Identification of environmental parameters and risk mapping of visceral leishmaniasis in Ethiopia by using geographical information systems and a statistical approach

Teshome Tsegaw1*, Endalamaw Gadisa1*, Ahmed Seid1, Adugna Abera1, Aklilu Teshome1, Abate Mulugeta2, Merce Herrero2, Daniel Argaw3, Alvar Jorge3, Abraham Aseffa1

1Armauer Hansen Research Institute, All-Africa Leprosy and TB Rehabilitation and Training Center, Jimma Road, P.O. Box 1005, Addis Ababa, Ethiopia; 2Disease Prevention and Control Programmes, World Health Organization, Ethiopia Country Office, Addis Ababa, Ethiopia; 3Department for the Control of Neglected Tropical Diseases, Leishmaniasis Control Programme, World Health Organization, Geneva, Switzerland; *Contributed equally

Abstract. Visceral leishmaniasis (VL), a vector-borne disease strongly influenced by environmental factors, has (re)-emerged in Ethiopia during the last two decades and is currently of increasing public health concern. Based on VL incidence in each locality (kebele) documented from federal or regional health bureaus and/or hospital records in the country, geographical information systems (GIS), coupled with binary and multivariate logistic regression methods, were employed to develop a risk map for Ethiopia with respect to VL based on soil type, altitude, rainfall, slope and temperature. The risk model was subsequently validated in selected sites. This environmental VL risk model provided an overall prediction accuracy of 86% with mean land surface temperature and soil type found to be the best predictors of VL. The total population at risk was estimated at 3.2 million according to the national population census in 2007. The approach presented here should facilitate the identification of priority areas for intervention and the monitoring of trends as well as providing input for further epidemiological and applied research with regard to this disease in Ethiopia.

Keywords: visceral leishmaniasis, neglected tropical diseases, risk map, geographical information systems, Ethiopia.

Introduction

Visceral leishmaniasis (VL), also known as Kala-azar is a zoonotic, infectious disease transmitted by the bite of infected sand flies (Phlebotomus in the old world and Lutzomyia in the Americas), which leads to death if untreated. It is caused by a parasite of the Leishmania donovani species complex, reported to consist of various strains belonging to the three species L. donovani, L. infantum and L. archibaldi (Zemanova et al., 2004). VL is a neglected tropical disease (NTD), which is strongly associated with poverty. It affects poor communities in 79 countries with an estimated annual incidence of 0.2 to 0.4 million new cases per year (Alvar et al., 2006a, 2012). Over 90% of the annual incidence occurs in just seven countries: Bangladesh, India, Nepal, Sudan, South Sudan, Ethiopia and Brazil (Alvar et al., 2012). Eastern Africa, the second largest VL focus after the Indian subcontinent, accounts for 17% of the global burden, with Ethiopia and the Sudans contributing the largest proportion.

Changes in the environment (climate change, deforestation and unplanned urbanization), population movements between endemic and non-endemic zones, appearance of therapy-resistant strains and immune suppression, mainly due to malnutrition and co-infection with the human immunodeficiency virus (HIV), are among the factors believed to contribute to (re)-emergence of VL (Alvar et al., 2006a; Chappuis et al., 2007). Unfortunately, the drugs available for the treatment of VL are expensive and relatively toxic (Alvar et al., 2006b).

Effective planning of control measures requires adequate understanding of the transmission dynamics (WHO, 2010). The increasing availability of satellite-derived products and the recent technological development in geographical information systems (GIS) have given scientists the opportunity to develop spatial and spatio-temporal risk maps of vectors and vector-
borne diseases, including malaria, schistosomiasis, trypanosomiasis, onchocerciasis and leishmaniasis, thereby contributing to better planning of control activities and the establishment of early warning systems (de la Rocque et al., 2001; Bavia et al., 2005; Gebre-Michael et al., 2005; Leonardo et al., 2007; Shirayama et al., 2009).

In Ethiopia, the estimated incidence of VL ranges from 3,700 to 7,400 cases per year (Alvar et al., 2012). The first reported case of VL in the country dates back to the 1940s (Coles et al., 1942) but the exact burden of disease, distribution and environmental determinants are poorly understood. From sporadic surveys, transmission is known to occur in three different and well-delimited ecological regions, one on the highland plateau and two others in the lowland areas (below 1,500 m) in the North and the South.

The foci in the semi-arid northern part of the lowlands are found on the Metema and Humera plains in the Tigray and Amhara regional states that border the Sudan and Eritrea. The local vector *Phlebotomus orientalis* is found associated with red acacia trees and black cotton soil (Gebre-Michael et al., 2010). These foci account for over 60% of the total burden of VL in Ethiopia with the highest reported co-infection rate (~38%) with HIV (Ayele and Ali, 1984). However, VL cases are relatively rare in these areas and the prevalence of persons with positive skin tests is low (Berhe et al., 1998). The disease has been reported further eastward in the Moyale area and Genale River basin in the Oromia regional state near the Kenyan border (Ayele and Ali, 1984). More recently, it has also been reported in the south-eastern semi-arid lowlands of the Ander and Liben areas of the Somali regional state (Marlet et al., 2003). The majority of cases occurred in the nomadic tribes and all the new cases were children, suggesting endemic transmission (Marlet et al., 2003).

The highland plateau focus of Libo-Kemkem and Fogera *woredas* in the Amhara regional state was first reported as an outbreak by the MSF in 2005 (Alvar et al., 2007). According to the studies assessing the outbreak, it began in Bura Kebele in 2003 with the number of cases peaking in 2005, after which the area was transformed into a low-incidence endemic zone (Herrero et al., 2009). The outbreak affected over 2,450 people and hundreds died (Alvar et al., 2007; Herrero et al., 2009). In response, a control programme was launched by the Ministry of Health in 2011 (Alvar et al., 2007; Herrero et al., 2009). The objective of the present work was to develop a risk map of VL in Ethiopia based on a set of environmental factors to estimate the total geographical surface of the endemic areas and the percentage of total population at risk.

**Materials and methods**

**Study area**

Ethiopia is a land-locked country between latitudes 3° and 15° N and longitudes 33° and 48° E in the Horn of Africa with a population of 74 million (according to the 2007 census) covering an area of 1,133,380 km². After Nigeria, Ethiopia is the most populous nation in Africa. The predominant climate type is tropical monsoon with wide topographic variations. The highlands have a considerably cooler climate than other regions at similar proximity to the equator. Administratively, Ethiopia is divided into nine autonomous states and two city administrations established under the 1994 constitution (Fig. 1). The administrative levels under the Federal Government are regional state (or city administration), zone, *woreda* (district), *kebele* (neighbourhood of about 500 households or 3,500-4,000 persons) and *gott* (village).
Epidemiological data

A structured questionnaire was used to document both epidemiological and VL presence and absence data at the kebele-level. Reported VL cases and information about their travel history before the onset of symptoms were gathered from the medical records at health facilities in the known VL foci and government reports at different administrative levels in the states of Amhara, SNNPR, Tigray and Oromia. Global positioning system (GPS) data were collected on site at the kebele-level using a hand-held GPS device (Garmin GPS 60©). One GPS point per kebele corresponds to a horizontal distance between VL cases of ≤300 m, a change in elevation between cases within ≤100 m in places with the same soil type.

A VL case was defined as an individual, with a clinical sign-symptom complex suggestive of VL, who also tested serologically positive with the rK39 immunochromatographic test (Diro et al., 2007) or parasitological examination of specimens according to the national health guidelines. An experienced clinician (member of the study team) examined suspected patients on site and appropriate specimens were collected at the health facility for further confirmation at the Leishmania laboratory at the Armauer Hansen Research Institute (AHRI).

Meteorological data

Annual average temperatures and annual rainfall for 20 consecutive years (1989-2009), derived from >100 meteorological stations spread across the country, were received from the Ethiopian Meteorological Service Agency (NMA, 2010). Kriging, applied for the annual average temperature and rainfall was used to obtain an interpolated continuous surface with a grid areal size of 0.81 km².

Soil data

The Food and Agricultural Organization of the United Nations (FAO) soil database (FAO, 1998) was clipped to the boundary of Ethiopia and por-
Trayed to assess which soil types were most associated with VL cases. The vector soil database was changed into raster format with the same pixel size as stated above. The soil type at each location was determined on site by professional geographers in the study team.

**Elevation data**

A digital elevation model (DEM) for the study area was obtained from the Shuttle Radar Topography Mission, US Geological Survey (SRTM; http://www.landcover.org (accessed on December 2011)). DEM from the SRTM are provided at a minimal spatial resolution of three arc-second (approximately 90 m) for areas outside the USA, but considering that the risk model was intended for the country level, we worked with a 30 arc-second (900 m²) version. A slope surface (in degrees) was subsequently generated from the DEM.

**Statistical method**

Binary and multivariate logistic regression approaches were used to assess the relative importance of environmental parameters with respect to presence or absence of VL cases in an area. The odds ratio (OR), including P-value of the binary logistic regression, was used to determine the relative importance of each environmental parameter with respect to VL presence. The OR of the multivariate analysis was used to assign weight to each environmental parameter. Then, the assigned weights were used in the spatially weighted overlay analysis at the pixel size of 0.81 km² grid for the country. Statistical analysis was done using Stata/SE version 11 (College Station, TX, USA). A total of 1,800 GPS points were considered for a step-wise binary logistic regression (689 marking VL presence and 1,111 marking VL absence).

**Weighted overlay**

The environmental parameters used in the analysis were changed into raster format with a pixel size of 0.81 km². The OR of the multivariate analysis result of all environmental parameters versus VL case presence/absence was used to assign weight for spatially weighted overlay and the disaggregated value of each parameter with the relative OR of the binary logistic value was used to run a model to produce the risk map of VL in Ethiopia using ArcGIS 9.3 (ESRI, Redlands, CA, USA).

**Results**

**Soil**

The soil digital data with the country’s boundary produced a clipped map of 20 major soil types. The overlay of the soil map with the GPS points of the VL presence/absence data showed that 74.4% of the areas with VL presence fell within the vertisol soil type and 9.4% within the lixisols type, while luvisol soil accounted for only 6.8%. The remaining 9.3% was shared by cambisol (5.4%) and fluvisol (3.9%). The binary logistic regression analysis (Table 1) indicates that vertisols, lixisols, cambisols and luvisols are significantly associated (P <0.05) with the presence of VL in an area.

**Rainfall**

Map overlays combining rainfall and VL data showed that 48% the endemic kebeles received, on average, less than 766 mm of annual rainfall. As the amount of annual average rainfall increased, the likelihood of presence of VL in a kebele decreased (Table 1).

**Temperature**

Map overlays combining the average annual temperature and the VL data indicated that the majority of VL cases (94.7%) occurred in the 20-37 °C annual average temperature category. The stepwise binary logistic regression with respect to annual average temperature showed that the number of VL increases with the average temperature (Table 1).

**Altitude**

Map overlays combining DEM and the VL data identified that the majority of the cases (95.3%) were located at altitudes below 1,872 m. The stepwise binary logistic regression with respect to altitude showed the occurrence of VL in decreases at higher altitudes (Table 1).

**Slope**

When slope and data points indicating VL cases were combined, 85.3% of all VL cases were found to be located in areas with slope values <1.53. A binary logistic regression comparing slope with presence/absence of VL cases indicated an inverse relationship between slope and presence of VL (Table 1).
Spatial overlay analysis

The multivariate logistic regression showed the average temperature followed by soil as having the strongest association with VL risk (Table 1). The risk map (Fig. 2) showed that high-risk areas were found in the lowlands including the Humera, Metema, Mirab Armachiho, Quara, Sheraro, Metema, Tahtay-Adiabo, Tsegede and Tegede woredas in the Amhara and Tigray states; Abobo, Akobo, Dima, Goge, Gambela Zuria, Itang, Jikawo, Jore, Lare and Wantawo woredas in the Gambela state; Assosa, Guba, Kurmuk, Mao-Komo Pawe Special and Sherkole woredas in the Beneshangul Gumuz state; Anfilo and Gidame woredas in the Oromia state; Dasenech, Dire, Gngatrom, Hamer, Konso, Meyo, Moyale, Selamago and Tetele woredas in SNNPR; and Barey, Chereti, Danan, Dolo-Addo, Dolo Bay, Debe-Woin, East Eme, Elkere, Filtu, Ferfer, Godey, Kelafo, Mustahil, Salahad and West Eme woredas in the Somali state (Fig. 2). The predicted risk levels of VL presence in terms of area coverage (and the corresponding percentages) are presented in Table 2. The proportion of areas at high risk was 33.6%, accounting for 30.2% of the total.

Model validation

Validation was conducted in areas predicted by the model in which the team had not previously collected GPS point data. The team visited a total of 15 woredas in Amhara, Beneshangul Gumuz, and Tigray states in order to check the accuracy of the predictions. Ten positive cases were identified in areas predicted by the model as being of high risk in Raya-Azebo, Guba, Kurmuk and Jawi woredas in Tigray, Beneshangul Gumuz and Amhara, respectively. In addition, data of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Bivariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds ratio</td>
<td>P-value</td>
</tr>
<tr>
<td>Soil type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alisol</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Cambisol</td>
<td>0.64</td>
<td>0.13</td>
</tr>
<tr>
<td>Lixisol</td>
<td>8.16</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Luvisol</td>
<td>0.62</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Nitosol</td>
<td>0.04</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Vertisol</td>
<td>24.32</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Rainfall (mm)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;766</td>
<td>2.73</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>766-1,057</td>
<td>1.26</td>
<td>0.04</td>
</tr>
<tr>
<td>1,057-1,142</td>
<td>3.64</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>1,142-1,224</td>
<td>0.51</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>&gt;1,224</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Average temperature (°C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;16</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>16-21</td>
<td>0.15</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>21-26</td>
<td>0.53</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>26-31</td>
<td>5.43</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>&gt;31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altitude (m)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;2,305</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>1,872-2,305</td>
<td>0.12</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>1,794-1,872</td>
<td>19.39</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>&lt;1,794</td>
<td>77.68</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Slope (degree)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;6.40</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>3.78-6.40</td>
<td>0.10</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>1.53-3.78</td>
<td>0.34</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>0.56-1.53</td>
<td>5.58</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>&lt;0.56</td>
<td>30.28</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

Table 1. Bivariate and multivariate analysis of environmental variables used in the development of the risk map of visceral leishmaniasis in Ethiopia.
VL presence in east Imey and Liben zones of Somali regional state, predicted to be at high risk was received from the World Health Organization (WHO) country office for Ethiopia (Abate M., personal communication). These areas were not visited by the team because of security concerns (Fig. 2). Most parts of Gambela were not surveyed directly in this study but were identified as being of medium to very high risk by the model. In this case, literature review through a PubMed search indicated that there is previous evidence of possible VL endemicity in Gambela (Hailu et al., 1996). According to the study, positive skin test rates varied from 7.6% to 1.7% among the local population (the Nuer and Anuak people of Gambela) and re-settlers from the highlands, respectively. Sero-prevalence rates were 7.3% and 8.1% for these populations. Moreover, no case of VL was encountered by the team in areas predicted as VL-free by the model when these were assessed as part of a parallel cutaneous leishmaniasis mapping project.

**Country-wide risk**

Based on the 2007 census and assuming the distribution of the population in a *kebele* stays relatively even, the total population at risk was estimated combining the VL risk map with that showing the *kebele* boundaries. Those parts of the *kebele* that fell in the high and very high risk categories were digitised and divided with the total area to determine the proportion. The resulting number for people at risk in the country as a whole was 3,275,338 but they were not evenly distributed geographically but were found to be living mainly in the Amhara, Tigray and SNNPR states (Table 3).

**Discussion**

In the current study, the annual average temperature shows a high leverage on VL occurrence (33%), which is in line with a previous study done in Gedaref State in Sudan by Elnaiem (2003). Soil was found to
be the second most important determinant with a percentage influence of 27%. Vertisol, the black-cotton soil type, was associated with 74.4% of the VL cases. According to FAO data, vertisol has a high content of the expanding clay type known as montmorillonite that forms cracks when dried but which absorbs water and swells causing self-mulching during the rainy season. The deep cracks provide a sand fly microhabitat suitable for breeding and resting (Elnaiem, 2003). This soil type is found associated with the red acacia and *P. orientalis*, which are common in areas endemic for *L. donovani* transmission (Gebre-Michael et al., 2010).

Altitude produced a relative percentage of influence of 20%. The logistic regression analysis of the categorised elevation value revealed that an altitude category value of <1,794 m is the most favourable elevation range for VL. The range of 1,794-1,872 m was found to be the second most favourable category for VL in this study. The Elnaiem et al. (2003) study in Gedaref state of Sudan, using a probabilistic model, concluded that altitude is one of the best predictors for the presence of VL in a village. The long-known endemic foci established through epidemiological surveys and self-reported cases of VL in Ethiopia are from the lowlands and the recently reported highland foci range from 1,800 to 2,000 m (Ali et al., 1994; Hailu et al., 1996; Alvar et al., 2007; Herrero et al., 2009; Sordo et al., 2012).

Rainfall was estimated to have a relative percentage of influence of 13%. Values between 1,057 and 1,142 mm showed the maximum rank with an OR of 3.64 followed by the 0-766 mm rainfall category (OR = 2.73). A study in Sudan VL foci on the influence of annual average rainfall on VL distribution documented that up to 1,200 mm is suitable for the growth of red acacia trees known to be preferred by sand flies associated with *L. donovani* (Thomson et al., 1999).

The risk model developed based on spatial weighted overlay analysis and multivariate logistic regression showed that the lowlands of north-western Ethiopia: namely the Metema, Mirab Armachiho, Quaf虐待Humera, Quara, Tegede and Tsegede *woredas* fell in the high and very high risk areas of VL presence (Fig. 3a). These areas are among the well-established and most studied foci of VL in the country (Haile and Anderson, 2006; Diro et al., 2007; Gebre-Michael et al., 2010). The north-western central foci (the Fogera, Libo-Kemkem, Misrak Belesa and Mirab Belesa *woredas*) in Amhara state mostly fell in the high and very high risk category of the model (Fig. 3b). The Fogera and Libo Kemkem foci are known in association with recent outbreak (Alvar et al., 2007; Sordo et al., 2012). The presence of VL transmission in Misrak (east) and Mirab (west) Belesa was indicated by Herrero et al. (2009) in their analysis of the natural history of the Libo-Kemkem and Fogera outbreak.

In the *West*, *woredas* such as Abobo, Akobo, Dima, Gambela Zuria, Goge, Itang, Jikawo, Jore, Lane and Wantawo in Gambela state and Mae-Komo, Assosa, Kurmuk, Sherkole, Menge, Guba, Mankush and Pawe in the state of Beneshangul Gumuz are predicted by the model to have high and very high risk for VL (Fig. 3c). We identified three positive cases of VL (positive by the rK39 serological method and on treatment during the visit) each in the Guba and Jawi *woredas* confirming the prediction by the model. In addition, seropositive cases for VL have been described previously by Hailu et al. (1996) in Gambela.

In the southern parts of the country, *woredas* like Konso, Hamer, Dasenech, Teltele and Nyangatom areas in SNNPR were predicted by the model to have high and very high risk for VL (Fig. 3d). These are all, except for the Hammer area, previously well described major VL foci in Ethiopia (Ayele and Ali, 1984; Ali and Ashford, 1993; Gebre-Michael and Lane, 1996; Hailu et al., 2009).

Other areas predicted by the risk model with high and very high risk for VL in the south-eastern parts of Ethiopia include Danan, Elkere, Debe-Woin, Hargele, Chereti, Filtu, Salahad and Sagag *woredas* (Fig. 3d). However, to our knowledge, except for Filtu (Abate, personal communication, WHO-Ethiopia), no VL cases have been reported from these areas so far. However, Marlet et al. (2003) reported an outbreak of VL at the border of Ethiopia, Somali and Kenya that might have affected the Afder, Liben Denan and Hagele areas of Ethiopia.
Conclusions

The findings of this study provide a useful and potentially valuable tool for leishmaniasis control. The results for VL based on the multivariate logistic regression and weighted overlay analyses can be replicated in other disease mapping projects. This would facilitate implementation of evidence-based integrated disease control activities at the country and lower administrative levels.

Acknowledgements

We thank the World Health Organization for financing this project and the Ethiopian Federal Ministry of Health for facilitating its smooth conduct. We also thank the Central Statistical Agency for their unreserved support in provision of kebele-level 2007 census data with their boundary shape file, the Ethiopian Mapping Agency for providing us the SPOT 5 metre satellite imagery, the Ethiopian Meteorology Service Agency for provid-
ing data on annual rainfall and temperature and the health bureaus and hospitals in the various parts of Ethiopia for providing us with VL case data. The Armauer Hansen Research Institute is a Ministry of Health research facility supported by the Ethiopian government and the Swedish and Norwegian Development Agencies (SIDA and NORAD).

References


Marlet MV, Sang DK, Rimecker J, Muga RO, Onsongo J, Davidson RN, 2003. Emergence or reemergence of visceral leishmaniasis in areas of Somalia, north eastern Kenya, and...