METHOD ARTICLE

Identifying the spatial patterning characteristics of HIV positive clients not linked to care using a geographic information system [version 1; peer review: 1 approved, 1 not approved]

Kevin Rombosia1,2, Nelly Rangara2, Elizabeth Oele2, Nancy Kirui3, Douglas Olaka3, Evans Ondura4, Bernard Mitto4

1School of Public health, Makerere University, Kampala, Kampala, 40123, Uganda
2Department of Health, Kisumu County, Kisumu, Kisumu, 40100, Kenya
3MOH - PEPFAR Kisumu West HIV Program, Kisumu, Kisumu, 40100, Kenya
4Health Information, Governance and Data Analytics (HIGDA) Project, Nairobi, Nairobi, Kenya

First published: 04 Jun 2020, 9:515
Latest published: 04 Jun 2020, 9:515

https://doi.org/10.12688/f1000research.23967.1

v1

Abstract

Background: Linkage to care is a crucial early step in successful HIV treatment. This study sought to identify the spatial patterning characteristics of HIV positive clients that are not linked to care in the Kisumu West HIV program using a geographic information system.

Methods: The geocodes of HIV positive, non-linked clients’ residences were exported to ArcGIS software. The spatial patterning characteristics of HIV clients that are testing positive and not linked to care was described using Global Moran’s I statistic, which is a measure of spatial autocorrelation.

Results: A total of 14,077 clients were tested for HIV. Of clients testing positive for HIV, 10% (n=34) were not yet linked to care two weeks after the diagnosis of HIV. Of the HIV positive non-linked clients, most (65%; n= 32) had spatially identifiable data about where they resided. Regarding the spatial patterning characteristics of the clients who tested HIV positive but were not linked to care and with spatially identifiable residence information, the Global Moran I statistic for autocorrelation was 0.435 (z score 1.383, p-value 0.167).

Conclusion: By using age as an attribute value, the spatial distribution of clients testing HIV positive and not being linked to care is random. Geographical information systems can be used to identify the spatial patterning characteristics of HIV positive clients that are not linked to care. A key requirement to achieving this would require the collection of precise and accurate spatially identifiable locator information but without compromising patient confidentiality.
**Keywords**
Geographic Information Systems, HIV, Linkage to care, Spatial patterning

<table>
<thead>
<tr>
<th>Corresponding author:</th>
<th>Kevin Rombosia (<a href="mailto:kevinrombosia@gmail.com">kevinrombosia@gmail.com</a>)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author roles:</strong></td>
<td>Rombosia K: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review &amp; Editing; Rangara N: Conceptualization, Project Administration; Oele E: Conceptualization; Kirui N: Data Curation, Formal Analysis, Investigation, Validation; Olaka D: Data Curation, Formal Analysis, Investigation, Validation; Ondura E: Visualization; Mitto B: Conceptualization, Data Curation, Methodology, Project Administration, Resources, Software, Supervision, Visualization</td>
</tr>
<tr>
<td><strong>Competing interests:</strong></td>
<td>No competing interests were disclosed.</td>
</tr>
<tr>
<td><strong>Grant information:</strong></td>
<td>This work was supported through a USAID grant to the Health Informatics Governance and Data Analytics Project (HIGDA) which was a five-year project mandated to support the health sector in Kenya to strengthen national and county organizational and management capacity in governance, health informatics, data analytics, monitoring, evaluation, learning and accountability. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.</td>
</tr>
<tr>
<td><strong>Copyright:</strong></td>
<td>© 2020 Rombosia K et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.</td>
</tr>
</tbody>
</table>

**How to cite this article:** Rombosia K, Rangara N, Oele E et al. Identifying the spatial patterning characteristics of HIV positive clients not linked to care using a geographic information system [version 1; peer review: 1 approved, 1 not approved] F1000Research 2020, 9:515 https://doi.org/10.12688/f1000research.23967.1

**First published:** 04 Jun 2020, 9:515 https://doi.org/10.12688/f1000research.23967.1
Introduction

For many years, the analysis and use of spatial data has guided public health efforts, and Dr. John Snow, who is often credited as the founder of epidemiology, is a classic example of using this type of data. Geographic Information Systems (GIS) are information management systems for spatial data. Spatial data is data that is associated with specific location features. GIS enables the user to process, analyse and present geographically referenced data. The use of GIS is increasingly being appreciated in the health sciences. For example, it can be used to map the distribution and accessibility of health facilities and assess the distribution of diseases, hence allowing for planning of their control.

UNAIDS has an ambitious 90-90-90 strategy. This strategy stipulates that by the year 2020, 90% of all people with HIV infection will know their HIV status, 90% of HIV positive people will be on antiretroviral therapy (ART) and 90% of all people on treatment will have achieved viral suppression. However, Kenya still has low rates of linking HIV positive persons to HIV care.

Investigating the spatial patterning of diseases is an epidemiology tool that provides greater understanding of their possible causes and the prevention steps that can be undertaken. According to Ward et al, using spatial statistics and maps to study diseases can provide better intuition in the study of diseases.

There are many barriers to linkage to care following HIV positive results. Previous studies have reported factors such as being male, younger in age, having fears of drug side effects, busy schedules, transport costs, distance to clinics, stigma and fear of disclosure of HIV status, staff shortages at health facilities, and delays in getting services at health facilities as contributing to poor linkage and retention in care. For these barriers to be effectively addressed and a move made towards the UNAIDS 90-90-90 goal, it requires, firstly, that any non-linked clients are identified and enrolled into care.

This study aimed to identify the spatial patterning characteristics of clients testing HIV positive who are not linked to care in the Kisumu West HIV program. If spatial clustering of clients is demonstrated for those who test HIV positive but are not linked to the Kisumu West HIV program, it would be an entry point for initiating targeted approaches for enrolling and retaining these clients in care. Such approaches include targeted community sensitization sessions on the need for linking into care following HIV positive results through targeted community health education in specific geographic areas that have a high prevalence of non-linkage. Also, it would help in obtaining linkage officers and case managers to target areas that have high densities of non-linked clients that are HIV positive. This targeted approach would be more cost effective, and should be combined with continuous HIV campaigns to address community stigma towards HIV/AIDS.

Similarly, if it is demonstrated that the spatial patterning characteristics of clients testing HIV positive and not linked care is non-clustered but geographically randomly distributed, it would reinforce the need to establish rigid client tracking mechanisms post-HIV testing at health facilities.

Methods

Study area

Kisumu County (Figure 1) is one of 47 counties in Kenya. It lies within longitudes 33° 20'E and 35° 20'E and latitudes 0° 20'South and 0° 50'South. The total land area is approximately 2,086 Km² and around 567 Km² is covered by water. The County borders the expansive fresh water lake (Lake Victoria) and has 7 sub-counties namely; Kisumu Central, Kisumu East, Kisumu West, Seme, Nyando, Nyakach and Muhoroni. The sub-counties are further divided into administrative wards. The Kisumu West HIV program covers 6 wards namely North-west Kisumu, West Kisumu, Central Seme, East Seme, North Seme and West Seme.

This HIV care, treatment and prevention program in Seme Sub County and part of Kisumu West Sub County utilizes a “hub and spoke” model. Kombewa County Hospital, being the largest facility, is the “hub” wherein 24 peripheral health facilities (22 facilities in Seme Sub County and 2 facilities in Kisumu West Sub County) pivot for supply of commodities and drugs, training, viral load and HIV PCR sample processing.

Data collection

A rapid results initiative (RRI) is a management tool that is used to mobilize teams to rapidly achieve tangible results. During the Kisumu West HIV Program RRI of 1st June 2017 to 31st August 2017, HIV counselling and testing was conducted in the 24 health facilities of the Kisumu West HIV Program. The eligibility criteria of the present study included clients who had never tested for HIV, had tested negative in the past 3 to 12 months or had tested negative in the past 3 months or less but the negative result could not be verified by any available health record. Clients who declined consent to participate or clients whose guardians declined consent were excluded from the study.

Primary data on clients’ name, age, gender, residence and HIV status was entered into HIV Testing Services (HTS) registers and fed into an information management system, DATIM (Data for Accountability, Transparency and Impact). Data for this study was extracted from DATIM. HIV positivity was calculated as:

\[
HIV \text{ positivity} = \left(\frac{\text{Number of clients testing } HIV \text{ positive}}{\text{Total Number of clients tested}}\right) \times 100\%
\]

This HIV positivity was calculated for each testing health facility and aggregated at ward level wherein the health facility is located.
Data management and analysis
The data mined was entered into Microsoft Excel version 16.35 (RRID:SCR_016137) and then exported to STATA statistical software version 15 (RRID:SCR_012763) for processing and analysis.

The Kisumu County shapefile and geocodes of HIV positive, non-linked clients’ residences were exported to ArcGIS (ArcGIS for Desktop Basic, RRID:SCR_011081). An alternative open source GIS software for this spatial analysis is QGIS (RRID:SCR_018507). The locations of the clients’ residences, represented by the nearest known landmarks to their homes, were mapped on ArcGIS as points. These landmarks included features such as schools, churches, market centres etc. The geocodes of these landmarks were extracted from Google Maps (Version 2019 by Google Incorporated) by right clicking the “What’s here” Scalable Vector Graphic (SVG) marker and copy – pasting to Microsoft Excel.

The spatial patterning characteristics of HIV+ clients not linked to care was described using the Global Moran’s I statistic. Global Moran’s I statistic is a measure of spatial autocorrelation developed by Patrick Alfred Pierce Moran.

In the Global Moran’s I statistic, spatial autocorrelation is characterized by a correlation of a point among nearby points in a given area and is given by:

\[
I = \frac{n}{\sum_{i<j} W_{ij}} \sum_{i} \sum_{j} W_{ij} (X_i - \bar{X})(X_j - \bar{X}) \frac{1}{\sum_{j} (X_j - \bar{X})^2}
\]

where \( n \) is the total number of observations (points or polygons), \( i \) and \( j \) represent different locations, \( X_i \) and \( X_j \) are values of the variable in the \( i \)th and \( j \)th locations, \( \bar{X} \) is the mean of the variable and \( W_{ij} \) is the weight in the spatial weight matrix. The clients’ age was used as the variable of interest because previous studies have demonstrated that younger age is a significant causative factor for non – linkage to care. The spatial weighted matrix that was used was the inverse distance weighting function in ArcGIS. This is also available in QGIS.

Ethical considerations
Written informed consent for publication of clients’ details was obtained from the clients/parents/guardians of the clients as appropriate. The data was de-identified by the use of client numbers (not names). The geocodes of the clients’ residences were represented using their nearest landmarks such as nearby landmarks.
schools, churches and market centres. The exact geocodes of the clients’ residences were not used in this study in order to protect confidentiality.

Ethical and institutional approvals for this study were obtained from the Makerere University Research and Ethics Committee and the Kisumu County Department of Health.

Results
A total of 14,077 clients were tested for HIV during the RRI period (men 38%, n=5,304; women 62%, n=8,773). Of the total clients tested for HIV, most (32%, n=4,552) belonged to the 25–49 year age group, within which the majority were women (62%, n=2,811) (Figure 2).

A total of 345 (2%) clients tested positive for HIV (range = 0–5%; mean = 1.5±1.6). Most clients that were HIV+ were in the 25–49 years age group (5%, n=226), of which more clients were men (52%, n=118; women, 48%, n=108). The <1 year old age group had the lowest number of HIV+ clients (0%, n=0). In the 1 – 4-year age band, West Seme Ward had the HIV highest positivity (89%, n=8) (Figure 3). The health facilities with the highest number of HIV+ clients were Kombewa (4.5%, tested n=2,505, HIV+ n=112), Manyuanda (4.2%, tested n=970, HIV+ n=41), Bodi (3.9%, tested n=431, HIV+ n=17) and Oriang Alwala (3.6%, tested n=248, HIV+ n=9).

Of all clients testing positive for HIV, 10% (n=34) were not yet linked to care two weeks after the diagnosis of HIV. Of these HIV+ non-linked clients, 50% were women (n=17), and most (65%, n=32) had spatially identifiable data about their residences’ nearest landmarks.

The geocodes of the clients’ residences’ nearest landmarks were obtained from Google Maps. The geocodes are as shown in Table 1.

The geocodes were used to map the distribution of HIV positive non-linked clients as shown in Figure 4.

Regarding the spatial patterning characteristics of the clients who tested HIV positive but were not linked to care and with spatially identifiable residence information, the Global Moran I statistic for autocorrelation was 0.435 (z score 1.383, p-value 0.167) (Figure 5).

![Full Cascade RS](image)

*Figure 2. Graph showing number of HIV positive clients, number enrolled into care and number started on ART.*
### Table 1. HIV positive, non-linked clients’ locator details.

<table>
<thead>
<tr>
<th>CLIENT ID</th>
<th>AGE (YEARS)</th>
<th>CLIENT’S RESIDENCE LATITUDE</th>
<th>CLIENT’S RESIDENCE LONGITUDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35</td>
<td>0.039082</td>
<td>34.217449</td>
</tr>
<tr>
<td>2</td>
<td>46</td>
<td>-1.290486</td>
<td>36.822754</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>-0.061824</td>
<td>34.540436</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>-0.045282</td>
<td>34.537509</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>-0.080819</td>
<td>34.41359</td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>-0.073843</td>
<td>34.417863</td>
</tr>
<tr>
<td>7</td>
<td>26</td>
<td>-0.139538</td>
<td>34.437759</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>-0.139538</td>
<td>34.437759</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>-0.139538</td>
<td>34.437759</td>
</tr>
<tr>
<td>10</td>
<td>47</td>
<td>-0.045282</td>
<td>34.537509</td>
</tr>
<tr>
<td>11</td>
<td>40</td>
<td>-0.109602</td>
<td>34.482296</td>
</tr>
<tr>
<td>12</td>
<td>23</td>
<td>-0.102394</td>
<td>34.49898</td>
</tr>
<tr>
<td>13</td>
<td>23</td>
<td>-0.102394</td>
<td>34.49898</td>
</tr>
<tr>
<td>14</td>
<td>36</td>
<td>-0.061824</td>
<td>34.540436</td>
</tr>
<tr>
<td>15</td>
<td>40</td>
<td>-0.108606</td>
<td>34.550417</td>
</tr>
</tbody>
</table>

### Figure 3. HIV positivity by age bands.
Figure 4. Distribution of non-linked clients.

Figure 5. Spatial autocorrelation test.
Discussion
In this study, ArcGIS, a GIS software, was used to incorporate spatial attributes of clients’ residence to the HIV RRI data. This use of maps helped to better visualize the meaning of the numbers to provide more intuition about the distribution of HIV positivity and the spatial patterning of clients who tested HIV positive but were not linked to care two weeks after the diagnosis in the Kisumu West HIV program.

Having mapped the HIV positivity by different age groups and by different wards, it was established that most clients who were HIV+ were in the 25-49 year age groups (5%, n=226), which included more men (52%, n=118; women 48%, n=108). This shows that greater preventive efforts are required, and the establishment of targeted interventions for this cohort, such as viremia clinics, that seek to increase retention to care and viral suppression. Also, it was established that there was high positivity in children below 4 years in the West Seme Ward, which necessitates a need to audit the Prevention of Mother To Child HIV Transmission (PMTCT) programs in the health facilities in West Seme Ward in order to establish why there is a high number of HIV exposed infants seroconverting to active HIV infection. It was noted that while Oriang Alwala is a low volume health facility, it had the fourth highest HIV positivity rate (3.6%, tested=248, HIV positive n= 9); hence a need to redeploy more testers and adherence counsellors to this facility.

Regarding linkage, it was established that 10% of the clients who tested positive during the RRI period for HIV in the Kisumu West HIV program had not been enrolled into care. Through the use of GIS, it was possible to describe the spatial patterning characteristics of these non-linked clients via their residences’ nearest landmarks. The Global Moran’s I statistic for autocorrelation was 0.435 (z score 1.383, p-value 0.167) indicating that the pattern does not appear to be significantly different than random. This means that the distribution of the clients testing HIV positive and not getting linked to care, using age as an attribute value, does not follow the path of spatial clustering. Therefore, the most effective strategy for getting such clients linked to care is not to administer blanket community sensitization sessions on the need for linking into care following HIV positive results through community health education as is routinely done in the Kisumu West HIV program. Instead, through the use of distribution maps, linkage officers can generate a better work plan of navigating the program area as they conduct home visits to these clients.

In calculating the Global Moran’s Autocorrelation statistic, the clients’ age was used as the y attribute value. It would have been more accurate to use a composite score that quantifies the ‘level of severity of a client’s non linkage’. However, even after an extensive literature review, we could not find such a scale.

There are some concerns or issues with use of GIS tools for public health efforts. Chief among those is a concern for privacy and confidentiality of individuals. It would be useful to collect more precise and accurate locator details, e.g. exact residence details, to allow for a more targeted approach towards getting these clients enrolled into care. However, this needs to be balanced with the need to protect clients’ confidentiality. There are certain strategies that can be employed to protect patient confidentiality. For example, data may need to be aggregated to cover larger areas such as a zip code or county, helping to mask individual identities during dissemination of results. Residential geocodes can as well be displaced. Also, maps can be constructed at smaller scales so that less detail is revealed. In this study, we represented clients’ residences using the landmarks nearest to the residences in order to conceal the pinpoint location of the clients’ residence’s information.

Conclusion
Geographical information systems can be used identify spatial patterning characteristics of HIV positive clients that are not linked to care. A key requirement to achieving this would require the collection of precise and accurate spatially identifiable locators information. More research needs to be done on how geographical information systems can be used to describe the spatial patterning characteristics of HIV positive clients that are not linked to care without compromising patient confidentiality.

Data availability
Underlying data

This project contains the following underlying data:
- 2017_RRI_ALL_Facilities.csv; used to calculate HIV positivity per facility
- HIV_Positivity.csv; used to calculate HIV positivity for different age bands
- Non_Linkage_To_HIV_Care.csv; used to calculate non-linkage to HIV care for different age bands
- Number_Newly_Enrolled_To_HIV_Care.csv; used to calculate non-linkage to HIV care for different age bands
- Number_Newly_Started_On_ART.csv; used to display number started on ART
- Number_Positive.csv; used to calculate HIV positivity for different age bands
- Number_Tested.csv; used to calculate HIV positivity for different age bands
- Residence_nearest_landmarks.csv; used for spatial analysis

Data are available under the terms of the Creative Commons Zero “No rights reserved” data waiver (CC0 1.0 Public domain dedication).
References


Open Peer Review

Current Peer Review Status: ✔️ ✗

Review Report 07 August 2020

https://doi.org/10.5256/f1000research.26441.r66050

© 2020 Montague B. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Brian T. Montague

Department of Internal Medicine, Division of Infectious Diseases, University of Colorado School of Medicine, Aurora, CO, USA

Thank you for the opportunity to review this interesting work. The authors describe use of GIS testing to looking at patterns of linkage to care in Western Kenya with the goal of identifying notable areas of less successful linkage that may suggest potential opportunities for program innovation.

The work is framed as a demonstration of the potential benefits of GIS mapping tools to support linkage to care. There is significant literature on this topic already, some of which I have added in citations here. If the authors wish to focus on this as the rationale for the work, more clarity would be needed regarding what their methods add to the existing literature regarding the use of GIS in evaluation of HIV linkage and care in similar settings.

For their particular results here, the distribution of those not linked was noted to be random which would suggest that the GIS provided relatively less information in this setting. The conclusion drawn from this was a little bit unclear, namely that given the distribution was random blanked distributions of messages was not helpful. One could interpret this in the opposite manner, namely that since there is no clustering, widespread distribution of messages is needed. Targeting interventions to clients identified as out of care wherever they are does not leverage the information from the GIS analysis.

While a potentially helpful tool, in this case it appears that the tool did not yield clearly actionable information and some clarification is needed with regard to the focus of the manuscript and what it adds to the existing literature.

References

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** HIV, infectious disease, epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

**Author Response 18 Aug 2020**

**Kevin WW Rombosia**, Makerere University, Kampala, Uganda

Dear Sir,

Thank you for taking your time to review our work.
Regarding the conclusions drawn using the GIS analysis, this work influenced a programmatic shift in the Kisumu West HIV program from blanket strategies such as outreaches in areas deemed to have low linkage to care to a more focused approach of messaging (on the need for linking to care) through health education at the health facility level.

Also, this may be viewed alongside the CDC revised approach of HIV testing at facility level as opposed to outreach testing (as previously widely practiced) because the former has been shown to have higher yield.

This is an area we are continuing to look into and perhaps our next work should evaluate the improvement on linkage to care after adopting this approach.

Kind regards,

Kevin Rombosia

**Competing Interests:** No competing interests.

---

**Reviewer Report 13 July 2020**

https://doi.org/10.5256/f1000research.26441.r65484

© 2020 Kourkouta L. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The author(s) is/are employees of the US Government and therefore domestic copyright protection in USA does not apply to this work. The work may be protected under the copyright laws of other jurisdictions when used in those jurisdictions.

Lambrini Kourkouta
Nursing Department, International University of Greece, Thessaloniki, Greece

It is a very interesting and novel manuscript.

HIV positive patients are often marginalised and a geographic information system would enable them to reach such a system. However, the references are not sufficient in number when such a subject is being examined.

To my way of thinking, more sources have to be found out in order to support the novelty of a geographic information system.

**Is the rationale for developing the new method (or application) clearly explained?**
Yes
Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Nursing, Hiv, Public Health. History

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

• Your article is published within days, with no editorial bias
• You can publish traditional articles, null/negative results, case reports, data notes and more
• The peer review process is transparent and collaborative
• Your article is indexed in PubMed after passing peer review
• Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com