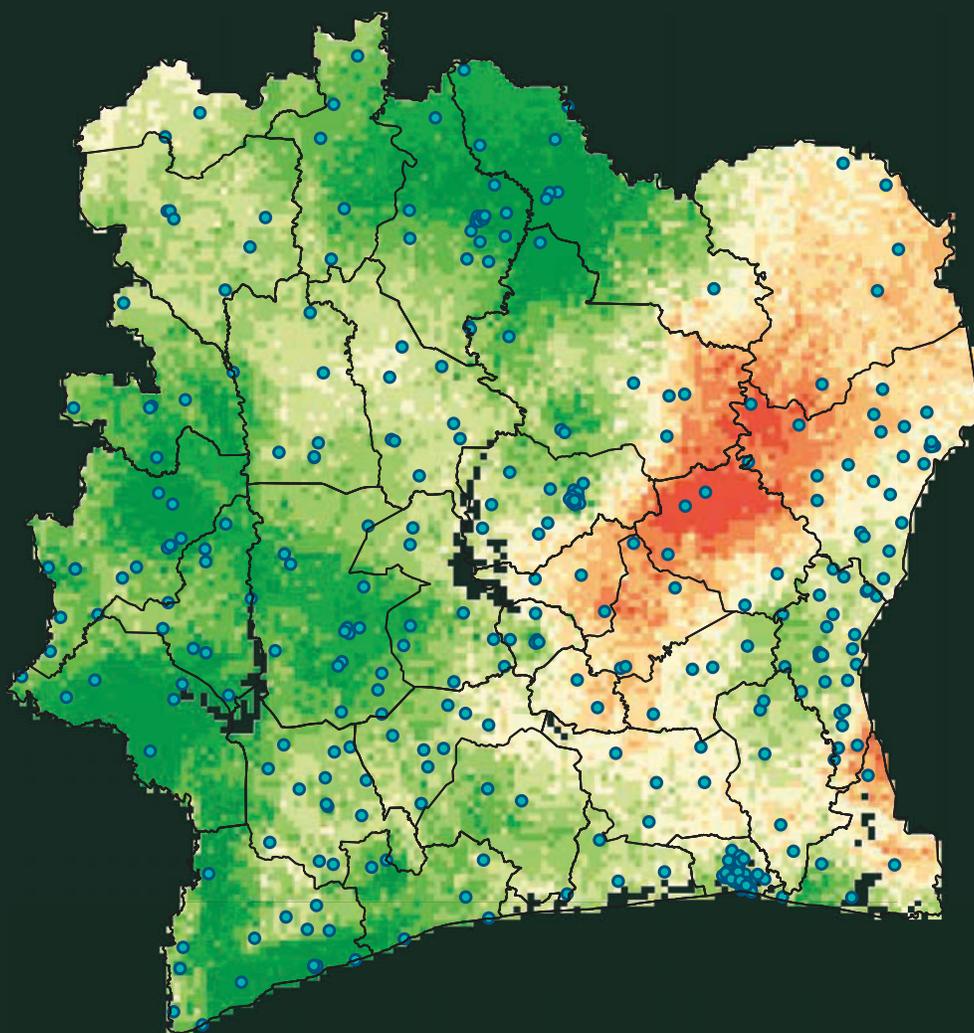


Réduction des Coûts et des Obstacles aux Evaluations grâce aux données géospatiales :

Nouvelles méthodes appliquées en Côte d'Ivoire pour le VIH/SIDA

Livre Blanc pour l'USAID | PEPFAR | Côte d'Ivoire
AidData \ William & Mary

Ariel BenYishay et Katherine Nolan
Juillet 2020



USAID
FROM THE AMERICAN PEOPLE



AIDDATA
A Research Lab at William & Mary

**Réduction des Coûts et des Obstacles aux Evaluations grâce aux
données géospatiales :
Nouvelles méthodes appliquées en Côte d'Ivoire pour le VIH/SIDA**

***Livre Blanc pour l'USAID | PEPFAR | Côte d'Ivoire
AidData \ William & Mary***

Ariel BenYishay et Katherine Nolan

Juillet 2020

Résumé

Nos nouvelles méthodes ont pour objectif d'aider les évaluateurs, les partenaires techniques et financiers et les maîtres d'œuvre des projets, à comprendre qu'ils ont la possibilité de combiner plusieurs séries d'enquêtes avec des données spatiales dans le cadre de l'évaluation des projets, notamment dans le secteur de la santé. En effet, ces nouvelles méthodes permettent d'utiliser les données géospatiales interpolées (« prédites ») à la place d'une ou de plusieurs séries de données d'enquêtes primaires. Nous stimulons la puissance statistique dans les cas les plus courants ce qui entraîne d'autres options de configuration de ces méthodes. Dans leur application, nous nous concentrons sur le cas des indicateurs du VIH/SIDA en Côte d'Ivoire, dérivés des données géolocalisées EDS et MCIS, ainsi que sur les couches interpolées de l'Institut pour l'Evaluation et les Mesures de Santé (IHME). Nous trouvons que la combinaison des données de base sur les surfaces prédites avec les données d'enquête de suivi donne la précision la plus statistique qui permet aux évaluateurs de détecter les effets les plus minimes des traitements. Ces configurations découlent d'évaluations rétroactives et sont ainsi utilisables dans divers contextes. Par ailleurs, elles peuvent aider les évaluateurs à éliminer les obstacles à la réalisation d'évaluations d'impact tels que les contraintes de coûts qui entravent souvent la collecte des données spécifiques aux projets.

Table des matières

Résumé	1
Introduction	4
Paysage des données relatives au VIH/SIDA	6
Littérature connexe	8
Vue d'ensemble des données	9
Analyse des données	13
Leçons sur l'intégration des données	17
Résultats	18
Discussion	19
Bibliographie	23

Tableaux et graphiques

Résumé	1
Introduction	4
Paysage des données relatives au VIH/SIDA	6
Littérature connexe	8
Vue d'ensemble des données	9
Analyse des données	13
Leçons sur l'intégration des données	17
Résultats	18
Discussion	19
Bibliographie	23

Liste des sigles et acronymes

CDC - Centre de contrôle des maladies

DPT - Diphtérie, Coqueluche et Tétanos

DHIS2 - Système d'information sur la santé au niveau des districts

SIGDEP 2.0 - Système d'information de gestion des dossiers électroniques des patients

eLIMS - Système d'information de gestion logistique

OpenELIS - Système ouvert de gestion de l'information des laboratoires

OEV - Orphelins et enfants vulnérables

ENSEA - École nationale supérieure de statistique et d'économie appliquée

MICS - Enquêtes par grappes avec indicateurs multiples

FFP - Food for Peace

EDS - Enquêtes démographiques et de santé

IHME - Institut pour l'évaluation et les métriques de santé

WALA - The Wellness and Agriculture for Life Advancement program

VIH/SIDA - Virus de l'immunodéficience humaine et Syndrome d'immunodéficience acquise

ONUSIDA - Programme commun des Nations Unies sur le VIH/SIDA.

FFM- Facilité de financement mondiale

Remerciements

Ce projet a été financé par USAID Côte d'Ivoire aux termes d'un accord de coopération (AID-OAA-A-12-00096) entre le Laboratoire de développement mondial de l'USAID et AidData du College of William & Mary dans le cadre du programme Réseau de recherche de solutions dans l'enseignement supérieur (HESN). Nous apprécions vivement leur soutien. Nous voudrions également adresser nos remerciements à William Ouattara et Anastase Mian pour leur travail acharné et leurs contributions à l'élaboration de ce document.

Introduction

Pour comprendre les conditions de vie des populations et évaluer les impacts des programmes visant à les améliorer, les gouvernements et les partenaires au développement se sont souvent appuyés sur les enquêtes quantitatives et qualitatives d'échantillon de la population. Cet échantillon porte parfois sur des types d'individus spécifiques, ou est censé représenter l'ensemble de la population. Ces enquêtes sont considérées comme étant le meilleur moyen d'évaluer le bien-être global des populations et les effets éventuels d'un projet donné.

Les ressources d'aide au développement sont limitées et, même si ces enquêtes sont importantes, elles peuvent être très coûteuses. Par exemple, le coût de la réalisation d'une série type d'Enquêtes démographiques et de santé (EDS), même dans un pays de taille moyenne tourne autour d'un million de dollars. Dans les pays très peuplés, comme que la Côte d'Ivoire, ce coût dépasse généralement 1,2 million de dollars (Jerven 2017). Le coût élevé de ces enquêtes justifie le fait qu'elles soient menées ponctuellement (généralement tous les 3 à 5 ans seulement) sur des échantillons représentatifs au niveau national. De plus, ces coûts peuvent être prohibitifs pour les évaluations d'impact potentiel des programmes de développement, où seule une infime part de ces programmes est couverte par les collectes de données d'enquêtes. Ainsi, il apparaît crucial de trouver d'autres moyens de recueillir des données de qualité auprès des population, à moindre coût.

Une première suggestion de réduction des coûts des enquêtes est s'appuye sur les données géolocalisées (secondaires) existantes qui sont de plus en plus disponibles, notamment les enquêtes auprès des ménages au niveau national ou dans le cadre d'un projet, les données administratives et les données de télédétection (par satellite et d'autres détecteurs). Ces données peuvent servir à améliorer la précision des estimations effectuées à partir des nouvelles données primaires, réduisant ainsi les tailles des échantillons requises dans le cadre des initiatives de collecte de données et, par conséquent, le coût et l'effort qui y sont liés. De plus, dans les cadres d'évaluation, ces sources de données supplémentaires peuvent, parfois, être

utilisées à la place des activités de collecte de données primaires si celles-ci n'existent pas, particulièrement en tant que données de base qui aident à constituer des groupes de comparaison qui sont similaires aux groupes traités.

Cependant, l'utilisation des données secondaires existantes à ces fins pose un certain nombre de difficultés. D'abord, les enquêtes de qualité, telles que les EDS et les Enquêtes par grappes avec indicateurs multiples (MICS), sont disponibles uniquement pour des échantillons de la population. Ces échantillons sont souvent éparpillés dans un pays ou une zone donnée et sont seulement représentatifs à l'échelle nationale. Par conséquent, plusieurs localités d'un pays ne sont pas directement couvertes par une série EDS. Ensuite, les données administratives (telles que les données sur les établissements de santé) et les données de télédétection (telles que l'imagerie satellite) pourraient ne pas saisir les résultats d'intérêt exacts dans une étude donnée, particulièrement dans les secteurs de la santé, de l'éducation et des secteurs sociaux connexes. Cependant, les récents progrès réalisés pour relever ces deux défis permettent d'accéder à des solutions : par exemple, les nouvelles couches interpolées produites par l'Institut pour l'évaluation et les métriques de santé (IHME) fournissent une estimation des principales mesures de santé à l'échelle infranationale, ce qui nous permet de compenser les populations manquantes non couvertes dans les enquêtes par sondage.

Notre objectif est d'aider les évaluateurs, les partenaires techniques et financiers, et les maîtres d'œuvre des projets, à comprendre qu'ils ont la possibilité de combiner plusieurs séries d'enquêtes avec des données spatiales dans le cadre de l'évaluation des projets. Nous présentons les nouvelles méthodes géospatiales qui permettent l'interpolation géospatiale de ces sous-ensembles et leur attachement aux données des enquêtes de collecte des données primaires. Nous stimulons l'efficacité statistique dans les cas les plus courants qui entraînent d'autres options de configuration de ces méthodes. Nous nous concentrons sur les cas où la collecte de données nouvelles et primaires pour les séries de base et de suivi n'est pas réalisable ou n'entre pas dans le budget. Nous procédons, ensuite, à une évaluation globale et une comparaison de la puissance statistique qu'offrent les autres configurations. Dans le cadre de

l'application de ces méthodes, nous nous concentrons sur le cas des indicateurs du VIH/SIDA en Côte d'Ivoire, dérivés des données géolocalisées EDS et MCIS, ainsi que sur les couches interpolées de l'Institut pour l'évaluation et les métriques de santé (IHME). Ce cadre est particulièrement important pour notre application car il offre un ensemble de configurations de données et une gamme d'initiatives de développement visant à réduire le VIH/SIDA au sein d'une large population.

Les résultats de nos simulations laissent penser que la combinaison des mesures des résultats de base à partir des surfaces prédites avec des surfaces de suivi tirées des données primaires nouvellement collectées donne une plus grande précision statistique. De toute évidence, ces projets sont plus faciles à réaliser que les autres qui s'appuient sur la collecte des données d'enquêtes de base, ouvrant la porte à beaucoup plus de projets d'évaluation réalisés rétrospectivement. Ces résultats montrent également comment les surfaces prédites nouvellement créées peuvent être utilisées pour aider à récupérer les scénarios « de base manquants » (dans le cas où aucune donnée d'enquête de base n'a été réalisée dans les sites traités et de comparaison) - un aspect important des évaluations d'impact géospatiales.

Paysage des données relatives au VIH/SIDA

Au cours de ces deux dernières décennies, des données variées et détaillées sur le secteur de la santé ont été collectées et compilées dans un certain nombre de plateformes riches en données, en Côte d'Ivoire. Ces sources de données locales sont entre autres, SISD2, SIGDEP 2.0, eLIMS, OpenELIS et la base de données OEV. Elles portent sur un large éventail de thèmes, notamment le VIH/SIDA, la santé maternelle, les informations cliniques, les stocks de médicaments, les informations sur les orphelins et les enfants vulnérables, etc. Il existe également des organisations privées ou des projets qui collectent des données auprès des ménages. Malgré la grande quantité de données qui sont collectées, la qualité des données reste une préoccupation (il est très difficile d'avoir des données exhaustives et fiables) et une bonne partie des données reste difficilement accessible pour les besoins des programmes et des évaluations.

Le Système d'information de gestion des dossiers électroniques des patients (SIGDEP 2.0), le Système d'information de gestion logistique (eLMIS), le Système ouvert de gestion de l'information des laboratoires (OpenELIS) et la base de données Orphelins et enfants vulnérables (OEV) font face à des problèmes de plusieurs ordres : (a) la faible capacité et le temps insuffisant pour la collecte des données dans les établissements ; (b) un processus de collecte ad hoc des données non-cliniques ; (c) des procédures de validation inefficaces et coûteuses, particulièrement lors des ateliers de validation des données ; et (d) l'absence de TI dans les zones reculées (Desai, 2018b). Face à ces défis, il peut être difficile pour les évaluateurs d'accéder à cette masse de données notamment en raison des préoccupations liées à la publication de données inexactes et/ou incomplètes (Desai, 2018b). Parmi les autres obstacles à l'accès des données, il y a la question liée à la protection de la vie privée pour les patients, les médecins et les cliniques.

Les principaux acteurs du système de données sur la santé sont, entre autres, les ministères, les établissements privés, et les universités telles que l'École nationale supérieure de statistique et d'économie appliquée (ENSEA) qui aident au renforcement des capacités des professionnels de la santé et des ministères en matière d'alphabétisation pour les données de base. Tous les grands partenaires au développement investissent également des ressources dans le renforcement des systèmes de santé et le développement de nouveaux outils (par ex, CDC et openEllis, ONUSIDA et la Situation Room ou la Banque mondiale et la carte sanitaire) (Desai, 2018b). Selon les récentes notes conceptuelles adressées au Fonds Mondial et après le lancement de la Facilité de Financement Mondiale (FFM), le financement des activités de production de données devrait s'améliorer.

Malgré le nombre d'acteurs, l'accroissement des financements et des investissements dans le système de santé de la Côte d'Ivoire, la volonté politique du gouvernement d'ouvrir les données, d'importants défis doivent être relevés par les organisations qui cherchent à rendre publics les ensembles de données sur la santé et pour les évaluateurs qui veulent y accéder (Desai, 2018b).

Actuellement, l'usage des données est également limité en raison des insuffisances liées à leur qualité, l'opportunité et l'utilisabilité des informations géospatiales (Barlow et al. 2017). Dans l'ensemble, la Côte d'Ivoire possède une masse de données sur la santé qui pourraient être utiles pour les évaluateurs des projets de santé. Cependant, malgré ces activités de collecte de ces données abondantes, il y a relativement peu de possibilités offertes aux évaluateurs qui cherchent à réaliser des évaluations d'impact sur les programmes de lutte contre le VIH/SIDA.

Littérature connexe

Au cours des cinq dernières années, les surfaces (interpolées) prédictives sont devenues de plus en plus répandues, aussi bien en termes de pays couverts que de diversité des résultats qu'elles produisent sur la santé et la population. Ces surfaces incluent maintenant les estimations de la population (Wardrop et al., 2018; Lloyd et al., 2019), la mortalité infantile (Gething et al., 2011; Akachi et al. 2011; Golding et al., 2017; Utazi et al., 2018; Bird et al., 2016), le temps du trajet pour aller aux services de santé et leur utilisation (Broer et al., 2018; Ruktanonchai et al. 2018; Ruktanonchai et al., 2016), les naissances et les grossesses (James et al., 2018), les inégalités en matière de santé chez les enfants (Youkavitch et al., 2018), les indicateurs de l'enquête EDS (Gething et al., 2015), les troubles de croissance chez les enfants (Osgood-Zimmerman et al., 2018), le niveau scolaire (Graetz et al., 2018), les maladies (Giorgi et al., 2018), les vaccins (Utazi et al. 2019) et d'autres indicateurs généraux en matière de santé publique (Diggle et Giorgi, 2019; Diggle et Giorgi, 2016). Il y a eu également un riche débat à travers plusieurs disciplines sur l'histoire (Mahy, et al. 2017), les méthodes (Aids, 2016), l'exactitude (Dube et al. 2019) et les forces et faiblesses (Carter, 2018; Leyk et al. 2019) de ces surfaces.

Par ailleurs, un certain nombre d'ouvrages se concentrent plus sur le VIH/SIDA et portent sur la variation de la prévalence du VIH au niveau infranational, dans des pays particuliers ou dans des ensembles de pays (Anderson et al., 2014; Kleinschmidt et al., 2007; Kandala et al., 2012; Larmarange et Bendaud, 2014; Okana et Blowe, 2016; Carrel et al., 2016; Coburn et al., 2017; Cuadros et al., 2017; Meyer-Rath et al., 2018; TB et al., 2016). En plus de ces travaux spécifiques

aux pays, les chercheurs de l'IHME ont développé des couches de prédiction pour un ensemble complet d'indicateurs pour toute l'Afrique subsaharienne (Dwyer-Lindgren et al. 2019).

Au-delà de la création des surfaces et de leur utilisation pour le suivi des indicateurs spécifiques en matière de santé, les surfaces prédictives n'ont pas été beaucoup utilisées dans les évaluations, particulièrement dans le domaine de la lutte contre le VIH/SIDA. Une utilisation éventuelle des surfaces prédictives dans les activités d'évaluation permet de tirer rétrospectivement les groupes de comparaison. Par exemple, dans le sud du Malawi, BenYishay et al. (2019) a utilisé une surface prédictive qui a été créée sur la base des estimations de l'enquête EDS dans le but de sélectionner les villages qui n'ont pas bénéficié du programme Wellness and Agriculture for Life Advancement (WALA) financé par Food For Peace (FFP), mais qui avaient des similitudes au départ dans la mesure où ces villages n'avaient pas bénéficié du programme. Par la suite, tous ces villages de comparaison et traités ont fait l'objet d'une enquête cinq ans après la fin du programme, ce qui a permis de comprendre l'impact durable du programme. Au-delà de ces utilisations, des évaluations limitées se sont servies des surfaces prédictives pour étudier les résultats des programmes. Étant donné la durée des évaluations et l'augmentation récente des surfaces prédictives disponibles et des méthodes d'évaluation géospatiales connexes, des changements pourraient être notés dans le futur. Il y a un domaine connexe et de plus en plus large consacré aux méthodes géospatiales d'évaluation d'impact qui utilisent la mesure des résultats géospatiaux liée aux données d'intervention pour récupérer les estimations quasi-expérimentales des répercussions causales (c.f. [BenYishay et al 2017](#)). Comme le montre ce document, les évaluateurs ont plusieurs possibilités d'utiliser les surfaces, soit seules soit combinées avec des données de l'enquête pour étudier les résultats du projet.

Vue d'ensemble des données

Pour répondre à ces questions, nous avons utilisé une combinaison des données 2015 de MICS, des données 2012 de l'EDS et des données de l'IHME pour la Côte d'Ivoire. Nous avons utilisé

ces sources de données car elles sont axées sur la santé, géolocalisées et accessibles au public. Nous nous sommes concentrés sur un ensemble de mesures communes disponibles à travers plusieurs sources, y compris le taux de prévalence du VIH fourni par les données EDS, ainsi que les premiers taux de vaccination contre la diphtérie, la coqueluche et le tétanos (DPT 1) et les questions liées au VIH dans les données MICS. Aussi bien les estimations du taux de vaccination DPT1 que les estimations sur le VIH ont été tirées des ensembles de données entre 2012 et 2015.

Les données de l'EDS ont été recueillies en 2011/2012 par le Programme d'enquête démographique et de santé. Cette enquête a couvert un large éventail de thèmes sur la santé, tels que les vaccins, la violence domestique, les connaissances et le test sur le VIH, et la mortalité infantile. Cette enquête a été géoréférencée au niveau du groupe. Afin d'éviter le ré-identification des répondants de l'enquête, les administrateurs de l'enquête ont déplacé de manière aléatoire les positions latitude/longitude GPS pour toutes les enquêtes. Afin d'utiliser les données sur le VIH, l'équipe a créé un certain nombre d'individus séropositifs au niveau du groupe. De même, nous avons calculé un taux de vaccination DPT1 au niveau du groupe.¹

Les données MICS étaient collectées en 2015/2016 à la place de l'enquête EDS et géocodées au niveau du groupe. Cette enquête comportait également un large éventail de questions relatives à la santé, allant de la nutrition et la vaccination infantile aux pratiques liées à la santé de la mère et des ménages. Même si l'équipe de recherche n'a pas réalisé de test du VIH dans le cadre de l'enquête, elle a posé un certain nombre de questions sur le VIH/SIDA. Pour estimer la prévalence du VIH, nous avons utilisé un index des questions relatives au VIH. Il s'agissait de savoir si le/la répondant(e) était marié(e) ou vivait avec un(e) partenaire en tant que mari et femme, s'ils ont utilisé un préservatif lors de leur dernière relation sexuelle, s'ils avaient

¹ Le calcul a été effectué en divisant le nombre d'enfants dans le ménage qui ont reçu le vaccin DPT1 par le nombre total d'enfants interrogés par les recenseurs au cours de la séance de vaccination. Ce calcul est différent de ceux concernant le vaccin DPT1 de l'IHME car l'institut a utilisé uniquement les enfants âgés de 12 à 23 mois. En calculant le taux de vaccination DPT1 à l'aide des enfants âgés de 12 à 23 mois uniquement, les taux de vaccination DPT1 de l'IHME et de l'EDS ont été mal corrélés (.266). La corrélation a considérablement augmenté lorsque nous avons utilisé tous les enfants qui ont participé à la séance de vaccination (.733).

des rapports et s'il le/la répondant(e) a eu plusieurs partenaires l'année écoulée, pour les hommes et pour les femmes. Les données de l'IHME utilisaient les mêmes variables dans leurs estimations de la prévalence du VIH, comme il est expliqué ci-dessous.

La troisième source de données utilisée dans cette analyse est constituée par les estimations de la prévalence du VIH de l'IHME et la couverture du vaccin DPT1. Les estimations de la prévalence du VIH de l'IHME sont des estimations du nombre d'adultes âgés de 15 à 49 ans qui vivent avec le VIH dans une cellule de grille de 5x5 km à travers 47 pays en Afrique (IHME, 2019a). Ces estimations ont été créées à l'aide d'un certain nombre de sources différentes, y compris les covariantes tirées des microdonnées de l'Enquête par grappes avec indicateurs multiples de 2000 (UNICEF, 2000), des microdonnées de l'enquête EDS (ICF International, 2005), des microdonnées de l'enquête EDS 2011-12 (ICF International, 2011-12), et des microdonnées de l'enquête MICS de 2016 (UNICEF, 2018). Les covariantes tirées des deux enquêtes EDS incluent la prévalence de :

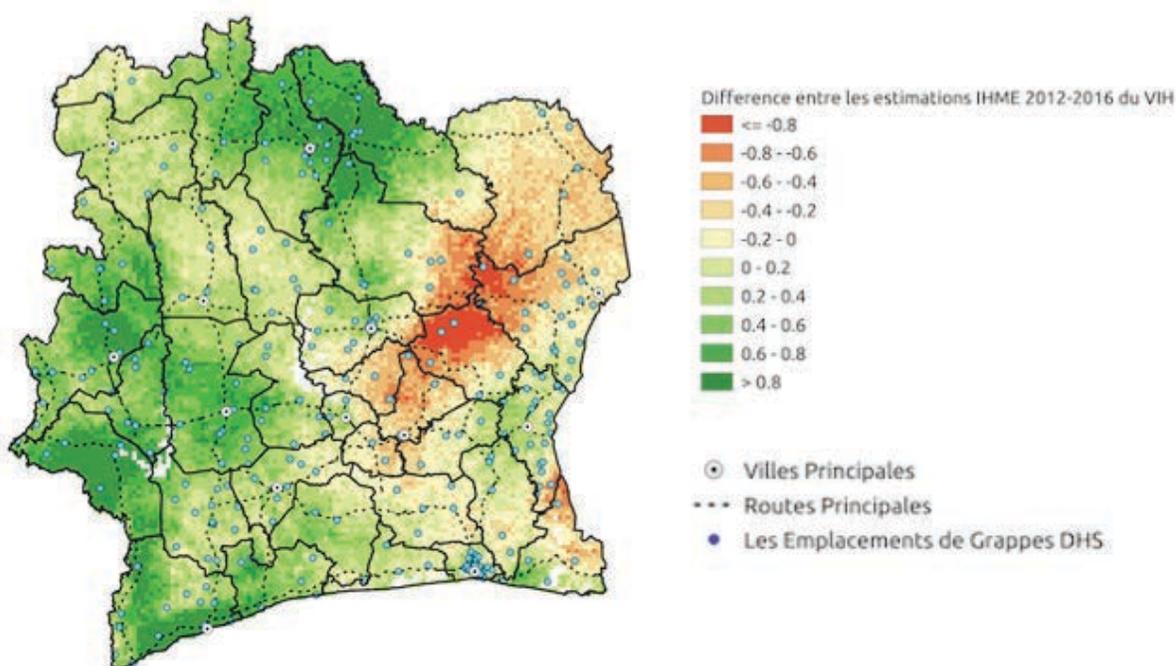
- la circoncision masculine (sous toutes ses formes)
- les symptômes des MST auto-déclarés
- le mariage ou la vie avec un partenaire en tant mari et femme
- le partenaire actuel qui vit ailleurs
- l'utilisation de préservatif lors de la dernière relation sexuelle
- déclaré avoir déjà eu des rapports entre jeunes adultes
- partenaires multiples l'année écoulée pour les hommes
- partenaires multiples l'année écoulée pour les hommes et les femmes.

Les covariantes utilisées dans les enquêtes EDS incluent la prévalence de :

- ont eu des rapports
- le mariage ou la vie avec un partenaire en tant mari et femme
- l'utilisation de préservatif lors de la dernière relation sexuelle
- déclaré avoir déjà eu des rapports entre jeunes adultes
- partenaires multiples l'année passée pour les hommes et les femmes

L'IHME a également utilisé ICAP 2017, le rapport d'enquête de Columbia University Mailman School of Public Health (ICAP, 2018), les estimations sur le VIH figurant dans le fichier des estimations de l'ONUSIDA 2017 et 2018 (UNAIDS, 2018) qui comportent des informations datant des années 2004-2005, 2008, 2013 et 2017.

Carte 1 : Changements des estimations sur le VIH de l'IHME, 2012-2016 (avec le recouvrement des sites de l'EDS)



Les estimations de la prévalence réalisées par l'IHME sur la couverture du vaccin DPT1 étaient basées sur le nombre d'enfants âgés de 12 à 23 mois qui ont reçu le vaccin DPT1 à travers 52 pays africains dans une cellule de grille de 5x5 km (IHME, 2019b). Pour effectuer ces estimations, l'IHME s'est basée sur l'enquête EDS 2011-2012 (ICF International, 2012), l'enquête MICS 2000 (UNICEF, 2000), et l'enquête MICS 2006 (UNICEF, 2006) pour créer un taux de prévalence dans

une cellule de grille de 5x5 km. À partir de ces enquêtes, les estimations de la prévalence ont été construites à l'aide des covariantes suivantes : âge, sexe et couverture du vaccin DPT1. Les chercheurs se sont servis de méthodes spatio-temporelles bayésiennes existantes pour analyser une base de données géolocalisées sur la prévalence du VIH à travers l'Afrique. Ils ont créé une grille d'estimation de la prévalence du VIH pour l'Afrique. Les chercheurs ont effectué des estimations de la prévalence du VIH au niveau de la cellule de grille et des estimations au niveau de la cellule de grille de la population âgée de 15 à 49 ans pour estimer, en fin de compte, le nombre de personnes vivant avec le VIH. Après avoir effectué ces estimations, ils « ont calculé les moyennes pondérées par la population des estimations au niveau de la cellule de grille pour générer des estimations pour les subdivisions administratives de premier niveau (par ex, les provinces ou les régions) et les subdivisions administratives de deuxième niveau (par ex, districts ou départements) dans chaque pays » (Dwyer-Lindgren, 2019).

Les données sur les estimations relatives au vaccin DPT1 et à la prévalence du VIH créées par l'IHME ont été extraites pour tous les sites des enquêtes EDS et MICS à l'aide de l'outil d'extraction de point dans QGIS. Ensuite, un identifiant unique a permis de faire correspondre ces données à celles des enquêtes EDS et MICS. QGIS a été également utilisé pour faire correspondre les sites des enquêtes EDS et MICS à l'aide de Hub Distance Analysis qui a permis d'associer deux par deux les villages les plus proches entre eux. Puisqu'il y avait plus de sites MICS que de sites EDS, tous les sites MICS qui n'ont pas eu de correspondance ont été abandonnés dans l'échantillon de l'enquête.

Analyse des données

Ce document étudie trois cas simulés différents qui reflètent les situations courantes auxquelles les chercheurs sont confrontés.

- **Cas 1 : Situation de base de l'enquête, suivi prédit**

Les données de l'enquête sont disponibles au départ, mais pas les données de l'enquête de suivi. Alors les chercheurs s'appuient une surface prédictive pour refléter les conditions de suivi.

- **Cas 2 : Situation de base prédite, suivi de l'enquête**

Aucune donnée de l'enquête n'est disponible au départ. Par conséquent, les chercheurs s'appuient sur la surface prédictive en plus des données de l'enquête disponibles dans la phase de suivi.

- **Cas 3 : Situation de base et suivi prédits**

Aucune donnée de l'enquête n'est disponible pour aucune des séries. Par conséquent, les chercheurs doivent s'appuyer sur les surfaces prédictives aussi bien pour la situation de départ que pour la fin.

Dans le Cas 1, les données de l'enquête EDS en 2012 et la surface prédictive de l'IHME en 2016 ont été utilisées pour étudier les vaccins DPT1. Nous avons procédé à une simulation qui permet d'étudier un changement hypothétique du nombre d'enfants vaccinés avec le DPT1 entre 2012 (à l'aide des données EDS) et 2016 (à l'aide des estimations IHME basées sur les sites de l'enquête EDS) pour les villages qui ont été affectés de manière aléatoire à un groupe de traitement ou de contrôle. Nous avons procédé à l'estimation des résultats en réalisant 20.000 simulations qui ont permis d'affecter de manière aléatoire des groupes de traitement ou de contrôle à tous les sites. La simulation a alors permis de créer un effet du traitement estimé en multipliant l'écart type de l'estimation du vaccin DPT1 2016 de l'IHME pour ce site par un effet du traitement standardisé affecté qui a augmenté les résultats d'un écart type de 0.01-0.20 à travers les simulations. Dans chaque itération, nous avons simulé les nouvelles données de résultats en incluant l'effet du traitement pour le groupe de traitement affecté de manière aléatoire. Nous avons, ensuite, régressé cette variable de résultat sur un indicateur de traitement et le nombre, au départ, des enfants vaccinés avec DPT1 en 2012 tirés des données EDS. Enfin, nous avons comparé les coefficients de résultats à travers chacune des 1.000 itérations pour

chacune des 20 valeurs de l'effet du traitement. Les statistiques sommaires des principales variables que nous avons utilisées sont incluses dans le tableau ci-dessous :

Tableau 1 : Statistiques sommaires des variables du cas 1

VARIABLES	N	Moyenn e	SD	Min	Max
Taux de vaccination DPT1 au niveau du groupe EDS	10.060	0,693	0,192	0,0256	1
Estimations du vaccin DPT1 2016 de l'IHME*	9.488	0,945	0,0278	0,859	0,991

*Extrait des sites du groupe EDS

Le Cas 2 a utilisé la surface prédictive de 2012 de l'IHME comme base et l'enquête MICS pour le suivi. Au lieu d'utiliser les données sur le vaccin DPT1 comme dans le Cas 1, cette simulation a utilisé un indice VIH créé à partir des données sur les hommes et les femmes par MICS en 2016 (décrites plus haut), ainsi que les estimations de la prévalence du VIH de l'IHME à partir de 2012 qui ont été extraites des sites de l'enquête MICS. Au cours de cette simulation, les mêmes méthodes d'analyse ont été adoptées, notamment en affectant un effet de traitement de plus en plus grand qui est mesuré en termes d'écart type de la variable substitutive du VIH tiré des données MICS de 2016 en vue de créer un effet du traitement simulé. Nous avons, ensuite, régressé ce résultat sur l'état du traitement généré de manière aléatoire et les estimations du VIH de l'IHME en 2012, en le répétant 1.000 fois pour chacune des 20 valeurs potentielles de l'effet du traitement (0.01-0.20).

Tableau 2 : Statistiques sommaires des variables du Cas 2

VARIABLES	N	Moyenn e	SD	Min	Max
Estimations du VIH de IHME en 2012*	17.823	3,473	0,762	1,196	5,093
Estimations du VIH de l'IHME 2016*	17.823	3,199	0,796	0,896	4,733
Indice VIH de MICS	13.972	0,494	0,131	0,250	1

*Ces estimations ont été extraites des sites MICS

Dans le Cas 3, les surfaces prédites de l'IHME sont utilisées aussi bien pour la situation de départ que pour la fin pour examiner les changements des résultats des estimations du VIH dans une situation hypothétique où certains groupes ont été traités grâce à un programme de lutte contre le VIH/SIDA alors que certains ne l'étaient pas. Le proxy VIH pour les données MICS a été remplacé par l'estimation du VIH de l'IHME en 2016 et des contrôles de base ont été ajoutés à la régression. Ces contrôles de base ont inclus deux variables tirées de l'enquête EDS sur le matériau du plancher et le matériau du toit des maisons des individus. Nous avons suivi la même procédure comme les deux premières simulations et créé une nouvelle variable de résultat avec l'effet du traitement estimée qui a été ajouté. La régression incluait la variable du traitement, les estimations du VIH de l'IHME en 2012, les estimations du VIH de IHME en 2016, ainsi que les variables de contrôle EDS et MICS. Les estimations de l'IHME ont été extraites des sites de l'enquête EDS

Tableau 3 : Statistiques sommaires pour les variables du Cas 3

VARIABLES	N	Moyenne	SD	Min	Max
Estimations du VIH de IHME en 2016 *	994	3,171	0,754	0,896	4,733
Estimations du VIH de IHME en 2012*	994	3,443	0,733	1,196	5,551
Qualité du sol	1.022	30,36	5,823	11	47
Qualité du toit	1.022	30,09	6,645	13	88

*Ces estimations ont été extraites des sites MICS et DHS

Leçons sur l'intégration des données

La conduite de ces trois analyses s'est heurtée à plusieurs défis, liés principalement à l'intégration des données. Puisque chaque ensemble de données a été créé par des organisations différentes, des difficultés ont été rencontrées dans la création des indicateurs qui ont correspondu à travers les données EDS, MICS et IHME. Par exemple, le taux de vaccination du DPT1 dans chaque groupe devait être calculé à l'aide des données de l'enquête. Ce calcul devait correspondre à celui effectué par l'IHME pour calculer les statistiques du vaccin DPT1. Pour l'IHME, le taux de vaccination DPT1 a été calculé à l'aide du nombre d'enfants âgés de 12 à 23 mois qui ont reçu le vaccin. Après avoir effectué ce même calcul pour les données EDS, il y a eu une faible corrélation entre les taux de vaccination DPT1 de l'enquête EDS et de l'IHME (.266). Compte tenu de cette situation, le calcul de l'enquête EDS a été effectué en utilisant tous les enfants à qui il a été posé des questions sur le vaccin DPT1 (la section indique que c'est pour les cinq derniers enfants qui sont nés du répondant). Ce calcul qui en a résulté a permis de corrélérer beaucoup mieux avec le calcul de l'IHME (.733).

Si chaque ensemble de données avait ses propres indicateurs et structure uniques, de manière générale, l'utilisation d'une combinaison de QGIS et d'un ensemble logiciel (Stata dans ce cas)

a été un moyen efficace de mener une analyse sur trois ensembles de données différents. La clé est capable de les relier à travers les points géolocalisés collectés par chaque organisation. Les maîtres d'œuvre et les évaluateurs du projet auront plus de souplesse et de possibilités pour évaluer les programmes s'ils arrivent à collecter les coordonnées de leurs propres sites de collecte de données et de mise en œuvre du projet. Ces coordonnées peuvent être utilisées pour correspondre avec d'autres données géolocalisées de l'enquête et les surfaces prédictives qui peuvent aider les maîtres d'œuvre et les évaluateurs à comprendre la manière dont les projets cherchent à améliorer les moyens d'existence des individus.

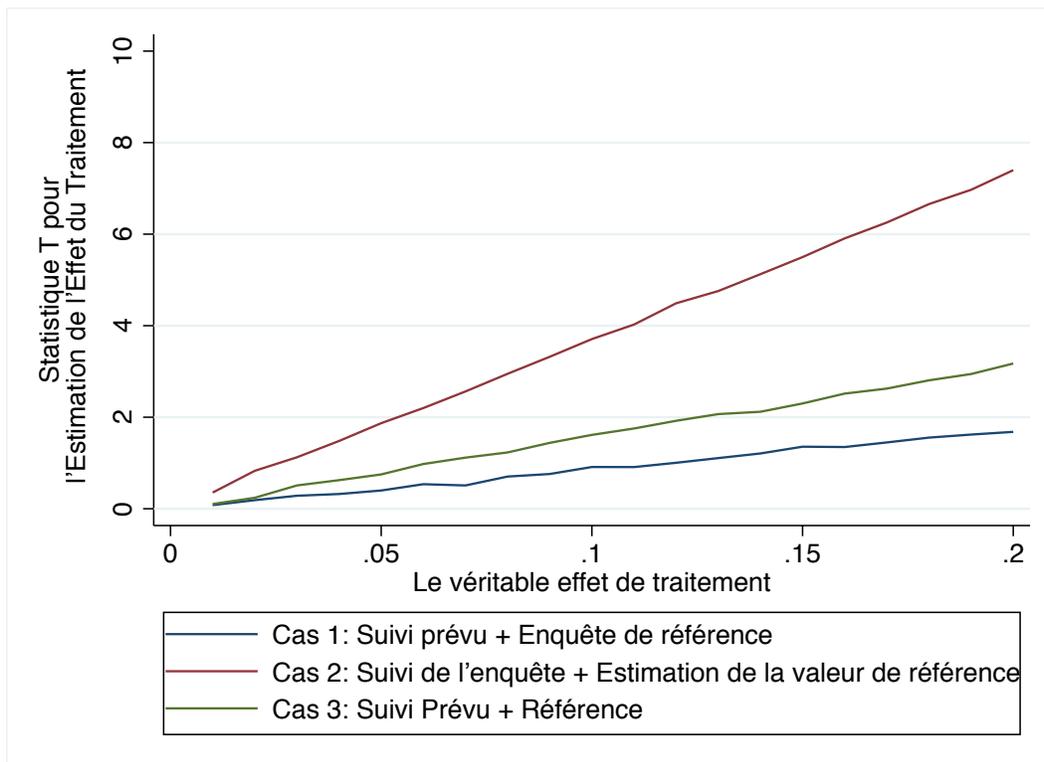
Résultats

Nous présentons les résultats des trois cas simulés dans la Figure 1 ci-dessous. Cette figure montre les statistiques liées à chaque simulation sur la taille de l'effet d'un traitement donné. Pour toutes les valeurs de l'effet du traitement, nous trouvons que le Cas 2 (Suivi de l'enquête + Base prédite) offre une signification statistique plus que n'importe lequel des deux autres cas. En effet, dans le Cas 2, il est possible de détecter statistiquement même les effets minimes du traitement, avec son effet minimal détectable à des unités d'écart type d'environ 0.05 (détectables à une statistique t de ~ 2 , avec un $p=0.05$). Ce résultat est largement tiré par la variation substantielle des résultats qui se reflète dans les données de l'enquête de suivi, qui incluent une variation à travers et au sein des groupes.

En particulier, la situation est la moins reluisante dans le Cas 1 (suivi prévu + base de l'enquête), aucun effet du traitement ne se distinguant statistiquement de zéro, même pour les effets ayant la plus grande ampleur. Cela est dû, en partie, à la variation plus étroite offerte par la surface prédite qui ne reflète aucune variation au sein du groupe des résultats obtenus dans l'enquête de suivi. De plus, le fait de dépendre des données de l'enquête au départ limite naturellement l'échantillon aux seuls sites du groupe de base. À cause de ces deux contraintes, la variance globale de l'échantillon dans les résultats limite la capacité à détecter les effets du traitement dans ce scénario.

Le Cas 3 (suivi prévu + base) offre une performance intermédiaire avec des effets de traitement de moyenne à grande ampleur (qui dépassent des écarts types de 0.15) statistiquement détectables. Alors que le Cas 3 fait face à une variation limitée dans les mesures des résultats des surfaces prédites, la capacité à tirer ces sites de l'échantillon en dehors des séries de base et de fin des données de l'enquête permet d'avoir un échantillon global assez grand pour détecter certains effets du traitement. En d'autres termes, l'utilisation des mesures des résultats de base dérivés de la surface prédite plutôt que les données de l'enquête permet de surmonter la principale limite dans ces contextes.

Figure 1 : Résultats des Cas simulés



Discussion

Le présent document a pour objet d'aider les évaluateurs, les donateurs et les maîtres d'œuvre à comprendre comment les surfaces prédites tirées des données géospatiales sur la santé

peuvent être utilisées pour lever les contraintes de l'enquête telles que les problèmes de budget et de temps. Dans ce contexte, nous trouvons que le fait de joindre les surfaces prédites reflétant les conditions de base avec les données de l'enquête de suivi offre la plus grande puissance statistique. Dans cette configuration, même les effets minimes du traitement peuvent être détectés (inférieurs aux unités d'écart type de 0.05). Plus important encore, cette configuration peut être facilement utilisée dans l'évaluation de l'impact géospatial -- une démarche de plus en plus utilisée pour fournir des estimations d'impact rigoureuses sans design randomisé ou collecte de données de base. Grâce à cette méthode, les évaluateurs pourraient éventuellement utiliser une surface interpolée pour leurs estimations de base et utiliser leur propre situation finale, ce qui pourrait permettre de réduire de moitié le coût de l'enquête. En effet, s'il y avait des circonstances où les évaluateurs font face à certaines restrictions, la capacité à utiliser ce type de méthodologie d'évaluation d'impact avec une seule édition de collecte de données primaires crée des opportunités d'apprentissage et d'évaluation là où il n'en existerait pas.

Nous avons particulièrement utilisé les données MCS et EDS dans cette analyse car ces enquêtes sur la santé sont géolocalisées et accessibles au public. Cependant, d'autres ensembles de données qui ne sont pas actuellement accessibles au public (tels que EDS2) peuvent donner des idées supplémentaires sur le fonctionnement des établissements de santé, les types de maladies qu'ils traitent, les médicaments qu'ils utilisent et dont ils ont besoin, etc. Les initiatives comme PEPFAR ou d'autres bailleurs de fonds ou maîtres d'œuvre de projets pourraient avoir accès à ces données géolocalisées supplémentaires des projets et à d'autres données et pourraient utiliser ces ensembles de données au lieu des enquêtes accessibles au public utilisées dans le présent document ou être combinés à ces données. La méthodologie que nous décrivons ne se limite pas seulement aux données MICS, EDS et IHME. De plus, si l'on creuse un peu les chemins ou mécanismes menant aux impacts de projets spécifiques, cet accès supplémentaire aux ensembles de données pourrait toujours être nécessaire au-delà de ceux qui sont utilisés directement dans nos simulations.

Nous n'évaluons pas la capacité de ces ensembles de données à tirer les groupes de comparaison similaires - c'est-à-dire à dépasser les préjugés dus aux différences préexistantes à travers les sites traités et non traités (cela se fait généralement en limitant le groupe de comparaison suffisamment tôt avant les interventions). Puisque l'affectation du traitement dans notre simulation se fait au niveau du site (plutôt qu'au niveau individuel ou des ménages), nous nous attendons à ce que les surfaces prédites et les données de l'enquête compensent suffisamment les différences spécifiques au site. D'autres interventions qui ciblent plus précisément les individus à l'intérieur des sites pourraient voir une plus grande importance accordée aux données de l'enquête, dans la mesure où les surfaces prédites pourraient ne pas être suffisantes pour surmonter ces différences.

Les surfaces prédites sont relativement nouvelles dans le monde du développement, mais les estimations provenant de ces sources pourraient ouvrir une nouvelle suite d'analyses et d'apprentissages à l'intention des praticiens et des évaluateurs. Ces sources de données sont particulièrement utiles dans les cas où les équipes ou les évaluateurs des projets étaient en mesure de collecter une série de données, mais une deuxième collecte n'était pas réalisable. Certains cas entraînent des projets avec des situations de base mais pas de vagues d'enquêtes de suivi, alors que d'autres incluent uniquement les séries de suivi (ou même aucune série d'enquête du tout). Nous trouvons que les résultats présentant une signification statistique sont certainement réalisables, même dans le dernier cas où seules les surfaces prédites pourraient être disponibles. La plus grande puissance statistique est atteinte dans le cas où les données de la surface prédite de base sont couplées avec les données de suivi de l'enquête.

Les principaux compromis concernent le fait de dépendre du centre de données de l'enquête et de la surface prédite sur les échantillons plus petits des enquêtes et la variation plus limitée inhérente aux surfaces prédites. C'est parce qu'elles sont disponibles pour les pays tout entiers (ou même les régions) que les surfaces prédites peuvent offrir des échantillons de sites plus grands. Cependant, en raison du fait que les prédictions derrière ces surfaces sont uniquement significatives par rapport aux groupes de sites (par ex, des cellules de grille de 5 km), elles

fournissent moins de variation au niveau des individus ou des ménages. Dans notre application en Côte d'Ivoire, ces compromis entraînent un mélange spécifique de ces ensembles de données pour servir de configuration optimale. Il est possible qu'avec des surfaces prédites plus précises qui seront disponibles à l'avenir, cette configuration optimale pourrait finalement changer pour être entièrement dépendante des surfaces prédites. La technique de recherche, nouvelle et innovante, présentée dans ce document offre aux évaluateurs des possibilités de réduire le coût d'une évaluation d'impact sans pour autant la rendre moins rigoureuse.

Bibliographie

Anderson, S.-J. et al. Maximising the effect of combination HIV prevention through prioritisation of the people and places in greatest need: a modelling study. *Lancet* 384, 249–256 (2014).

Akachi Y, Atun R. Effect of investment in malaria control on child mortality in sub-Saharan Africa in 2002-2008. *PLoS One*. 2011;6(6):e21309. doi:10.1371/journal.pone.0021309

Austin Carter. Comparative Analysis of HIV Estimation Methods and Results in Locations with Case Reports and High-Quality Vital Registration Data. 2018. Thesis, University of Washington. <https://digital.lib.washington.edu/researchworks/handle/1773/42899>

Barlow, S., Spencer, J., and Wilkes, B. (2017). Innovations in Geographic Information System Mapping Technology. MEASURE Evaluation. <https://www.measureevaluation.org/resources/publications/ws-17-41>.

BenYishay, Ariel, Daniel Runfola, Rachel Trichler, Carrie Dolan, Seth Goodman, Bradley Parks, Jeffery Tanner, Silke Heuser, Geeta Batra, and Anupam Anand. 2017. A Primer on Geospatial Impact Evaluation Methods, Tools, and Applications. AidData Working Paper #44. Williamsburg, VA: AidData at William & Mary.

BenYishay, A., Velyvis, K, Nolan, K., Khatiwada L. K., Dolan, C., Guzmán, D. B., Purekal, T., Mamun, A., & Wilf, S. (2019). Long-term Impact Evaluation of the Malawi Wellness and Agriculture for Life Advancement Program. Williamsburg, VA and Washington, DC: AidData at William & Mary and Mathematica Policy Research.

Bird, Tomas & Pezzulo, Carla & Bird, Tomas & Utazi, Chigozie & Sorichetta, Alessandro & Tatem, Andrew & Yourkavitch, Jennifer & Burgert, Clara. (2016). Geospatial Modeling of Child Mortality across 27 Countries in Sub-Saharan Africa.

Broer, P. Niclas, Klug, Stefanie J., Snow, Rachel C., Okiro, Emelda A., Ouma, Paul O., Snow, Robert W., Tatem, Andrew J., Meara, John G., Alegana, Victor A. and Juran, Sabrina (2018) Geospatial mapping of access to timely essential surgery in sub-Saharan Africa. *BMJ Global Health*, 3 (4), 1-10, [e000875]. (doi:10.1136/bmjgh-2018-000875).

CDC Retro-CI, Ministry of the Fight Against AIDS (Côte d'Ivoire), National Institute of Statistics (Côte d'Ivoire), ORC Macro. Côte d'Ivoire AIDS Indicator Survey 2005. Fairfax, United States: ICF International.

Carrel, M. et al. Changing spatial patterns and increasing rurality of HIV prevalence in the Democratic Republic of the Congo between 2007 and 2013. *Health Place* 39, 79–85 (2016).

Coburn, B. J., Okano, J. T. & Blower, S. Using geospatial mapping to design HIV elimination strategies for sub-Saharan Africa. *Sci. Transl. Med.* 9, eaag0019 (2017).

“Côte d’Ivoire.” PHIA Project. ICAP at Columbia University. 2018.
<https://phia.icap.columbia.edu/countries/cote-divoire/>. Accessed on December 2019.

Cuadros, D. F. et al. Mapping the spatial variability of HIV infection in sub-Saharan Africa: effective information for localized HIV prevention and control. *Sci. Rep.* 7, 9093 (2017).

Desai, Harsh (2018a). Open DCH Project Synthesis Report.

Desai, Harsh (2018b). Trip Report - Cote D’Ivoire.

Diggle, P. and E. Giorgi (2016). Model-based geostatistics for prevalence mapping in low-resource settings. *Journal of the American Statistical Association* 111, 1096– 1120.

Diggle, P. J. and E. Giorgi (2019). *Model-based Geostatistics for Global Public Health: Methods and Applications*. Boca-Raton: Chapman and Hall/CRC

Dube, Yolisa Prudence, Ruktanonchai, Corrine W, Saco, Charfudin, Tatem, Andrew, Munguambe, Khatia, Boone, Helena, Vilanculo, Faustino Carlos, Sevene, Esperanca, Matthews, Zoe, von Dadelszen, Peter and Makanga, Prestige T (2019) How accurate are modelled birth and pregnancy estimates? Comparison of four models using high resolution maternal health census data in southern Mozambique. *BMJ Global Health*, 4 (5), 1-12. ([doi:10.1136/bmjgh-2018-000894](https://doi.org/10.1136/bmjgh-2018-000894)).

Dwyer-Lindgren, L, Cork MA, Sligar A, Steuben KM, Wilson KF, Provost NR, Mayala BK, VanderHeide JD, Collison ML, Hall JB, Biehl MH, Carter A, Frank T, Douwes-Schultz D, Burstein R, Casey DC, Deshpande A, Earl L, El Bcheraoui C, Farag TH, Henry NJ, Kinyoki D, Marczak LB, Nixon MR, Osgood-Zimmerman A, Pigott D, Reiner RC, Ross JM, Schaeffer LE, Smith DL, Weaver ND, Wiens KE, Eaton JW, Justman JE, Opio A, Sartorius B, Tanser F, Wabiri N, Piot P, Murray CJL, Hay SI. Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature*. 15 May 2019. doi:10.1038/s41586-019-1200-9.

Gething, P. W., D. C. Casey, D. J. Weiss, D. Bisanzio, S. Bhatt, E. Cameron, K. E. Battle, U. Dalrymple, J. Rozier, P. C. Rao, M. Kutz, R. Barber, C. Huynh, K. Shackleford, M. Coates, G. Nguyen, M. Fraser, R. Kulikoff, H. Wang, M. Naghavi, D. Smith, C. Murray, S. Hay, and S. Lim (2016). Mapping plasmodium falciparum mortality in Africa between 1990 and 2015. *New England Journal of Medicine* 375, 2435–2445.

Gething, P., A. Tatem, T. Bird, and C. Burgert-Brucker (2015). Creating spatial interpolation surfaces with DHS data. Technical report, ICF International. DHS Spatial Analysis Reports No. 11.

Giorgi, E., P. J. Diggle, R. W. Snow, and A. M. Noor (2018). Geostatistical methods for disease mapping and visualization using data from spatio-temporally referenced prevalence surveys. *International Statistical Review* 86, 571–597.

Graetz, N., J. Friedman, A. Osgood-Zimmerman, R. Burstein, M. H. Biehl, C. Shields, J. F. Mosser, D. C. Casey, A. Deshpande, L. Earl, R. Reiner, S. Ray, N. Fullman, A. Levine, R. Stubbs, B. Mayala, J. Longbottom, A. Browne, S. Bhatt, D. Weiss, P. Gething, A. Mokdad, S. Lim, C. Murray, E. Gakidou, and S. Hay (2018). Mapping local variation in educational attainment across Africa. *Nature* 555, 48

Golding, N., R. Burstein, J. Longbottom, A. Browne, N. Fullman, A. OsgoodZimmerman, L. Earl, S. Bhatt, E. Cameron, D. Casey, L. Dwyer-Lindgren, T. Farag, A. Flaxman, M. Fraser, P. Gething, H. Gibson, N. Graetz, L. Krause, X. Kulikoff, S. Lim, B. Mappin, C. Morozoff, R. Reiner, A. Sligar, D. Smith, H. Wang, D. Weiss, C. Murray, C. Moyes, and S. Hay (2017). Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. *The Lancet* 390, 2171–2182.

Harsh Desai, Open DCH Project Synthesis Report, 2018a

Harsh Desai, Trip Report - Côte D'Ivoire, Jan 2018b

ICF International, Ministry of the Fight Against AIDS (Côte d'Ivoire), National Institute of Statistics (Côte d'Ivoire). Côte d'Ivoire Demographic and Health Survey 2011-2012. Fairfax, United States of America: ICF International.

Institute for Health Metrics and Evaluation (IHME). Africa HIV Prevalence Geospatial Estimates 2000- 2017. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2019a.

Institute for Health Metrics and Evaluation (IHME). Africa Diphtheria-Pertussis-Tetanus Vaccine Coverage Geospatial Estimates 2000-2016. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2019b.

James, William H.M., Tejedor Garavito, Natalia, Hanspal, Sophie E., Sutton, Andrew, Hornby, Graeme, Pezzulo, Carla, Nilsen, Kristine, Sorichetta, Alessandro, Ruktanonchai, Corrine W., Carioli, Alessandra, Kerr, David, Matthews, Zoe and Tatem, Andrew (2018) Gridded birth and pregnancy datasets for Africa, Latin America and the Caribbean. *Scientific Data*, 5, [180090]. ([doi:10.1038/sdata.2018.90](https://doi.org/10.1038/sdata.2018.90)).

Kandala, N.-B., Campbell, E. K., Rakgoasi, S. D., Madi-Segwagwe, B. C. & Fako, T. T. The geography of HIV/AIDS prevalence rates in Botswana. *HIV AIDS* 4, 95–102 (2012).

Kleinschmidt, I., Pettifor, A., Morris, N., MacPhail, C. & Rees, H. Geographic distribution of human immunodeficiency virus in South Africa. *Am. J. Trop. Med. Hyg.* 77, 1163–1169 (2007)

Larmarange, J. & Bendaud, V. HIV estimates at second subnational level from national population-based surveys. *AIDS* 28, S469–S476 (2014).

Leyk, Stefan, Gaughan, Andrea E., Adamo, Susana B., de Sherbinin, Alex, Balk, Deborah, Freire, Sergio, Rose, Amy, Stevens, Forrest R., Blankespoor, Brian, Frye, Charlie, Comenetz, Joshua, Sorichetta, Alessandro, MacManus, Kytt, Pistolessi, Linda, Levy, Marc, Tatem, Andrew and Pesaresi, Martino (2019) The spatial allocation of population: a review of large-scale gridded population data products and their fitness for use. *Earth System Science Data*, 11 (3). ([doi:10.5194/essd-11-1385-2019](https://doi.org/10.5194/essd-11-1385-2019)).

Lloyd, Christopher T., Chamberlain, Heather, Kerr, David, Yetman, Greg, Pistolessi, Linda, Stevens, Forrest R., Gaughan, Andrew E., Nieves, Jeremiah, Joseph, Hornby, Graeme, MacManus, Kytt, Sinha, Parmanand, Bondarenko, Maksym, Sorichetta, Alessandro and Tatem, Andrew (2019) Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. *Big Earth Data*, 3 (2), 108-139. ([doi:10.1080/20964471.2019.1625151](https://doi.org/10.1080/20964471.2019.1625151)).

Maina, J., Ouma, P.O., Macharia, P.M. et al. A spatial database of health facilities managed by the public health sector in sub Saharan Africa. *Sci Data* 6, 134 (2019). <https://doi.org/10.1038/s41597-019-0142-2>

Mary Mahy, Tim Brown, John Stover, Neff Walker, Karen Stanecki, Wilford Kirungi, Txema Garcia-Calleja & Peter D. Ghys (2017) Producing HIV estimates: from global advocacy to country planning and impact measurement, *Global Health Action*, 10:sup1, DOI: [10.1080/16549716.2017.1291169](https://doi.org/10.1080/16549716.2017.1291169)

Meyer-Rath, G. et al. Targeting the right interventions to the right people and places: the role of geospatial analysis in HIV program planning. *AIDS* 32, 957–963 (2018).

Morten Jerven (2017) How Much Will a Data Revolution in Development Cost?, *Forum for Development Studies*, 44:1, 31-50, DOI: [10.1080/08039410.2016.1260050](https://doi.org/10.1080/08039410.2016.1260050)

“National HIV estimates file.” UNAIDS. 2018.

<https://www.unaids.org/en/dataanalysis/datatools/spectrum-epp>. Accessed on January 2020.

National School for Statistics and Economics Applied (ENSEA), United Nations Children's Fund (UNICEF), United Nations Educational, Scientific and Cultural Organization (UNESCO). Côte d'Ivoire Multiple Indicator Cluster Survey 2000. New York, United States of America: United Nations Children's Fund (UNICEF).

National Institute of Statistics (Côte d'Ivoire), United Nations Children's Fund (UNICEF). Côte d'Ivoire Multiple Indicator Cluster Survey 2016. New York, United States: United Nations Children's Fund (UNICEF), 2018.

Okano, J. T. & Blower, S. Sex-specific maps of HIV epidemics in sub-Saharan Africa. *Lancet Infect. Dis.* 16, 1320–1322 (2016)

Osgood-Zimmerman, A., A. I. Millier, R. W. Stubbs, C. Shields, B. V. Pickering, L. Earl, N. Graetz, D. K. Kinyoki, S. E. Ray, S. Bhatt, A. Browne, R. Burstein, E. Cameron, D. Casey, A. Deshpande, N. Fullman, P. Gething, H. Gibson, N. Henry, M. Herrero, L. Krause, I. Letourneau, A. Levine, P. Liu, J. Longbottom, B. Mayala, J. Mosser, A. Noor, D. Pigott, E. Piwoz, P. Rao, R. Rawat, R. Reiner, D. Smith, D. Weiss, K. Wiens, A. Mokdad, L. S.S., C. Murray, N. Kassebaum, and S. Hay (2018). Mapping child growth failure in Africa between 2000 and 2015. *Nature* 555, 41.

Paige, John & Fuglstad, Geir-Arne & Riebler, Andrea & Wakefield, Jon. (2019). Design- and Model-Based Approaches to Small-Area Estimation in a Low and Middle Income Country Context: Comparisons and Recommendations.

Pezzulo, C., T. Bird, C. Edson, C. Utazi, A. Sorichetta, A. Tatem, J. Yourkavitch, and C. Burgert-Brucker (2016). Geospatial modeling of child mortality across 27 countries in sub-Saharan Africa. Technical report, ICF International. DHS Spatial Analysis Reports No. 13.

Ruktanonchai, Corrine, Warren, Nilsen, Kristine, Alegana, Victor, Bosco, Claudio, Ayiko, Rogers, Kajeguka, Andrew C. Seven, Matthews, Zoe and Tatem, Andrew (2018) Temporal trends in spatial inequalities of maternal and newborn health services among four East African countries, 1999 - 2015. *BMC Public Health*, 18, [1339]. ([doi:10.1186/s12889-018-6241-8](https://doi.org/10.1186/s12889-018-6241-8)).

Ruktanonchai, Corrine W., Ruktanonchai, Nick W., Nove, Andrea, Lopes, Sofia, Pezzulo, Carla, Bosco, Claudio, Alegana, Victor A., Burgert, Clara R., Ayiko, Rogers, Charles, Andrew S.E.K., Lambert, Nkurunziza, Msechu, Esther, Kathini, Esther, Matthews, Zoe and Tatem, Andrew J. (2016) Equality in maternal and newborn health: modelling geographic disparities in utilisation of care in five East African countries. *PLoS ONE*, 11 (8), 1-17. ([doi:10.1371/journal.pone.0162006](https://doi.org/10.1371/journal.pone.0162006)).

Subnational Estimates Working Group of the HIV Modelling Consortium. Evaluation of geospatial methods to generate subnational HIV prevalence estimates for local level planning. *Aids* (London, England). 2016 Jun;30(9):1467-1474. DOI: 10.1097/QAD.0000000000001075.

TB, Hallett & S-J, Anderson & Asante, Cynthia & N, Bartlett & Bendaud, Victoria & Bhatt, Samir & Burgert, Clara & Cuadros, Diego & Dzangare, Janet & Fecht, Daniela & Gething, P. & Ghys, Peter & JM, Guwani & Heard, Nathan & Kalipeni, Ezekiel & Kandala, Ngianga-Bakwin & AA,

Kim & ID, Kwao & Larmarange, Joseph & Zulu, Leo. (2016). Evaluation of geospatial methods to generate subnational HIV prevalence estimates for local level planning. *AIDS*.

United Nations Children's Fund (UNICEF), National Institute of Statistics (Côte d'Ivoire). Côte d'Ivoire Multiple Indicator Cluster Survey 2006. New York, United States: United Nations Children's Fund (UNICEF).

Utazi, C. Edson, Thorley, Julia, Alegana, Victor A., Ferrari, Matthew J., Takahashi, Saki, Metcalf, C. Jessica E., Lessler, Justin, Cutts, Felicity T. and Tatem, Andrew J. (2019) Mapping vaccination coverage to explore the effects of delivery mechanisms and inform vaccination strategies. *Nature Communications*, 10 (1), [1633]. ([doi:10.1038/s41467-019-09611-1](https://doi.org/10.1038/s41467-019-09611-1)).

Wardrop, N., W. Jochem, T. Bird, H. Chamberlain, D. Clarke, D. Kerr, L. Bengtsson, S. Juran, V. Seaman, and A. Tatem (2018). Spatially disaggregated population estimates in the absence of national population and housing census data. *Proceedings of the National Academy of Sciences* 115, 3529–3537.

Yourkavitch, Jennifer & Burgert, Clara & Assaf, Shireen & Delgado, Stephen. (2018). Using geographical analysis to identify child health inequality in sub-Saharan Africa. *PLoS ONE*. 13. e0201870. [10.1371/journal.pone.0201870](https://doi.org/10.1371/journal.pone.0201870).