

RESEARCH

Open Access



# Capturing the spatial variability of HIV epidemics in South Africa and Tanzania using routine healthcare facility data

Diego F. Cuadros<sup>1,2\*</sup>, Benn Sartorius<sup>3</sup>, Chris Hall<sup>4</sup>, Adam Akullian<sup>5</sup>, Till Bärnighausen<sup>6,7,8</sup> and Frank Tanser<sup>6,9</sup>

## Abstract

**Background:** Large geographical variations in the intensity of the HIV epidemic in sub-Saharan Africa call for geographically targeted resource allocation where burdens are greatest. However, data available for mapping the geographic variability of HIV prevalence and detecting HIV ‘hotspots’ is scarce, and population-based surveillance data are not always available. Here, we evaluated the viability of using clinic-based HIV prevalence data to measure the spatial variability of HIV in South Africa and Tanzania.

**Methods:** Population-based and clinic-based HIV data from a small HIV hyper-endemic rural community in South Africa as well as for the country of Tanzania were used to map smoothed HIV prevalence using kernel interpolation techniques. Spatial variables were included in clinic-based models using co-kriging methods to assess whether cofactors improve clinic-based spatial HIV prevalence predictions. Clinic- and population-based smoothed prevalence maps were compared using partial rank correlation coefficients and residual local indicators of spatial autocorrelation.

**Results:** Routinely-collected clinic-based data captured most of the geographical heterogeneity described by population-based data but failed to detect some pockets of high prevalence. Analyses indicated that clinic-based data could accurately predict the spatial location of so-called HIV ‘hotspots’ in > 50% of the high HIV burden areas.

**Conclusion:** Clinic-based data can be used to accurately map the broad spatial structure of HIV prevalence and to identify most of the areas where the burden of the infection is concentrated (HIV ‘hotspots’). Where population-based data are not available, HIV data collected from health facilities may provide a second-best option to generate valid spatial prevalence estimates for geographical targeting and resource allocation.

**Keywords:** Spatial analysis, HIV prevalence, High resolution maps, Healthcare facilities, Sub-Saharan Africa

## Background

Human immunodeficiency virus (HIV) prevalence in sub-Saharan Africa (SSA) is characterized by large geographical variation [1, 2]. The overall HIV epidemic has been shown to be concentrated across clustered micro-epidemics of different geographical scales [3–6]. This evidence has been aligned with the Joint United Nations Programme on HIV/AIDS (UNAIDS) concept “know your epidemic, know your response” for the identification

of geographical populations at higher risk and burden of HIV [4]. Previous studies have explored the impact of focusing resources and control interventions using a spatially-targeted allocation strategy [7–9]. Results from these studies have supported this strategy and shown that spatial targeting of interventions can substantially improve the efficiency of resource allocation, compared to a homogeneous distribution strategy [8, 10–12]. Based on this evidence, programs such as The United States President’s Emergency Plan for AIDS Relief (PEPFAR) and UNAIDS Fast-Track strategy have gradually shifted its strategy towards optimization of resource allocation including geographically relevant data [13, 14].

\*Correspondence: diego.cuadros@uc.edu

<sup>1</sup> Department of Geography and Geographic Information Science, University of Cincinnati, Cincinnati, OH 45221, USA

Full list of author information is available at the end of the article



The implementation of spatially-targeted intervention strategies faces numerous challenges. Population-based spatial data are scarce and gathering spatial HIV data for identifying areas of high burden of the infection can be costly to implement in resource-limited settings. Some international agencies such as USAID's Demographic and Health Survey (DHS) collect nationally representative population-based epidemiologic data from resource limited settings [1], but the surveys are not routinely conducted, and spatial data are not available for several countries where the surveys are implemented. Other surveillance systems such as the Africa Centre Demographic Information System (ACDIS), or the Centre for the AIDS Programme of Research in South Africa (CAPRISA) also include spatial information, but they are conducted in selected micro-geographical areas, limiting the generalizability of their findings to other settings or to larger geographical scales.

Alternatively, there is a wealth of clinic-based data collected from different healthcare facilities that conduct routine HIV testing and other HIV services. The feasibility of using such sources of data to explore the spatial structure of the HIV epidemic in a given setting, however, is unknown. Against this background, we address the following question: can routinely collected and readily available HIV testing data, such as those collected from healthcare facilities, be used to accurately map the broad spatial structure of the HIV epidemic? To assess whether clinic-based HIV data accurately capture the spatial structure of HIV prevalence and to identify the so-called 'hotspots' of infection, we conducted a series of spatial statistical analyses at two different geographical scales (national and local level), thereby offering a potentially rapid and inexpensive approach to understanding the spatial structure of HIV epidemics across differently geographic scales.

## Methods

For comparison, we conducted the analysis using data from two different geographical scales, local and national scale.

### Data sources: South Africa (local level)

Population-based local level data come from one of the most comprehensive demographic surveillance systems in Africa: the ACDIS [5, 15, 16], which is located in Hlabisa subdistrict, one of the five subdistricts in the rural district of Umkhanyakude in northern KwaZulu-Natal, South Africa (Additional file 1: Figure S1 A). This population-based surveillance system has routinely collected socio-demographic, behavioral and epidemiological information on a population of approximately 90,000 participants within a circumscribed geographic

area (438 km<sup>2</sup>) for over a decade (Additional file 1: Figure S1 A). Along with the ACDIS is a population-based HIV surveillance and sexual behavior survey which takes place annually. We included the population-based HIV surveillance conducted in 2014 in our analysis. A total of 5174 homesteads (georeferenced to <2 m) were included in the survey (Additional file 1: Figure S2 A). The estimated HIV prevalence using the population-based survey data ( $pHIV$ ) at a sample location  $i$  was defined to be  $pHIV_i = H_{HIVi}/N_i$ , where  $H_{HIVi}$  denotes the number of sampled people from location  $i$  who were HIV positive and  $N_i$  denotes the total number of sampled individuals at location  $i$ .

Data sources for the clinic-based data for this study area come from the district health information system (DHIS). Antenatal clinic data from DHIS collected in 2014 from 10 healthcare facilities located in the area where the population-based surveillance is conducted were included in the analysis (Additional file 1: Figure S2 B). The antenatal healthcare facility data are collected among pregnant women attending the healthcare facilities. These data have provided invaluable information for tracking HIV prevalence and trends in most countries with generalized HIV epidemics [17]. The estimated HIV prevalence using the healthcare facility data ( $cHIV$ ) at a healthcare facility  $j$  was defined to be  $cHIV_j = f_{HIVj}/N_j$ , where  $f_{HIVj}$  denotes the number of pregnant woman from healthcare facility  $j$  who were HIV positive and  $N_j$  denotes the total number of pregnant woman tested at healthcare facility  $j$ .

### Data sources: Tanzania (national level)

National level data come from the DHS conducted in Tanzania in 2011–2012 [18]. Subjects were enrolled in DHS surveys via a two-stage sampling procedure to select households. A total of 568 sampling geo-located randomly selected community clusters was included in the survey (Additional file 1: Figure S2 C). The global positioning system was used to identify and record the geographical coordinates of each DHS sample location. A total of 17,745 individuals (9756 women and 7989 men) from the selected households were eligible for the study. Further details related to the DHS methodology, study design, and data can be found elsewhere [18–20]. The method to estimate the HIV prevalence at each DHS sample location was the same as the method previously described to estimate HIV prevalence using the population-based survey (ACDIS) for the local level data. Clinic-based data from Tanzania come from antenatal clinic surveillance where HIV testing was routinely conducted to pregnant women attending these clinics in 2010 [9, 21]. Data from 132 healthcare facilities in Tanzania were included in the analysis (Additional file 1:

Figure S2 D). The method to estimate the HIV prevalence at each healthcare facility was the same as the method previously described to estimate the HIV prevalence using the DHIS antenatal healthcare facility data in South Africa. Further description of data sources and maps illustrating the sampling locations are included.

### Spatial analysis

The prevalence of HIV was estimated at each sample location (for ACDIS, DHS, or healthcare facility). We used ESRI ArcGIS Desktop 10.3 [22] to generate continuous surface maps of HIV prevalence from each of the four datasets using a kriging interpolation technique, a methodology widely used in spatial mapping [23–26]. Kriging is a geostatistical method that generates an estimated continuous surface from a scattered set of points with z-values (i.e. HIV prevalence) implemented with the Geostatistics tool in ArcGIS. Kriging assumes that the distance between sample points reflects a spatial correlation that can be used to explain the observed variation in the surface. The method fits a mathematical function to the data points to determine the output value for each location. We used ordinary kriging to predict values of HIV prevalence at unmeasured locations by estimating a variogram of weighted averages of the data [27]. A secondary analysis using cokriging method was conducted to improve spatial estimations by including covariates of population density distribution, distance to the closest main road, and distance to the closest healthcare facility.

### Data sources comparisons

HIV prevalence estimations from the population-based data and clinic-based data derived from both models, kriging and cokriging, were extracted at each data point included for the ACDIS study area and Tanzania (Additional file 1: Figure S2 A, C). Partial rank correlation coefficient (PRCC) to compare population-based data

and clinic-based data estimations were conducted for both ACDIS study area and Tanzania. Likewise, residuals between the two continuous surface maps of HIV prevalence generated using both types of data sources (population-based and clinic-based data) were estimated from the two different scales (local and national). A third method included the estimation of the spatial correlation between HIV prevalence calculated from the two types of data sources using bivariate local indicators of spatial autocorrelation (LISA) included in the GeoDa environment [25]. This method identifies significant spatial clustering based on the degree of linear association between HIV prevalence at a given location estimated using the population-based and the clinic-based data [26]. Maps were generated illustrating the locations with statistically significant associations along with the type of spatial association between both HIV prevalence estimations (i.e. high–high HIV, low–low, low–high, and high–low). Finally, HIV ‘hotspots’ (areas with HIV prevalence in the upper quintile estimated independently for each data source, population-based and clinic-based data [2]) were identified, then the HIV hotspots identified using both sources of data were compared, and the percentage of area in which both sources of data consistently identified HIV hotspots was estimated.

## Results

### South Africa (local level)

Table 1 summarizes the estimated measures for map comparisons generated using both sources of data. Figure 1 illustrates the association between HIV prevalence estimated from population-based and clinic-based data for the ACDIS study area (A, B) and Tanzania (C, D). Continuous surface maps of HIV prevalence in the ACDIS study area are illustrated in Fig. 2. Semi-variograms and histograms for pixel density distribution of HIV prevalence are included in Additional file

**Table 1 Summary of the estimated measures at local and national level estimated using kriging and cokriging methods**

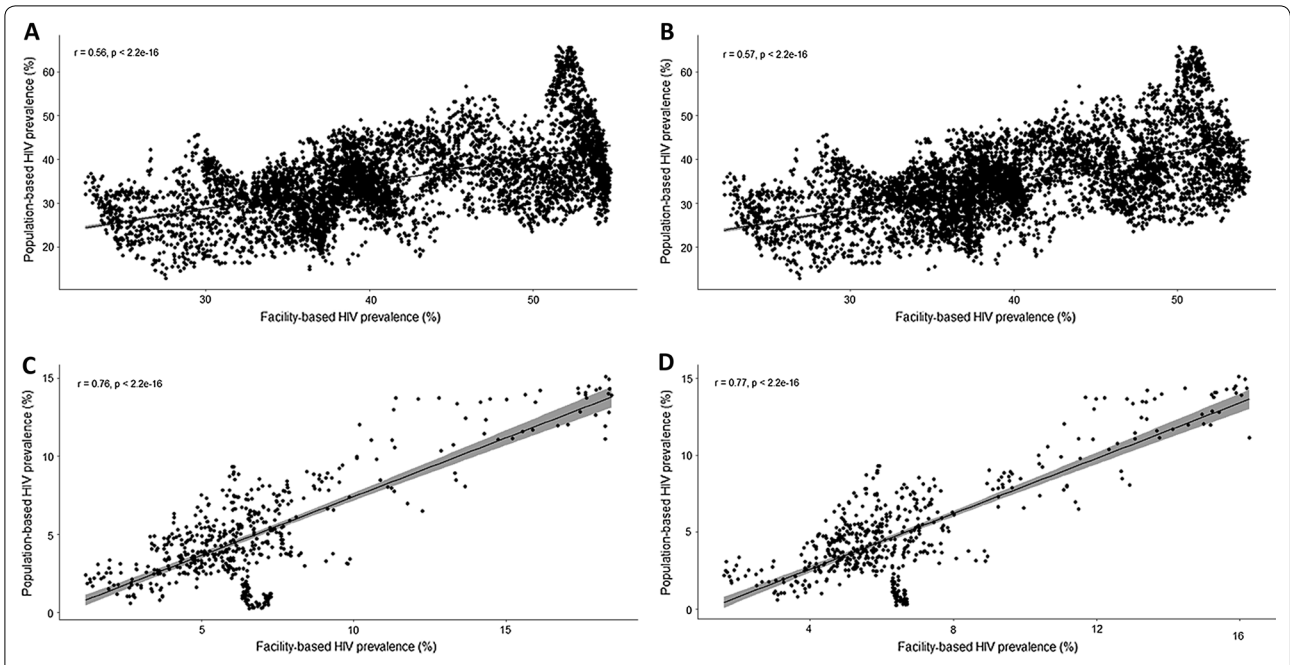
Measure	Local level (ACDIS study area)		National level (Tanzania)	
	Kriging	Cokriging	Kriging	Cokriging
PRCC <sup>a</sup> (p value)	0.56 (p < 0.005)	0.57 (p < 0.005)	0.76 (p < 0.005)	0.77 (p < 0.005)
LISA consistency areas <sup>b</sup>	59%	61%	71%	77%
LISA inconsistency areas <sup>c</sup>	32%	30%	16%	10%
HIV ‘hotspot’ areas detected <sup>d</sup>	54%	56%	77%	84%

<sup>a</sup> Partial rank correlation coefficient (PRCC) comparisons between population-based data and clinic-based data estimations

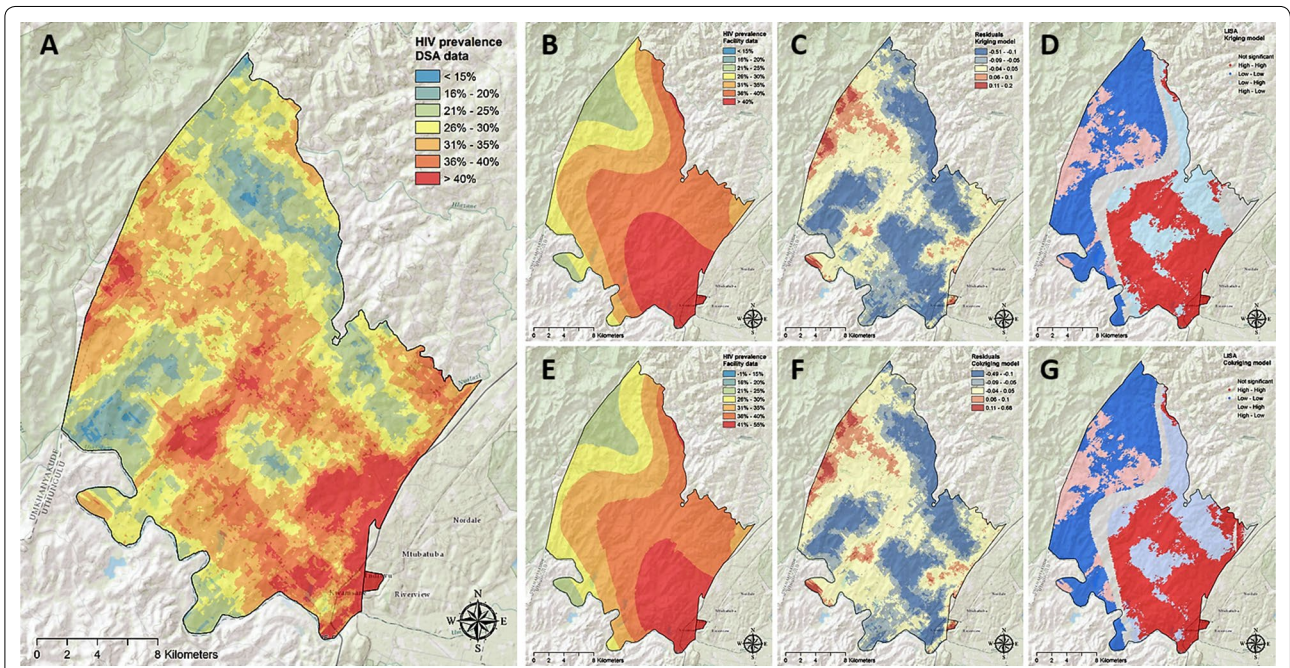
<sup>b</sup> Percentage of area where the local indicator of spatial autocorrelation (LISA) estimations comparisons between population-based data and clinic-based data estimations were consistent with statistical significance

<sup>c</sup> Percentage of area where the local indicator of spatial autocorrelation (LISA) estimations comparisons between population-based data and clinic-based data estimations were inconsistent with statistical significance

<sup>d</sup> Percentage of areas with HIV prevalence in the upper quintile detected by the mode



**Fig. 1** Comparisons between HIV prevalence estimations using population-based HIV prevalence data and clinic-based prevalence data. HIV prevalence estimations from the population-based data and clinic-based data derived from both models, kriging and cokriging, were extracted at each data point (black dots) included for the ACDIS study area and Tanzania. In **A** Africa Centre Demographic Information System study area using kriging, **B** Africa Centre Demographic Information System study area using cokriging, **C** Tanzania using kriging, **D** Tanzania using cokriging. The line in all figures indicates the fitted line between HIV prevalence estimations using population-based HIV prevalence data and clinic-based data



**Fig. 2** Continuous surface maps of **A** the estimated HIV prevalence using the Africa Centre Demographic Information System data, **B** kriging model for the estimated HIV prevalence using clinic-based data, **C** residuals from the kriging model, **D** LISA analysis of the kriging model, **E** cokriging model for the estimated HIV prevalence using clinic-based data, **F** residuals from the cokriging model, **G** LISA analysis of the cokriging model. Maps were created using ArcGIS® software by Esri version 10.3 [22] (<http://www.esri.com>)

(Additional file 1: Figures S5, S6 and S7). Mean pixel-level HIV prevalence in this area estimated using population-based data was 30.2% [95% confidence interval (CI) 16.5–43.9%]. Kriging interpolation maps revealed substantial geographical variation of the HIV epidemic in this small area of study (Fig. 2A), and identified areas with high HIV prevalence, particularly at the center, north-western, south-eastern parts of the study region.

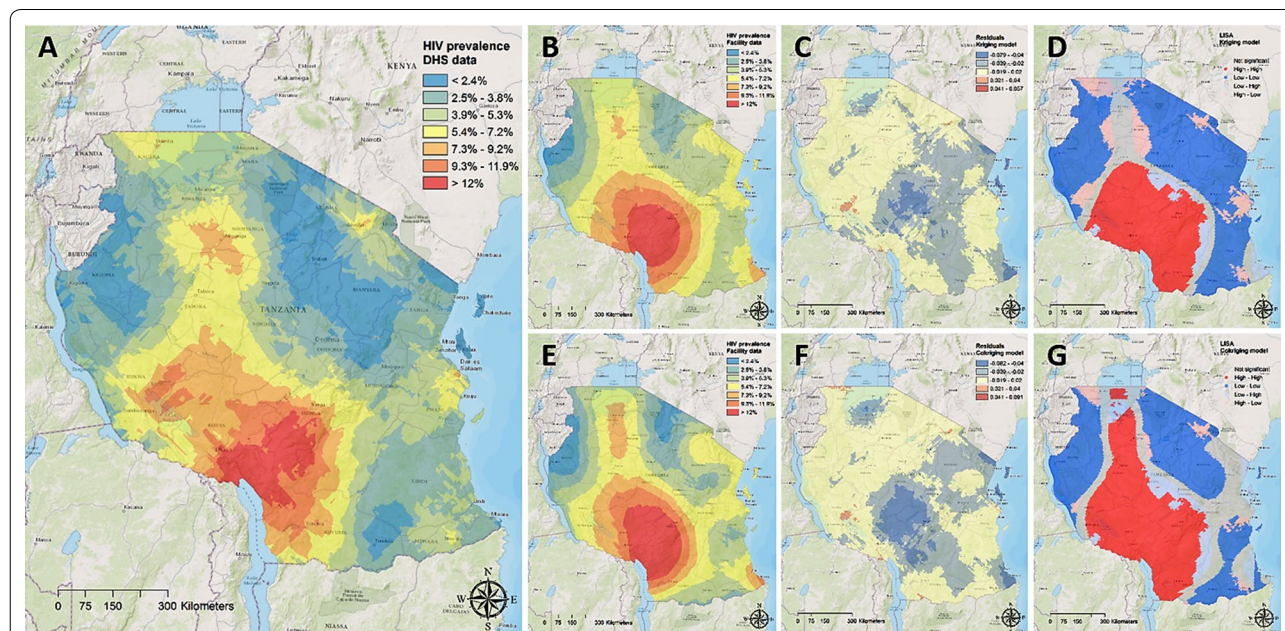
Mean pixel-level HIV prevalence in this area estimated using clinic-based data was 35.6% (95% CI 21.9–49.3%). Kriging interpolation of clinic-based data identified the high burden areas of HIV infection located at the south-eastern part of the study area but failed to identify some other areas with high HIV prevalence, particularly at the north-western part of the study area (Fig. 2B). Residual analysis was consistent with this result and indicated that HIV prevalence interpolation using clinic-based data underestimated the HIV prevalence in the north-western part of the study area, but also overestimated the HIV prevalence in the high burden areas (Fig. 2C). LISA analysis indicated that estimations from kriging interpolation using these two types of data sources were consistent in identifying high or low burden areas in 59% of the study area (Fig. 2C). The estimations from these two models diverged in 32% of the study area: kriging interpolation using clinic-based data predicted high HIV burden areas in low HIV prevalence areas in 12% of the study area, and

low HIV burden areas in high HIV prevalence areas in 20% of the study area (Additional file 1: Figure S6). There was a non-statistical significant spatial association in the remaining 9% of the study area. Inclusion of cofactors in the cokriging model moderately increased the accuracy of predictions (Fig. 2E, F), correctly identifying high or low burden areas in 61% of the study area (Fig. 2G, Additional file 1: Figure S6).

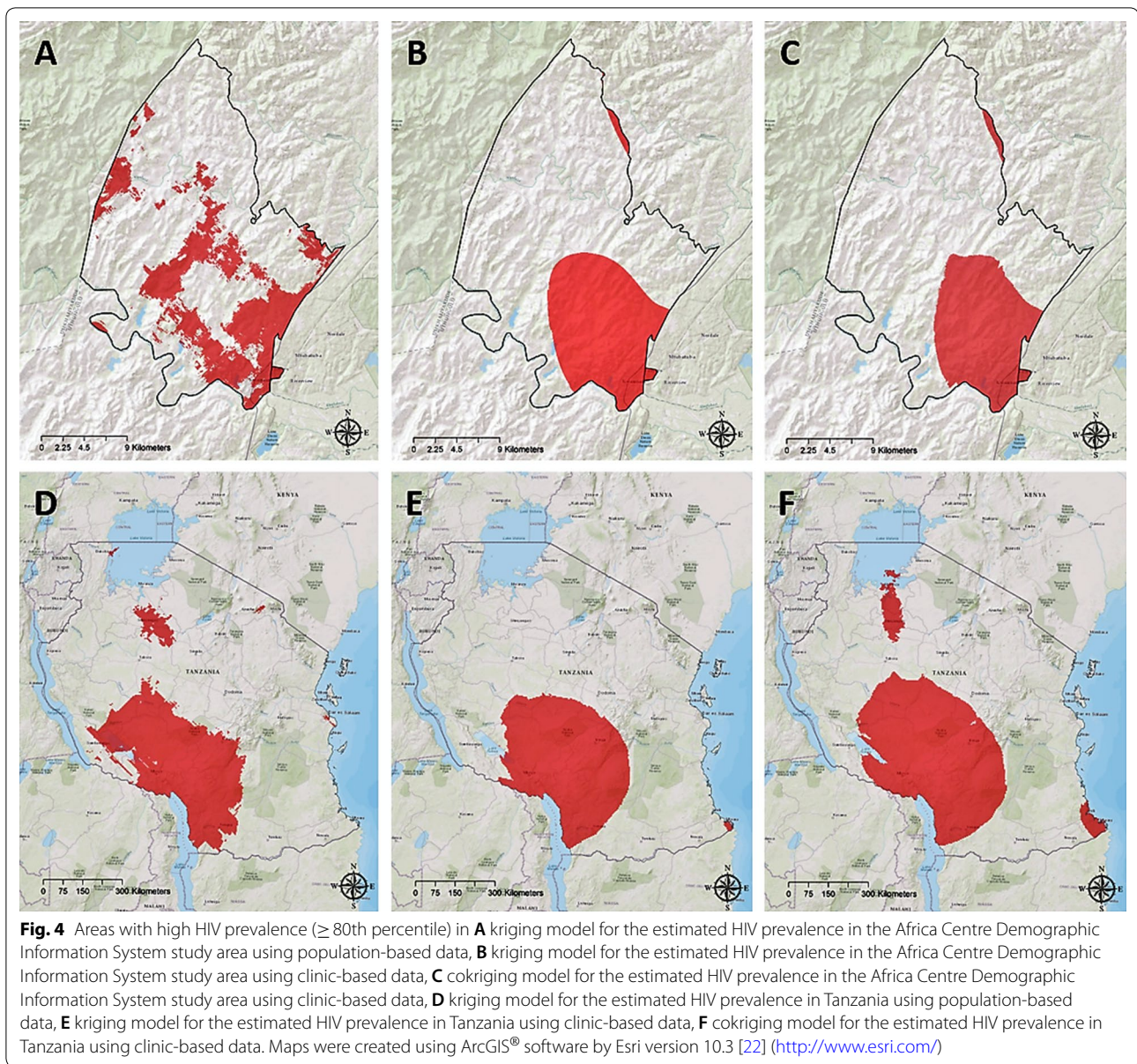
Map in Fig. 4A illustrates the location of the HIV ‘hot-spots’ (areas with HIV prevalence in the upper quintile) in the ACDIS study area identified using population-based data. Kriging model map generated using clinic-based data (Fig. 4B) accurately predicted 54% of these high HIV burden areas, whereas cokriging model map (Fig. 4C) accurately located 56% of these high HIV prevalence areas.

### Tanzania (national level)

Continuous surface maps for the HIV prevalence in Tanzania are illustrated in Fig. 3. Mean pixel-level HIV prevalence in this area estimated using population-based data was 5.2% (95% CI 1.0–9.4%). The burden of the infection appears to be concentrated in the south-western part of the country, between the districts of Mbeya and Iringa, where the HIV prevalence can reach more than 9% (Fig. 3A).



**Fig. 3** Continuous surface maps of **A** the estimated HIV prevalence using the Tanzania Demographic and Health Survey data, **B** kriging model for the estimated HIV prevalence using clinic-based data, **C** residuals from the kriging model, **D** LISA analysis of the kriging model, **E** cokriging model for the estimated HIV prevalence using clinic-based data, **F** residuals from the cokriging model, **G** LISA analysis of the cokriging model. Maps were created using ArcGIS® software by Esri version 10.3 [22] (<http://www.esri.com/>)



Mean pixel-level HIV prevalence in Tanzania estimated using clinic-based data was 6.5% (95% CI 1.6–11.3%). Kriging interpolation of clinic-based data identified all high burden areas of HIV infection detected using population-based data (Fig. 3B). However, residual analysis indicated that clinic-based data could overestimate the HIV prevalence in higher burden areas (Fig. 3C). LISA analysis indicated that estimations from kriging interpolation using DHS and clinic-based data were consistent identifying high or low burden areas in 71% of the area in Tanzania (Fig. 3D). The estimations from these two models diverged in 16% of the country area. Kriging interpolation using clinic-based data predicted high

HIV burden areas in low HIV prevalence areas in 9% of the study area, and low HIV burden areas in high HIV prevalence areas in 7% of the study area (Additional file 1: Figure S7). There was a non-statistical significant spatial association in the remaining 13% of the study area. Similar to the results from the ACDIS study area, inclusion of cofactors in the cokriging model moderately increased the accuracy of predictions (Fig. 3E, F), consistently identifying high or low burden areas in 77% of the study area (Fig. 3G, Additional file 1: Figure S7).

Figure 4D maps the location of the HIV ‘hotspots’ (areas with HIV prevalence in the upper quintile) in Tanzania identified using population-based data. Kriging

model map generated using clinic-based data (Fig. 4E) accurately predicted 77% of these high HIV burden areas, whereas cokriging model map (Fig. 4F) accurately located 84% of these high HIV prevalence areas.

## Discussion

Our results suggest that clinic-based data are able to capture the broad spatial structure of HIV epidemics in these hyperendemic settings. Analysis of this information accurately identified the high HIV burden areas (HIV 'hotspots'), thereby offering a less expensive and readily available alternative source of data for the geographical identification of vulnerable populations at high risk of infection.

Accuracy of these predictions varied depending on the geographical scale in which the comparisons were conducted. Clinic-based HIV prevalence data successfully captured the broad spatial structure of the HIV epidemic in the areas studied, but the accuracy was reduced to some extent when the resolution of the geographical scale was increased (i.e. from national to local scale). For example, LISA results showed consistency of the HIV estimations in 59% of area in the ACDIS study area (local level), whereas it showed consistency of the HIV estimations in 71% of the area in Tanzania (national level). This discrepancy could be the result of reducing the number of healthcare facilities (data-points) in the analysis when the resolution of the scale is increased. For example, the HIV prevalence map generated for Tanzania included data from more than 100 healthcare facilities distributed across the country. In contrast, mapping the spatial distribution of HIV at sub-district level, focusing on a single town, only included 10 healthcare facilities. As a result, the statistical power and geographic resolution of the spatial interpolation could be reduced, amplifying discrepancies between population-based and clinic-based estimations. However, it is important to note that ACDIS has the unusual aspect of being bordered on the west by an uninhabited area, the Hluhluwe–Imfolozi national park. For that reason, there were no facility data points to improve boundary estimates on these areas. Despite this limitation, clinic-based data still captured most of the geographical variation of the HIV epidemic at local level, and located most of the high burden areas identified in previous studies, where the HIV epidemic is largely concentrated in the ACDIS study area [4, 5, 15, 16].

Inclusion of cofactors using cokriging method moderately increased the accuracy of spatial HIV prevalence predictions, particularly at national level. Nevertheless, it is important to note that this was an exploratory analysis, and only few cofactors were included in our study. Inclusion of more behavioral and biological cofactors associated with the risk of HIV infection such as male

circumcision, condom use, life time number of sexual partners, and wealth index among others could effectively improve model predictions as it has been shown in previous studies [2, 9].

The primary limitation of the approach proposed here is the comparability across different sources of data collected from different sampled populations. Community-level population surveillances target the general population, whereas clinic-based surveillance systems collect data from specific subpopulations who seek care, such as pregnant women, or individuals at high risk of infection that are frequently tested or seeking treatment. Furthermore, while clinic-based data usually contain geographically representative of nearby populations, catchment areas are variable and HIV positive populations may be more likely to seek care at specific medical facilities with specialized services [28]. As expected, the clinic-based data over-estimated prevalence [29, 30]. Despite the fact that clinics are located following some decision rule, which is not completely spatially random regarding epidemic burdens, clinic-based data are still able to capture the spatial patterns, which is the ultimate goal of the approach proposed here. Lastly, we conducted our analysis using ESRI ArcGIS software, but alternative open access software such as QGIS (<https://qgis.org/en/site/>), GRASS (<https://grass.osgeo.org/>), and R (<https://www.r-project.org/>), among others, are suitable to conduct similar spatial analyses as the ones conducted in this study.

## Conclusions

Our results suggest that analysis of clinic-based data could provide robust estimations of the broad spatial structure of the HIV epidemic in hyperendemic settings. However, similar analyses should be conducted in other African settings to assess whether these patterns are observed elsewhere. Our study illustrates the potential utility of routine HIV testing data collected from different healthcare facilities to identify and monitor the high HIV burden areas. This methodology would overcome the challenging methodological and economic issues that accompany collecting community-level population-based data. Routine HIV testing data collected in different healthcare facilities as well as other sources of data from local small HIV sample surveys would allow for a rapid and cost-effective visualization of the epidemic, facilitating decision making to redefine resource allocation and surveillance systems, focused on the geographical HIV 'hotspots' that could be fueling the epidemic [5, 8, 12]. This approach may provide valid spatial prevalence estimates for geographical targeting where the burden of the infection is concentrated and where resources are needed the most.

## Additional file

Additional file 1. Supplementary materials.

### Abbreviations

HIV: human immunodeficiency virus; SSA: sub-Saharan Africa; UNAIDS: Joint United Nations Programme on HIV/AIDS; PEPFAR: United States President's Emergency Plan for AIDS Relief; ACDIS: Africa Centre Demographic Information System; CAPRISA: AIDS Programme of Research in South Africa; DHIS: district health information system; DHS: Demographic and Health Survey; PRCC: partial rank correlation coefficient; LISA: local indicators of spatial autocorrelation; CI: confidence interval.

### Authors' contributions

DFC contributed the study and its design, conducted the statistical and spatial modeling analyses, and wrote the first draft of the paper. BS and CH contributed to study conception and design, conduct of the statistical modeling analyses, interpretation of the results, and writing of the manuscript. AA, TB and FT contributed to study conception and design, interpretation of the results, and writing of the manuscript. All authors read and approved the final manuscript.

### Author details

<sup>1</sup> Department of Geography and Geographic Information Science, University of Cincinnati, Cincinnati, OH 45221, USA. <sup>2</sup> Health Geography and Disease Modeling Laboratory, University of Cincinnati, Cincinnati, USA. <sup>3</sup> Department of Public Health Medicine, School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa. <sup>4</sup> Geographical Information Systems and Science Program, Kingston University, London, UK. <sup>5</sup> Institute for Disease Modeling, 3150 139th Ave SE, Bellevue, USA. <sup>6</sup> Africa Health Research Institute, University of KwaZulu-Natal, Durban, South Africa. <sup>7</sup> Heidelberg Institute for Public Health, University of Heidelberg, Heidelberg, Germany. <sup>8</sup> Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Boston, USA. <sup>9</sup> School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa.

### Acknowledgements

The authors thank Measure Demographic and Health Surveys (Measure DHS) for releasing these national surveys in the service of science, and the United States Agency for International Development and other donors supporting these initiatives.

### Competing interests

The authors declare that they have no competing interests.

### Availability of data and materials

The data that support the findings of this study are available from the Demographic and Health Surveys (<http://www.measuredhs.com>) and the Africa Centre Demographic Information System (<https://www.ahri.org/research/>) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Demographic and Health Surveys.

### Ethics approval and consent to participate

Procedures and questionnaires for standard Demographic Health Surveys (DHS) have been reviewed and approved by the ICF International Institutional Review Board (IRB). Additionally, country-specific DHS survey protocols are reviewed by the ICF IRB and typically by an IRB in the host country. The ICF International IRB ensures that the survey complies with the U.S. Department of Health and Human Services regulations for the protection of human subjects, while the host country IRB ensures that the survey complies with laws and norms of the nation. In the original primary data collection for each DHS, informed consent was sought from all participants prior to serological testing for HIV (<http://dhsprogram.com/What-We-Do/Protecting-the-Privacy-of-DHS-Survey-Respondents.cfm#sthash.Ot3N7n5m.dpuf>). We sought and were granted permission to use the core dataset for this analysis by MEASURE

DHS. Ethics approval for data collection and use from the Africa Centre Demographic Information System data was obtained from the biomedical and ethics committee of the University of KwaZulu-Natal (Durban, South Africa).

### Funding

DC and FT were supported by the South African Medical Research Council (SA MRC) Flagship Grant (MRC-RFA-UFSP-01–2013/UKZN HIVEPI). FT was supported by two National Institute of Health (NIH) Grants (R01HD084233 and R01AI124389) as well as a UK Academy of Medical Sciences Newton Advanced Fellowship (NA150161). TB is funded by the Alexander von Humboldt Foundation through the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research. He is also supported by the Wellcome Trust, the European Commission, the Clinton Health Access Initiative and NICHD of NIH (R01-HD084233), NIAID of NIH (R01-AI124389 and R01-AI12339) and FIC of NIH (D43-TW009775).

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 23 March 2018 Accepted: 2 July 2018

Published online: 11 July 2018

### References

- Zulu LC, Kalipeni E, Johannes E. Analyzing spatial clustering and the spatiotemporal nature and trends of HIV/AIDS prevalence using GIS: the case of Malawi, 1994–2010. *BMC Infect Dis*. 2014;14(1):285.
- Cuadros DF, Li J, Branscum AJ, Akullian A, Jia P, Mziray EN, Tanser F. Mapping the spatial variability of HIV infection in sub-Saharan Africa: effective information for localized HIV prevention and control. *Sci Rep*. 2017;7(1):9093.
- Cuadros DF, Awad SF, Abu-Raddad LJ. Mapping HIV clustering: a strategy for identifying populations at high risk of HIV infection in sub-Saharan Africa. *Int J Health Geogr*. 2013;12(1):28.
- Tanser F, de Oliveira T, Maheu-Giroux M, Bärnighausen T. Concentrated HIV sub-epidemics in generalized epidemic settings. *Curr Opin HIV AIDS*. 2014;9(2):115.
- Tanser F, Bärnighausen T, Cooke GS, Newell M-L. Localized spatial clustering of HIV infections in a widely disseminated rural South African epidemic. *Int J Epidemiol*. 2009;38(4):1008–16.
- Cuadros DF, Branscum AJ, Mukandavire Z. Temporal stability of HIV prevalence in high burden areas regardless of declines in national HIV prevalence in Malawi and Zimbabwe. *London: AIDS (Lond, Engl)*; 2018.
- Meyer-Rath G, McGillen JB, Cuadros DF, Hallett TB, Bhatt S, Wabiri N, Tanser F, Rehle T. Targeting the right interventions to the right people and places: the role of geospatial analysis in HIV program planning. *AIDS*. 2018;32(8):957–63.
- Anderson S-J, Cherutich P, Kilonzo N, Cremin I, Fecht D, Kimanga D, Harper M, Masha RL, Ngongo PB, Maina W. Maximising the effect of combination HIV prevention through prioritisation of the people and places in greatest need: a modelling study. *The Lancet*. 2014;384(9939):249–56.
- Hallett T, Anderson S-J, Asante CA, Bartlett N, Bendaud V, Bhatt S, Burgert C, Cuadros DF, Dzangare J, Fecht D. Evaluation of geospatial methods to generate subnational HIV prevalence estimates for local level planning. *AIDS*. 2016;30(9):1467–74.
- Aral SO, Torrone E, Bernstein K. Geographical targeting to improve progression through the sexually transmitted infection/HIV treatment continua in different populations. *Curr Opin HIV AIDS*. 2015;10(6):477–82.
- Grantham KL, Kerr CC, Wilson DP. Local responses to local epidemics for national impact need advanced spatially explicit tools. *AIDS*. 2016;30(9):1481–2.
- Dobra A, Bärnighausen T, Vandormael A, Tanser F. Space-time migration patterns and risk of HIV acquisition in rural South Africa. *AIDS (Lond, Engl)*. 2017;31(1):137.
- Controlling the epidemic: delivering on the promise of an AIDS-free Generation, PEPFAR 3.0. 2014. <https://www.pepfar.gov/documents/organization/234744.pdf>.



14. Joint United Nations Programme on HIV/AIDS. UNAIDS 2016–2021 strategy: on the fast-track to end AIDS. Geneva: Joint United Nations Programme on HIV/AIDS; 2015.
15. Tanser F, Bärnighausen T, Grapsa E, Zaidi J, Newell M-L. High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. *Science*. 2013;339(6122):966–71.
16. Tanser F, Vandormael A, Cuadros D, Phillips AN, de Oliveira T, Tomita A, Bärnighausen T, Pillay D. Effect of population viral load on prospective HIV incidence in a hyperendemic rural African community. *Sci Transl Med*. 2017;9(420):eaam8012.
17. Dee J, Calleja JMG, Marsh K, Zaidi I, Murrill C, Swaminathan M. HIV surveillance among pregnant women attending antenatal clinics: evolution and current direction. *JMIR Public Health Surveill*. 2017;3(4):e85.
18. Tanzania Commission for AIDS ZAC, National Bureau of Statistics, Office of Chief Government Statistician, ICF International. Tanzania HIV/AIDS and Malaria Indicator Survey 2011–12. Calverton, MD: ICF International; 2013.
19. Tanzania Commission for AIDS ZAC, National Bureau of Statistics, Office of Chief Government Statistician, Macro International Inc. Tanzania HIV/AIDS and Malaria Indicator Survey 2007–08. Calverton, MD: Macro International Inc.; 2008.
20. Tanzania Commission for AIDS NBoS, ORC Macro. Tanzania HIV/AIDS Indicator Survey 2003–04. Calverton, MD: ORC Macro; 2005.
21. The United Republic of Tanzania, Ministry of Health. National guidelines for voluntary counseling and testing, 2005. National AIDS control programme (NACP) 2005.
22. ESRI. ArcGIS 10.x. Redlands, CA: ESRI; 2004.
23. Berke O. Exploratory disease mapping: kriging the spatial risk function from regional count data. *Int J Health Geogr*. 2004;3(1):18.
24. Goovaerts P. Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. *Int J Health Geogr*. 2005;4(1):31.
25. Carrat F, Valleron A-J. Epidemiologic mapping using the “kriging” method: application to an influenza-like epidemic in France. *Am J Epidemiol*. 1992;135(11):1293–300.
26. Naish S, Dale P, Mackenzie JS, McBride J, Mengersen K, Tong S. Spatial and temporal patterns of locally-acquired dengue transmission in northern Queensland, Australia, 1993–2012. *PLoS ONE*. 2014;9(4):e92524.
27. Oliver MA, Webster R. Kriging: a method of interpolation for geographical information systems. *Int J Geogr Inf Syst*. 1990;4(3):313–32.
28. Akullian AN, Mukose A, Levine GA, Babigumira JB. People living with HIV travel farther to access healthcare: a population-based geographic analysis from rural Uganda. *J Int AIDS Soc*. 2016;19(1):20171.
29. Zaba BW, Carpenter LM, Boerma JT, Gregson S, Nakiyingi J, Urassa M. Adjusting ante-natal clinic data for improved estimates of HIV prevalence among women in sub-Saharan Africa. *AIDS*. 2000;14(17):2741–50.
30. Boisson E, Nicoll A, Zaba B, Rodrigues LC. Interpreting HIV seroprevalence data from pregnant women. *JAIDS J Acquir Immune Defic Syndr*. 1996;13(5):434–9.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

