INDEPTH Mortality Clustering Supplement

Spatio-temporal clustering of mortality in Butajira HDSS, Ethiopia, from 1987 to 2008

Peter Byass[1]*#, Mesganaw Fantahun[2], Anders Emmelin[1], Mitike Molla[2] and Yemane Berhane[2,3]

[1]Department of Public Health and Clinical Medicine, Umeå Centre for Global Health Research, Umeå University, Umeå, Sweden; [2]School of Public Health, Addis Ababa University, Addis Ababa, Ethiopia; [3]Addis Continental Institute of Public Health, Addis Ababa, Ethiopia

***Background*:** Mortality in a population may be clustered in space and time for a variety of reasons, including geography, socio-economics, environment and demographics. Analysing mortality clusters can therefore reveal important insights into patterns and risks of mortality in a particular setting.

***Objective and design*:** To investigate the extent of spatio-temporal clustering of mortality in the Butajira District, Ethiopia, from 1987 to 2008. The Health and Demographic Surveillance System (HDSS) dataset recorded 10,696 deaths among 951,842 person-years of observation, with each death located by household, in which population time at risk was also recorded. The surveyed population increased from 28,614 in 1987 to 62,322 in 2008, in an area approximately 25 km in diameter. Spatio-temporal clustering analyses were conducted for overall mortality and by specific age groups, grouping the population into a 0.01° latitude–longitude grid.

***Results*:** A number of significantly high- and low-mortality clusters were identified at various times and places. Butajira town was characterised by significantly low mortality throughout the period. A previously documented major mortality crisis in 1998–1999, largely resulting from malaria and diarrhoea, dominated the clustering analysis. Other local high-mortality clusters, appreciably attributable to meningitis, malaria and diarrhoea, occurred in the earlier part of the period. In the later years, a more homogeneous distribution of mortality at lower rates was observed.

***Conclusions*:** Mortality was by no means randomly distributed in this community during the period of observation. The clustering analyses revealed a clear epidemiological transition, away from localised infectious epidemics, over a generation.

Keywords: *mortality clustering*; *Ethiopia*; *malaria*; *meningitis*; *diarrhoea*; *demographic surveillance*; *epidemiological transition*

Access the supplementary material to this article: animated graphic – sequence.ppsx – displaying the clusters (viewed over time) shown in Table 1 (see Supplementary files under Reading Tools online).

Received: 23 April 2010; Revised: 24 June 2010; Accepted: 26 June 2010; Published: 30 August 2010

Many factors can lead to clustering of mortality in space and time. These include geographical influences [topological variation as well as transient events and natural disasters (1)], social determinants [housing inequality, income gradients (2)] and anthropogenic factors [industrial contamination, civil unrest (3)], as well as access to quality health care. Different sectors within a population (e.g. age groups) may also have different susceptibilities to factors that influence mortality clustering.

Analysis, interpretation and presentation of mortality clustering is, however, not entirely straightforward, and the process may or may not add significant understanding to analysis using other epidemiological tools (4). Spatio-temporal clustering methods can be applied to many disease processes, with disease incidence or mortality as endpoints. Particular diseases can be studied,

---

[#]The Deputy Editor Peter Byass, has not had any part in the review and decision process for this paper.

Global Health Action 2010. © 2010 Peter Byass et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 3.0 Unported License (http://creativecommons.org/licenses/by-nc/3.0/), permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Global Health Action Supplement 1, 2010. DOI: 10.3402/gha.v3i0.5244

such as malaria, where geography can reasonably be hypothesised to affect transmission (5). The reasons for assessing clustering may simply be part of overall epidemiological descriptions or may relate to specific hypotheses such as the emergence of high-risk areas consequent on interventions, e.g. malaria transmission around newly constructed irrigation dams (6).

Because spatio-temporal mortality clustering requires detailed information not only about deaths, but also their time, location and the population at risk in that location, it can only be carried out in relatively well-documented populations that are followed longitudinally. In Africa and Asia, this effectively restricts the technique to defined research locations, e.g. the Health and Demographic Surveillance Systems (HDSS) of the INDEPTH Network (7) or other well-documented populations. HDSSs in Burkina Faso (8) and Ghana (9) have previously reported this approach in relation to childhood mortality. In this study, we are looking at clustering in all-cause mortality in a defined population over an extended period of time.

In the Butajira District of Ethiopia, we previously demonstrated highly clustered mortality associated with particular weather events and episodes of infectious disease occurring over a short period of time during 1998–1999 (1). Naturally, this acute event forms an integral part of this wider consideration of the Butajira District over a much longer time period and including a semi-urban area, but would nevertheless be expected to feature in the spatio-temporal cluster results. We have also previously shown that the population structure changed considerably over this period, reflecting changing epidemic patterns and generally falling mortality rates (10). However, the aim of this paper is to take a longer-term perspective on how mortality patterns in this population might be clustered in time and space, and to see how the patterns of clustering vary between different age groups.

Methods

The Butajira HDSS was established in 1986 and commenced regular surveillance from the start of 1987. The HDSS was established in 10 local communities (*kebeles*) of the Meskan and Mareko District, located with the Southern Nations, Nationalities and People's Region (SNNPR) to the south-west of Addis Ababa (Fig. 1). The communities were randomly selected using probability proportional to size sampling at the outset. One of the communities comprised a sector of Butajira town (8.12°N, 38.38°E, approximately 2,050 m above sea level), which is the district headquarters and is characterised by a more urban, though still basic, lifestyle. The other nine communities were rural in character, with those to the west typified by a higher altitude farming lifestyle with permanent cropping. The area to the west of 38.3°E is a steep mountainous environment with habitation ranging to 2,700 m above sea level. To the east, the altitude is lower (to around 1,800 m above sea level) and those communities generally practice more arable agriculture on flatter land. In rural Ethiopia, rural communities are defined as areas of land, containing a mixture of scattered housing and agricultural land, rather than the typically more concentrated village settlements found in many other African countries. A health centre was located in Butajira town from the outset of the surveillance, and this was supplemented in 2003 by the opening of a newly built district hospital at the northern edge of the town.

A household census was undertaken at the start of the surveillance and this has since been regularly updated on a continuous basis. The initial population surveyed was 28,614, rising to 62,322 by the end of 2008, in an area approximately 25 km in diameter. With the advent of GPS technology, the geo-location of every household was recorded. Details of residents within each household were also updated during regular household visits (initially monthly, later quarterly). As part of these visits, any deaths that had occurred in the household were also recorded. Although verbal autopsies were not performed for most of the period in question, all deaths were followed up with a question to the household as to the perceived cause of death, in relatively few basic categories. These included common diseases relatively familiar to the population, such as malaria, diarrhoea, meningitis, etc. These responses should not, however, be considered to be definitive cause of death data.

Taking local patterns of habitation into account, we considered using households or communities as potential units of spatial analysis. Although we based previous analyses on households (1), they are generally too small for stratification by age group. Communities as defined in this locality, however, can vary from about 1 to 10 km across, and as the entire surveillance population was located in just 10 communities, a spatial analysis at that level could lead to a loss of important detail. Therefore, for these analyses we divided the overall HDSS area (approximately 8.00–8.20°N, 38.25–38.55°E) into a 0.01° grid (which is roughly equivalent to 1 km squares) and used this as a framework for spatial analyses. This gave 140 populated grid squares. The range of person-years observed within the grid squares was 50–109,000 (median 4,014) and deaths 0–660 (median 49).

Spatio-temporal clustering analyses were carried out using SatScan™ v8 (11, 12) with a discrete Poisson model and 999 replications for the entire population and by age group (under 5 years, 5–14 years, 15–49 years, 50–64 years and 65+ years). By default, SatScan™ allows spatial clusters of up to 50% of population at risk and temporal clusters of up to 50% of the study period. However, pilot analyses showed that these defaults lumped interesting variations in mortality into a cluster covering a large proportion of the area and a large part of the study period.

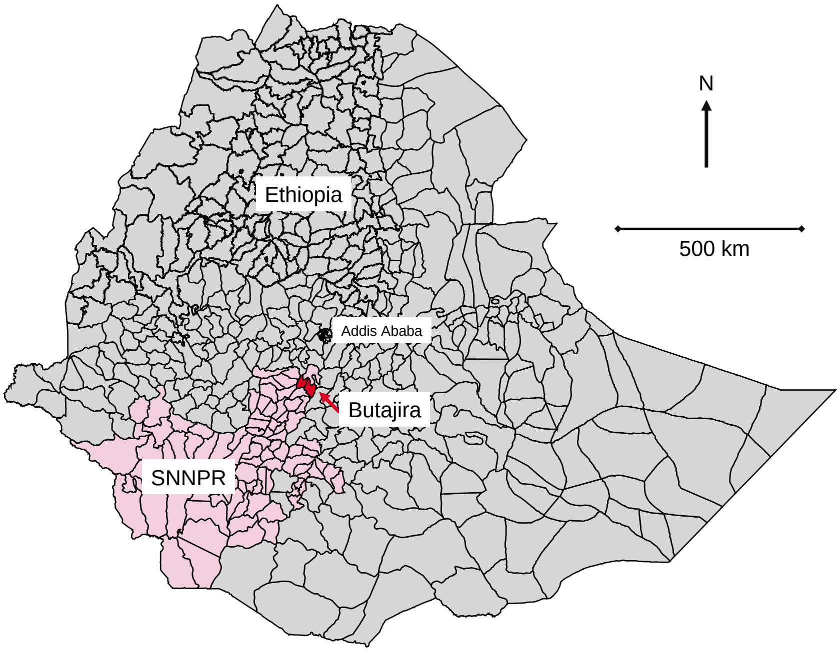

Fig. 1. Map of Ethiopia showing the location of Butajira, within the Southern Nations, Nationalities and People's Region (SNNPR).

Therefore, we adjusted the maximum allowed size of spatial clusters to a radius of 4 km, corresponding to an area of approximately 50 km^2, and the maximum period for a time cluster to 3 years, so that details could emerge through a range of smaller clusters. Mortality clusters as identified by SatScan™ are by definition centred on one of the grid squares used as spatial units of observation. Temporal clusters covering the entire area and spatial clusters covering the entire period were allowed and clusters could either relate to high- or low-mortality entities, which is important for revealing decreases in mortality over time and possibly advantaged locations (such as those close to health facilities). Mapping used Diva-GIS open-source software (http://www.diva-gis.org).

Ethical approval for the overall Butajira surveillance was received from the National Ethical Clearance Committee, and individual informed consent was obtained from participants.

Results

During the period 1987–2008, a total of 10,696 deaths were recorded in 951,842 person-years of community observation in the Butajira District. This corresponded to an overall crude mortality rate of 11.2 per 1,000 person-years. Mortality rates within age groups were 35.3 per 1,000 person-years for those aged under 5 years, 5.0 for 5–14 years, 4.6 for 15–49 years, 15.6 for 50–64 years and 39.4 for 65 years and over. The spatial mortality pattern for the entire 22-year period is shown geographically on the 0.01° grid in Fig. 2, together with the location of Butajira hospital and the main roads in the area. This figure simply shows the distribution of crude mortality by grid squares, not involving any analysis of clustering.

Spatio-temporal clustering of overall mortality was analysed including age group as a covariate, revealing a number of statistically significant clusters 1.1–1.6, as detailed in the upper section of Table 1. Because of the dominant effect of mortality clustering across the whole area during 1998–1999, a second model was run that excluded data from those two years. These results, with significant clusters 2.1–2.6, are shown in the lower section of Table 1. The clusters shown in Table 1 can be viewed over time as an animated graphic (see Supplementary files under Reading Tools online – sequence.ppsx). The predominance of malaria and diarrhoea as stated causes of death for the very high mortality rates recorded in 1998–1999 has been detailed previously (1). For the high-mortality clusters shown in Table 1, we looked at the perceived causes of death in the cluster compared with overall mortality, and these findings are presented in Table 2.

Table 3 shows the statistically significant clusters of mortality identified within the age groups under 5 years, 5–14 years, 15–49 years, 50–64 years and over 65 years. The majority of the age-specific high-mortality clusters were associated with the 1998–1999 mortality crisis, though some constituted part of other clusters. For the under-5 age group, cluster 6 was part of overall cluster 2.6 with a high rate of meningitis, and cluster 7 was part of overall cluster 2.4, dominated by malaria. That same malaria cluster is also reflected in cluster 5 among the 5–14 age group. Cluster 6 for the 15–49 year age group

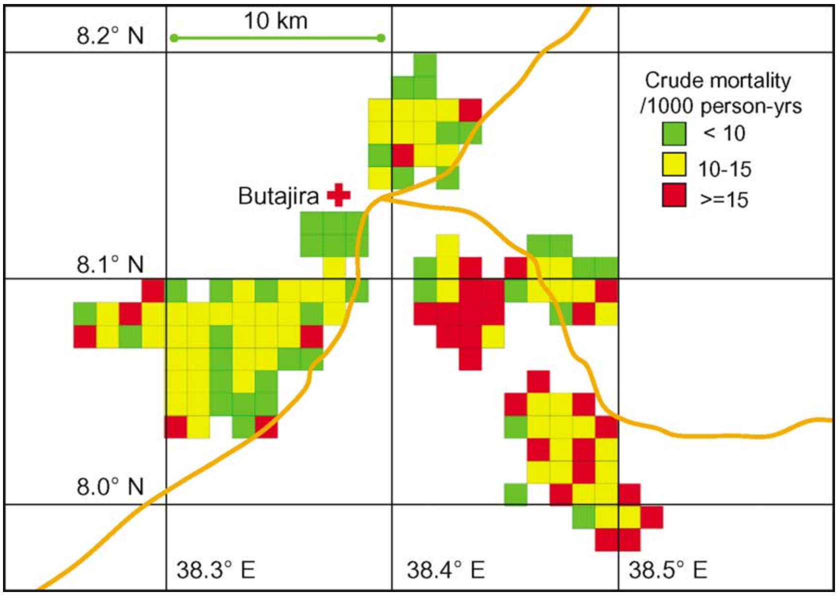

Fig. 2. Mapped distribution of crude mortality rates in Butajira District, as divided into a 0.01° grid of latitude and longitude. The overall data consist of 10,696 deaths in 951,842 person-years over the period 1987–2008.

was part of overall cluster 2.3, dominated by meningitis. All the age groups, apart from the under-5s, showed some low-mortality clusters in the period from 2000 onwards.

Discussion

The overall mapping of mortality, as shown in Fig. 2, clearly indicates a spatially heterogeneous pattern, and our previous analyses of the 1998–1999 crisis provide an evident example of strong temporal heterogeneity. Both of these concepts are reflected in the clustering analysis of the overall mortality data as shown in Table 1. Three important points emerge: cluster 1.1 reflects the overwhelming overall effect of the 1998–1999 crisis, with a relative risk of 2.40 compared with the overall time period; cluster 1.2 shows the considerable mortality advantage of living in Butajira town, with a relative risk of 0.5 throughout the period studied when compared with the overall area; and three of the remaining four clusters in the overall model reflect especially intense foci of mortality during 1998–1999, on top of the overall increased risk for that period.

Since this overall analysis was evidently dominated by the effect of the 1998–1999 crisis, the lower section of Table 1 reflects the overall clustering analysis with the omission of the 1998–1999 data. Without the overriding effect of the 1998–1999 crisis, a number of other

Table 1. Spatio-temporal clustering analysis of mortality in Butajira, 1987–2008, covering 10,696 deaths in 951,842 person-years

Cluster	Location (cluster centre)	Radius (km)	Period	Cluster type	Deaths obs/exp	Relative risk	*p*-Value
Entire data set (1987–2008)							
1.1	Entire	–	1998–1999	High	2,379/1,005	2.40	0.001
1.2	8.125°N, 38.365°E	2.46	Entire	Low	1,179/2,135	0.50	0.001
1.3	8.085°N, 38.435°E	3.51	1998–1999	High	462/122	3.93	0.001
1.4	8.015°N, 38.475°E	3.99	1998–1999	High	424/156	2.78	0.001
1.5	8.065°N, 38.325°E	3.99	1998–1999	High	566/248	2.35	0.001
1.6	8.155°N, 38.415°E	3.13	1988–1989	High	153/49	3.14	0.001
Data from 1987 to 1997 and 2000 to 2008 only							
2.1	Entire	–	2006–2008	Low	760/1,759	0.38	0.001
2.2	8.125°N, 38.385°E	2.48	Entire	Low	1,030/1,861	0.49	0.001
2.3	8.095°N, 38.455°E	3.48	1988–1990	High	498/179	2.89	0.001
2.4	8.165°N, 38.415°E	2.46	1991	High	155/39	4.04	0.001
2.5	8.075°N, 38.335°E	3.99	2005–2007	Low	158/383	0.40	0.001
2.6	8.025°N, 38.485°E	3.99	1988–1989	High	235/88	2.72	0.001

features emerge. Cluster 2.1 shows a substantial reduction in mortality during the last 3 years of the period observed, with a relative risk of 0.38; the beneficial effect in the urban area is still reflected throughout the period in cluster 2.2; clusters 2.3, 2.4 and 2.6 reflect other high-mortality foci and cluster 2.5 reflects reduced mortality in the later part of the period in the highland area.

It is interesting to note that all the high-mortality clusters occurred before 2000 and, as can be seen from Table 2, were all associated with high rates of particular infectious causes. The period from 2000 onwards, by contrast, was characterised by a lack of high-mortality clusters and reducing overall mortality, and this was reflected in all age groups with the exception of the under-5s.

Table 2. Reported perceived causes of 10,705 deaths in Butajira from 1987 to 2008, overall and within significant high-mortality clusters

Cluster	Meningitis, *n* (%)	Malaria, *n* (%)	Diarrhoea and malnutrition, *n* (%)	Other causes, *n* (%)
Overall	97 (0.9)	1,440 (13.5)	1,949 (18.2)	7,219 (67.4)
1.3	2 (0.4)	109 (23.6)	163 (35.3)	188 (40.7)
1.4	0	69 (16.3)	92 (21.7)	263 (62.0)
1.5	0	94 (16.6)	193 (34.1)	279 (49.3)
1.6	11 (6.8)	17 (10.6)	12 (7.5)	121 (75.2)
2.3	33 (6.6)	67 (13.4)	44 (8.8)	356 (71.2)
2.4	0	41 (26.4)	27 (17.4)	87 (56.1)
2.6	22 (9.3)	46 (19.5)	29 (12.3)	139 (58.9)

Table 3. Spatio-temporal clustering analysis by age group of mortality in Butajira, 1987–2008, covering a total of 10,696 deaths in 951,842 person-years

Age group (deaths/1,000 person-years)	Cluster	Location (cluster centre)	Radius (km)	Period	Cluster type	Deaths obs/exp	Relative risk	*p*-Value
Under 5 years (5,040/142,889)	1	Entire	–	1998–1999	High	999/499	2.25	0.001
	2	8.125°N, 38.385°E	2.48	Entire	Low	1,211/449	0.43	0.001
	3	8.085°N, 38.435°E	2.22	1998–1999	High	674/220	4.63	0.001
	4	8.015°N, 38.485°E	3.99	1998–1999	High	227/83	2.81	0.001
	5	8.065°N, 38.315°E	3.99	1998–1999	High	224/108	2.12	0.001
	6	8.105°N, 38.485°E	2.47	1988–1989	High	105/35	3.04	0.001
	7	8.155°N, 38.425°E	2.48	1991	High	60/17	3.46	0.001
5–14 years (1,287/255,923)	1	Entire	–	2005–2007	Low	34/214	0.14	0.001
	2	8.115°N, 38.365°E	1.56	Entire	Low	90/270	0.28	0.001
	3	8.095°N, 38.455°E	3.99	1997–1999	High	98/30	3.40	0.001
	4	8.015°N, 38.475°E	3.99	1998–1999	High	68/17	4.23	0.001
	5	8.175°N, 38.425°E	3.30	1989–1991	High	61/16	3.85	0.001
	6	8.065°N, 38.345°E	3.99	2004–2006	Low	4/46	0.08	0.001
15–49 years (2,095/458,023)	1	Entire	–	2006–2008	Low	128/434	0.25	0.001
	2	8.145°N, 38.395°E	3.99	2006–2008	Low	36/165	0.20	0.001
	3	8.095°N, 38.425°E	3.48	1997–1999	High	107/29	3.78	0.001
	4	8.065°N, 38.305°E	3.99	1998–1999	High	112/45	2.56	0.001
	5	8.025°N, 38.475°E	3.98	2006–2008	Low	17/61	0.27	0.001
	6	8.085°N, 38.485°E	3.13	1988–1990	High	50/17	2.96	0.001
50–64 years (962/61,724)	1	Entire	–	1999	High	157/45	4.02	0.001
	2	8.085°N, 38.445°E	3.98	1999	High	48/8	6.32	0.001
	3	8.065°N, 38.345°E	3.99	1999	High	49/11	4.59	0.001
	4	8.095°N, 38.365°E	2.47	1998–1999	High	27/7	3.69	0.003
	5	8.025°N, 38.375°E	2.47	1998–2000	High	38/14	2.73	0.006
	6	8.145°N, 38.395°E	3.99	2006–2008	Low	18/48	0.37	0.012
65+ years (1,312/33,282)	1	Entire	–	1997–1999	High	313/166	2.15	0.001
	2	8.065°N, 38.335°E	3.99	1999	High	56/15	3.78	0.001
	3	8.175°N, 38.415°E	2.46	1997–1999	High	49/19	2.67	0.001
	4	8.095°N, 38.425°E	1.56	2003–2005	Low	0/14	0	0.006
	5	8.035°N, 38.495°E	3.51	1998–2000	High	43/17	2.54	0.008

From a methodological perspective, dividing the population into a 0.01° grid was an effective way of handling the clustering analyses and also facilitated mapping to clearly visualise mortality patterns and clusters. It also helped to avoid misinterpreting village effects as clustering, in the hypothetical case that particular villages might have their own characteristic mortality patterns. Fig. 2 clearly shows within-village heterogeneity in mortality rates, which would not be captured in an analysis by village. Limiting the radius and time span of clusters was methodologically important for distinguishing discrete local phenomena such as the disease epidemics characterised in Table 2.

These spatio-temporal analyses of mortality in Butajira do not reveal major new findings that have not been previously seen using other approaches, but do show that mortality was far from randomly distributed over space and time. However, they do facilitate the depiction of mortality in geographic terms, and show the clear trend from intense, local mortality hot spots (associated with epidemic outbreaks of particular diseases) in the earlier years, towards lower and more homogeneous mortality in the later years. The significant mortality advantage of living in Butajira town also emerges very clearly, though attributing causes to this is not so easy. Overall, this is a graphic depiction of the epidemiological transition that this population has experienced over a single generation.

Acknowledgements

We particularly wish to thank the Butajira Demographic Surveillance System population for their continued participation over an entire generation, and the large number of field staff who have contributed to this long-term data collection. SaTScan™ is a trademark of Martin Kulldorff. The SaTScan™ software was developed under the joint auspices of (a) Martin Kulldorff, (b) the National Cancer Institute and (c) Farzad Mostashari of the New York City Department of Health and Mental Hygiene.

Conflict of interest and funding

Financial support from Sida/SAREC, Stockholm, during this long-term programme is gratefully acknowledged. The Umeå Centre for Global Health Research is supported by FAS, the Swedish Council for Working Life and Social Research (Grant No. 2006-1512).

References

1. Emmelin A, Fantahun M, Berhane Y, Wall S, Byass P. Vulnerability to episodes of extreme weather: Butajira, Ethiopia, 1998–1999. Glob Health Action 2008; 2. DOI: 10.3402/gha.v2i0.1829.
2. Chen JT, Rehkopf DH, Waterman PD, Subramanian SV, Coull BA, Cohen B, et al. Mapping and measuring social disparities in premature mortality: the impact of census tract poverty within and across Boston neighborhoods, 1999–2001. J Urban Health 2006; 83: 1063–84.
3. Kurumatani N, Kumagai S. Mapping the risk of mesothelioma due to neighborhood asbestos exposure. Am J Respir Crit Care Med 2008; 178: 624–9.
4. Olsen SF, Martuzzi M, Elliott P. Cluster analysis and disease mapping – why, when, and how? A step by step guide. BMJ 1996; 313: 863–6.
5. Yeshiwondim AK, Gopal S, Hailemariam AT, Dengela DO, Patel HP. Spatial analysis of malaria incidence at the village level in areas with unstable transmission in Ethiopia. Int J Health Geogr 2009; 8: 5.
6. Ghebreyesus TA, Byass P, Witten KH, Getachew A, Haile M, Yohannes M, et al. Appropriate tools and methods for tropical microepidemiology: a case-study of malaria clustering in Ethiopia. Ethiop J Health Dev 2003; 17: 1–8.
7. Evans T, AbouZahr C. INDEPTH@10: celebrate the past and illuminate the future. Glob Health Action 2008; 2. DOI: 10.3402/gha.v1i0.1899.
8. Sankoh OA, Yazoumé Y, Sauerborn R, Müller O, Becher H. Clustering of childhood mortality in rural Burkina Faso. Int J Epidemiol 2001; 30: 485–92.
9. Awini E, Mattah P, Sankoh O, Gyapong M. Spatial variations in childhood mortalities at the Dodowa Health and Demographic Surveillance System site of the INDEPTH Network in Ghana. Trop Med Int Health 2010; 15: 520–8.
10. Berhane Y, Wall S, Fantahun M, Emmelin A, Mekonnen W, Hogberg U, et al. A rural Ethiopian population undergoing epidemiological transition over a generation: Butajira from 1987 to 2004. Scand J Public Health 2008; 36: 436–41.
11. Kulldorff M and Information Management Services, Inc. SaTScan™ v8.0: software for the spatial and space-time scan statistics. Available from: http://www.satscan.org/ [cited 25 March 2010].
12. Kulldorff M. A spatial scan statistic. Commun Stat Theory Methods 1997; 26: 1481–96.

***Peter Byass**
Department of Public Health and Clinical Medicine
Umeå Centre for Global Health Research
Umeå University
SE-90185 Umeå, Sweden
Tel: +46 90 785 3345
Email: peter.byass@epiph.umu.se