Research Journal of Applied Sciences, Engineering and Technology 13(6): 448-453, 2016
DOI:10.19026/rjaset.13.3005
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: July 7, 2015 Accepted: August 30, 2015 Published: September 15, 2016

Research Article

Correspondence Analysis for the Trend of Human African Trypanosomiasis

H.I. Okagbue, S.O. Edeki and A.A. Opanuga
Department of Mathematics, Covenant University, Canaanland, Otta, Nigeria

Abstract: The aim of this research work is to give a graphical picture of the declining trend of the Human African Trypanosomiasis (T.b. gambiense) in 12 selected endemic countries (based on 10 years data) via the application of correspondence analysis. Grouping the countries into three regions affects the model but reveals that the disease is most endemic in Central Africa but least in West Africa. Hence, we therefore recommend that efforts must be intensified by the countries in the Central Africa to reduce the menace of the disease since graphically; there have been reported cases mostly in this region for the past 10 years.

Keywords: Correspondence analysis, endemic, Human African Trypanosomiasis

INTRODUCTION

According to the World Health Organization (WHO, 2014), African Trypanosomiasis (sleeping sickness) is a parasitic disease of humans and other animals. It is caused by protozoa of the species *Trypanosoma brucei*. There are two known types that affect human beings: *Trypanosoma brucei gambiense* (T.b.g) and *Trypanasoma brucei rhodesciense* (T.b.r).

This research focuses on *Trypanosoma brucei gambiense* (T.b.g) which causes over 98% of reported cases and have high fatality rate of 100% if left untreated (Kennedy, 2013). Both types of the diseases are usually transmitted by the bite of an infected tsetse fly *Glossina spp*.

The symptoms manifest in two phases: the haemolymphatic phase which can last a day to a week or a month is often characterized by fever, headaches, fatigues, pains in the joint and itching. If the disease is left untreated, it overwhelms and defeats the host's immune defenses and can cause extensive damage with more symptoms like anaemia, endocrine, cardiac and kidney dysfunction (Kennedy, 2013; Brun *et al*., 2010). The second-neurological phase begins when the parasite invades the host central nervous system by passing through the blood-brain barrier (Brun *et al*., 2010). Disruptions of the sleep cycle are key symptoms of this phase. The host daytime performances suffer because of daytime sleep episodes and at night, wakefulness or insomnia while other neurological symptoms include confusion, tremor, general muscle weakness, hemiparesis and paralysis (Lundkvist *et al*., 2004). According to Uganda (2008), damage done in the neurological phase is irreversible.

T.b.gambiense is endemic in 24 countries of West, Central and East Africa. Many epidemiological data and analysis are available but this research was done using the data from the World Health Organization (country data of reported cases of Human African Trypanasomiasis T.b.gambiense).

In literature, some of the numerous works done on HAT include the following: the history of sleeping sickness in Africa, the epidemiology of HAT, model for predicting the effect of climate change on African Trypanosomiasis, human activities and demographic variables related to HAT, estimation of persons at risk and so on (Cox, 2004; Cattand, 2001; Moore *et al*., 2012; Tongué *et al*., 2011; Hide, 1999; Simarro *et al*., 2011).

RESEARCH OBJECTIVES

- To examine the manifestations of the disease with respect to each endemic country using correspondence analysis
- To examine the manifestation of the disease when the endemic countries are collapsed to (3) regions of Central, East and West Africa using correspondence analysis and
- To compare the result with the mortality rate and life expectancies of the 12 selected countries.

Geospatial Analysis and Geographic Information System GIS are some of the data analytics tools that have been used in analyzing HAT data (Franco *et al*., 2013). This research employed the use of correspondence analysis.

CA is an exploratory data technique used to analyze categorical data (Benzeci, 1992). Epidemiologists frequently collect data on multiple categorical variables with the goal of examining

Corresponding Author: S.O. Edeki, Department of Mathematics, Covenant University, Canaanland, Otta, Nigeria

associations among these variables (Panagiotakos and Pitsavos, 2004). The outcome of a CA is a geographical display of the rows and columns of a contingency table that is designed to permit visualization of the key connections among the variable responses in a low-dimensional space (Panagiotakos and Pitsavos, 2004). Such a representation reveals a more global picture of the relationships among row-column pairs, which would otherwise not be detected through pairwise analysis (Sourial *et al*., 2010). The row and column points are shown on the same graphical display allowing for easier visualization of the association among variables (Yelland, 2010).

The detailed procedures, analysis and practical use of the correspondence analysis can be seen in the works of Higgs (1991), Hill (1974) and Doey and Kurta (2011). The key feature of CA is that it uses chi-square statistics to measure the distance between points in the biplot (Clausen, 1998). The goal of the CA is to explain the most inertia or variance or variability in the model in the least number of dimensions (Garson, 2008).

METHODOLOGY

The data used for this research is obtained from the WHO website. In what follows, data from the countries that have recorded at least one case of HAT will be considered via correspondence analysis.

Basic assumption of the C.A.:

- CA assumes that the data analyzed is discrete in nature (Yelland, 2010; Doey and Kurta, 2011)
- CA also recommends that the data should be made up of several categories typically more than three (Franco *et al*., 2013; Doey and Kurta, 2011). Hence, the data for the research is in 12 rows and 10 columns
- CA also assumes that the values in the frequency table must be non-negative (Doey and Kurta, 2011; Clausen, 1998).

The concerned countries are those that have recorded at least one case of HAT for the past 10 years. This is done to avoid violating the assumptions of correspondence analysis. Since violations of any of the assumptions will yield misleading results (Doey and Kurta, 2011).

Data for *Trypanosoma brucei gambiense* for the past 10 years are collected except those countries whose data are unavailable. The countries are Gambia, Guinea-Bissau, Liberia, Niger Republic, Senegal and Sierra Leone. Also, the countries that have reported at least one zero case are excluded. These include Benin Republic, Burkina Faso, Ghana, Mali, Nigeria and Togo.

Table 1: Correspondence table[a]

	Year										
Country	2013	2012	2011	2010	2009	2008	2007	2006	2005	2004	Active Margin
Angola	69	70	154	211	247	517	640	1105	1727	2280	7020
Cote d'ivoire	7	9	10	8	8	14	13	29	42	74	214
Cameroun	6	7	15	16	24	13	7	15	3	17	123
Central African Republic	62	301	132	395	1054	1194	854	460	666	738	5856
Chad	193	197	276	232	510	196	97	276	198	403	2578
Congo PR	20	39	61	87	87	102	189	300	398	873	2156
Congo DR	5647	5983	5595	5629	7183	7326	8162	8023	10269	10369	74186
Equatorial Guinea	3	2	1	8	7	11	15	13	17	22	99
Gabon	16	9	17	22	14	24	30	31	53	49	265
Guinea	78	70	57	68	79	90	69	40	94	95	740
South Sudan	117	317	272	199	373	623	469	709	1853	1742	6674
Uganda	9	20	44	101	89	198	120	298	311	378	1568
Active Margin	6227	7024	6634	6976	9675	10308	10665	11299	15631	17040	101479

Table 2: The correspondence table (Summary)

					Proportion of Inertia		Confidence singular value	
Dimension	Singular Value	Inertia	Chi Square	Sig.	Accounted for	Cumulative	Standard deviation	Correlation 2
1	0.259	0.067			0.700	0.700	0.003	0.031
2	0.135	0.018			0.189	0.889	0.003	
3	0.074	0.006			0.058	0.947		
4	0.057	0.003			0.034	0.981		
5	0.037	0.001			0.014	0.995		
6	0.016	0.000			0.003	0.998		
7	0.011	0.000			0.001	0.999		
8	0.008	0.000			0.001	1.000		
9	0.003	0.000			0.000	1.000		
Total		0.096	9730.405	0.000[a]	1.000	1.000		

Table 3: Summary table-a 99 degrees of freedom {Confidence Row Points}

	Standard Deviation in Dimension		Correlation
Country[a]	1	2	1-2
Angola	0.018	0.024	0.144
Cote d'ivoire	0.081	0.096	0.115
Cameroun	0.132	0.203	0.042
Central African Republic	0.033	0.018	0.118
Chad	0.043	0.089	0.005
Congo PR	0.045	0.051	0.159
Congo DR	0.004	0.005	0.415
Equatorial Guinea	0.164	0.193	0.185
Gabon	0.096	0.131	0.009
Guinea	0.071	0.109	-0.123
South Sudan	0.025	0.032	0.167
Uganda	0.049	0.062	0.304

Table 4: Confidence column points

	Standard Deviation in Dimension		Correlation
Year	1	2	1-2
2013	0.017	0.020	0.227
2012	0.016	0.028	0.042
2011	0.019	0.022	0.030
2010	0.018	0.032	-0.002
2009	0.017	0.038	0.050
2008	0.020	0.031	-0.017
2007	0.018	0.035	0.016
2006	0.020	0.024	-0.039
2005	0.017	0.022	-0.027
2004	0.015	0.020	-0.041

The valid data represents more than 95% of the total reported cases. This is important because the homogeneity of variance across row and column variables must be met (Garson, 2008). Any attempt to analyze the missing values (in the case of data unavailability) yields inconclusive and misleading results (Doey and Kurta, 2011; Garson, 2008).

Data analysis: Table 1 to 4 and Fig. 1 to 5 display the results of the analysis.

The research also looked at the trend of HAT when the countries are arranged across to their geographical proximity. The numbers of reported cases of each country in each region are added together.

RESULTS AND DISCUSSION

Table 1 and 5 are the correspondence tables for the countries and regions respectively. It was observed from Table 2 that the model explains only 9.6% of the variance. The high Chi-square value of 9730.405 showed that there is high correspondence between the countries and the reported years. The model is significant at 0.05 with dimension 1 and 2 accounted for most of the total variance. And as such the Fig. 1 to 5 were drawn in 2D while the other dimensions were dropped. Table 3 to 9 are the Confidence Row Points and Confidence Column points which provide the standard deviation of row and column scores in each dimension. This is used to assess the precision of the estimates of points on their axes, just like the

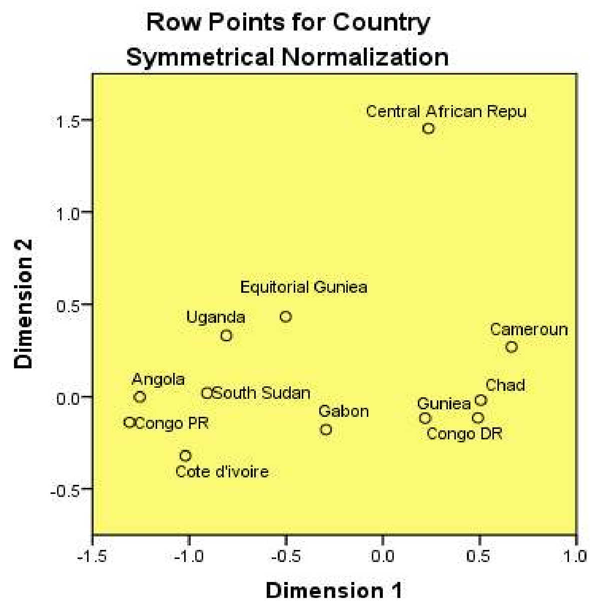

Fig. 1: Row points for the countries

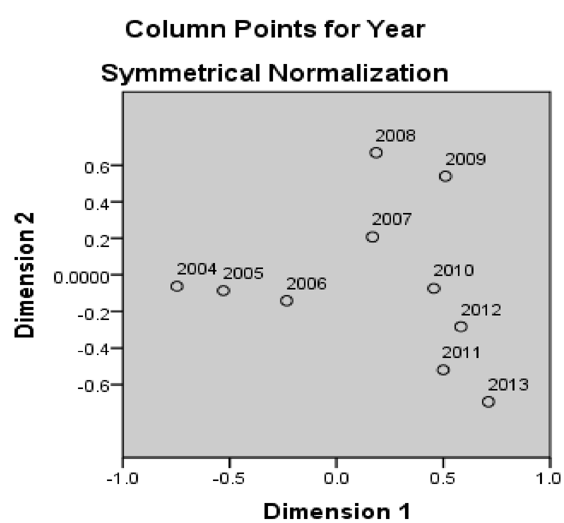

Fig. 2: Column points for the years

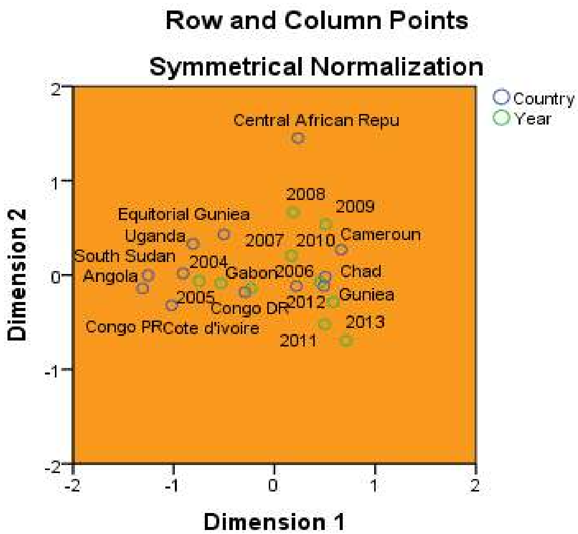

Fig. 3: Row and column points

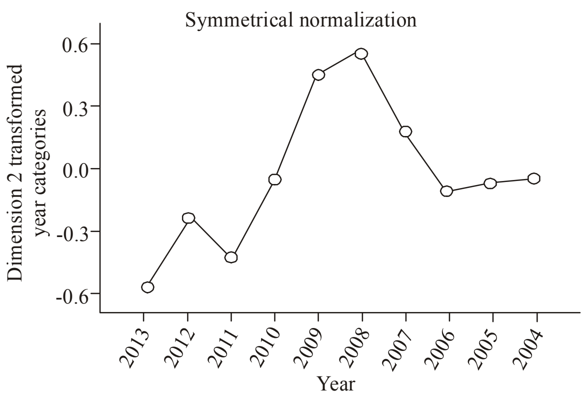

Fig. 4: The general trend

confidence intervals in statistical analyses (Doey and Kurta, 2011).

In comparing Fig. 1 with correspondence Table 1, it is observed that Central African Republic is located

Table 5: Countries and regions

Region	Countries
Central Africa	Angola Cameroun Central African Republic Chad Congo PR Congo DR Equatorial Guinea Gabon
East Africa	South Sudan Uganda
West Africa	Cote D' Voire Guinea

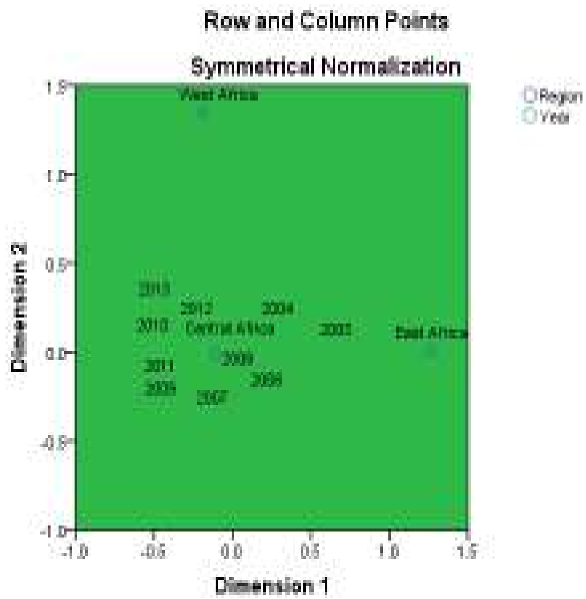

Fig. 5: Rows and column points

a far from others because of the dissimilarity of the data compared with others. The reported numbers of cases are on steady decline in Congo DR, Guinea, Chad and moderate in Cameroun. Analysis from Fig. 2 shows that the reported number of cases was on the decline from 2004-2006, increase from 2007-2009 and have been on a steady decline thereafter. The combination of Fig. 1 and 2 reflect in Fig. 3.

Table 5 is the summary table when the 12 countries are grouped under 3 regions. The model explains only

Table 6: The correspondence table

	Year[a]										
Region	2013	2012	2011	2010	2009	2008	2007	2006	2005	2004	Active margin
Central Africa	6016	6608	6251	6600	9126	9397	9994	10223	13331	14751	92297
East Africa	126	337	316	300	462	821	589	1007	2164	2120	8242
West Africa	85	79	67	76	87	104	82	69	136	169	954
Active margin	6227	7024	6634	6976	9675	10322	10665	11299	15631	17040	101493

Table 7: Summary

					Proportion of inertia		Confidence singular value	
Dimension	Singular value	Inertia	Chi Square	Sig.	Accounted for	Cumulative	Standard deviation	Correlation 2
1	0.141	0.020			0.985	0.985	0.003	0.008
2	0.017	0.000			0.015	1.000	0.003	
Total		0.020	2046.291	0.000[a]	1.000	1.000		

Table 8: Confidence row points

	Standard Deviation in Dimension		Correlation[a]
Region	1	2	1-2
Central Africa	0.002	0.002	-0.346
East Africa	0.014	0.010	-0.060
West Africa	0.090	0.117	-0.065

Table 9: Confidence column points[a]

	Standard Deviation in Dimension		Correlation
Year	1	2	1-2
2013	0.015	0.022	0.164
2012	0.006	0.009	0.205
2011	0.004	0.003	-0.070
2010	0.005	0.006	0.244
2009	0.007	0.007	-0.278
2008	0.003	0.004	0.021
2007	0.011	0.015	-0.125
2006	0.015	0.022	0.017
2005	0.007	0.005	-0.364
2004	0.008	0.009	-0.279

2% of the total variance because only 3 rows were used against the recommended four categories or more (Benzecri, 1992). Analysis from Fig. 5 indicates that most of the reported cases were recorded in the Central African countries of Angola, Cameroun, Central African Republic, Chad, Congo PR, Congo DR, Equatorial Guinea and Gabon.

CONCLUSION

Generally, the trend of HAT (T.b. gambiense) is decreasing fastest in West Africa but still very endemic in Central Africa. Correspondence analysis has proven useful in exploring the data and as such reveals the hidden features that ordinarily may be complex to analyze. Prediction using this model can be possible if the model explains a reasonable amount of variability of the categorical data. Efforts should be intensified in the various African countries for effective data management; the research should have included more countries if their data were available.

It can also be verified that the most endemic countries have low life expectancy, high mortality rate and high illiteracy, lack of portable and clean drinking water, inadequate health facilities and dearth of medical personnel.

ACKNOWLEDGEMENT

WHO and other related bodies: DFID, USAID, UNICEF and SWISSAID are appreciated for their efforts in tackling the spread of HAT in Africa. The authors wish to sincerely thank Covenant University for financial support.

REFERENCE

Benzecri, J.P., 1992. Correspondence Analysis Handbook. Marcel Decker, New York.

Brun, R., J. Blum, F. Chappuis and C. Burri, 2010. Human African trypanosomiasis. Lancet, 375(9709): 148-159.

Cattand, P., 2001. The epidemiology of human African trypanosomiasis: A complex multifactorial history. Med. Trop. (Mars.), 61(4-5): 313-322. (In French)

Clausen, S.E., 1998. Applied Correspondence Analysis: An Introduction. Sage Publications, Thousand Oaks, CA, US.

Cox, F.E., 2004. History of sleeping sickness (*African trypanosomiasis*). Infect. Dis. Clin. N. Am., 18(2): 231-245.

Doey, L. and J. Kurta, 2011. Correspondence analysis applied to psychological research. Tutorials Quant. Method. Psychol., 7(1): 5-14.

Franco, J.R., M. Paone, G. Cecchi, A. Diarra, J.A. Ruiz, R.C. Mattioli, J. Jannin and P.P. Simarro, 2013. The Atlas of human African trypanosomiasis. Progress status and prospects. Proceeding of the 32st ISCTRC Conference. Khartoum, Sep. 8th-12th.

Garson, D., 2008. Correspondence from Statistics. Topics in Multivariate Analysis. Retrieved from: http://faculty.chass.ncsu.Edu/garson/pa765/statnote.htm.

Hide, G., 1999. History of sleeping sickness in East Africa. Clin. Microbiol. Rev., 12(1): 112-125.

Higgs, N.T., 1991. Practical and innovative uses of correspondence analysis. J. Roy. Stat. Soc. D-Sta., 40(2): 183-194.

Hill, M.O., 1974. Correspondence analysis: A neglected multivariate method. J. Roy. Stat. Soc. C-App., 23(3): 340-354.

Kennedy, P.G., 2013. Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). Lancet Neurol., 12(2): 186-194.

Lundkvist, G.B., K. Kristensson and M. Bentivoglio, 2004. Why trypanosomes cause sleeping sickness. Physiology, 19: 198-206.

Moore, S., S. Shrestha, K.W. Tomlinson and H. Vuong, 2012. Predicting the effect of climate change on African trypanosomiasis: Integrating epidemiology with parasite and vector biology. J. R. Soc. Interface, 9(70): 817-830.

Panagiotakos, D.B. and C. Pitsavos, 2004. Interpretation of epidemiological data using multiple correspondence analysis and log-linear models. J. Data Sci., 2: 75-86.

Simarro, P.P., G. Cecchi, J.R. Franco, M. Paone, E.M. Fèvre, A. Diarra, J.A.R. Postigo, R.C. Mattioli and J.G. Jannin, 2011. Risk for human African trypanosomiasis, Central Africa, 2000-2009. Emerg. Inf. Dis., 17(12): 2322-2324.

Sourial, N., C. Wolfson, B. Zhu, J. Quail, J. Fletcher, S. Karunananthan, K. Bandeen-Roche, F. Béland and H. Bergman, 2010. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. J. Clin. Epidemiol., 63(6): 638-646.

Tongué, L.K., M.P. M'eyi, R.G. Kamkuimo, D. Kaba, J.F. Louis and R. Mimpfoundi, 2011. Transmission of human African trypanosomiasis in the Komo-Mondah focus, Gabon. Pan Afr. Med. J., 8: 36.

Uganda, 2008. Sleeping Sickness Reaching Alarming Levels: New Vision.

WHO, 2014. Media Center, Fact Sheet on Trypanosomiasis Human African (Sleeping Sickness). World Health Organization, Washington.

Yelland, P.M., 2010. An introduction to correspondence analysis. Math. J., 12: 1-23.