

## Spatial clustering of endemic Burkitt's lymphoma in high-risk regions of Kenya

Jeanette J. Rainey<sup>1\*</sup>, Dorine Omenah<sup>2</sup>, Peter O. Sumba<sup>2</sup>, Ann M. Moormann<sup>3</sup>, Rosemary Rochford<sup>4</sup> and Mark L. Wilson<sup>5</sup>

<sup>1</sup>Department of Epidemiology, University of Michigan, Ann Arbor, MI

<sup>2</sup>Centre for Vector Biology and Control Research, Kenya Medical Research Institute

<sup>3</sup>Centre for Global Health and Diseases, Case Western Reserve University

<sup>4</sup>Department of Microbiology and Immunology, SUNY Upstate Medical University, Syracuse, NY

<sup>5</sup>Department of Epidemiology, University of Michigan, Ann Arbor, MI

Endemic Burkitt's lymphoma (eBL), the most common childhood cancer in sub-Saharan Africa, occurs at a high incidence in western Kenya, a region that also experiences holoendemic malaria. Holoendemic malaria has been identified as a co-factor in the etiology of this cancer. We hypothesized that eBL may cluster spatially within this region. Medical records for all eBL cases diagnosed from 1999 through 2004 at Nyanza Provincial General Hospital were reviewed for case residential information to examine this hypothesis. Two cluster detection methods, Anselin's Local Moran test for spatial autocorrelation and a spatial scan test statistic, were applied to this residential data to determine whether statistically significant high- and low-risk areas were present in the Province. During the 6-year study period, 272 children were diagnosed with eBL, with an average annual incidence of 2.15 cases per 100,000 children. Using Empirical Bayes smoothed rates, the Local Moran test identified 1 large multi-centered area of low eBL risk ( $p$ -values < 0.01) and 2 significant multi-centered clusters of high eBL risk ( $p$ -values < 0.001). The spatial scan detected 3 small independent low-risk areas ( $p$ -values < 0.02) and 2 high-risk clusters ( $p$ -values = 0.001), both similar in location to those identified from the Local Moran analysis. Significant spatial clustering of elevated eBL risk in high-malaria transmission regions and of reduced incidence where malaria is infrequent suggests that malaria plays a role in the complex eBL etiology, but that additional factors are also likely involved.

© 2006 Wiley-Liss, Inc.

**Key words:** endemic Burkitt's lymphoma; Kenya; Nyanza Province; GIS; spatial clustering

Endemic Burkitt's lymphoma (eBL) is the most common childhood cancer in sub-Saharan Africa. The distribution of this cancer has geographic and climate-associated features. Cases are most frequently observed in regions roughly 10° north and south of the equator with minimum annual rainfall and mean temperatures of 50 cm and 15.6°C, respectively.<sup>1,2</sup> These specific features have been linked to chronic and intense malaria transmission (e.g. holoendemic malaria).<sup>3–5</sup> Within these holoendemic malaria transmission regions, however, eBL occurrence is not uniform,<sup>5</sup> suggesting that other environmental factors may contribute to the pattern of eBL.<sup>4,6,7</sup>

Many studies have been conducted in an effort to explain this apparent nonrandom eBL distribution. Pike *et al.* first detected spatial-temporal eBL clustering in Uganda,<sup>8</sup> where cases within the West Nile District during 1961–1965 occurred more closely in space and time than could be expected by chance alone. Other investigations identified spatial-temporal clustering in Bwamba County of Uganda during 1966–1968<sup>9</sup> and more recently, among older children living in communities bordering Lake Malawi.<sup>10</sup> Studies in Ghana, Northern Tanzania and other districts in Uganda, however, failed to detect eBL case clustering.<sup>11–13</sup> It is unclear whether these conflicting findings result from eBL risk-factors that are present in certain places only at certain times, bias in eBL case detection or the impact of demographic changes on the distribution of at-risk populations.<sup>13</sup> Despite its potential link to underlying causal mechanisms, the question of whether eBL clusters within high-risk regions remains unanswered.

Although Kenya lies within the well-defined "lymphoma belt", the country was not included in these early statistical investiga-

tions of the spatial-temporal distribution of eBL. More recently, however, information from 2 studies suggests that eBL hot-spots do exist in Nyanza Province, a high-risk region in western Kenya. Makata *et al.*<sup>14</sup> noted elevated rates of pediatric cancers, including eBL, in communities directly bordering Lake Victoria, and data published by Mwanda *et al.*<sup>15</sup> described greater eBL incidence among children living in Siaya District, in the north of the Province. These findings, if confirmed, could provide important leads in identifying potential eBL co-factors. However, spatial statistics were not used to determine whether the distribution of eBL observed in these studies reflected significant clustering or chance variability in disease occurrence alone.

Our investigation was conducted to test whether these previous findings (*i.e.*, spatial clustering in Lakeside communities and/or Siaya District) are statistical and epidemiologically valid and determine whether other areas in the region also experience elevated eBL risk. Accordingly, we analyzed medical records of eBL cases diagnosed at Nyanza Provincial General Hospital from 1999 through 2004 and used 2 cluster detection methods to statistically evaluate the spatial distribution of eBL in Nyanza Province. A new look at the spatial distribution of this old disease in a high-risk Province in Kenya, especially given significant demographic changes in the region, may not only confirm recognized high-risk areas but also elucidate previously unidentified spatial patterns. A better understanding of eBL disease patterns will ultimately generate new information regarding the multi-factorial etiology of this pediatric cancer.

### Material and methods

#### Study population

Nyanza Province is located in western Kenya, bordered to the west by Lake Victoria, east by the Rift Valley highlands, north by Western Province and south by Tanzania (Fig. 1). Most of the Province lies just south of the equator and has a semi-tropical to tropical climate. Malaria transmission varies from holoendemic in lowland areas around Lake Victoria to seasonal (April to July) in the highlands (elevation ~1,600–2,000 m). Although rainfall occurs throughout the year, precipitation is typically elevated during the "long" (March to June) and "short" (October to December) rainy seasons.

The population of Nyanza Province in 1999 was ~4.4 million, of which ~50% was <15 years old.<sup>16</sup> In 1999, the Province was divided into 12 Districts, 65 Divisions, 324 Locations and 933

Grant sponsor: The Comprehensive Cancer Center (Munn Idea award), the Center for the Education of Women and the Global Health Program, The University of Michigan; Grant sponsor: Case Western Reserve University; Grant number: NIH K08 AI 51565.

\*Correspondence to: Centers for Disease Control and Prevention, National Immunization Program MS E-05, Global Immunization Division, 1600 Clifton Road N.E., Atlanta, GA, 30333, USA. Fax: +404-639-8474. E-mail: jkr7@cdc.gov

Received 11 October 2005; Revised 5 May 2006; Accepted 19 May 2006

DOI 10.1002/ijc.22179

Published online 3 October 2006 in Wiley InterScience (www.interscience.wiley.com).



FIGURE 1 – Map of Kenya and District-level map of Nyanza Province.

Sublocations. This analysis focuses on clustering using 4th administrative level data, that is Location-specific eBL incidence rates (IR) (1999 median childhood population estimate = 5,569, range 421–31,776; median surface area = 34.4 km<sup>2</sup>, range 1.4–154.9 km<sup>2</sup>). Under-five childhood mortality was 116 per 100,000, with malaria and respiratory infections contributing to the majority of these deaths.<sup>17</sup> In 2000, more than a quarter of the urban adult population in Nyanza was seropositive for HIV, and up to 40% in certain high-risk groups.<sup>18</sup> Most people in the Province are of the Luo tribe, with the Kisii, Suba and Kuria tribal groups representing smaller proportions of residents.

Nyanza Provincial General Hospital is located in Kisumu city. As the largest hospital in the Province, it provides out-patient and in-patient services, and maintains its own pathology and radiology laboratories. The hospital is the referral center for childhood

cancer cases and is the only medical facility that provides chemotherapeutic treatment for eBL in the region.

#### *eBL case data and human subjects approval*

We reviewed pediatric medical records of all patients  $\leq 15$  years old who had been admitted to Nyanza Provincial General Hospital from 1999 through 2004 with a physician diagnosis of Burkitt's lymphoma. A standardized abstraction form was used to collect demographic, clinical and place of residence information. All eBL cases hospitalized at Nyanza Provincial Hospital during 2003 and 2004 were included in an immunologic study on eBL and were histologically confirmed. Although all cases diagnosed and treated for eBL between 1999 and 2002 were also histologically confirmed, documentation was not always available for patients hos-

pitalized during this time period. When documentation was not available, only cases treated for eBL and responding to treatment with no potential competing diagnosis (*e.g.* non-eBL cancer) were included in this study.

Our study protocol was approved by Internal Review Boards at the University of Michigan, Case Western Reserve University and University Hospitals of Cleveland, as well as the Ethical Review Committee at the Kenya Medical Research Institute. The Chief Medical Director at Nyanza Provincial General Hospital granted access to eBL medical records.

#### Population and geographic data

We calculated childhood (0–15 years of age) population estimates for each Location in Nyanza Province from 1999 through 2004. These estimates were based on 1999 census data obtained from the Kenya Medical Research Institute/Wellcome Trust Collaborative Programme (KEMRI/Wellcome Trust) in Nairobi and 2000–2004 age-specific population projections generated by the Kenyan Central Bureau of Statistics.<sup>16</sup> Population projections were based on algorithms incorporating the impact of HIV/AIDS, other causes of mortality, births and migration. Digitized maps of 1999 Location administrative boundaries in Nyanza Province were also obtained from the KEMRI/Wellcome Trust Collaborative Programme.

#### Data management

Data were entered into an excel database and converted to SAS (version 8.2, SAS Institute, Cary, NC) for preliminary analysis. We excluded cases residing outside Nyanza Province as well as duplicate records. If we detected an exact match on all demographic variables, with age differences reflecting only changes in admission dates, then a case was considered a duplicate. The entry with the later hospital admission date was excluded. To minimize error in case mapping, we cross-referenced place of residence with the administrative place names from the 1999 census. If inconsistencies were noted between 2 administrative levels (*e.g.*, Location and Sublocation), the smaller unit (Sublocation) was assumed to be correct, as parents and guardians were most likely to be familiar with more local governmental administrative location names. ArcGIS (version 9.1, ESRI, Redlands, CA) was used for mapping and visualization.

#### Statistical analysis

We used 2 cluster detection programs, GeoDa (version 0.9.5-i, Urbana-Champaign, IL) and SaTScan (version 6, Boston, MA), to fully examine our hypothesis regarding the presence of eBL spatial clustering in the Province. Both these programs test for spatial clustering using area (case and population at-risk) data, and output the location, approximate size and significance level of identified clusters. The programs are publicly available from the internet. Different spatial analytical methods may identify different underlying spatial patterns.<sup>19</sup> Consistent findings using more than 1 method, therefore, would suggest robust results.

#### GeoDa–Anselin's Local Moran test

We used the GeoDa cluster detection software program for generating Anselin's Local Moran (LISA) test statistics. The Local Moran statistic assesses spatial autocorrelation and identifies administrative units (*i.e.*, Locations) with disease rates statistically similar to and dissimilar from their neighbors. The null hypothesis is that there is no association in eBL IRs among neighboring administrative units. The alternative hypothesis is that spatial clustering exists (neighboring Locations have similar eBL rates).

IRs based on relatively uncommon health events such as eBL can be unstable (large variances). Therefore, we applied Empirical Bayes (EB) smoothing methods to our eBL disease data. This smoothing method adjusted disease rates, especially in administrative units with small population estimates, toward the overall mean of the study area. We defined neighbors using queen conti-

guity relationships (single point) and obtained Monte Carlo *p*-values based on 9,999 conditional randomizations. Because of multiple comparisons and chances for overlapping clusters, we used an alpha level of 0.01 to assess statistical significance.

#### SaTScan–Spatial scan

We repeated our cluster detection analysis using the spatial scan test in SaTScan. This test uses a circular window in which the center of the circle moves across the study area. At each center position, the radius of the window varies between 0 and a user defined upper limit, thereby encompassing different sets of neighboring administrative units (along with disease and population data for that unit). An administrative unit is included in the circle if the centroid for that unit lies in the circle. An infinite number of distinct circles varying in size and location are created. For each circular window, the spatial scan calculates the likelihood of observing the reported and expected number of eBL cases inside the circle given the number of reported and expected cases outside the circle. The circular window with the maximum likelihood is defined as the most likely cluster, indicating that it is least likely to have occurred by chance alone. The scan statistic evaluates the null hypothesis that the eBL disease rate is the same throughout the study area. The alternative hypothesis is that there is at least 1 circular window in which the disease rate inside the circle varies significantly from remainder of the Province.

We used raw eBL data in SaTScan to avoid potential bias in estimating the likelihood ratio test statistic. A data file containing raw eBL case and population data for the centroid (longitude and latitude coordinate) of each 4th level administrative unit in the Province was generated in GeoDa. This data file was imported into SaTScan assuming a Poisson probability disease model (case and at-risk population data). Although we set the maximum cluster size at 50% of the study population for purposes of statistical inference (avoiding preselection bias), we limited the size of reported clusters to those representing 20% or less of the population to focus on localized clustering. Presence of high-risk clusters and low-risk areas was assessed simultaneously. *p*-Values for maximum likelihood ratios were based on 9,999 Monte Carlo randomizations. An alpha level of 0.05 was used to assess statistical significance. Likelihood-ratio based test statistics and reported *p*-values account for multiple testing. Results from the spatial scan and tests for spatial autocorrelation were compared with each other as well as with the smoothed eBL rate map.

## Results

#### Study population

From 1999 through 2004, 372 eBL patients were diagnosed with eBL at Nyanza Provincial General Hospital. Cases residing outside of Nyanza Province ( $n = 76$ ), duplicated case reports ( $n = 12$ ) and cases lacking a laboratory or clinical diagnosis ( $n = 29$ ) were excluded (categories not mutually exclusive), resulting in a study population of 272 cases. The 6-year average annual IR was 2.15 cases per 100,000 children (Table I). Place of residence information to the 4th administrative level (*i.e.*, Location) was available for 235 (86.0%) of these cases. The median onset age of these cases was 7.0 years (interquartile range = 5.0–9.0), and 58.7% were male (95% CI = 52.3–65.0). Cases were reported from 5 dif-

TABLE I – ANNUAL AND 6-YEAR AVERAGE ANNUAL eBL INCIDENCE RATES PER 100,000 CHILDREN, NYANZA PROVINCE 1999–2004 ( $n = 272$ )

Year	No.	At-risk population	IR
1999	33	2,132,186	1.54
2000	47	2,098,970	2.23
2001	52	2,103,266	2.47
2002	17	2,102,056	0.80
2003	60	2,107,032	2.84
2004	63	2,105,241	2.99
6-year average	45.3	2,108,125	2.15



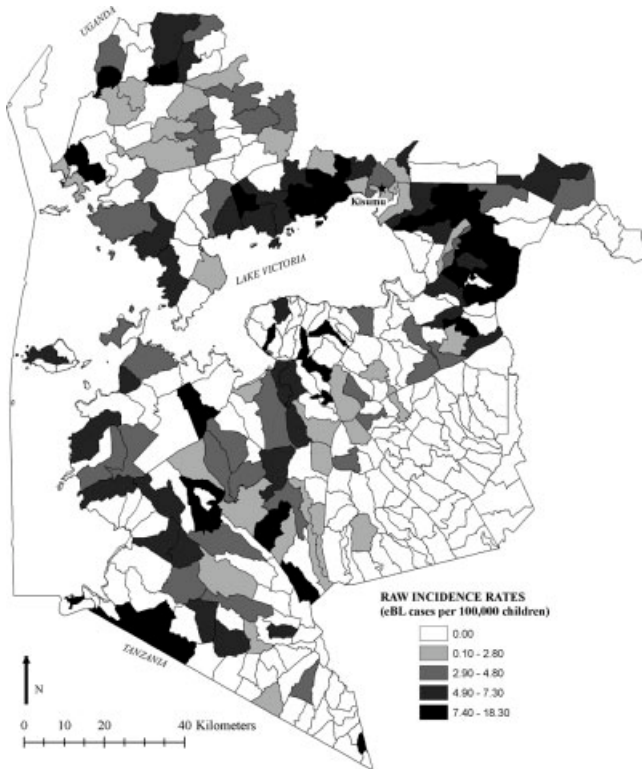


FIGURE 2 – Raw IRs (eBL cases per 100,000 children) by Location ( $n = 324$ ), Nyanza Province, 1999–2004. Breakpoints generated using a quintile classification scheme.

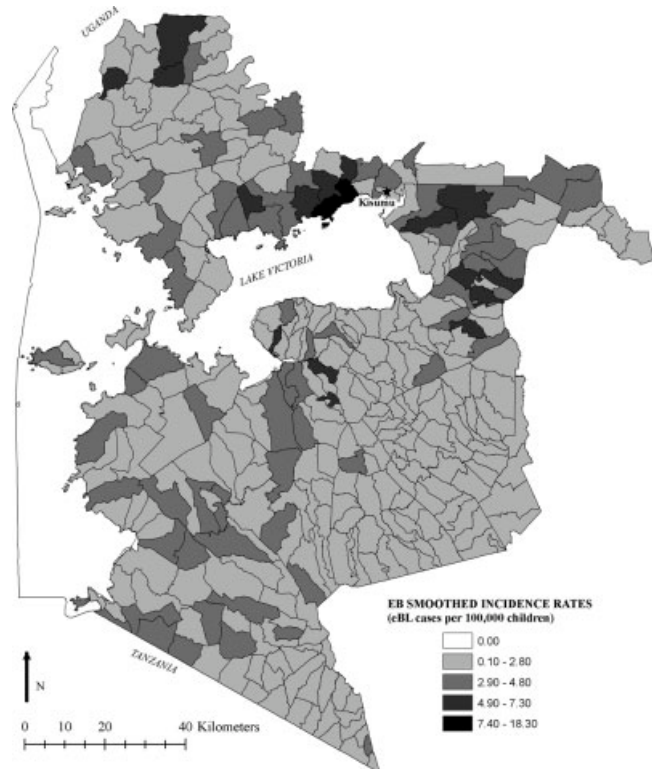


FIGURE 3 – EB smoothed rates (eBL cases per 100,000 children) by Location ( $n = 324$ ), Nyanza Province, 1999–2004. Breakpoints based on quintile classification scheme generated from the distribution of raw IRs (same as used in raw rate map).

ferent tribes. By far, the majority of these were Luo (93.0%), followed by Kisii (3.3%) and Kuria (2.2%). Individual cases also occurred in Suba, Samia and Kalenjin children.

#### Rate maps

Raw and EB smoothed IRs are presented in Figures 2 and 3. Raw rates are presented using a quintile classification scheme. Breakpoints from this classification scheme were applied to the distribution of EB rates to better visualize the impact of smoothing. According to the EB smoothed map, Locations on the north shore of Lake Victoria appeared to have elevated risk when compared to the surrounding areas, as did several administrative units in Nyando District as well as in the northern most region of Siaya District. The apparent area of high incidence on the southwest border with Tanzania was less pronounced following smoothing. Low-risk units were concentrated in the highland Districts of Nyamira, Central Kisii and Gucha.

#### Spatial clustering of cases using GeoDa–Anselin's Local Moran test (LISA)

Formal cluster analysis identified significant eBL clustering in the Province. Two multi-centered clusters of high eBL risk were identified using the Local Moran test for spatial autocorrelation (Table II and Fig. 4). As suggested in the EB smoothed rate map, the first large cluster, located in the western region of Kisumu District, was centered on 5 Locations: Central Kisumu, East Seme, South Central Kisumu, South West Kisumu and West Kisumu. EB rates for these Locations ranged from 2.1 to 8.0 cases per 100,000 children. The second high-risk cluster was detected in Nyando District, comprising 2 Location centers: Kikolo/East Kano and North Nyakach. EB smoothed IRs of these Locations were 3.1 and 4.4 cases per 100,000 children, respectively. A single large

and significant area of Locations with low eBL risk was detected. This area of low risk was centered on 45 of the 65 Locations in the highland Districts of Nyamira, Kisii and Gucha (data not presented). During the 6-year period, only 1 case was reported from the 45 cluster centers.

#### Spatial clustering of cases using SaTScan–spatial scan test statistic

Table III and Figure 5 present the results from the spatial scan test for spatial clustering. The most likely high-risk eBL cluster included 3 Locations in the western region of Kisumu District: East Kisumu, South West Kisumu and West Kisumu (High-Risk Cluster No. 1,  $p = 0.0001$ ). The second most likely high-risk cluster included a total of 20 administrative units; 5 located in the eastern region of Kisumu District and 15 located in Nyando District (High-Risk Cluster No. 2,  $p = 0.0001$ ). By comparing these clusters to the EB smoothed rate map, we find that the rates for Western Kisumu cluster are all above 4.90 cases per 100,000 children, while the EB smoothed rates for the Nyando/Eastern Kisumu cluster were highly variable, ranging from 1.1 to 5.5 cases per 100,000 children. The 15 cluster Locations in Nyando District generally had higher smoothed rates compared to the 5 Locations in Eastern Kisumu District. Both high-risk clusters detected in SaTScan were either included in or overlapped those identified using GeoDa. Three areas of significantly low eBL risk were identified with the spatial scan (Low Risk Area No. 1,  $p = 0.0001$ ; Low Risk Area No. 2,  $p = 0.0004$ ; Low Risk Area No. 3,  $p = 0.0170$ ). The most likely low-risk area was located in the east-central highland region of Nyamira, Kisii and Gucha Districts. All 3 areas were consistent in location with the low-risk area identified in GeoDa, with the exception of Low Risk Area No. 3, which also included 3 adminis-

**TABLE II – STATISTICALLY SIGNIFICANT HIGH-RISK CLUSTERS IDENTIFIED WITH ANSELIN'S LOCAL MORAN TEST FOR SPATIAL AUTOCORRELATION IN GeoDa, NYANZA PROVINCE, 1999–2004**

Cluster	Center	Raw IR <sup>1</sup>	EB smoothed IR <sup>1</sup>	Local Moran	p-value
Western Kisumu	Central Kisumu	5.3	3.8	2.19	0.005
	East Seme	13.0	7.0	8.88	0.001
	S.C. Kisumu	13.2	8.0	11.89	<0.001
	S.W. Kisumu	16.9	7.2	6.15	0.005
	West Kisumu	2.4	2.1	0.17	0.001
Nyando	N. Nyakach	6.7	3.1	0.86	0.001
	Kikolo/E. Kano	10.2	4.4	0.86	<0.001

<sup>1</sup>Rates reflect cases per 100,000 children.

**TABLE III – STATISTICALLY SIGNIFICANT HIGH-RISK CLUSTERS AND LOW-RISK AREAS OF eBL IDENTIFIED WITH THE SATSCAN SPATIAL SCAN TEST STATISTIC, NYANZA PROVINCE, 1999–2004**

Type	No.	Reported <sup>1</sup> cases	Expected cases	Relative risk	No. of locations	p-value
High-Risk Cluster	1	18	2.64	7.30	3	0.0001
	2	34	10.99	3.48	20	0.0001
Low-Risk Area	1	0	23.45	0.00	20	0.0001
	2	0	13.89	0.00	11	0.0004
	3	3	18.33	0.15	14	0.017

<sup>1</sup>Cumulative case totals over 6-year study period.

trative units from Migori District. SaTScan also excluded several highland administrative units bordering Homa Bay District to the west.

Neither cluster detection method identified significant high-risk areas within Siaya District. Only the Western Kisumu District cluster included administrative units bordering Lake Victoria in both GeoDa and SaTScan. Although several administrative units on the southwest border with Tanzania appeared to have elevated risk, these were not identified as a statistically significant cluster in either cluster detection method.

**Discussion**

*Study population*

Our study detected significant spatial clustering of eBL in Nyanza Province, a high-risk eBL region in Western Kenya. These findings were based on place of residence information obtained from children who reached Nyanza Provincial General Hospital and were diagnosed with eBL. The dramatic nature of eBL presentation suggests that most cases will be referred to the Provincial hospital when treatment is available. Moreover, that our study included 4 cases from Kuria District in the southeastern-most region of Nyanza Province, as well as 9 cases from Muhuru Division in Migori District with some of the most isolated communities in southwest Nyanza, indicates that we have captured the majority of cases occurring in the Province during the study period. Similar case demographics between our study and others support this assertion; the median disease onset age (7.0 years), approximate equal gender ratio and similar ethnic distribution (93% Luo) in our study were also consistent with those from other studies in Kenya and a follow-up study in Northern Tanzania.<sup>15,20</sup> The exception may be for the year 2002 when treatment was not available and for the few cases for which residence information at the 4th administrative level was not available. Our conclusions, therefore, assume a similar geographic distribution for undiagnosed eBL cases which may have occurred during 2002 and for cases without high resolution residence information.

*Spatial clustering*

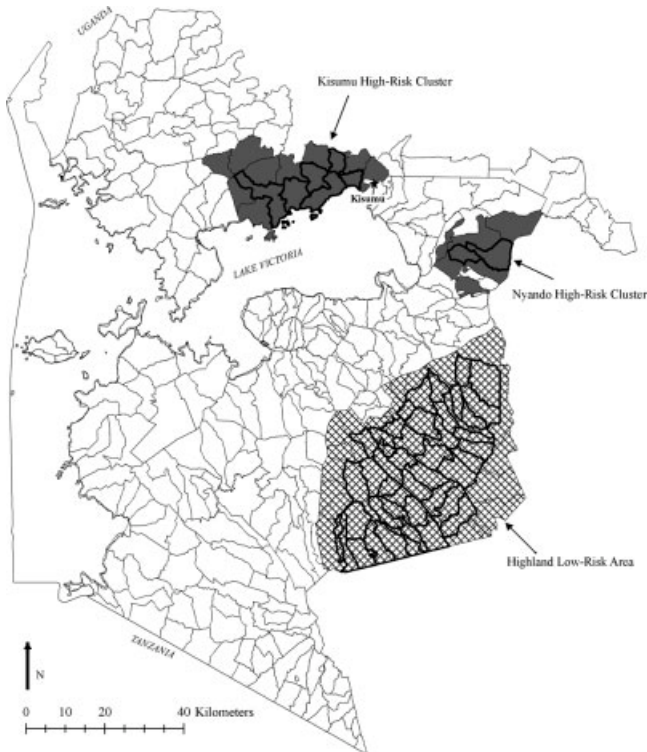
Two cluster detection methods were used to test the hypothesis that eBL spatially clusters in Nyanza Province. Both methods detected similar and significant high-risk clustering and low-risk areas, suggesting that our results are robust. One important eBL

co-factor, Epstein-Barr virus infection, is evenly distributed in the population and infection occurs early in childhood.<sup>21,22</sup> Our consistent findings, therefore, may imply clustering of other environmental or socio-cultural eBL co-factors, although potential reporting bias must also be assessed. Neither high-risk cluster included administrative units in Siaya District, as initially hypothesized. Whether eBL clusters near Lake Victoria was more difficult to assess. Only 1 high-risk Lakeside cluster was identified by both cluster detection methods.

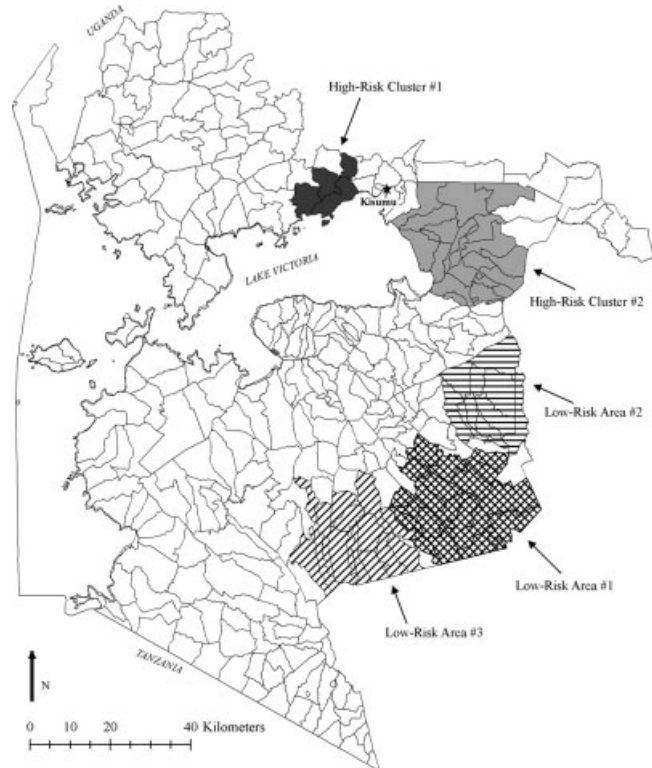
Morrow proposed that geographic variability in eBL incidence reflects differences in malaria transmission intensity.<sup>5</sup> Because of the lack of local-level transmission data for our analysis, we cannot rule this out. Detecting highly significant areas of low eBL risk in the highlands does support this hypothesis. Elevation for Locations in the highlands ranges from 1,475 to 2,205 m. Although malaria can occur throughout the region, transmission is generally more seasonal (March to July) in the highlands when compared to the lowland areas near Lake Victoria. Transmission above 1,600 m is rare.<sup>23,24</sup> Locations included in the most likely low-risk area identified by SaTScan had minimum elevations of roughly 1,580 m or more.

The most significant high-risk cluster identified with both methods involved administrative units situated in the western region of Kisumu District, bordering but not including Locations in urban downtown Kisumu. This cluster includes the major transportation route from Siaya and Bondo Districts to Kisumu city and Nyanza Provincial General Hospital. Proximity to this transportation route may have increased case recognition from these Locations. Elevated eBL incidence, however, was not observed in administrative units directly east of Kisumu city, despite similar road conditions. Although the Kenya Medical Research Institute is also located within this cluster, it is unlikely that non-eBL related research conducted in neighboring communities influenced eBL recognition and health service utilization. These factors (*i.e.*, major transportation routes and the Kenya Medical Research Institute) are not present in the majority of high-risk cluster administrative units located in Nyando District.

Statistically significant clustering in the Province, including in areas where malaria is holoendemic, provides evidence that other co-factors may be involved. These findings may be either linked to unique geographic features of the area, related to shared socio-cultural practices or both. Nyando District, for example, is a rural and economically poor region prone to seasonal flooding. While flooding may result in increased mosquito breeding habitat and



**FIGURE 4** – Statistically significant ( $p < 0.01$ ) high-risk clusters and low-risk area of eBL identified by Anselin's Local Moran test for spatial autocorrelation in GeoDa, Nyanza Province, 1999–2004. Bolded boundaries reflect high-risk cluster and low-risk area centers.



**FIGURE 5** – Significant high-risk clusters and low-risk areas of eBL detected with the SaTScan spatial scan test statistic, Nyanza Province, 1999–2004.

malaria transmission, it may also be associated with temporary displacement of people and crop destruction. Both frequent relocation<sup>25</sup> and poor nutrition (including protein deficiency)<sup>26</sup> have been associated with increased risk in previous studies.

Additionally, *Euphorbia tirucalli*, a plant with tumor promoting properties,<sup>27</sup> grows abundantly throughout the Province, but its use in household, medicinal and play time activities may vary.<sup>28,29</sup> This plant is frequently used for a variety of purposes, including firewood, in and near the Kisumu and Nyando District high-risk clusters. The fact that both these cluster areas have also been greatly impacted by the HIV/AIDS epidemic cannot be ignored. All eBL cases included in our study were HIV negative. Nevertheless, parental death due to HIV/AIDS or other cause may have a significant impact on family structure and health status. No sibling cases were observed during the 6-year study period, suggesting that genetic factors are unlikely to explain eBL clustering in the Province.

#### *Local Moran and spatial scan test statistics*

This analysis used 2 cluster detection methods. The spatial scan test statistic assesses disease distribution using centroids and a circular scan. Certain geographic features, such as Lake Victoria, may not be considered, and cluster boundaries, due the number of overlapping circular windows, may not be well defined.<sup>30</sup> Anselin's Local Moran test for spatial autocorrelation, on the other hand, is less sensitive to the unique geographic features of Nyanza Province, but examines clustering according to user defined neighbor relationships. Despite these different attributes, both cluster detection methods identified significant high-risk clusters and low-risk areas in similar geographic locations. Consistent results using these 2 methods, in addition to 6-years of eBL case data, and a rate smoothing technique suggest that these results are robust. High resolution malaria transmission data will soon be available. Determining whether these spatial clusters occur independently from

malaria transmission intensity, as well as over different time periods, is an important next step. Improving our understanding of the geographic distribution of eBL will ultimately provide crucial information on the role of environmental and/or socio-cultural factors in the complex causal mechanisms underlying eBL etiology.

Nyanza Province continues to represent a high-risk region for eBL with an average annual IR of more than 2 cases per 100,000 children. Our study was the first to rigorously investigate a fine-scale spatial distribution of eBL in this region. Two different cluster detection methods identified statistically significant high-risk clusters and a large low risk region. Most noteworthy are the low-risk areas observed in the highland region and the 2 high-risk clusters situated in Kisumu and Nyando Districts. The etiology of eBL is complex and multi-factorial. Significant spatial clustering of elevated eBL risk in high-malaria transmission regions and of reduced incidence where malaria is infrequent suggests that malaria plays a role in eBL etiology, but that additional co-factors may also be involved. Future studies are needed to evaluate these eBL clusters, and to determine if they are associated with specific environmental and/or social-cultural characteristics.

#### **Acknowledgements**

We thank the Kenya Medical Research Institute in Kisumu, Kenya, for the opportunity to conduct this study. The success of this project also relied on the dedicated work of Dr. Juliana Otieno, chief pediatrician, and Dr. Margaret Odour, Department of Pathology, both of Nyanza Provincial General Hospital. Dr. Otieno diagnosed and treated the Burkitt's lymphoma cases included in this study, and also provided essential support with patient records and case histories. Dr. Odour worked diligently to assure quality examination of all laboratory specimens. We also thank Dr. Robert Snow at the Kenya Medical Research Institute/



Wellcome Trust Collaborative Programme in Nairobi for access to population and digitized map data for Nyanza Province, and Mr. Scott Swan at the University of Michigan Center for Statistical

Consultation and Research for assistance with importing data into a geographic information system. This research was approved by the Director of the Kenya Medical Research Institute.

### References

- Burkitt DA. "Tumor Safari" in East and Central Africa. *Br J Cancer* 1962;16:379-86.
- Haddow AJ. An improved map for the study of Burkitt's lymphoma. *East Afr Med J* 1963;40:429-32.
- Burkitt D. Etiology of Burkitt's lymphoma-an alternative hypothesis to a vectored virus. *J Natl Cancer Inst* 1969;42:19-28.
- Evans AS. Epidemiology of Burkitt's lymphoma: other risk factors. *IARC Sci Publ* 1985;60:197-204.
- Morrow RH. Epidemiological evidence for the role of falciparum malaria in the pathogenesis of Burkitt's lymphoma. *IARC Sci Publ* 1985;60:97-105.
- Williams CK. Clustering of Burkitt's lymphoma and other high-grade malignant lymphoproliferative diseases, but not acute lymphoblastic leukaemia among socio-economically deprived Nigerians. *East Afr Med J* 1988;65:253-63.
- van den Bosch C, Lloyd G. Chikungunya fever as a risk factor for endemic Burkitt's lymphoma in Malawi. *Trans R Soc Trop Med Hyg* 2000;94:704-705.
- Pike MC, Williams EH, Wright B. Burkitt's tumor in the West Nile District of Uganda 1961-1965. *Br Med J* 1967;2:395-9.
- Morrow RH, Pike MC, Smith PG, Ziegler JL, Kisuule A. Burkitt's lymphoma: a time-space cluster of cases in Bwamda County of Uganda. *Br Med J* 1971;2:491,492.
- van den Bosch C, Hills M, Kazembe P, Dziweni C, Kadzamia L. Time-space case clusters of Burkitt's lymphoma in Malawi. *Leukemia* 1993;7:1875-8.
- Biggar RJ, Nkrumah FK. Burkitt's lymphoma in Ghana: urban-rural distribution, time-space clustering and seasonality. *Int J Cancer* 1979;23:330-6.
- Brubaker G, Geser A, Pike M. Burkitt's lymphoma in the North Mara District of Tanzania 1964-1970: failure to find evidence of time-space clustering in a high risk isolated rural area. *Br J Cancer* 1973;28:469-72.
- Siemiatycki J, Brubaker G, Geser A. Space-time clustering of Burkitt's lymphoma in East Africa: analysis of recent data and a new look at old data. *Int J Cancer* 1980;25:197-203.
- Makata AM, Toriyama K, Kamidigo NO, Eto H, Itakura H. The pattern of pediatric solid malignant tumors in western Kenya, east Africa, 1979-1994: an analysis based on histopathologic study. *Am J Trop Med Hyg* 1996;54:343-7.
- Mwanda OW, Rochford R, Moormann A, MacNeil A, Wilson ML. Burkitt's lymphoma in Kenya: geographical, age, gender and ethnic distribution. *East Afr Med J* 2004;8:S68-S77.
- Kenya Ministry of Planning and National Development, Central Bureau of Statistics. The 1999 Kenya population census. Republic of Kenya, 2002.
- Kenya Ministry of Planning and National Development, Central Bureau of Statistics. The 1999 population and housing census, the popular report. Republic of Kenya, August 2002.
- United States Census Bureau, International Programs Center, Population Division. HIV/AIDS surveillance data base, HIV/AIDS Profile. Kenya, 2000.
- Jacquez GM, Greiling DA. Local clustering in breast, lung, and colorectal cancer in Long Island. *Int J Health Geogr* 2003;2:1-12.
- Geser A, Brubaker G, Draper CC. Effect of a malaria suppression program on the incidence of African Burkitt's lymphoma. *Am J Epidemiol* 1989;129:740-52.
- Biggar RJ, Henle W, Fleisher G, Bocker J, Lennet ET, Henle G. Primary Epstein-Barr virus infections in African infants. Decline of maternal antibodies and time of infection. *Int J cancer* 1978;22:239-43.
- Moormann AM, Chelimo K, Sumba OP, Lutzke ML, Ploutz-Snyder R, Newton D, Kazura J, Rochford R. Exposure to holoendemic malaria results in elevated Epstein-Barr virus loads in children. *J Infect Dis* 2005;191:1233-8.
- Kenya Ministry of Health, Malaria Control Program. Epidemiology of malaria in Kenya. *Afr J Med Pract* 1994;1:5,6.
- Minakawa N, Sonye G, Mogi M, Githeko A, Yan G. The effects of climatic factors on the distribution and abundance of malaria vectors in Kenya. *J Med Entomol* 2002;39:833-41.
- Morrow RH. Introduction to Burkitt's lymphoma in Africa. *Cancer Res* 1974;34:1211-15.
- Clifford P, Singh S, Stjernsward J, Klein G. Long-term survival of patients with Burkitt's lymphoma: an assessment of treatment and other factors which may relate to survival. *Cancer Res* 1967;27:2578-615.
- Osato T, Mizuno F, Imai S, Aya T, Koizumi S, Kinoshita T, Tokuda H, Ito Y, Hirai N, Hirota M. African Burkitt's lymphoma and an Epstein-Barr virus-enhancing plant *Euphorbia tirucalli*. *Lancet* 1987;1:1257-8.
- van den Bosch C, Griffin BE, Kazembe P, Dziweni C, Kadsamira L. Are plant factors a missing link in the evolution of endemic Burkitt's lymphoma? *Br J Cancer* 1993;68:1232-5.
- MacNeil A, Sumba OP, Lutzke ML, Moormann AM, Rochford R. The latex of the plant *Euphorbia tirucalli* activates Epstein-Barr virus lytic cycle. *Br J Cancer* 2003;88:1566-9.
- Kulldorf M, Feuer EJ, Miller BA, Freedman LS. Breast cancer clusters in the northeast United States: a geographic analysis. *Am J Epidemiol* 1997;146:161-70.