OPEN

Finding Hidden HIV Clusters to Support Geographic-Oriented HIV Interventions in Kenya

Anthony Waruru, MPhil,† Thomas N. O. Achia, PhD,† James L. Tobias, MSc,‡ James Ng'ang'a, MSc,§ Mary Mwangi, PhD,† Joyce Wamicwe, MBChB,‖ Emily Zielinski-Gutierrez, PhD,† Tom Oluoch, PhD,‡ Evelyn Muthama, MSc,† and Thorkild Tylleskär, PhD**

Background: In a spatially well known and dispersed HIV epidemic, identifying geographic clusters with significantly higher HIV prevalence is important for focusing interventions for people living with HIV (PLHIV).

Methods: We used Kulldorff spatial-scan Poisson model to identify clusters with high numbers of HIV-infected persons 15–64 years old. We classified PLHIV as belonging to either higher prevalence or lower prevalence (HP/LP) clusters, then assessed distributions of sociodemographic and biobehavioral HIV risk factors and associations with clustering.

Results: About half of survey locations, 112/238 (47%) had high rates of HIV (HP clusters), with 1.1–4.6 times greater PLHIV adults observed than expected. Richer persons compared with respondents in lowest wealth index had higher odds of belonging to a HP cluster, adjusted odds ratio (aOR) 1.61 [95% confidence interval (CI): 1.13 to 2.3], aOR 1.66 (95% CI: 1.09 to 2.53), aOR 3.2 (95% CI: 1.82 to 5.65), and aOR 2.28 (95% CI: 1.09 to 4.78) in second, middle, fourth, and highest quintiles, respectively. Respondents who perceived themselves to have greater HIV risk or were already HIV-infected had higher odds of belonging to a HP cluster, aOR 1.96 (95% CI: 1.13 to 3.4) and aOR 5.51 (95% CI: 2.42 to 12.55), respectively; compared with perceived low risk. Men who had ever been clients of female sex worker had higher odds of belonging to a HP cluster than those who had never been, aOR 1.47 (95% CI: 1.04 to 2.08); and uncircumcised men vs circumcised, aOR 3.2 (95% CI: 1.74 to 5.8).

Conclusions: HIV infection in Kenya exhibits localized geographic clustering associated with sociodemographic and behavioral factors, suggesting disproportionate exposure to higher HIV risk. Identification of these clusters reveals the right places for targeting priority-tailored HIV interventions.

Key Words: HIV/AIDS, clustering, geographic differences, Kulldorff spatial-scan statistics, Kenya

(*J Acquir Immune Defic Syndr* 2018;78:144–154)

BACKGROUND

Kenya accounts for 6% of people living with HIV (PLHIV) in sub-Saharan Africa (SSA) and has the fourth highest adult HIV prevalence in the world and fifth in new HIV infections.[1] In 2015, Kenya was estimated to have over 1.5 million HIV-infected adults and children and an annual incidence of about 0.3% among adults aged 15–49 years.[2,3] In Kenya, HIV care and treatment efforts in the past 2 decades have led to reduced HIV prevalence from 6.8% in 2007 to 5.6% in 2012 among adults 15–49 years old.[4] However, 65% of new infections occur in just 9 of the country's 47 counties.[5] At county level, estimated adult HIV prevalence ranges from 24.8% in Siaya to 0.4% in Wajir.[3] These disparate burdens present an impetus to plan for better use of resources in HIV epidemic control.

Traditionally, Kenya's national HIV program uses UNAIDS estimation and projections package (EPP) models to estimate national and county HIV prevention and treatment needs.[6] Mathematical models provide projections of HIV epidemics, resource allocation, and interventions at subnational levels.[7] For example, based on modeling, it is estimated that 14% more infections in Kenya could be averted over a 15-year period (2014–2029) if resources were targeted to the most effective interventions and regions most in need.[8] Processes such as commodities estimates and quantification of HIV care and treatment depend heavily on these kind of estimates. However, rarely do such processes focus on more granular units beyond the subnational (county) level. With a worldwide commitment to end the HIV epidemic by 2030,[9,10] a more granular versus broadly generalized spatial epidemiological analysis to identify hidden clusters and HIV infection patterns is important because it helps to better target

Received for publication September 12, 2017; accepted January 29, 2018.

From the *Centre for International Health, University of Bergen, Bergen, Norway; †Division of Global HIV & TB, U.S Centers for Disease Control and Prevention (CDC), Nairobi, Kenya; ‡Division of Global HIV & TB, U.S Centers for Disease Control and Prevention, Atlanta, GA; §National Bureau of Statistics, Nairobi, Kenya; and ‖National AIDS and STI Control Program (NASCOP), Ministry of Health, Nairobi, Kenya.

Supported by the US President's Emergency Plan for AIDS Relief (PEPFAR) through cooperative agreements (#PS001805, GH000069-05, and PS001814) from the US Centers for Disease Control and Prevention (CDC), Division of Global HIV &TB (DGHT). KAIS 2012 was also funded in part by support from the Global Fund, World Bank, and the Joint United Nations Team for HIV/AIDS.

Presented as a poster in CROI 2017; February 13–16, 2017, Seattle, WA.

The authors have no conflicts of interest to disclose.

Correspondence to: Anthony Waruru, MPhil, P.O Box 606, Nairobi, 00621, Kenya (e-mail: awaruru@cdc.gov).

interventions. Conversely, spatially overgeneralized HIV estimates can mask the true pattern of HIV and may lead to inefficient allocation of resources and missed opportunities for prevention and treatment. Pin-pointing cluster differences that impact on HIV diagnosis and linkage to care continuum outcomes could be performed.[11] Other assessments of geographical features such as transportation accessibility may show cluster variations impacting linkage to care and outcomes such as viral load suppression.[12]

Kulldorff and Nagarwalla have described a general method for disease cluster detection using a scan window of a specified size and scans through space to identify clusters.[13] The method tests if the number of individuals with a disease occurs at random over space or if any clusters can be detected. This method has been applied in multiple studies owing to its efficiency in identifying clusters. Spatial clustering patterns may exist in relation to physical geographical features: for example, a road or railway network, commercial activity such as large farms and their positioning or proximity of clusters to such features. Cluster analysis and mapping may be used to identify "micro-epidemics" in certain geographic areas such as rural communities,[14] populations at higher risk of HIV infection,[15] demonstrate decline in HIV prevalence,[16] key populations such as female sex workers (FSWs),[17] and for assessing access to HIV-related services such as HIV testing and counseling[18,19] or provision of antiretroviral treatment.[20] In addition, mapping may be used for allocation of targets,[7] thereby improving efficiencies in resource allocation.

HIV spatial clustering studies have rarely used nationwide population-based survey data. Yet, there is opportunity to explore associations between spatial clustering, individual, and ecological characteristics. Population-based surveys provide high quality demographic, socioeconomic, and behavioral data at individual level and additional ecological data that relate to where individuals live which may determine risk of exposure including sexual partnerships. In Kenya, HIV epidemic is generally considered to be concentrated in 5 high-burden counties (Homabay, Kisumu, Migori, Siaya, and Nairobi) with 50% of the unmet antiretroviral treatment need.[21] We conducted geospatial analysis on data from a nationally representative Kenya AIDS Indicator Survey (KAIS) 2012 to identify clusters with higher rates of HIV-infected persons 15–64 years old. In this article, we have described spatial-epidemic clustering of HIV prevalence in Kenya beyond the well-known subnational pattern. We have also explored relationships between HIV clustering and sociodemographic and behavioral risk indicators, awareness of HIV status and access to HIV services such as HIV testing services (HTS), and voluntary medical male circumcision.

METHODS

Study Design and Population

The methods used in KAIS 2012 have been previously described.[22] In brief, KAIS is a cross-sectional household survey whose target population was adults aged 15–64 years and children aged 18 months to 14 years. The survey was conducted from October 2012 to February 2013 using a stratified 2-stage cluster (survey locations) sample to identify households and within households, eligible respondents were interviewed. The North Eastern region was not surveyed because of regional insecurity. KAIS 2012 study locations were sampled from the National Sample Survey and Evaluation Programme (NASSEP V) sampling frame which is developed and maintained by the Kenya National Bureau of Statistics (KNBS). All households in each sampled cluster were geocoded using Global Information System (GIS), at the time of sampling frame development. However, to provide for confidentiality, cluster geocentroids were used for data aggregated at cluster level. This analysis was restricted to respondents aged 15–64 years with a confirmed HIV-positive test result from the national testing laboratory.

Data Collection Methods

Participants were interviewed using a standardized questionnaire regarding household and demographic characteristics, biobehavioral factors, and use of HIV-related services such as HTS and voluntary medical male circumcision. Use of Information and Communication Technologies (ICT) for data collection in KAIS has been described elsewhere.[23] Data were collected on tablet computers (Mirus Innovations, Mississauga, Ontario, Canada) and securely transmitted electronically to a central database in Nairobi. Blood was obtained and tested for HIV antibodies at the National HIV Reference Laboratory (NHRL) using the Vironostika HIV-1/2 UNIF II Plus O Enzyme Immunoassay (bioMérieux, Marcy d'Etoile, France) as the screening assay and the Murex HIV.1.2.O HIV Enzyme Immunoassay (DiaSorin, SpA, Saluggia, Italy) as the confirmatory assay.

Selection of Variables

Commonly, determinants of HIV epidemic are grouped into 3: sociocultural, socioeconomic, and epidemiological. Behavioral factors may include condom use, age at the onset of sexual activity, and sexual intercourse with multiple and nonregular partners; others include condom use, age at the onset of sexual activity, and sexual intercourse with multiple and nonregular partners.[24] We included sex, age, region, residence, and wealth index (calculated based on household characteristics and ownership of assets measures captured in the household questionnaire). The wealth index was generated using factor analysis calculated using standard methods,[25] and the resulting indices were grouped into 5 quintiles from lowest to highest. Other key variables included: marital status, education level, employment, travel, awareness of HIV status, ever tested, being sexually active in the past year, consistent condom use, number of lifetime sex partners, and among men, circumcision and ever having been a client of FSWs.

Spatial-Scan Methodology

We determined clusters with significant numbers of HIV-infected individuals by using a Poisson-based model performed through spatial-scan statistics program SaTScan

version 9.4,[26] (downloadable from http://www.satscan.org). The number of HIV-positive persons in a cluster was assumed to be Poisson-distributed according to an underlying population at risk. The population at risk was determined from the number of respondents tested and inverse calculation based on survey weights. The spatial-scan statistic was calculated using likelihood ratio test and Monte Carlo simulations to calculate the maximum likelihood ratio.[13,27] We set the maximum number of standard Monte Carlo replications to 999 and considered a cluster to be statistically significant when its log likelihood ratio (LLR) was greater than the standard Monte Carlo critical values at $P < 0.05$. The most likely clusters were reported alongside their log likelihood ratio, relative risks, and P values.

The Kulldorff spatial cluster detection looped over all the 358 survey locations. Using the observed cases, the likelihood of each cluster being a high prevalence (HP) or low prevalence (LP) was computed using a purely spatial Poisson model. For the purpose of classifying HP clusters and grouping the PLHIV as belonging to HP clusters, we took all study clusters that had a significant P value of <0.05 identified after cluster analysis and grouped respondents as belonging to HP clusters and conversely for LP clusters. To avoid detection of large clusters, we assumed a maximum of 30% of the population were at risk and defined a sizeable scan window with a maximum diameter of 100 km for 3 reasons; comparison with other studies, for example,[15,28] reasonable HIV program implementation reach and to avoid biasing our analysis to smaller clusters. Consideration for a high proportion of population at risk has been suggested by Kulldorff et al.[13] Conventionally, if the proportion at risk is unknown, 50% is set as the default but may lead to identification unnecessarily large and less informative clusters.[29] The proportion of adults at highest risk of HIV infection is unknown because in general population surveys, it is not possible to segregate most-at-risk populations. Given that 34% of the population in Kenya are aged 15–24 years,[30] to be modest in our estimation, we assumed 30% of the adult population were at greatest risk in a generalized epidemic such as Kenya's. This age category corresponds to those who have highest HIV prevalence.[4]

Characterizing Respondents in HP Clusters

Results of identified clusters were imported into Statistical Analysis System (SAS) version 9.3 for statistical analyses.[31] We classified persons as belonging to HP vs LP clusters. Using this classification, we assessed distributions and associations of clustering with sociodemographic and biobehavioral HIV risk factors. We used PROC SURVEYFREQ in SAS to do χ^2 tests to compare weighted proportions. We tested for associations for social demographic, behavioral, male circumcision, and HTS utilization to belonging to a HP cluster using PROC SURVEYLOGISTIC in SAS and presented both unadjusted and adjusted odds ratios (aORs). Mapping of identified clusters and related spatial features was performed using Quantum GIS (QGIS) version 2.16.

Characterizing HP Clusters

After mapping, we visually reviewed the clusters of interest (HP clusters) regarding proximity to or located within features of interest such as trade centers, commercial activities such as large tea plantations or flower farms, near an informal settlement, etc. We added an Open Street Map layer and overlaid the centroid coordinates displayed as circles of varying sizes depending on the estimated number of HIV-infected to qualitatively characterize HP clusters.

Ethics Approval and Consent to Participate

This study was approved by the Kenya Medical Research Institute's (KEMRI) Ethical Review Committee (ERC), the US Centers for Disease Control and Prevention's (CDC) Institutional Review Board (IRB), and the Committee on Human Research of the University of California, San Francisco (UCSF).

Consent for Publication

At household level, the head of household consented to the household questionnaire; the heads of households were adults aged 18–64 years or emancipated individuals with no parent or guardian or not living with their parent/guardian. Individual consent or assent was sought by the field interviewer for all eligible household members to participate in the individual questionnaires. In the case of participants aged 10–17 years, consent was obtained from a parent/guardian or other adult responsible for the child/youth health and welfare before the child/youth was asked for his/her assent. Oral informed consent for HIV testing was required for adults and emancipated minors. Verbal informed consent with a signature of the interviewer on the consent form served as documentation of the consenting.

RESULTS

HIV Spatial Clustering Levels

Of 358 survey clusters, 238 (66.5%) had at least 1 HIV-infected person (Fig. 1). Of those, about half, 112/238 (47%) had high rates of HIV (HP clusters), with 1.1–4.5 times greater PLHIV 15–64 years old observed than expected. These were grouped into significant HP and LP clusters; 43 of 47 and 35 of 36, respectively (Table 1). Clusters were identified in multiple regions, with the larger clusters in Nyanza region and several in Nairobi, but also in central-Rift Valley, Central, and Coast regions (Fig. 1). The cluster with highest relative risk was near the Indian Ocean, Coast region; followed by 1 near Lake Victoria, Nyanza region; Mathare slums in Nairobi; 2 more in Nairobi region; 1 in a rural area near a tea plantation in Central Kenya; and another cluster in a rural area near a trade center, central-Rift Valley region (Fig. 2). The HP clusters had a median radius of 7.2 km, interquartile range (IQR) 3.3–10.9 km, whereas LP clusters had a median radius of 40.9 km, IQR 21.8–73.0 km (Table 1).

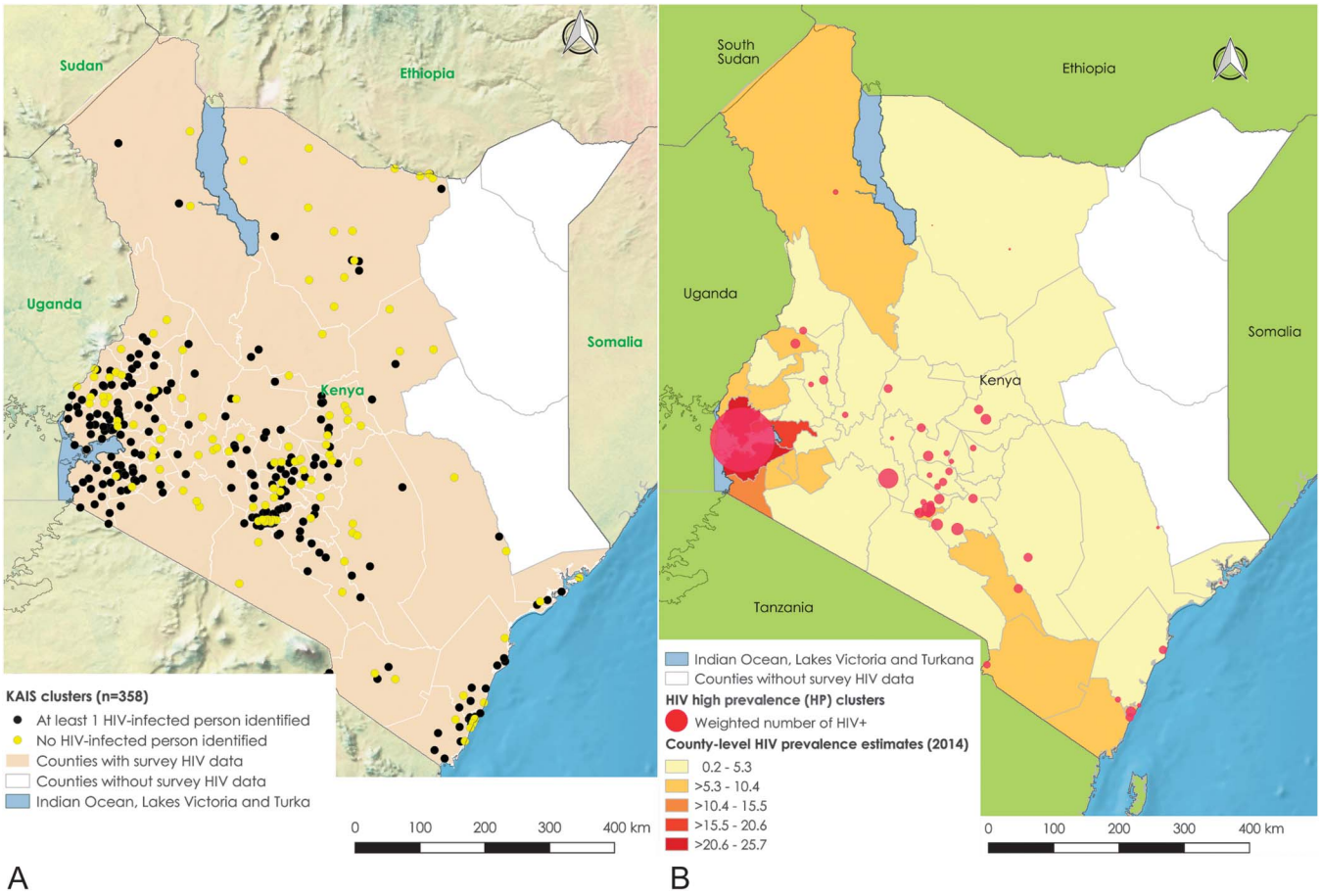


FIGURE 1. A, Distribution of 358 sampled survey locations, Kenya AIS 2012. The figure shows the distribution of 353 sampled survey locations in Kenya. The North eastern Kenya was not included in the survey. B, Distribution of HP clusters, Kenya AIS 2012, compared with well-known spatial distribution of HIV prevalence. The figure shows the HP clusters in Kenya. The base layer choropleth shows well-known spatial distribution of HIV prevalence at county level.

Characterizing HIV Clustering

Sociodemographic Factors and HIV Clustering

Nearly two-thirds of the respondents in LP clusters were from rural areas 66.7% [95% confidence interval (CI): 62.2 to 71.2] vs 53.3% (95% CI: 44.7 to 62.0) in HP clusters ($P = 0.025$). More respondents in LP than HP clusters were categorized as belonging to lowest wealth quintile 22.1% (95% CI: 18.4 to 25.9) vs 12.1% (95% CI: 9.3 to 14.9), $P = 0.002$. There were more widows/widowers living in HP clusters than in LP clusters; 8.5% (95% CI: 7.2 to 9.7) vs 6.3% (95% CI: 5.6 to 7.0), $P = 0.015$. Fewer respondents in HP clusters had no primary education, 4.6% (95% CI: 3.4 to 5.8) vs 8.0% (95% CI: 6.3 to 9.7) in LP clusters, $P = 0.008$. More respondents had traveled away from home for more than 1 month in HP vs LP clusters; 39.1% (95% CI: 36.0 to 42.1) vs 32.6% (95% CI: 30.6 to 34.5), $P < 0.001$.

Biobehavioral Factors and HIV Clustering

There were fewer circumcised men in HP vs LP clusters; 81.5% (95% CI: 76.8 to 86.2) vs 94.9% (95% CI: 93.7 to 96.1), $P < 0.001$. More men had ever been clients of FSWs in HP vs LP clusters; 23.3% (95% CI: 19.8 to 26.9) vs 15.3% (95% CI: 12.9 to 17.7), $P = 0.0002$.

Respondents were distributed similarly across HP and LP clusters by sex ($P = 0.070$), age ($P = 0.113$), employment status ($P = 0.195$), awareness of HIV status ($P = 0.185$), reported sexual activity in the past year ($P = 0.112$), and number of lifetime sex partners ($P = 0.151$) (Table 2).

Associations of Sociodemographic and Biobehavioral Factors With HIV Clustering

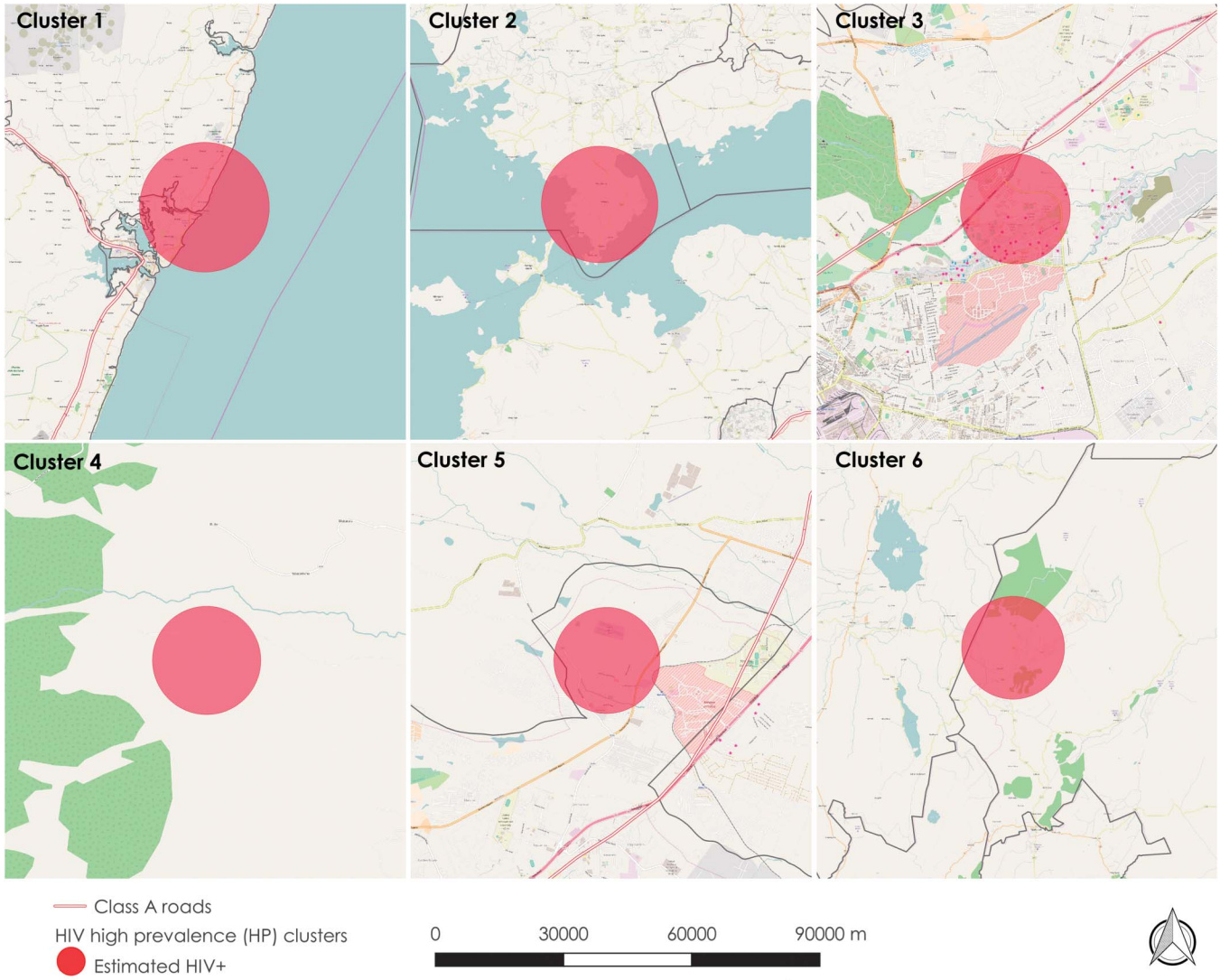
In adjusted analysis, persons in the second, middle, fourth, and highest wealth quintiles compared with those belonging to the lowest wealth index had higher odds of belonging to a HP cluster. Respondents who perceived themselves to have greater risk or already had HIV had higher odds of belonging to a HP cluster compared with those who perceived themselves as having no risk, aOR 1.96 (95% CI: 1.13 to 3.4) and aOR 5.51 (95% CI: 2.42 to 12.55), respectively; men who had ever been clients of FSW had higher odds of belonging to a HP cluster than

TABLE 1. Distribution of Cases in Significant HP and LP Clusters, Kenya AIDS Indicator Survey 2012

Clusters*	Type	Radius (km)†	Locations	*P*	Observed‡	Expected
All HP§	HP	—	112		771,136	360,082
1	HP	74.41	48	0.001	448,005	165,718.69
2	HP	38.43	6	0.001	40,806	25,558.14
3	HP	12.29	2	0.001	8336	6544.09
4	HP	10.48	2	0.001	15,326	7534.81
5	HP	8.91	2	0.001	7666	6085.42
6	HP	8.6	2	0.001	6173	3953.56
7	HP	5.83	2	0.001	6672	5415.06
8	HP	5.75	2	0.001	10,758	9455.37
9	HP	3.52	3	0.001	11,190	7964.56
10	HP	2.47	2	0.001	11,502	9423.98
11	HP	2.23	3	0.001	10,082	7853.92
12	HP	1.67	3	0.001	21,166	6435.71
13–43	HP	0	31	0.001	165,880	91,041.99
44	HP	0	1	0.32	3151	2960.14
45	HP	0	1	0.726	3546	3365.38
46	HP	0	1	0.742	191	151.67
47	HP	0	1	0.992	686	619.8
Median (IQR)		7.2 (3.3–10.9)				
All LP‖			185		225,634	622,225
1	LP	95.3	7	0.001	1399	9071.32
2	LP	93.7	5	0.001	0	3968.61
3	LP	93.62	2	0.001	0	1602.46
4	LP	91.54	5	0.001	191	1367.38
5	LP	89.55	43	0.001	61,947	141,623.52
6	LP	81.38	2	0.001	2218	3499.46
7	LP	70.25	2	0.001	0	6569.79
8	LP	67.97	11	0.001	25,058	52,656.06
9	LP	60.33	6	0.001	179	3264.32
10	LP	57.86	24	0.001	25,766	84,288.47
11	LP	51.74	2	0.001	0	265.07
12	LP	45.83	4	0.001	0	12,615.09
13	LP	35.94	5	0.001	6443	14,806.86
14	LP	35.48	3	0.001	0	15,054.23
15	LP	27.55	27	0.001	49,707	80,482.57
16	LP	24.44	5	0.001	0	25,184.3
17	LP	23.3	3	0.001	1952	5672.86
18	LP	22.45	2	0.001	0	10,623.45
19	LP	19.94	3	0.001	0	16,901.58
20	LP	15.09	2	0.001	8034	15,076.77
21	LP	14.72	6	0.001	2301	14,652.94
22	LP	6.32	2	0.001	4832	10,408.77
23	LP	5.66	10	0.001	23,988	44,431.4
24	LP	3.28	2	0.001	0	9074.82
25–35	LP	0	18	0.001	9944	37,265.91
36	LP	0	1	0.868	1675	1796.68
Median (IQR)	Median	40.9 (21.8–73)				

*47 HP clusters and 36 LP clusters identified. Analysis to detect HP and LP clusters were performed independent of each other.
†Radius of scan window set to 100 km for both HP and LP clusters.
‡Estimated number of cases in HP clusters (weighted n).
§HP, high prevalence.
‖LP, low prevalence.

those who had never been, aOR 1.47 (95% CI: 1.04 to 2.08); persons who had ever had an HIV test had higher odds of belonging to a HP than LP cluster aOR 1.45 (95% CI: 1.14 to 1.84) and uncircumcised men had higher odds of belonging to a HP cluster than circumcised, aOR 3.2 (95% CI: 1.74 to 5.8) (Table 3).



Cluster 1, relative risk 4.6, near Indian ocean, Coast region; **Cluster 2**, relative risk 3.7, near Lake Victoria, Nyanza region; **Cluster 3**, relative risk 3.3, Mathare slums, Nairobi, Nairobi region; **Cluster 4**, relative risk 3.2, near a tea plantation, Central region; **Cluster 5**, relative risk 3.1, near the Northern bypass, Nairobi region; **Cluster 6**, relative risk 2.9, near a trade center, Central-Rift region

FIGURE 2. Highest burden HP clusters, Kenya AIS 2012, in relation to geographic features of interest. The figure shows the distribution of top 4 highest burden HP clusters in order of relative risk. The base layer shows geographic features of interest where these clusters are located.

DISCUSSION

Why These Analyses Are Important

The results of our analyses indicate that it is possible to obtain more granular geographical variation in HIV prevalence, expanding our insights into the spatial distribution of HIV in Kenya. In our analyses, we have demonstrated that survey data can be used to show HP clusters in unexpected LP regions in a generalized epidemic. Our findings strengthen the move to support geographically targeted intervention packages for HIV-infected persons and/or most-at-risk. Information about these smaller foci in otherwise LP areas is especially important with the programmatic focus on the subnational units with the highest HIV burden and prevalence.

HIV Spatial Clustering in Kenya

Our observed spatial variation in HIV prevalence highlights clustered HIV prevalence across Kenya within microepidemics of varying magnitudes. Such microepidemics have been observed in other SSA countries.[32–34] High HIV prevalence clusters were found in medium generalized epidemic regions; for example, in some parts of Central Kenya and Central-Rift Valley that have considerably lower HIV prevalence of 3.8% and 4.3%, respectively, compared with the national prevalence of 5.6%.[4] Most of the HP clusters were near a lake/river, major road highway, economic hub, or in highly productive agricultural zones such as tea-growing areas and flower farms. HP clusters in Nairobi region were in informal settlements. Our study indicates that there

TABLE 2. Sociodemographic and Behavioral Characteristics by HIV Clustering, Kenya AIDS Indicator Survey 2012, N = 11,626

	From Clusters With High HIV Rates*			From Clusters With Low HIV Rates†			
Characteristic	**n**	**Weighted %**	**95% CI**	**n**	**Weighted %**	**95% CI**	***P***
Total	3305			8321			
Sex							0.0695
Men	1347	47.5	45.5 to 49.5	3489	49.7	48.5 to 51.0	
Women	1958	52.5	50.5 to 54.5	4832	50.3	49.0 to 51.5	
Age (yrs)							0.1132
Age 15–24	1115	33.8	31.5 to 36.1	2703	32.5	31.2 to 33.8	
Age 25–34	874	26.9	24.8 to 28.9	2336	27.8	26.3 to 29.2	
Age 35–44	597	17.6	16.2 to 19.0	1611	19.9	18.8 to 20.9	
Age 45+	719	21.7	19.5 to 23.9	1671	19.9	18.5 to 21.2	
Region							<0.0001
Nairobi	370	10	5.1 to 14.9	944	11.5	9.1 to 13.9	
Central	479	13.1	7.5 to 18.7	944	13.1	9.9 to 16.3	
Coast	228	4.6	1.6 to 7.7	1234	10.8	8.5 to 13.2	
Eastern	308	12.6	5.7 to 19.4	2013	16.3	13.1 to 19.5	
Nyanza	1030	31.8	24.6 to 39.0	601	6.7	3.5 to 9.9	
Rift valley	308	13.2	6.0 to 20.4	1759	32.1	27.4 to 36.9	
Western	582	14.6	8.1 to 21.1	826	9.5	6.8 to 12.1	
Residence							0.0248
Rural	1842	53.3	44.7 to 62.0	5659	66.7	62.2 to 71.2	
Urban	1463	46.7	38.0 to 55.3	2662	33.3	28.8 to 37.8	
Wealth quintiles							0.0022
Lowest	433	12.1	9.3 to 14.9	2001	22.1	18.4 to 25.9	
Second	761	21.7	17.4 to 26.1	1736	20.9	18.4 to 23.3	
Middle	706	19.9	16.8 to 23.0	1612	19.7	17.3 to 22.2	
Fourth	797	24.6	20.5 to 28.8	1380	17.2	14.6 to 19.8	
Highest	608	21.6	15.1 to 28.1	1592	20.1	16.6 to 23.6	
Marital status							0.0152
Never married or single	970	29.8	27.5 to 32.1	2501	31.5	29.8 to 33.2	
Widowed	299	8.5	7.2 to 9.7	595	6.3	5.6 to 7.0	
Separated or divorced	146	4.3	3.5 to 5.1	412	4.8	4.2 to 5.3	
Married or cohabiting	1889	57.4	54.9 to 59.8	4810	57.4	55.7 to 59.1	
Education level							0.0076
No primary	161	4.6	3.4 to 5.8	1177	8.0	6.3 to 9.7	
Incomplete primary	286	7.3	5.6 to 9.1	702	7.7	6.6 to 8.8	
Complete primary	1120	33.1	30.5 to 35.7	2574	32.3	30.6 to 34.0	
Secondary and above	1738	55	52.1 to 57.8	3868	52.0	49.7 to 54.2	
Currently employed							0.1945
Yes	1629	50.3	46.8 to 53.7	3563	47.4	45.1 to 49.8	
No	1672	49.7	46.3 to 53.2	4752	52.6	50.2 to 54.9	
Traveled in past 12 mo							<0.0001
Never traveled away	1477	45.6	42.2 to 48.9	4463	54.7	52.4 to 57.1	
Less than 1 month	458	15.4	13.2 to 17.5	1062	12.7	11.6 to 13.9	
One month or longer	1256	39.1	36.0 to 42.1	2587	32.6	30.6 to 34.5	
Aware of HIV status							0.1854
No	1672	49.7	46.3 to 53.2	4752	52.6	50.2 to 54.9	
Yes	1629	50.3	46.8 to 53.7	3563	47.4	45.1 to 49.8	
Ever had an HIV test							<0.0001
Yes	1629	50.3	46.8 to 53.7	3563	47.4	45.1 to 49.8	
No	1672	49.7	46.3 to 53.2	4752	52.6	50.2 to 54.9	
Circumcised (men)							<0.0001
Circumcised	1098	81.5	76.8 to 86.2	3297	94.9	93.7 to 96.1	
Uncircumcised	247	18.5	13.8 to 23.2	181	5.1	3.9 to 6.3	
Sexually active in past year							0.1122

TABLE 2. (*Continued*) Sociodemographic and Behavioral Characteristics by HIV Clustering, Kenya AIDS Indicator Survey 2012, N = 11,626

	From Clusters With High HIV Rates*			From Clusters With Low HIV Rates†			
Characteristic	**n**	**Weighted %**	**95% CI**	**n**	**Weighted %**	**95% CI**	***P***
Sexually active	2427	74.2	71.6 to 76.7	5780	71.7	70.1 to 73.2	
Not sexually active	878	25.8	23.3 to 28.4	2541	28.3	26.8 to 29.9	
Consistent condom use last sex							<0.0001
Yes	1629	50.3	46.8 to 53.7	3563	47.4	45.1 to 49.8	
No	1672	49.7	46.3 to 53.2	4752	52.6	50.2 to 54.9	
Lifetime sex partners							0.1512
One partner	376	11.5	9.4 to 13.5	1130	13.2	12.2 to 14.3	
More than 1	2911	88.5	86.5 to 90.6	7098	86.8	85.7 to 87.8	
Ever been a client of an FSW							0.0002
Never	899	76.7	73.1 to 80.2	2516	84.7	82.3 to 87.1	
Ever	266	23.3	19.8 to 26.9	418	15.3	12.9 to 17.7	

*Data are for adults from 47 HP clusters, 43 of them significant at ($P < 0.05$).
†Data are for adults from 36 LP clusters, 35 of them significant at ($P < 0.05$).

are pockets of higher HIV infection that otherwise may not be well described in a generalized and spatially diffused epidemic. Such pockets of higher HIV rates have been identified in other SSA countries even in the context of generalized epidemics.[14,15]

Population Distribution Within Clusters

In our analyses, we identified HP clusters that were about half in size compared with LP clusters. This may indicate more intense and localized pockets of infection, as opposed to diffused infections. We had similar number of men and women in both HP and LP clusters and similar age distributions by cluster types. This indicates that population demographics may not play a role in clustering as do behavioral or structural factors. Expectedly, a higher proportion of study clusters categorized as HP were in Nyanza region which has the highest HIV prevalence in Kenya, at 14.9% vs the 5.6% national prevalence.[4] We found that two-thirds of respondents in LP clusters were from rural areas compared with half of those from HP clusters indicating that degree of urbanization has a big role to play in clustering. This is corroborated by the higher income among residents in HP clusters and the smaller number of respondents in HP vs LP clusters reporting no education. There were more widows/widowers living in HP clusters than in LP clusters, perhaps an indication of higher HIV-associated mortality as has been described elsewhere in SSA.[35] The disproportionately high number of widows/widowers found in HP clusters may additionally mean a higher HIV-associated mortality of spouses in areas with disproportionate higher HIV rates.

Clustering, Social, and Behavioral Patterns

We found that there were more respondents in HP clusters who had traveled away from home for ≥1 month as compared to LP clusters. Although people travel for various reasons, most extended travel is work related. Work-related migration may mean access to disposable income and sexual partners during travel, hence higher potential for exposure to HIV infection. In our analysis, more residents in HP clusters had ever had an HIV test compared with those in LP clusters may affirm these results because use of HTS may be related to perception of risk. It has been observed in SSA that seeking HIV testing is associated with perception of higher risk,[36] and the converse may imply that perception of low risk may deter HIV testing. There was an overall pattern of greater sexual risk in HP clusters compared with LP clusters, with a higher proportion of respondents being sexually active, fewer persons reporting using condoms consistently, more lifetime sexual partners on average and also more men in HP clusters that had ever been clients of FSWs. Rates of male circumcision were lower in HP clusters, with lower circumcision rates associated with higher risk of HIV acquisition and transmission in multiple studies including randomized control trials in SSA.[37–43]

Limitations

Our analysis is subject to a few limitations. First, the actual population size data per cluster was not available; hence, we worked back to estimate the cluster population size using the household weights. KAIS 2012 did not include the former North Eastern Province, which generally has very low HIV prevalence; hence, the findings presented here may not be nationally representative. It would have been useful to assess whether there are HP clusters in a region where HIV prevalence is very low. KAIS was not designed to capture key populations such as FSWs, men who have sex with men, or person who inject drugs; hence, this analysis of clustering focuses on risks in the general population and may not explicitly reveal patterning related to KP spatial distribution. There may be other individual characteristics which we did not include, yet they may confound the associations with HIV

TABLE 3. Factors Associated With Being in a High HIV-Prevalence Cluster*, Kenya AIDS Indicator Survey 2012

	From HP Clusters			Unadjusted OR		aOR	
Characteristic*	**n/N**	**%**	**95% CI**	**OR (95% CI)**	***P***	**aOR (95% CI)**	***P***
Region							
Nairobi	370/1314	26.6	13.4 to 39.9	2.12 (0.8 to 5.63)	0.13	0.76 (0.23 to 2.45)	0.642
Central	479/1423	29.5	15.6 to 43.4	2.44 (0.93 to 6.44)	0.071	1.63 (0.55 to 4.86)	0.376
Coast	228/1462	15.1	4.6 to 25.5	1.04 (0.35 to 3.04)	0.945	0.7 (0.21 to 2.33)	0.566
Eastern	308/2321	24.2	10.3 to 38.1	1.87 (0.66 to 5.25)	0.236	2 (0.64 to 6.26)	0.232
Nyanza	1030/1631	66.4	51.1 to 81.7	11.57 (4.34 to 30.84)	<0.001	7.9 (2.8 to 22.26)	<0.001
Rift Valley	308/2067	14.6	5.8 to 23.4	ref		ref	
Western	582/1408	39.1	21.9 to 56.3	3.76 (1.37 to 10.28)	0.01	4.34 (1.43 to 13.2)	0.01
Education level							
No primary	161/1338	19.3	13.0 to 25.5	ref		ref	
Incomplete primary	286/988	28.3	21.1 to 35.5	1.65 (1.09 to 2.52)	0.019	1.71 (0.84 to 3.45)	0.138
Complete primary	1120/3694	29.9	24.4 to 35.3	1.79 (1.21 to 2.63)	0.003	1.57 (0.81 to 3.07)	0.185
Secondary+	1738/5606	30.5	25.0 to 36.1	1.84 (1.27 to 2.67)	0.001	1.54 (0.82 to 2.91)	0.179
Wealth quintiles							
Lowest	433/2434	18.5	13.1 to 24.0	ref		ref	
Second	761/2497	30.2	22.9 to 37.5	1.9 (1.43 to 2.53)	<0.001	1.61 (1.13 to 2.3)	0.008
Middle	706/2318	29.5	23.1 to 36.0	1.84 (1.27 to 2.68)	0.001	1.66 (1.09 to 2.53)	0.017
Fourth	797/2177	37.3	29.4 to 45.3	2.62 (1.65 to 4.16)	<0.001	3.2 (1.82 to 5.65)	<0.001
Highest	608/2200	30.9	20.9 to 40.9	1.96 (1.09 to 3.55)	0.025	2.28 (1.09 to 4.78)	0.029
Marital status							
Never married or single	970/3471	28.2	22.9 to 33.6	1.05 (0.81 to 1.35)	0.73	1.2 (0.76 to 1.9)	0.427
Widowed	299/894	35.9	29.1 to 42.6	1.49 (1.09 to 2.03)	0.013	1.2 (0.72 to 1.99)	0.486
Separated or divorced	146/558	27.3	20.8 to 33.9	ref		ref	
Married or cohabiting	1889/6699	29.3	24.2 to 34.5	1.1 (0.87 to 1.39)	0.404	1.43 (0.93 to 2.19)	0.102
Traveled in past 12 mo							
Never traveled away	1477/5940	25.5	20.6 to 30.4	ref		ref	
Less than 1 month	458/1520	33.3	26.5 to 40.1	1.45 (1.16 to 1.82)	0.001	1.36 (1.02 to 1.81)	0.036
One month or longer	1256/3843	33.1	27.3 to 38.8	1.44 (1.2 to 1.72)	<0.001	1.17 (0.9 to 1.51)	0.245
Risk perception							
No risk	1054/4212	25.8	20.7 to 30.9	ref		ref	
Small	1082/4241	26.4	21.1 to 31.7	1.03 (0.85 to 1.25)	0.762	1.13 (0.84 to 1.51)	0.426
Moderate	350/1023	36.2	29.1 to 43.2	1.63 (1.3 to 2.05)	<0.001	1.39 (0.96 to 2.0)	0.079
Great	184/457	39.3	30.8 to 47.8	1.86 (1.42 to 2.44)	<0.001	1.96 (1.13 to 3.4)	0.017
I already have HIV	170/247	68.6	58.3 to 79.0	6.29 (4 to 9.92)	<0.001	5.51 (2.42 to 12.55)	<0.001
Ever been a client of FSW							
Never	899/3415	26.8	21.6 to 31.9	ref		ref	
Ever	266/684	38.1	30.1 to 46.0	1.68 (1.28 to 2.21)	<0.001	1.47 (1.04 to 2.08)	0.029
Ever had an HIV test							
Yes	1629/5192	30.6	25.2 to 35.9	1.69 (1.4 to 2.04)	<0.001	1.45 (1.14 to 1.84)	0.003
No	1672/6424	28.2	22.9 to 33.5	ref		ref	
Circumcised (men)							
Circumcised	1098/4395	25.5	20.3 to 30.6	ref		ref	
Uncircumcised	247/428	59.3	49.4 to 69.1	4.26 (2.83 to 6.41)	<0.001	3.18 (1.74 to 5.8)	<0.001

*Excludes "don't know" category.

clustering. However, we believe that we captured the most important demographic and behavioral factors that can be easily collected in household interviews. Finally, the proportion of adults at risk of HIV infection is unknown because in general population surveys, it is not possible to segregate most-at-risk populations. Hence, we assumed 30% of the adult population were at greatest risk because we wanted to be modest in our estimation.

CONCLUSIONS

Our analyses provide information on finer geographic areas of focus in developing HIV prevention and treatment activities. Hence, resource allocation needs to be performed equitably as opposed to equally across regions and subnational units. The HIV program in our setting may need to rethink targeting of interventions for specific populations such

as workers at large commercial agricultural or transport corridors, or underserved population in urban informal settlements. Other considerations such as enhanced HTS for persons perceiving themselves as having a high risk of HIV acquisition, interventions targeting widows and widowers such as accelerated HTS and linkage to care and treatment. Finding hidden HIV clusters to support geographic-oriented HIV interventions in Kenya is an important consideration in developing country operational plans to improve resource allocation. We suggest the need to integrate more granular analyses, lest we overlook geographically smaller HP foci.

ACKNOWLEDGMENTS

The authors thank the survey teams for their work during KAIS data collection and all persons who participated in this national survey. The authors thank Angela Broad, Silas Mulwa, Serenita Lewis, Yakubu Owolabi, Timothy Kellogg, and Mike Grasso for their input on the development of the EDC system; Nicky Okeyo and Daniel Kwaro for their technical support; Wolfgang Hladik, Mike Grasso, and George Rutherford for reviewing and providing input on the article; and the KAIS Study Group for their contribution to the design of the survey and collection of the data set: Willis Akhwale, Sehin Birhanu, John Bore, Angela Broad, Robert Buluma, Thomas Gachuki, Jennifer Galbraith, Anthony Gichangi, Beth Gikonyo, Margaret Gitau, Joshua Gitonga, Mike Grasso, Malayah Harper, Andrew Imbwaga, Muthoni Junghae, Mutua Kakinyi, Samuel Mwangi Kamiru, Nicholas Owenje Kandege, Lucy Kanyara, Yasuyo Kawamura, Timothy Kellogg, George Kichamu, Andrea Kim, Lucy Kimondo, Davies Kimanga, Elija Kinyanjui, Stephen Kipkerich, Danson Kimutai Koske, Boniface O. K'Oyugi, Veronica Lee, Serenita Lewis, William Maina, Ernest Makokha, Agneta Mbithi, Joy Mirjahangir, Ibrahim Mohamed, Rex Mpazanje, Nicolas Muraguri, Patrick Murithi, Lilly Muthoni, James Muttunga, Jane Mwangi, M.M., Sophie Mwanyumba, Francis Ndichu, Anne Ng'ang'a, J.N., John Gitahi Ng'ang'a, Lucy Ng'ang'a, Carol Ngare, Bernadette Ng'eno, Inviolata Njeri, David Njogu, Bernard Obasi, Macdonald Obudho, Edwin Ochieng, Linus Odawo, Jacob Odhiambo, Caleb Ogada, Samuel Ogola, David Ojakaa, James Kwach Ojwang, George Okumu, Patricia Oluoch, T.O., Kenneth Ochieng Omondi, Osborn Otieno, Yakubu Owolabi, Bharat Parekh, George Rutherford, Sandra Schwarcz, Shanaaz Sharrif, Victor Ssempijja, Lydia Tabuke, Yuko Takanaka, Mamo Umuro, Brian Eugene Wakhutu, Celia Wandera, John Wanyungu, Wanjiru Waruiru, A.W., Paul Waweru, Larry Westerman, and Kelly Winter.

REFERENCES

1. Joint United Nations Programme on HIV/AIDS (UNAIDS). *The GAP Report 2015* UNAIDS, Geneva, Switzerland; 2015.
2. National AIDS and STI Control Program (NASCOP) Kenya Ministry of Health, National AIDS Control Council (NACC). *Kenya AIDS Response Progress Report 2016*. Nairobi, Kenya: National AIDS Control Council (NACC); 2016.
3. National AIDS and STI Control Programme Ministry of Health, National AIDS Control Council (NACC). *Kenya HIV Estimates 2015*. Nairobi, Kenya: National AIDS Control Council (NACC); 2015.
4. National AIDS and STI Control Programme (NASCOP). *Kenya AIDS Indicator Survey 2012: Final Report*; 2014. Available at: http://nacc.or.ke/wp-content/uploads/2015/10/KAIS-2012.pdf.
5. National AIDS/STI Control Program. *Kenya HIV Prevention Revolution Road Map*. Nairobi, Kenya: National AIDS and STI Control Program (NASCOP); 2014.
6. Ghys PD, Brown T, Grassly NC, et al. The UNAIDS Estimation and Projection Package: a software package to estimate and project national HIV epidemics. *Sex Transm Infect.* 2004;80(suppl 1):5–10.
7. Kerr CC, Stuart RM, Gray RT, et al. Optima: a model for HIV epidemic analysis, program prioritization, and resource optimization. *J Acquir Immune Defic Syndr.* 2015;69:365–376.
8. Anderson SJ, Cherutich P, Kilonzo N, et al. Maximising the effect of combination HIV prevention through prioritisation of the people and places in greatest need: a modelling study. *Lancet.* 2014;384:249–256.
9. Joint United Nations Programme on HIV/AIDS (UNAIDS). *Highlights from the High-level Side Event: Fast-track Ending the AIDS Epidemic by 2030*; 2014.
10. Joint United Nations Programme on HIV/AIDS (UNAIDS). UNAIDS 2016–2021 strategy: on the fast-track to end AIDS. 2015:1–120. Available at: http://www.unaids.org/sites/default/files/media_asset/20151027_UNAIDS_PCB37_15-18_EN_rev1.pdf.
11. Ransome Y, Dean LT, Crawford ND, et al. How do social capital and HIV/AIDS outcomes geographically cluster and which sociocontextual mechanisms predict differences across clusters? *J Acquir Immune Def Syndr.* 2017;76:13–22.
12. Goswami ND, Schmitz MM, Sanchez T, et al. Understanding local spatial variation along the care continuum: the potential impact of transportation vulnerability on HIV linkage to care and viral suppression in high-poverty areas, Atlanta, Georgia. 2016;77(7):616–623. doi: 10.1097/QAI.0000000000000914.
13. Kulldorff M, Nagarwalla N. Spatial disease clusters: detection and inference. *Stat Med.* 1995;14:799–810.
14. Wand H, Ramjee G. Targeting the hotspots: investigating spatial and demographic variations in HIV infection in small communities in South Africa. *J Int AIDS Soc.* 2010;13:41.
15. 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;12:28.
16. Cuadros DF, Abu-raddad LJ. Spatial variability in HIV prevalence declines in several countries in sub-Saharan Africa. *Health Place.* 2014;28:45–49.
17. UNAIDS Reference Group on Estimates Modelling and Projections. *Identifying Populations at Greatest Risk of Infection—Geographic Hotspots and Key Populations*; 2013:25–26.
18. Yao J, Agadjanian V, Murray AT. Spatial and social inequities in HIV testing utilization in the context of rapid scale-up of HIV/AIDS services in rural Mozambique. *Health Place.* 2014;28:133–141.
19. Hixson BA, Omer SB, Rio C, et al. Spatial clustering of HIV prevalence in Atlanta, Georgia and population characteristics associated with case concentrations. *J Urban Heal.* 2011;88:129–141.
20. Mee P, Collinson MA, Madhavan S, et al. Evidence for localised HIV related micro—epidemics associated with the decentralised provision of antiretroviral treatment in rural South Africa : a spatio–temporal analysis of changing mortality patterns (2007–2010). *J Glob Health.* 2014;4:010403. doi:10.7189/jogh.04.010403.
21. Kenya PEPFAR interagency team. Kenya country operational plan COP 2016: strategic direction summary. 2016. Available at: https://www.pepfar.gov/documents/organization/257644.pdf.
22. Waruiru W, Kim AA, Kimanga DO, et al. The Kenya AIDS indicator survey 2012: rationale, methods, description of participants, and response rates. *J Acquir Immune Defic Syndr.* 2014;66:3–12.
23. Ojwang' JK, Lee VC, Waruru A, et al. Using information and communications technology in a national population-based survey: the Kenya AIDS indicator survey 2012. *J Acquir Immune Defic Syndr.* 2014;66(suppl 1):S123–S129.
24. Temah CT. *What Drives HIV/AIDS Epidemic in Sub-Saharan Africa?* Vol. 17; 2009.
25. Rutstein SO, Johnson K. *The DHS Wealth Index. DHS Comp Reports No 6*. Calverton: ORC Macro; 2004:1–71.
26. Kulldorff M. SaTScan—software for the spatial, temporal, and space-time scan statistics. 2016. Available at: http://www.satscan.org/.

27. Kulldorff M. A spatial scan statistic. *Commun Stat Theor Meth.* 1997;26 (8):1481–1496. doi:10.1080/03610929708831995.
28. Messina JP, Emch M, Muwonga J, et al. Spatial and socio-behavioral patterns of HIV prevalence in the Democratic Republic of Congo. *Soc Sci Med.* 2010;71:1428–1435.
29. Han J, Zhu L, Kulldorff M, et al. Using Gini coefficient to determining optimal cluster reporting sizes for spatial scan statistics. *Int J Health Geogr.* 2016;15:1–11.
30. Kenya National Bureau of Statistics (KNBS). The 2009 Kenya population and housing census - population distribution by age, sex and administrative units. 2010;IC:546. Available at: http://www.knbs.or.ke/index.php?option=com_phocadownload&view=category&download=584:volume-1c-population-distribution-by-age-sex-and-administrative-units&id=109:population-and-housing-census-2009&Itemid=599.
31. SAS Institute Inc. *Statistical analysis system. SAS for windows*. Cary, NC.
32. Tanser F, Bärnighausen T, Cooke GS, et al. Localized spatial clustering of HIV infections in a widely disseminated rural South African epidemic. *Int J Epidemiol.* 2009;38:1008–1016.
33. 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;14:1–21.
34. Mayer JD. The geographical understanding of HIV/AIDS in Sub-Saharan Africa. *Nor J Geogr.* 2005;59:6–13.
35. Hosegood V. The demographic impact of HIV and AIDS across the family and household life-cycle: implications for efforts to strengthen families in sub-Saharan Africa. *AIDS Care.* 2009;21(suppl 1):13–21. doi:10.1080/09540120902923063.
36. Musheke M, Ntalasha H, Gari S, et al. A systematic review of qualitative findings on factors enabling and deterring uptake of HIV testing in Sub-Saharan Africa. *BMC Public Health.* 2013;13:220.
37. Sgaier SK, Reed JB, Thomas A, et al. Achieving the HIV prevention impact of voluntary medical male circumcision: lessons and challenges for managing programs. *PLoS Med.* 2014;11:e1001641.
38. Bailey RC, Plummer FA, Moses S. Reviews Male circumcision and HIV prevention: current knowledge and future research directions. *Lancet Infect Dis.* 2001;1:223–231.
39. Gray RH, Kigozi G, Serwadda D, et al. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. *Lancet* 2007; 369:657–666.
40. Puren A, Auvert B, Taljaard D, et al. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 trial. *PLoS Med.* 2005;2:e298. doi:10.1371/journal.pmed.0020298.
41. Bailey R, Gray R, Africa S. Male circumcision to cut HIV risk in the general population the printed journal includes an image merely for illustration. *Lancet.* 2007;369:617–619.
42. Wawer MJ, Makumbi F, Kigozi G, et al. Circumcision in HIV-infected men and its effect on HIV transmission to female partners in Rakai, Uganda: a randomised controlled trial. *Lancet.* 2009;374:229–237.
43. Bailey RC, Moses S, Parker CB, et al. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. *Lancet.* 2007;369:643–656.