmes, des enfants et des adolescents

Produire des statistiques sanitaires nationales et infranationales fiables sur la santé et la nutrition de la mère, du nouveau-né et de l'enfant

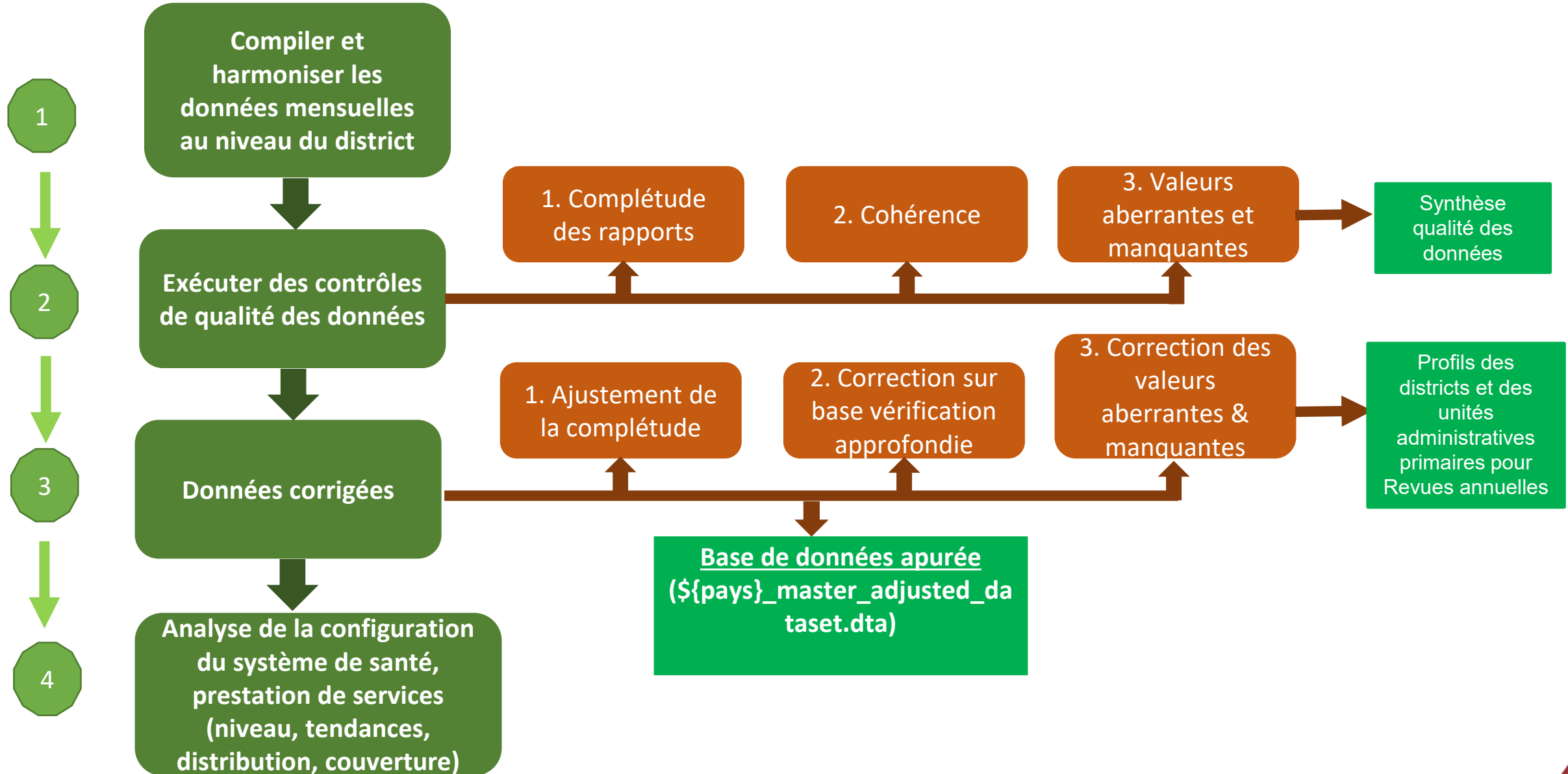
Atelier d'analyse pour l'Afrique subsaharienne
Countdown to 2030 / APHRC / GFF / UNICEF / OMS
13-17 juin 2022

Abdoulaye Maïga, PhD



SESSION #1 – PRÉPARATION DES DONNÉES

Étapes méthodologiques pour la préparation des données



Données et indicateurs du SISR

- Les données incluent les nombres mensuels d'utilisation des services par district de janvier 2017 à décembre 2021
 - Extrait de DHIS-2 dans chaque pays
- Informations supplémentaires sur le taux de complétude mensuel, la population, la région, la zone (urbaine/rurale), les données sur la force du système de santé

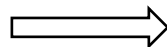
Données administratives

No.	District	Region	GFF priority region	Starting year of prioritization	Tanzania	Total number of hospitals (depending on country definition)	Total number of health facilities (all types/categories)
1	Arusha City Council	Arusha Region					
2	Arusha District	Arusha Region					
3	Babati District	Arusha Region					
4	Babati Town Council	Arusha Region					
5	Bagamoyo District	Arusha Region					

Indicateurs prioritaires

District name	Year	Month	Total number of ANC 1st visit	Total number of Pregnant women completing 4 ANC visits	Total number of IPT 2nd dose	Total number of deliveries in health facilities	Total number of deliveries with skilled birth attendants	Total number of Caesarian Sections
Arusha City Council	2017	January	49	49	100	18	21	87.5
Arusha City Council	2017	February	49	49	100	18	21	87.5
Arusha City Council	2017	March	49	49	100	20	22	91.7
Arusha City Council	2017	April	49	49	100	18	21	87.5
Arusha City Council	2017	May	49	49	100	18	21	87.5

Taux de complétude



District name	Year	Month	ANC reporting			Institutional delivery reporting		
			Expected number (#)	Received number (#)	Reporting completeness rate (%)	Expected number (#)	Received number (#)	Reporting completeness rate (%)
Arusha City Council	2017	January	49	49	100	18	21	87.5
Arusha City Council	2017	February	49	49	100	18	21	87.5
Arusha City Council	2017	March	49	49	100	20	22	91.7
Arusha City Council	2017	April	49	49	100	18	21	87.5
Arusha City Council	2017	May	49	49	100	18	21	87.5

country	urban_ru-1	adminlevel_1	district	year	month	anc1	rr_anc	
1	Tanzania	Urban	Arusha Region	Arusha City Council	2017	January	2342	100
2	Tanzania	Urban	Arusha Region	Arusha City Council	2017	February	2105	100
3	Tanzania	Urban	Arusha Region	Arusha City Council	2017	March	2376	100
4	Tanzania	Urban	Arusha Region	Arusha City Council	2017	April	1780	100
5	Tanzania	Urban	Arusha Region	Arusha City Council	2017	May	2209	100
6	Tanzania	Urban	Arusha Region	Arusha City Council	2017	June	2004	100
7	Tanzania	Urban	Arusha Region	Arusha City Council	2017	July	1964	100
8	Tanzania	Urban	Arusha Region	Arusha City Council	2017	August	2077	100
9	Tanzania	Urban	Arusha Region	Arusha City Council	2017	September	1860	100
10	Tanzania	Urban	Arusha Region	Arusha City Council	2017	October	2117	100

Contenu du fichier Excel des données

- ❑ Nom de la feuille de calcul Excel
 - ❑ DHIS2_dataset_Countryname
par exemple, "DHIS2_dataset_Tanzania"
- ❑ Service_data_1 (CPN, accouchement, CPoN, PF)
- ❑ Service_data_2 (Vaccination, consultation, hospitalisation)
- ❑ Service_data_3 (Mortalité)
- ❑ Reporting_completeness (Tous les services par formulaires de déclaration)
- ❑ Population_data (Total et par âge)

Indicateurs

- CPN 1ère visite
- CPN 4ème visite
- TPI 2nd dose
- Accouchement dans structure de santé
- Accouchement assisté
- Césarienne
- CPoN 48h
- Planning familial : nouvelles visites / utilisatrices
- Planification familiale : revisite

- BCG administré (nourrissons)
- Vaccination pentavalent 1^{ère} (nourrissons)
- Vaccination pentavalent 3^{ème} (nourrissons)
- Rougeole 1^{ère} dose (nourrissons)
- Nombre de consultations des moins de 5 ans
- Nombre total de consultations
- Nombre d'hospitalisations des moins de 5 ans
- Nombre total d'hospitalisations

Décès dans les établissements de santé

- Mort-nés (frais)
- Mort-nés (macérés)
- Décès des moins de 5 ans
- Décès maternels

Population

- Population totale
- Total des moins de 1 an
- Total moins de 5 ans
- Naissances totales
- Total des naissances vivantes
- Total femmes 15-49 ans
- Taux de croissance annuel de la population

Problèmes courant dans les données

- ❑ Apporter des modifications non demandées au modèle
 - ❑ Insertion de nouvelles feuilles de calcul/suppression de feuilles de calcul existantes
 - ❑ Insertion de nouvelles colonnes/suppression de colonnes existantes
 - ❑ Renommer des feuilles de calcul
 - ❑ Changement du format du fichier (ex. Format large)
 - ❑ Fusionner des cellules
- ❑ Différents nombres de districts selon les feuilles de calcul
- ❑ Différentes orthographes des noms de district dans les feuilles de calcul
 - ❑ par exemple, “Nairobi”, “ Nairobi”, “Nairoby”, “NAIROBI”, “nairobi”, “Nai-robi”, “Nai-robi”,
- ❑ Espaces supplémentaires (invisibles) pour les variables non numériques
- ❑ Format d'année et de mois différent
- ❑ Doublons
- ❑ Données manquantes
- ❑ Utiliser 0 à la place des données manquantes et vice versa
- ❑ Etc.

Suivez les conseils généraux dans le « READ-ME » ainsi que les instructions spécifiques à chaque feuille Excel

Remplir complètement et correctement les feuilles

Codes de traitement de données et de contrôles de qualité

- ❑ Fichiers do Stata
- ❑ 1_Code_RHIS_Data_Preparation
 - ❑ Importer et transformer les données Excel au format Stata
 - ❑ Fusionner des données et créer de nouvelles variables
- ❑ 2_Code_RHIS_DQA_Completeness
 - ❑ Évaluer les problèmes de complétude
 - ❑ Ajuster les reportages incomplets
- ❑ 3_Code_RHIS_DQA_internal_consistency
 - ❑ Évaluer la cohérence interne (valeurs aberrantes) et les valeurs manquantes
 - ❑ Corriger la cohérence interne et les valeurs manquantes
 - ❑ Générer un résumé des scores de qualité des données pour le rapport de qualité des données
 - ❑ Générer une base de données apurée et ajusté pour l'analyse (`{country}_master_adjusted_dataset.dta`)

Conseils avant d'utiliser les codes

- ❑ Créer un dossier dédié à l'analyse qui contiendra les données d'entrées et les résultats
- ❑ Renommez votre fichier Excel d'entrée : DHIS2_dataset_Countryname
 - ❑ par exemple, DHIS2_dataset_Kenya
- ❑ Lisez et suivez les instructions en haut de chaque do.file
- ❑ Effectuer les changements requis selon les instructions
 - ❑ Modifier le répertoire du dossier de travail en fonction de l'emplacement des données d'entrée sur votre ordinateur
 - ❑ Définir le nom du pays
 - ❑ Définir la première année (2017 par défaut)
 - ❑ Définir l'année dernière (2021 par défaut)
 - ❑ Définir le seuil pour les rapports faibles (90 par défaut)
 - ❑ Définir le facteur d'ajustement pour les rapports incomplets (valeur k) par service (0,25 par défaut)

Livrables escomptés de #1 pour le rapport/poster

Organisation administrative

Nombre de provinces / régions / départements
 Nombre of districts

Etablissements de santé

Nombre d'établissements de santé dans le pays
 Disponibilité des données sur le personnel de santé
 Disponibilité des données sur les lits dans les hôpitaux

Période des données sur les établissements de santé

Premier mois et année avec des données
 Dernier mois et année avec des données

Enquête ménage (3 plus récentes enquêtes)	
Nom de l'enquête	Année

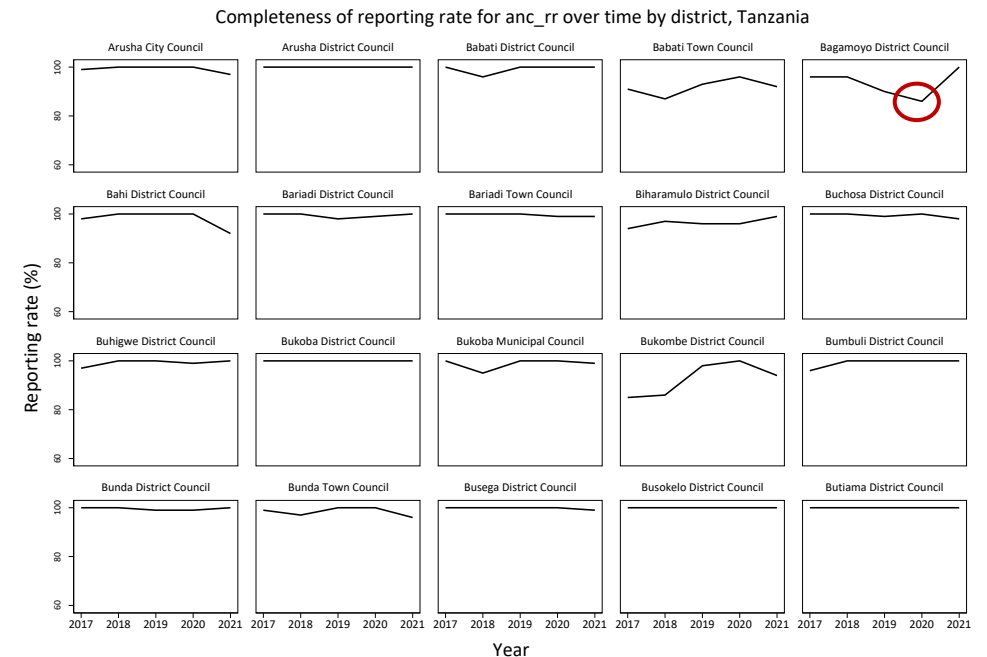
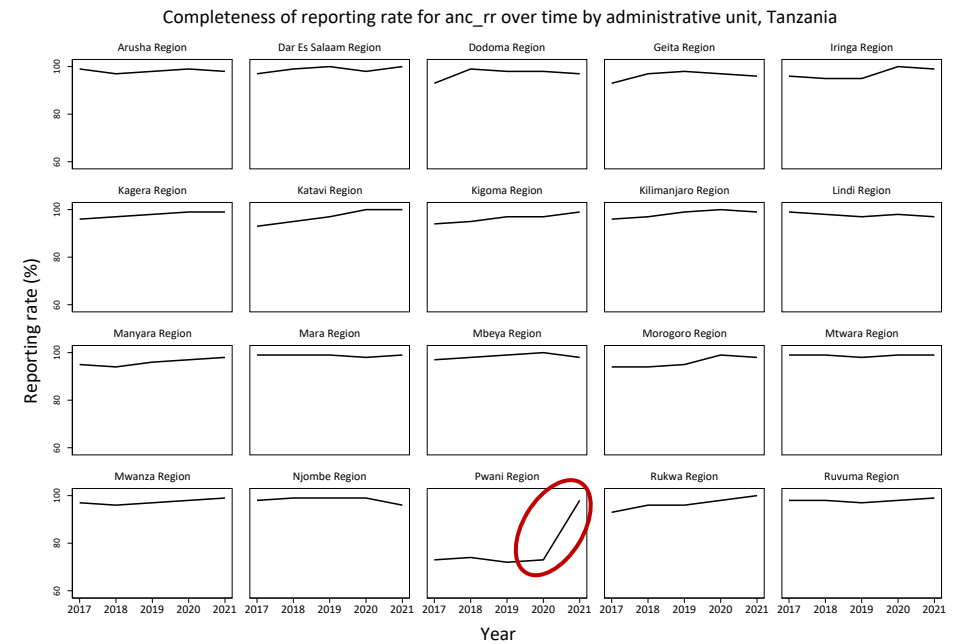
Indicateurs disponibles avec des données de routine pour les analyses

Consultation prénatale 1ère visite
 Consultation prénatale 4ème visite
 TPI 2nde dose (paludisme)
 Accouchement dans structure de santé ou assisté par personnel qualifié
 Césarienne
 Soins postnatal
 Planification familiale – nouvelle et révisite
 BCG vaccination
 Pentavalent / DTCoq 1ère dose
 Pentavalent / DTCoq 3ème dose
 Rougeole vaccination
 Morts-nés (frais / macéré)
 Décès maternels
 Consultations moins de 5 ans
 Hospitalisations moins de 5 ans
 Décès des moins de 5 ans dans établissement de santé

Population projection data in DHIS2	
Indicator	Year
Total population for every year	Yes / No
Live births for every year	Yes / No
Population under 1 year for every year	Yes / No

SESSION #2 – ÉVALUATION ET AJUSTEMENT DES RAPPORTAGES INCOMPLETS

Évaluation et ajustement des rapportages incomplets par les établissements de santé



Évaluation et ajustement des reportages incomplets

- ❑ Des reportages incomplets peuvent avoir un effet majeur sur les niveaux et les tendances de la couverture et d'autres statistiques provenant des établissements de santé
 - Si nous ne tenons pas compte de la complétude, nous supposons qu'aucun service n'est fourni par les établissements non déclarants
- ❑ Il y a deux niveaux d'incomplétude à considérer
 - Au niveau de l'établissement : établissements qui n'ont pas communiqué de données
 - Niveau de service : niveau de service attendu parmi les établissements qui n'ont pas communiqué de données
- ❑ L'ajustement doit tenir compte des deux niveaux d'incomplétude
- ❑ L'ajustement dépend des hypothèses sur le nombre d'extrants de services (soins de grossesse, vaccinations, etc.) fournis dans les établissements non déclarants par rapport à ceux qui ont déclaré.
 - Aucun? Quelques? Environ la moitié des installations déclarantes ? La même chose que les installations de déclaration ?

Évaluation et ajustement des rapportages incomplets

- La formule pour l'ajustement de rapportages incomplètes est comme suit :

$$N_{(ajusté)} = N_{rapporté} + N_{rapporté} * (1/(c) - 1) * k$$

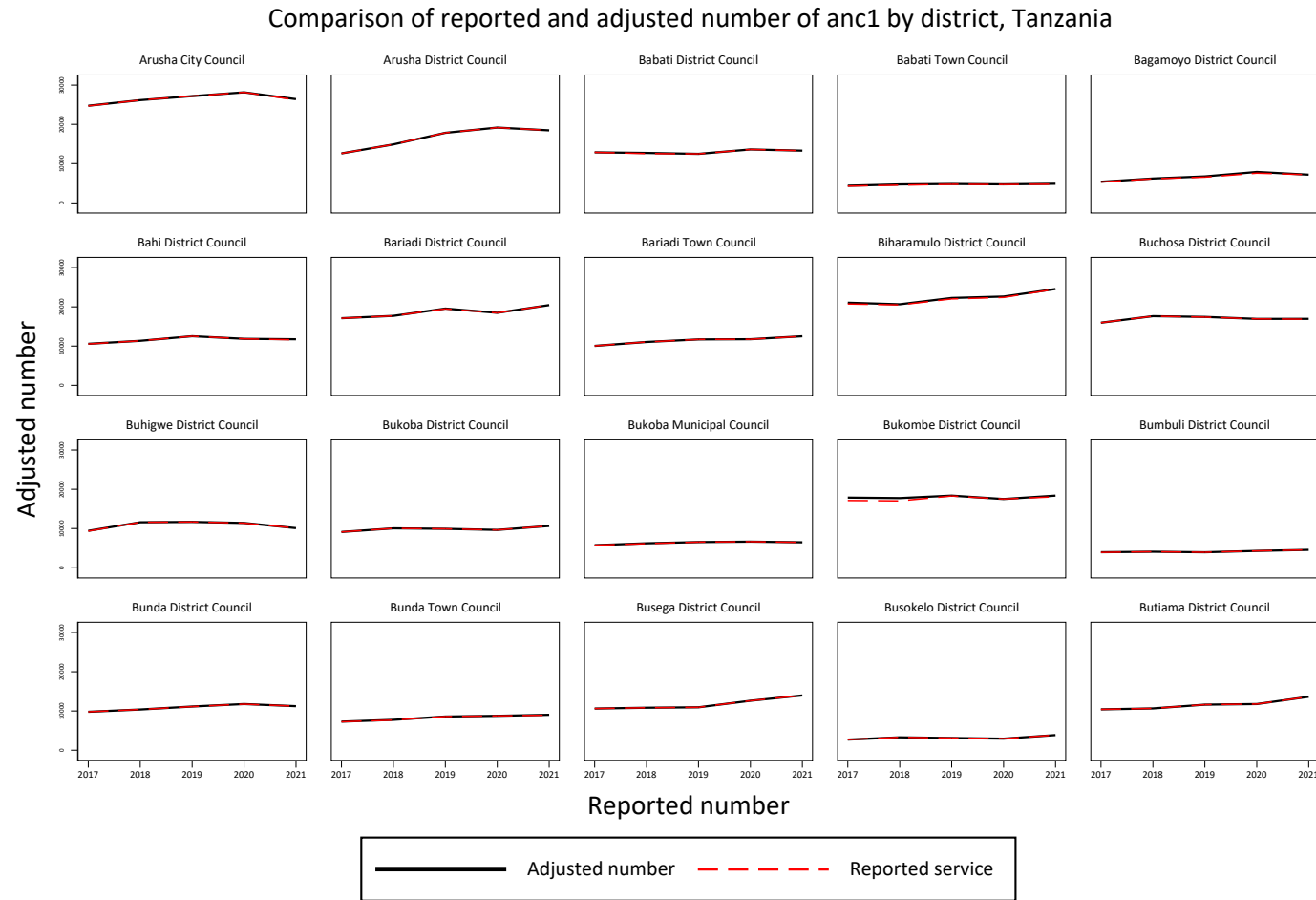
où N=nombre de services prestés, c=complétude des rapports, k=facteur d'ajustement.

- k=0 pas de services dans les établissements non déclarants
- k=0,25 certains services, mais beaucoup plus faible que les établissements ayant rapporté
- k=0,5 la moitié du taux par rapport aux établissements ayant rapporté
- k=0,75 presque autant que les établissements ayant rapporté
- k=1,0 même taux de services que les établissements ayant rapporté

Évaluation et ajustement des rapportages incomplets

- ❑ Le choix de k est basé sur des considérations d'experts en analyse dans le pays :
 - Le rapportage est-il lié aux approvisionnements ? (par exemple, les vaccins)
 - Le fait de ne pas rapporter reflète-il des ruptures de stocks ?
 - Quelle est la qualité des rapportages des établissements privés et quel rôle jouent-ils dans la prestation de services tels que les accouchements ou la vaccination ?
 - Les grands hôpitaux procèdent-ils aux rapportages et quel rôle jouent-ils dans la prestation de services ?

Évaluation et ajustement des reportages incomplets



- ❑ Revérifiez les effets de l'ajustement en comparant le nombre rapporté au nombre ajusté pour les reportages incomplets

**SESSION #3 – ÉVALUATION ET CORRECTIONS DES VALEURS
ABERRANTES EXTRÊMES ET COHÉRENCE INTERNE ; TABLEAU DE
BORD DE LA QUALITÉ DES DONNÉES**

Contrôles de cohérence interne dans le temps & ajustements des valeurs aberrantes extrêmes

- ❑ En général, nous nous attendons à une variation limitée d'une année à l'autre du nombre d'interventions signalées, en particulier pour les interventions à couverture élevée (par exemple, CPN1,DTCoq1)
- ❑ S'il y a beaucoup de variation et qu'il n'y a pas de bonne explication, les données sont considérées comme « polluées » et peuvent présenter de graves problèmes de qualité des données
- ❑ Pour les données annuelles, une valeur aberrante extrême est définie comme tout nombre dans l'ensemble de données supérieur ou inférieur à 5 écart-type par rapport à l'écart absolu médian (MAD) calculé à partir des 3 années précédentes. Ils ont été corrigés en imputant une valeur basée sur la valeur médiane de l'année civile. Imputation similaire pour les valeurs manquantes.

Valeur aberrante si

$$\text{Médiane} - 1.4826 * 5 * \text{MAD} < X_i > \text{Médiane} + 1.4826 * 5 * \text{MAD}^{\ast}$$

Borne inférieure = Médiane - 1,4826 * 3 * MAD

Borne supérieure = Médiane + 1,4826 * 3 * MAD

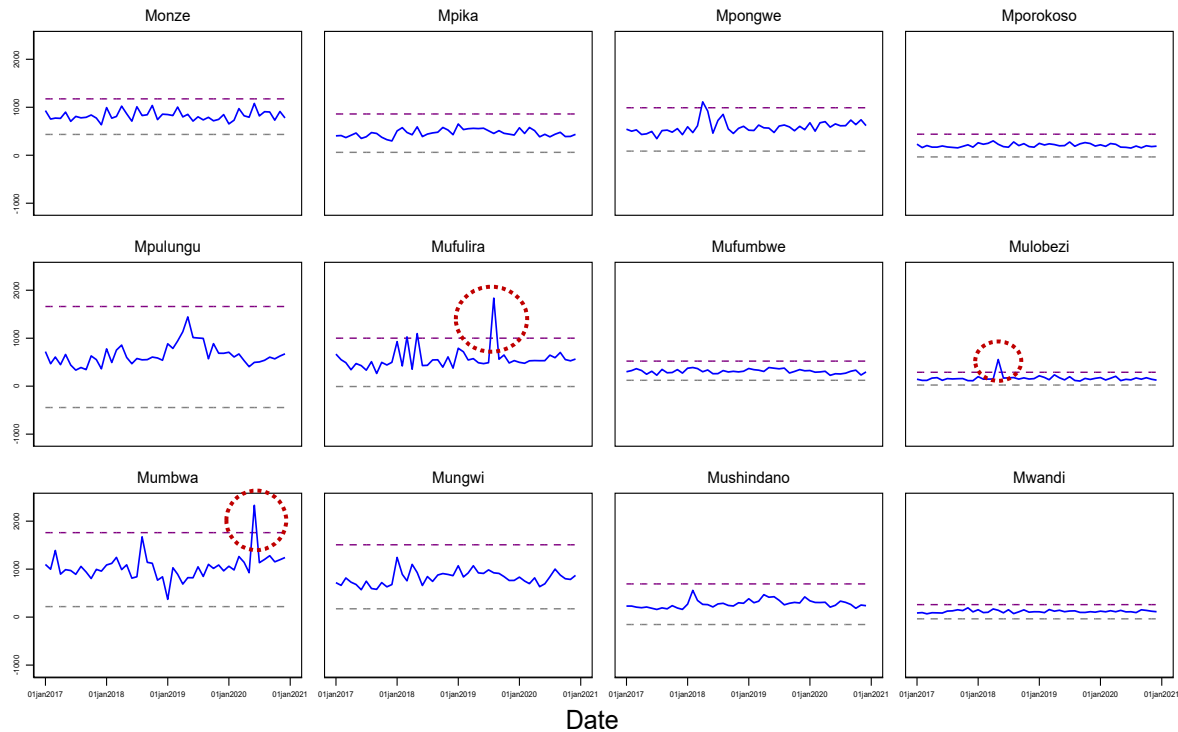
où X_j est la valeur de l'observation pour une période donnée (année) et le MAD est défini comme l'écart absolu median (MAD = médian(| $X_j - X_{\sim}$ |)), où X_{\sim} est la médiane des trois années précédentes).

✱

Pour une distribution normale, un écart type par rapport à la moyenne est d'environ 1,4826 MAD (règle de Hampel X84)

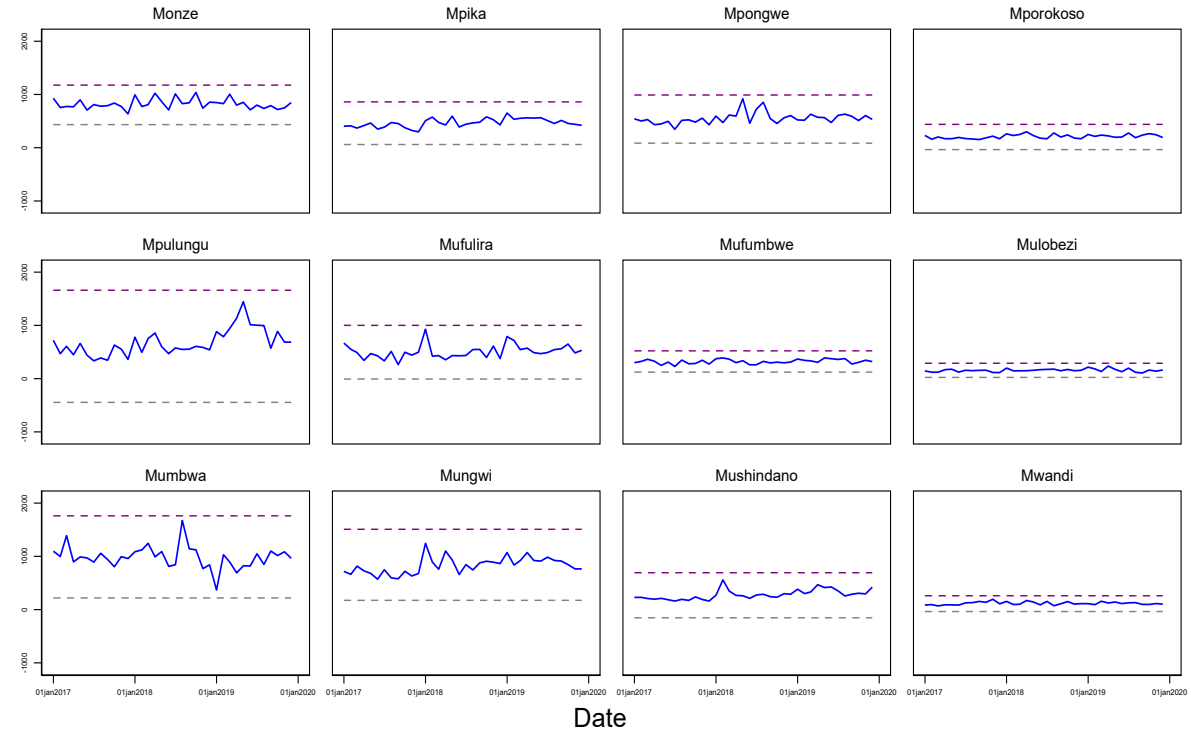
Contrôles de cohérence interne dans le temps & ajustements des valeurs aberrantes extrêmes

Assessment of anc1 outliers over time by district, Zambia



Outlier > ±1.482*5*MAD (5 std from median)

anc1 after correcting outliers over time by district, Zambia



Outlier > ±1.482*5*MAD (5 std from median)

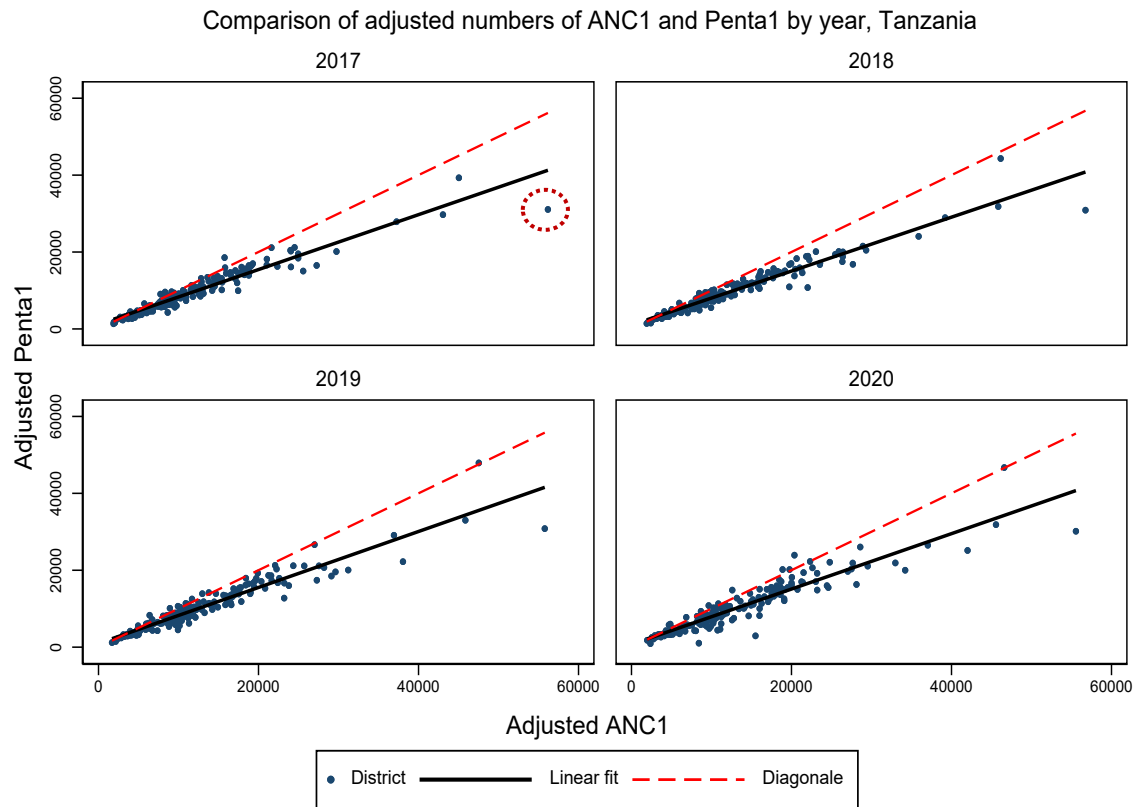
Cohérence interne entre services

- ❑ La cohérence interne des services est généralement évaluée entre le nombre d'événements pour CPN1 à Penta1 et pour Penta1 à Penta3
- ❑ Cohérence calculée comme le ratio de nombre de ANC1 / Penta1 ; et ratio de nombre de Penta1/Penta3
 - ❑ Qualité des données discutable lorsque le rapport est en dehors 1,0 et 1,5
 - ❑ % de districts avec un ratio CPN1/Penta1 (Penta1/Penta3) entre 1,0 et 1,25 (seuil utilisé de 85 % ou plus)
- ❑ Cette mesure peut également être calculée comme la différence absolue entre les ratios attendus et rapportés des deux indicateurs
 - Plus la différence est grande, plus la qualité des données déclarées est faible (inférieure à 5 → bonne qualité; 5-14.9 → modéré; 15+ → mauvaise qualité)
 - Données externes requises de l'enquête auprès des ménages pour calculer le ratio attendu
 - Une certaine variation attendue car l'enquête comporte une erreur d'échantillonnage qui affecte les valeurs Penta1 - Penta3, de sorte que le rapport attendu n'est qu'une indication générale

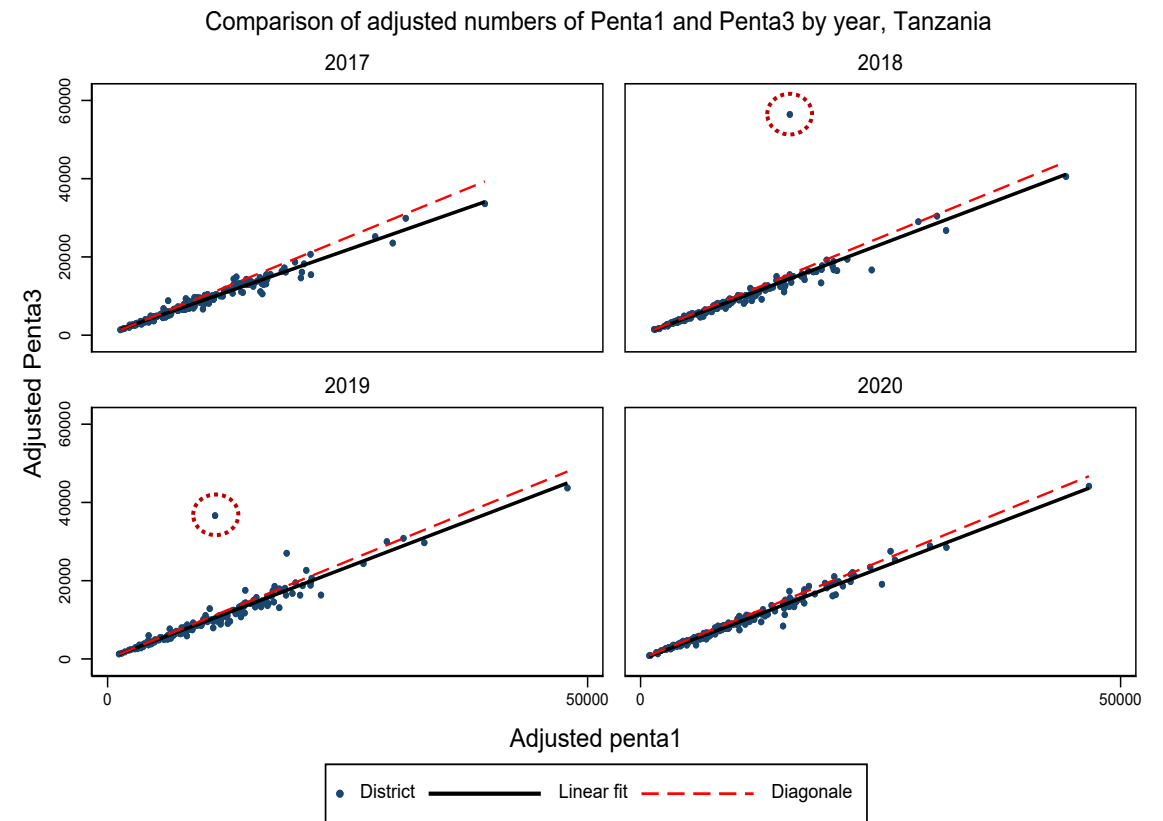
Cohérence interne entre services

- ❑ Les diagrammes de dispersion avec une intervention sur l'axe X et une sur l'axe Y peuvent également être utiles pour examiner les valeurs aberrantes du district
 - Fournit un bon visuel de la corrélation entre les données de service
 - Le coefficient de corrélation de Pearson peut être ajouté
 - La présence de valeurs aberrantes potentielles peut être facilement vérifiée à partir du graphique

Cohérence interne entre services



R-squared=0.9027



R-squared=0.8824

- ❑ Les points de dispersion montrent les districts pour lesquels la qualité des données est discutable car un bon alignement entre les points de données est attendu
- ❑ Une vérification approfondie des données est recommandée pour confirmer toute incohérence qui devrait être corrigée à partir d'une évaluation claire

Contrôles de cohérence interne dans le temps & ajustements des valeurs aberrantes extrêmes

- ❑ Les fluctuations de la prestation de services ne sont pas nécessairement le résultat de problèmes de qualité des données
- ❑ Important à considérer :
 - Changements dans les activités programmatiques telles que l'intensification de la campagne, les ruptures de stock (par exemple, la vaccination)
 - Caractère saisonnier des maladies pour les consultations ou les hospitalisations (par exemple, le paludisme)
 - Impact des facteurs contextuels tels que la migration de la population (par exemple, les déplacements récents ou les réfugiés)
- ❑ Bonne connaissance du contexte et des données du programme pour interpréter les données sur les tendances
- ❑ Tenir compte de l'augmentation annuelle prévue du service en raison de la croissance démographique et de l'amélioration de la couverture des services

Livrables escomptés de #2 & 3 pour le rapport/poster

Rapport sur les ajustements et les corrections

- Rapporter le seuil utilisé pour évaluer les faibles taux de rapportage
 - Si la valeur par défaut est utilisée, alors rapporter cela et expliquer ce que cela implique pour le lecteur.
- Rapporter les facteurs d'ajustements (appelé aussi factor k) qui ont été utilisés pour ajuster les reportages incomplets (par service).
 - Si le facteur par défaut a été utilisé, alors rapporter cela et expliquer ce que cela implique pour le lecteur.
- Fournir une brève description des corrections qui ont été faites pour les valeurs extrêmes aberrantes et les incohérences.
 - Quelles valeurs ont été utilisées pour corriger les valeurs extrêmes aberrantes et les incohérences?
- Cela pourrait inclure un graphique pour illustrer l'impact de la correction au niveau district ou région or national.

Livrables escomptés de #2 & 3 pour le rapport/poster

Tableau: Tableau synthèse des scores de qualité des données aux niveaux national et infranational, 2017-2021

- Résultats des codes Stata #3 peuvent être utilisés pour développer un tableau des scores annuels des indicateurs de qualité des données
- Modifications possibles du modèle de tableau pour améliorer sa présentation
- Des graphiques sélectionnées des qualité des données peuvent être ajoutés
- Ajouter des commentaires par rapport à l'interprétation des données de la table de qualité des données, par exemple :
 - Y-a-t-il une configuration quelconque de la qualité des données par année pour laquelle il y a une explication? (inclure l'explication)
 - Y-a-t-il certaines régions ou autres unités administratives infranationales pour lesquelles les données sont particulièrement problématiques?
 - Y-a-t-il certain formulaires de rapportage des services (ex. CPN, travail et accouchement, vaccination) qui sont problématiques?

Exemple de tableau synthèse de la qualité des données

		2017	2018	2019	2020	2021
1	Complétude des reportages mensuels des structures de santé (vert >90%)					
1a	% de reportages mensuels escomptés (moyenne, nationale)*	94	95	95	96	97
1b	% de districts avec des complétude de rapportage >= 90%*	82	86	88	91	91
1c	% de districts avec structure de santé sans valeurs mensuelles manquantes dans l'année *	100	100	100	100	99
2	Valeurs aberrantes extrêmes (vert > 95%)					
2a	% de valeurs mensuelles qui ne sont pas valeurs mensuelles aberrantes extrêmes (moyenne, national)*	100	100	100	99	97
2b	% de districts sans valeurs aberrantes extrêmes dans l'année*	97	98	97	96	86
3	Cohérence interne de rapportage annuelle (vert>85%)					
	Ratio nombre ANC1 – Penta1 (national)	1.06	1.10	1.09	1.20	1.18
3a	% de districts avec ratio ANC1-Penta1 entre 1,0 and 1,5	63	77	68	69	70
	Ratio nombre Penta1 – Penta3 (national)	1.11	1.10	1.09	1.09	1.10
3b	% de districts avec ratio Penta1-Penta3 entre 1,0 and 1,5	91	93	85	88	88
	Score annuel de qualité des données (moyenne indicateur 1a à 3b)	90	93	91	91	90

*Moyenne des services de CPN, accouchement, vaccination et consultations