**A Worldwide Investigation of Tuberculosis Epidemics**

**Christine S.M. Currie and Kathryn A. Hoad**

**This is a pre-review copy of the article, which can only be used for non-commercial and personal use.**

**Abstract:** We analyse the tuberculosis (TB) epidemics of 211 countries with a view to proposing more efficient and targeted TB control strategies. Countries are classified by how their TB case notification rates have evolved over time and the age distribution of those suffering from active TB disease in 2008. Further analysis of key statistics associated with each of the countries shows the impact of different indicators. As expected, HIV is a key driver of TB epidemics and affects their age-distribution and their scale. The level of development of a country and its wealth also vary with the shape and scale of a country’s TB epidemic. Immigration has an influence on the shape of TB epidemics, which is particularly pronounced in highly developed countries with low levels of TB disease in the native population. We conclude by proposing how the TB control programme in each country analysed should prioritise its efforts.

**Key words:** health service; statistics; resource management; tuberculosis

# Introduction

Tuberculosis (TB) epidemics have typically been classified by their severity, with the World Health Organization (WHO) providing a list of the top 22 high burden countries in terms of the absolute number of TB cases they report [1]. These high burden countries have received particular attention in recent years as they account for 81% of the TB cases worldwide [1]. However, this classification is not necessarily the most appropriate for deciding on the optimal resource allocation, most importantly because it ignores the population size of the country and so is heavily biased towards countries with large populations such as India and China. In fact only 6 of the WHO’s 22 high burden countries appear in the top 22 countries for case notification rates, i.e. number of TB cases reported per 100,000 population in 2008 [2].

TB is one of the leading causes of death in the world today [3]. The importance of TB as a public health problem is highlighted by its inclusion in the Millennium Development Goals (MDGs), which require that the incidence of TB should be falling by 2015 (MDG Target 6.c, http://www.un.org/millenniumgoals/). The Stop TB Partnership has set two further goals: that the mortality and prevalence rates of disease should be half of their 1990 levels by 2015 ([1], page 1). The first and second goals appear reachable globally if the current rates of decline in incidence and mortality are sustained ([1], page 1). However, mortality is currently not falling quickly enough in the African Region to reach this target regionally, largely a result of the severe HIV epidemic in Sub-Saharan Africa. The world is not on track to, nor does it appear currently possible to, reach the required reduction in prevalence ([1], page 1) by 2015. There is therefore still a need for sustained improvements in TB control to keep reducing the burden of this disease.

Efficient resource allocation for TB is difficult, partly as a result of the number of funding sources: four major donor institutions (the World Bank, the US Government, the Bill and Melinda Gates Foundation and the Global Fund for HIV/AIDS, TB and Malaria); and many more smaller donors, which are estimated to make up approximately two thirds of the total funding for health [4]. There have been relatively few modelling studies that consider resource allocation just for TB, although TB is one of the diseases Flessa [5] considers in the allocation of health care resources in a developing country setting. Zaric and Brandeau [6] set up a general method for resource allocation for infectious diseases, using a compartmental model to describe the disease process. The method optimises resource allocation between different populations, time periods and interventions, suggesting it could be adapted to finding the optimal allocation of resources for TB worldwide. However, the number of countries, and so populations, to consider may make the problem of worldwide resource allocation prohibitively large. An alternative to optimal allocation is to consider equity-based allocation, e.g. based on the TB case-notification rate in each country, and the analysis described here will be most useful if this approach is used. A study by Lasry et al. [7] based on the allocation of funds for HIV prevention, suggests that using an equity-based approach for the high-level allocation of funds, such as between countries, can give reasonable results, especially if an optimal allocation is then used at the lower level. For a problem of this size, equity-based allocation may be the most feasible.

Each year the World Health Organisation compiles a report on the state of global TB control which presents incidence, notification, case detection and mortality rates as well as other treatment outcome data for every monitored country. Despite this wealth of data now available there have been surprisingly few recent studies that utilise these data to analyse TB epidemics on a global scale [8], [9]. We here aim to provide a classification of countries that will enable coherent strategies for TB control to be rolled out across clusters on a global scale. We classify TB epidemics by their history – TB case notification rates from 1980 to 2008 – and by their shape – the distribution of TB cases across age groups, as recorded in 2008. We anticipate that the classification of countries by their history will provide a guide to the countries most in need of strengthening their TB programmes, whereas the classification of countries by the age distribution of reported TB cases should be particularly useful in reducing the prevalence and mortality from TB, in line with the Millennium Development Goals (MDGs). The age distribution of a country’s TB case notifications gives a clear indication of the age groups in which TB disease is most prevalent; therefore, targeting of case detection and cure activities at these age groups should be the most efficient means of reducing the numbers suffering and dying from TB disease. In recent years, the DOTS strategy (<http://www.who.int/tb/dots/en/>) has been the focus of TB control at the World Health Organization, which aims to provide a good underlying service for TB case detection and cure. The work described in this paper considers the next level of TB control: more effective allocation of resources to achieve higher case detection rates for TB disease.

We use cluster analysis to classify countries and three cluster analyses are described in Section 2 using as the input data (1) TB case notification rates from 1980 to 2008 (Table A-3, [2]); (2) the age distribution of TB case notifications in 2008 for males; and (3) the age distribution of TB case notifications in 2008 for females (Table A-9, [2]). In section 3 we analyse and interpret the resulting clusters of countries with regard to a number of indicators, e.g. HIV prevalence and human development index (HDI), which could affect TB incidence. Section 4 contains a final discussion and policy recommendations.

# Classification of Epidemics

In cluster analysis observations are classified into groups or clusters, where all of the observations in a group correspond to data points that are close together, based on the distance measure being used. Therefore, using historical case notification rates as our input to the cluster analysis allows us to group together countries with similar historical trends in their TB incidence; while considering the age distributions of TB should enable us to group together countries that have similar issues with regard to combating TB. (Detailed information on how to carry out Cluster analysis can be found in most good multivariate statistics texts, e.g. [10]. Analysis was carried out using the SPSS statistical package www.spss.com/uk/).

## Evolution of TB Case Notification Rate

We wish to classify the time series of TB case notification rates for each country in the world so that countries with similar trends in the TB case notification rate and similar scales for TB case notifications are grouped together. Given the non-linearity of the trends for many of the countries and the variability in the data, cluster analysis appears to be the most appropriate method for classification. Hierarchical clustering using the squared Euclidean distance with Ward linkage [11] is utilised to provide an initial clustering of the countries which is then used as the input for a K-Means (non-hierarchical) cluster analysis, which provides more coherent groupings for countries. The analysis is performed on 188 countries (see appendix A.2 for a list) with 29 input parameters, namely TB case notification rates in 1980 to 2008 inclusive.

The results of the cluster analysis suggest that there should be 5 clusters. The distribution of countries across these clusters can be seen in Figure 1 and a list of the countries placed in each of the clusters is given in the appendix (Table A.2). Figures 2 and 3 show how TB case notification rates vary over time for each of the clusters.

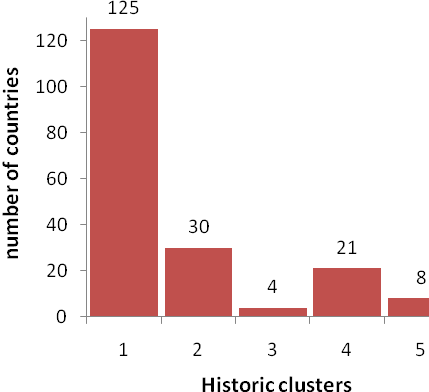


Figure 1: Distribution of countries across the 5 historic clusters.

(a) 

(b) 

(c) 

(d) 

(e) 

Figure 2: Case notification rates for the 5 historic clusters showing the medians (black), 5th, 10th, 90th and 95th percentiles (grey) and minimum and maximum rates (dashed). Cluster 3 contains only 4 countries and so the individual time lines are shown in graph c.

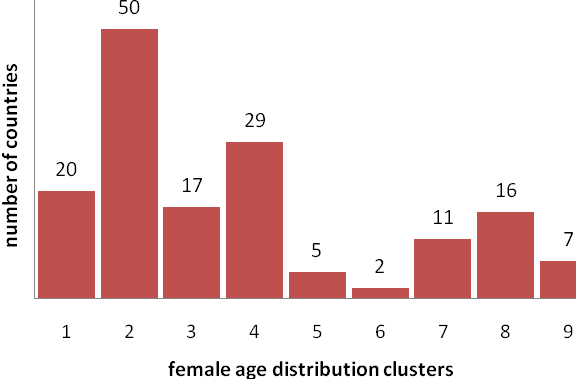


Figure 3: Plot of the variable centroids and the grand centroid for the K-Means clusters for the historic TB data.

Figure 3, which shows the centroids for the K-Means variables reiterates the plots of Figure 2 and gives a good indication of the different shapes and scales of the clusters. Clusters 1, 2 and 4 have relatively low TB levels that are static (cluster 1), increasing (cluster 2) and decreasing (cluster 4). Cluster 3 contains outliers with relatively high and increasing TB levels. Cluster 5 has severe, increasing TB levels and has particularly large distances between its centroid value and the other clusters’ centroids, underlining the difference in case notification rates for this cluster as compared with the others.

## Age-Specific TB Case Notifications

We carry out a cluster analysis for males and females separately, using as inputs the proportion of all smear-positive TB cases for the gender under consideration in each of seven age groups (0-14, 15-24, 35-44, 45-54, 55-64, 65+) in 2008. The distribution of countries across the various clusters can be seen in Figures 4a and b.



(b)

(a)

Figure 4a and b: distribution of countries across male (a) and female (b) age clusters

As a way of checking these clustering results, we carried out a K-Means analysis with these clusters as a starting point. The majority of clusters remained the same, but there was a noticeable mixing of countries between the clusters with peaks in the 15-34 age ranges. This could be due to the influence of values in the older age ranges on the automated clustering process.

The countries and their cluster memberships for the male and female age distributions are listed in the appendix (Table A.2). Figure 5 shows how active TB case proportions vary over age band for each of the male clusters, while figure 6 shows the same information for the female clusters. The male clusters show clear differences in the shape of the TB epidemic across age, dominated by the age band in which the peak proportion of active TB cases occurs. The female clusters also show clear differences in the shape of their TB epidemics across age and the clustering is strongly influenced by both the age band where the peak proportion of active TB cases occurs and the trend in the older age ranges.



(a) Male age cluster 1, (b) Male age cluster 2,

(c) Male age cluster 3, (d) Male age cluster 4,

(e) Male age cluster 5, (f) Male age cluster 6.

Figure 5: The median, percentiles, minimum and maximum of the proportion of smear positive TB cases in each age band for each male age cluster.



Figure 6: The median, percentiles, minimum and maximum of the proportion of smear positive TB cases in each age band for each female age cluster.

# Cluster interpretation and profiles

In this section we interpret the epidemic clusters created in the previous section and consider the impact on the TB epidemic of various key statistics (indicators) associated with each of the countries. We use principal component analysis (PCA) to investigate the differences between the historic and age clusters in relation to key ‘health and wealth’ indicators for each country. We then investigate further the impact of immigration (sec 3.3) on TB morbidity in more detail.

## Interpretation of historic and age-distributed clusters

We first consider the relationship between our historic TB clusters and the countries’ ‘wealth’ and HIV burden. Then, using the results of the clustering of countries by their age profiles, investigate the effect of HIV, ‘wealth’ and immigration on the age-profile of a TB epidemic.

A person infected with HIV has a greater risk of reactivation of a TB infection and will progress more quickly through TB disease [12]. Therefore, it is no surprise that the countries with the most severe HIV epidemics have seen dramatic increases in their TB case notification rates. Figure 7 shows this for Botswana, which has particularly high levels of HIV and TB. We would therefore expect HIV to vary between the historic TB clusters as these take into account the scale of the TB epidemic. Accordingly, there is indeed a clear distinction between the historic TB cluster 5, which includes only the countries in the south of Africa which have the most severe TB and HIV epidemics in the world, and the remaining historic TB clusters (see figure 8). Cluster 4 is made up of middle to low income countries with fairly low HIV prevalence whereas cluster 3 seems to consist of outliers with no HIV problems. Cluster 2 includes fairly poor countries with moderately high HIV prevalence and in terms of geography includes many of the remaining countries in Sub-Saharan Africa, as well as some of the poorer countries in Asia. Cluster 1 is a large cluster that contains countries that are, on the whole, fairly rich and well-developed with low levels of TB. However, there are exceptions: most of the West African countries are included in this cluster (Rwanda and Burundi being the only two clustered in a different group), and these countries are extremely poor.



Figure 7: TB case notifications per 100,000 (solid line) and HIV prevalence (dashed line) for Botswana.



Figure 8: HIV prevalence (2007) across the historic TB clusters.

The age-distributed clusters show some more interesting patterns. Countries with high HIV prevalence rates have a similar age distribution of TB cases. All of the countries with HIV prevalence in 2007 greater than 2.5% are included in the male age-distribution cluster 3, which has a sharp peak in TB case notifications in the 25-34 years age group. A similar pattern is seen in the female age-distribution clusters, with all of the countries with HIV prevalence greater than 6.0% in 2007 included in the female age-distribution cluster 2, which again has a sharp peak in the 25-34 years age group. This peak reflects the age distribution of HIV infection, which tends to peak between 30 and 35 years for men and 25 and 30 years for women in high HIV prevalence settings [13] and results in a corresponding increase in TB cases.

The female age distributions follow similar patterns to those of the males, with the exception of cluster 1, which has a shape that we have only observed among the females. In cluster 1, there is a peak in the 25-34 age band and an increase in the oldest age group. The countries included in this cluster tend to be richer nations. We suspect that the peak in the 25-34 age band is likely to be due to immigrants from countries with severe TB epidemics but, as there tend to be fewer female immigrants than male immigrants from non-industrialised nations to industrialised nations (Box 10, page 48, [14]), the rise in TB rates among the elderly that is more typical of richer nations with ageing epidemics, is still visible for females, while this signal has been obscured in the corresponding male data.

There is a wide peak in the female cluster 3 over the 15-34 age bands then a sharp decline in the older age groups. The countries included in this cluster tend to have severe TB epidemics and relatively low HIV prevalence. Female cluster 4 and male cluster 2 include mainly middle income countries and have a peak in the 15-24 age group. The female cluster 7, mainly made up of South American countries, shows a similarity with these two clusters, having a peak in the 15-24 age group, but the decline is slower. Male cluster 4 and the female clusters 5 and 8 all have profiles that increase with age, indicative of ageing TB epidemics, supported by their memberships, which consist of high HDI (richer) countries. Similarly, male cluster 1 also contains fairly rich countries, with low TB rates and little HIV, but maintains an age distribution that is constant over the adult population. Female cluster 9 has a symmetric profile, similar to that observed for cluster 6 in the males and is made up of low to middle income countries with little HIV. Male cluster 5 is similar in make-up but has its peak skewed towards the 45-54 age band.

Comparing the results for the three cluster analyses, we can see that in each case there is one group that contains all of the countries with severe HIV epidemics (2 for females, 3 for males and 5 for historic data). More developed countries also tend to be grouped together in one or more clusters (1, 5 and 8 for women; 4 for men; 1 for historic data). There are some anomalies in the age-distributed clusters. In particular, there are several countries with very low rates of HIV prevalence, which have been included in the clusters of countries with high HIV prevalence. This is most likely due to immigration from countries with high HIV prevalence, and we discuss this further in section 3.3.

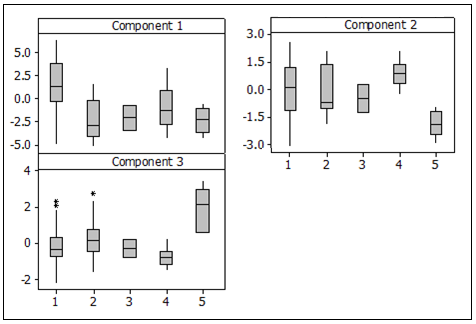
## Effect of ‘health and wealth’ indicators on TB

In order to further investigate differences between clusters we collected indicators relating to the economic conditions in each clustered country, e.g. gross domestic product per capita and the human development index (HDI); biological and behavioural risk factors, e.g. smoking rates and HIV prevalence; the quality of the health service, e.g. government health expenditure and TB case detection rates; and the distribution and life expectancy of their population. The indicators were chosen as factors that could directly or indirectly affect TB incidence, with reference to [8]. (A full list of the indicators considered in the analysis is given in the appendix (Table A.1).)

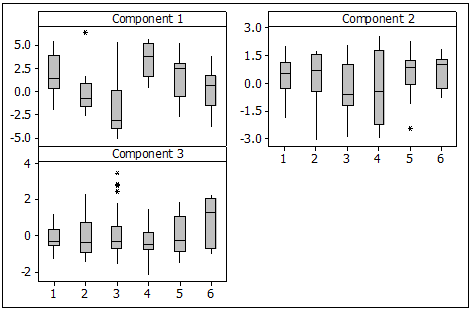
Unsurprisingly, there are significant correlations between many of the indicators. We therefore carried out principal component analysis (PCA) to reduce the dimensionality of the data set and to arrive at a set of uncorrelated variables that could be compared across clusters.

The PCA results suggest that three components can be used to describe over 75% of the variability in the indicator variables. Full results are given in the appendix (Table A.3) but broadly speaking, the first component contains all of the variables at significant levels except the growth in GDP, HIV prevalence, government health expenditure as a percentage of GDP and the smear-positive TB cure rate, suggesting that countries with high scores in the first principal component will be richer and more highly developed. The second component includes a measure of the level of corruption in the country, per capita growth in GDP, total health spend and government health spend; and the third component includes the smear-positive TB cure rate and has a minor contribution from the HIV prevalence with an opposite sign.

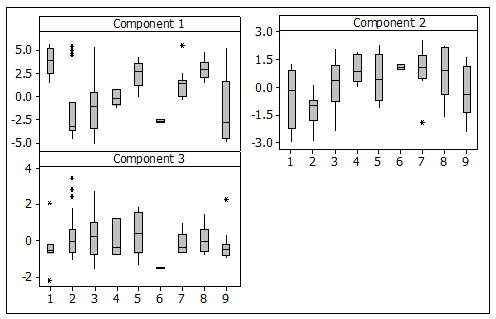
Figure 9 shows how the component scores vary for the different clusters. There is obviously some overlap between the clusters for all three of the components but differences are also apparent, especially in component 1 scores, which describe the level of development and wealth of a country. We carried out ANOVA analyses to compare the differences in component scores between clusters. The results back up the observation that all three sets of clusters can help to explain the variation in component 1 scores (p-values less than 0.001). This is the only component score that shows significant variation between clusters for the male age-distribution clusters, whereas there is some evidence of differences in the mean values of the score for component 2 for the female age-distribution and historic clusters (p-values 0.000 and 0.006 respectively). Interestingly, only the historic clusters show significant variability in the score for component 3 (p-value 0.000), which measures the level of the TB cure rate in a country in contrast to its HIV prevalence. The ANOVA test assumes normal data and similar variances, and the box plots show that some of the data are skewed and the variances of clusters can be very different. Therefore, although the ANOVA results give a good indication of how well the clusters can describe the variability in component scores, they should only be used as a guide. Full results of the ANOVA analyses described in this section are given in the appendix (Table A.4).



(a)



(b)



(c)

Figure 9: Box plots showing the variation in the PCA component scores for (a) Historic TB clusters; (b) Male age-distribution clusters; (c) Female age-distribution clusters.

## Immigration

The effect of immigration on TB epidemics in the more developed countries can be observed from the results of the clustering of countries by their age-specific TB case notifications. Several of the clusters generated have countries included in them that you would not expect based on the level of development of the country and its HIV prevalence. For example, Norway is included in cluster 2 in the men’s age-dependent TB clusters. On the whole, cluster 2 is made up of low to middle HDI countries, whereas Norway had the highest HDI in the world in 2007 [15]. Approximately 81% of the TB cases discovered in Norway were among foreign citizens [16], with immigrants from Somalia, Pakistan and Vietnam having a particularly high incidence of TB disease [17]. Interestingly, both Somalia and Pakistan are in the same cluster as Norway for age-dependent TB among males (there was insufficient data to classify Norway based on age-dependent TB among females). Sweden, the United Kingdom and Switzerland also have very high percentages of TB among foreign citizens (72%, 64% and 59% respectively [16]) and have age distributions for TB cases that are more akin to the countries where the majority of their immigrants come from than other countries of a similar level of development. Outside of Europe, the percentage of TB among foreign-born persons in the USA was 57% in 2006 [18], which seems to be low enough to allow the USA to be placed in a group with other well-developed nations such as Japan and New Zealand. Cain et al. [19] suggest that the majority of cases coming from immigrants in the USA are due to activation of a latent TB infection. Post-migration screening of migrants for active TB disease is routine in many industrialised countries, using similar techniques, and producing similar yield levels (number of cases found for each migrant screened). However, the USA also offers pre-entry screening for TB disease, in contrast to many European countries. Results from three studies suggest that this has a relatively high median yield (1.2% with an inter-quartile range of 0.85 – 1.25%, compared with a median yield from post-entry screening of 0.51% IQR [0.17 – 1.23%] for the USA, Canada, Australia and Japan [20]). This could provide some explanation for the different behaviour observed in the USA from many of the western European countries.

# Discussion

The Millennium Development Goal for TB seeks to achieve a reduction in TB incidence across the world, with the obvious constraint being resources. In working towards this goal, the first question to answer is which countries have most need of the available resources, i.e. how should the important donor organisations (WHO, World Bank, US Government, Gates Foundation) be proportioning their funding for TB control? Using an equity-based approach to resource allocation we could base the allocation of funds on the latest estimate of the TB case notification rate in a country. Under this approach, countries in historical TB cluster 5, with an average TB case notification rate of 522 per 100,000, would expect to receive almost 20 times more funding per head of population than countries in historical TB cluster 1, where the average TB case notification rate was 29 per 100,000 in 2007. This seems a reasonable first step at allocating resources.

The clustering gives us additional information however about the trends in TB case notification rates over time and it may be more beneficial for countries in clusters showing an increase in the TB case notification rate, such as those of clusters 2 and 5 (see Figure 2) to receive additional funding per head of population than countries where the TB case notification rate is steady or declining. Interestingly, out of the 22 high burden countries identified by WHO [1] only 10 are included in these two clusters. Without carrying out any modelling of the effect of funding on TB case notification rates, quantifying the amount of additional funding required is impossible. Nonetheless, the work that we describe here could provide a good first step to carrying out a full resource allocation study for TB around the world, as the classification of epidemics means that we would only need to build and parameterise 5 models of TB, rather than one for each country in the world.

The next question is how to allocate the resources within each country, which is more pertinent to local governments. All of the results above indicate that there is a significant positive correlation between HIV prevalence and TB case notification rates. The biology also suggests that HIV is one of the drivers of the TB epidemic and therefore reducing the HIV prevalence must be a priority both for itself and for reducing TB disease in countries where the HIV epidemic is particularly severe (e.g. countries in cluster 5 of the historic TB clusters). In countries with high HIV prevalence, the peak age group for TB tends to be the 25-34 years group, suggesting that this is the age group to focus on for TB control in these countries.

Results of the age-distribution clusters suggest that the focus of TB control in a country should depend on its level of development, rate of immigration and HIV prevalence, with the following general rules:

1. Rich countries with high levels of development and low immigration: case detection and cure should be focused on the 65+ age group.
2. Rich countries with high levels of immigration from countries with severe TB epidemics: case detection and cure should be focused on the immigrants, with the targeted age group dependent on the country of origin.
3. Low to middle income countries with low HIV prevalence: case detection and cure should be focused on the 15-24 age group
4. Countries with high HIV prevalence: TB case detection and cure should be focused on the 25-34 years age group.

The age-distribution clusters also give an insight into how TB epidemics might change over time. If we consider the male clusters, then the particularly severe TB epidemics tend to be classified in clusters 2 and 3, where the peak in the case notifications are in the 15-24 and 25-34 years age bands respectively, with the push to the older age group probably being a result of a reasonably severe HIV epidemic. Where treatment has been good for many decades, such as in the countries in male cluster 4, case notification rates tend to be very low and the distribution of cases among the age groups tends to increase with age. The high proportion of cases in the oldest age group is most likely due to reactivation of latent *mycobacterium tuberculosis* (*mTB*) infections, among the elderly population who have a much higher prevalence of *mTB* infections than the younger members of the population who are much less likely to have been exposed to TB disease. Between these two extremes, we have male clusters 5 and 6. Male cluster 5 is much more symmetrical than clusters 2 and 3, while male cluster 6 has its peak in the 45-54 years age group. One could argue that countries with good TB control are likely to have age-distributions for TB case notifications that progress from having a peak in the 15-24 age band (male cluster 2 shape) to the more symmetric shape of male cluster 5, with the peak moving to the 45-54 years age group (male cluster 6 shape) and finally an age distribution that increases over age (male cluster 4 shape) as the section of the population with the highest prevalence of *mTB* infections ages. This again has implications for detection and treatment of TB. As a country’s TB epidemic ages, the focus of the TB programme must adapt accordingly as strategies best suited to detecting TB among the 15-24 years age bands will not necessarily be appropriate for those in the 45-54 years age band.

In conclusion, we have shown how TB epidemics are similar and different across the world and given an indication of where funds should be targeted to best reduce TB case notifications.

**Acknowledgments**

There are no conflicts of interest for any authors of this paper.

# References

[1] World Health Organization. Global tuberculosis control: WHO report 2010.

[2] World Health Organization. Global tuberculosis control: a short update to the 2009 report.

[3] World Health Organization. World Health Statistics 2008.

[4] Sridhar, D. and Batniji, R., Misfinancing global health: a case for transparency in disbursements and decision making, The Lancet , 372 (2008) 1185-1191.

[5] Flessa, S., Priorities and allocation of health care resources in developing countries: a case study from the Mtwara region, Tanzania, European Journal of Operational Research, 150 (2003) 67-80.

[6] Zaric, G.S. and Brandeau, M.L., Dynamic resource allocation for epidemic control in multiple populations, IMA Journal of Mathematics Applied in Medicine and Biology, 19 (2002) 235-255.

[7] Lasry, A., Zaric, G.S. and Carter, M.W., Multi-level resource allocation for HIV prevention: a model for developing countries, European Journal of Operational Research, 180 (2007) 786-799.

[8] Dye, C., Lönnroth, K., Jaramillo, E., Williams, B.G. and Raviglione, M., Trends in tuberculosis incidence and their determinants in 134 countries, Bulletin of the World Health Organization, 87 (2009) 683-691.

[9] Corbett, E.L., Watt, C.J., Walker, N., Maher, D., Wiliams, B.G., Raviglione, M.C., Dye,C., The Growing Burden of Tuberculosis – Global Trends and Interactions with the HIV Epidemic, American Medical Association, 163 (2003) 1009-1021.

[10] Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W. Multivariate Data Analysis (5th Edition), Prentice-Hall (1998).

[11] Ward, J.H., Hierarchical grouping to optimize an objective function, Journal of the American Statistical Association, 58 (1963) 236-244.

[12] Morgan, D., Mahe, C., Mayanja, B. and Whitworth, J.A.G., Progression to symptomatic disease in people infected with HIV-1 in rural Uganda: prospective cohort study, British Medical Journal, 324 (2002) 193-197.

[13] Williams, B.G., Gouws, E., Colvin, M., Sitas, F., Ramjee, G. and Abdool Karim, S.S., Patterns of infection: using age prevalence data to understand the epidemic of HIV in South Africa, South African Journal of Science, 96 (2000) 1-9.

[14] United Nations High Commissioner for Refugees, UNHCR Statistical Yearbook 2009: trends in displacement, protection and solutions (2010) Available at http://www.unhcr.org/statistics.

[15] United Nations Development Programme. Human Development Report: overcoming barriers: human mobility and development. (2009) Available at http://hdr.undp.org.

[16] EuroTB and the national coordinators for tuberculosis surveillance in the WHO European Region, Surveillance of tuberculosis in Europe. Report on tuberculosis cases notified in 2006, Institut de veille sanitaire, Saint-Maurice, France (2008).

[17] Farah, M.G., Meyer, H.E., Selmer, R., Heldal, E. and Bjune, G., Long-term risk of tuberculosis among immigrants in Norway. International Journal of Epidemiology, 34 (2005) 1005-1011.

[18] Centers for Disease Control and Prevention, Reported tuberculosis in the United States, 2006. Atlanta, GA: US Dept of Health and Human Services, Centers for Disease Control and Prevention (2007).

[19] Cain, K.P., Benoit, S.R., Winston, C.A., Mac Kenzie, W.R., Tuberculosis among foreign-born persons in the United States. Journal of the American Medical Association, 300 (2008) 405-412.

[20] Klinkenberg, E., Manissero, D., Semenza, J.C. and Verver, S., Migrant tuberculosis: screening in the EU/EEA: yield, coverage and limitations. European Respiratory Journal, 34 (2009) 1180-1189.