Statistical methods in analysing health inequalities among the world citizens

Shanmuganathan S. ¹, and W. Claster ²

¹ Geoinformatics Research Centre, School of Computing and Mathematical Sciences
Auckland University of Technology, Auckland, New Zealand
² Ritsumeikan Asia Pacific University, Beppu, Oita, Japan
email: subana.shanmuganathan@aut.ac.nz

Abstract: While many international and national institutions world over, such as, World Health Organisation (WHO), the World Bank and national/ state health institutions, strive to promote health, prevent and control disease, formulate policies, programmes and evaluate interventions, and mobilise resources, all of these with an ultimate aim to prolong life of the world citizens as well as increase access to health care, health inequalities between advantaged and disadvantaged populations continue to increase at unprecedented rates and with greater complexity.

In this digital age there is no dearth for data on public health, health disparities and related factors, such as socioeconomic, income and GDP. However, lately, transforming this digital data into useful knowledge to learn more about human wellbeing and public health issues linked with appropriate factors is becoming a huge challenge even though it is seen as an inevitable requisite for the international and national health institutions to overcome the health disparities that are reaching alarming proportions. Public as well as private health care research institutions, such as, WHO and Pan-American Health Organisation, academic studies use different sets of core data and approaches for generating composite indices that best present the situation, progress and areas of urgent need in spending. A literature review on the subject shows that our ability to understand the processes as becoming increasingly complicated. Empirical studies into theorising health inequalities, linking relevant factors with inequities and understanding the underlying cause, reveal that we are now investigating into gaining more insights in the mechanisms that link poorer health to poorer circumstance. This is being carried out after having established that all health inequities are inherently inequitable thereby needing political intervention which is understandably seen as a troubling presumption.

In view of these factors, the paper presents an approach to group countries based on public health and related factors selected from an initial study by WHO using the World Bank and UN’s Millennium Ecosystem Assessment (MA) goals as well as tables with advanced (but simpler) statistical data analysis methodologies. The results portray the public health issues that dominate the national agenda, political intervention, evaluation measures and budget of these different country groups, especially the difference within countries are seen obvious and cannot be ignored even though they cannot be validated with standard statistical methods such as significance tests. For example, in developed countries the issues are focused on how to manage the escalating obesity and diabetes (type 2) by introducing more tax on fast food whereas, in African countries the focus is on reducing poverty to meet the basic needs of malnourished children, the younger generation, who are growing without their parents; victims of AIDS epidemic, starvation and natural disasters.

Keywords: health care indicators
1. INTRODUCTION

A recent report by The World Health Organisation (WHO) from an examination on primary health care systems, revealed of huge “inequalities and inefficiencies” in the healthcare sector around the world (Chan, 2008). Furthermore, the report stated that “Shortcomings in the systems” have left the health status of different populations, both within and among countries as “dangerously out of balance”. The current trend in the health sector is seen as increasingly becoming more focused on providing state-of-the-art technology and specialist care which is seen as less productive when compared with what could be achieved through preventive medicine, such as general practitioners. WHO estimates that better use of exiting preventive measures could reduce the global burden of disease as much as 70%. The following are the three major threats identified, and being attributed to the 21st century’s:

- unhealthy lifestyles spread across the world,
- unplanned urbanisation being carried out rapidly,
- ageing population

The WHO director-general warned that the above factors would contribute to a rise in chronic diseases, such as heart disease, stroke and cancer creating new demands for long term care. Striking inequalities have been observed in health outcomes, access to care and what people have to pay for care, the differences being stark today than it was 30 years ago. This situation would not naturally gravitate towards greater fairness and efficiency, but needs political intervention, such as concrete action to promote effective primary care hence the director-general in effect encouraged countries to go back to basics.

On a global scale, annual government expenditure on health care could be varying from as little as $20 to well over $6,000 per person. On the other hand, there are huge differences occurring even within cities, for example, in Nairobi when the under-five mortality rate was reported to be below 15 per 1,000 in the high-income areas, in a slum in the same city the rate was 254 per 1,000 (Chan, 2008). Similar differences can be seen discussed in (Burrows, 2008; Punam and Chuhan, 2007). The latter report posted by the World Bank further pointed out that the difference to be staggering presented with figures; one such case is the maternal mortality in childbirth which in many low-income African countries described to be 100 times higher than that of in the high-income countries, such as Europe. Rich countries have 3.7 physicians per 1,000 population compared with just 0.4 per 1,000 in low-income countries. The World Bank report as well as many other studies identify social economic factors as the root for health inequalities (Marmot, 2005).

The list of reports, research findings and warnings that recognise the importance of health inequalities grows at exponentially. Nonetheless, the critical question for public health practitioners on how to build the social movement to support burgeoning political efforts to address this unacceptable level of health inequalities in all societies remains unanswered (Beaglehole, 2009). Furthermore, the review report concluded that the major reason for any appropriate response for this as our poor understanding of the underlying causes and the dynamics of the socioeconomic determinants. Even more challenging is the development and implementation of a set plan and action to reduce the health inequalities by intervening on the determinants. Oliver, et al, (2002) as well support the argument and point out that there is a lack of clarity about where the inequality really matters and how this should be measured. Many people within the research community in health-inequality are seen as narrow focused on the extent and scope of differences in populations, for example health-inequality researchers tend to narrow down the scope of their work, but apply the findings to the whole society (Guy paper). Most of the empirical studies based on survey driven data are seemed to be focused on only a particular group of population defined by income, sex or race, hence suffer from serious methodological, theoretical and philosophical flaws and fail to look into aspects concerning the wider implications of any policy intervention (Forbes & Wainwright, 2001).

Epidemiological studies rely on the availability of information based on valid, reliable data. Hence, the availability of such information and data is considered as a sine qua non for analysing and evaluating the real situation of health status for making evidence based decisions and programming in health (Pan Americal Health Organisation, 2001). Developing a set of suitable and objective measures to keep a tap on the health state of a population is an old tradition in public health. Many such epidemiological studies, attributed to William Farr’s works of the 19th century, were for a long time based on mortality and survival alone. The need to include other factors such as socioeconomic and health care factors as an additional dimension is a more recent one and is discussed in the next section.

Finally, in this digital age, data on health care as well as related factors, mainly socioeconomic, among or within countries is in abundance. An interesting extension to the popular Moore’s law permits us to apply
Shanmuganathan and Claster, Statistical methods in analysing health inequalities among the world citizens.

the law to almost every measure of any digital electronic device/ functioning and its capability, such as processing speed, memory capacity, even the number and size of pixels in digital cameras and also the capacity of data storage devices. Therefore, the major challenge in computing and those rely on them is how this data that is described to be doubling in every two years, could be made available in a succinct and unambiguous way either with traditional statistical methods or other means, to an audience of decision/ policy makers in the health care sector (Rosling, 2007)

With that introduction to the current situation on health care status and issues within and among different countries section 2 outlines some recent approaches adopted by selected research and health care institutions in developing health indicators and their major findings. Section 3 presents the k-means clustering/ country group profiles, clustering being carried out using health care and socioeconomic factors chosen as appropriate in the current context to study the health inequalities. The approach as well could be applied to national data to study health related matters arising within a country.

2. HEALTH CARE RESEARCH INSTITUTIONS AND RECENT STUDIES

Currently there are many institutions in each and every country, actively collecting health data under different care systems, such as public/ epidemiological, primary health care, and analysing the data to see trends and issues arising from them that are of concern needing intervention, to inform decision and policy making management. The WHO is the main authoritative institution as far as global health care is concerned hence the section begins with a review on WHO’s recent findings, views and methods adopted to measure global health and care status.

2.1. WHO and health care

The most recent challenge faced by WHO is the “inequalities and inefficiencies” in the healthcare sector seen around the world (Chan, 2008), and is detailed in the Introduction.

Previously, “Confronting global health challenges and achieving health Millennium Development Goals (MDGs) require health systems to be strengthened” (World Health Organisation, 2006:1), was seen as the main theme in the WHO2006 report released at the opening of World Health Day 2006, at which World Health Workforce Decade (2006-2015) as well was declared. The report acknowledged the fact that primary purposes of health systems as improving the health of the world population even through these different systems obviously comprise many organisations, institutions and resources but all of their existence and functioning aimed at ensuring, delivering of services, generation of resources, financing and stewardship. Since its establishment in 1948, WHO has evolved into what it is today firmly stood up to the many challenges it was faced with and it continues to actively engage itself by changing the way it operates, its recent attempt being to meet the health related MDG goal targets to remove poverty from the face of this earth (World Health Organisation, 2006).

The WHO 2006 report described in great detail of the recent change in the institutional approach to health policy-making in health services. With refined government responsibilities and objectives the health sector has increasingly flooded with pluralist systems also involving the participation of different communities and societies, such as social services and nongovernmental organisations. The report found achieving balanced and adequate provision as well as efficient use of human resources as crucial and challenging due to this new role of the states as the stewards of the health systems with so many resources and systems to deal with. This change in the states’ role to become from being the major healthcare service provider to regulator can be seen as the reason for the development of some new health care indicator systems, such as the Millennium Development Goals (MDG) developed by Millennium Ecosystem Assessment (MA) program initiated by the UN member organisations as well as Organisation for Economic Corporation and Development (OECD).

2.2. WHO health care indicators

By the turn of this century in realisation of what is been described by the WHO as “a crisis point in global public health” (World Health Organisation, 2008:2), the world was faced with the stark reality in equalities in the health status within and between counties. The most startling one for the WHO was the millions of people dying from diseases, such as tuberculosis, malaria, measles, diarrhoea related and respiratory infections that were thought to be eradicated; this was happening amidst unprecedented new wealth and powerful new technology were introduced. This was also accompanied by falling life expectancy in poorer countries, and with new diseases, such as HIV/AIDS threatening to add further burden on the already stretched health systems. All in all, the situation led representatives from 189 countries to adopt a set of goals with set targets to be achieved within an agreed time frame to reduce poverty and promote human development. The main
goals relating to health care were to reduce child mortality, improve maternal health and combat HIV/AIDS, malaria and other disease by a set target date in year 2015. These indicators are used herein for analysing health inequalities within countries and are discussed in sections 2 and 3. Subsequent sections discuss the k-means results along with some conclusions and future work.

2.3. OECD health care indicators

The MDG were actually focused on resolving the health inequalities relating to third world or developing countries. A panel of OECD Health Care Quality Indicators Project with 21 countries and many international and inter-state institutions, selected indicators to benchmark care delivered in these countries through a detailed review process based on scientific soundness and policy importance for use in OECD countries. The report begins with

“Some context may be helpful in understanding the motivation for the Panel's work: "To improve care for their citizens and to realise...potential efficiency gains, policymakers are looking for methods to measure and benchmark the performance of their health care systems as a precondition for evidence-based health policy reforms. As published international health data sets such as OECD Health Data currently lack comparable measures for the technical quality of national health systems, there is, so far, little possibility of such international benchmarking. To fill this gap, the OECD Health Care Quality Indicators Project (HCQI) has brought together 21 countries, the World Health Organization (WHO), the European Commission (EC), the World Bank, and leading research organisations, such as the International Society for Quality in Health Care (ISQua) and the European Society for Quality in Healthcare (ESQH)." (Marshall, Leatherman, Mattke, & memebrs, 2004:1). The panel decided to find quality indicators to capture what it described as core components of care in three major sectors and they are;

- Health promotion; population-based strategies carried out through efforts to change health-related behaviour (Obesity prevalence, Physical activity, Smoking rate, Diabetes prevalence and Gonorrhoea/Chlamydia rates)
- Preventive medicine; organised, population-directed services (Blood typing and antibody screening for prenatal patients, HIV screen for prenatal patients, Bacteriuria screen for prenatal patients, Immunisable conditions, Low birth weight rate, Adolescent immunisation, Anaemia screening for pregnant women, Cervical gonorrhoea screening for pregnant women, Hepatitis B screen for pregnant women, Hepatitis B documentation in record at time of delivery, Hepatitis B immunisation for high-risk groups, Influenza vaccination for high-risk groups and Pneumococcal vaccination for high-risk groups
- Diagnosis and treatment; primary care; diagnostic and therapeutic activities that constitute the first line of organised personal medical care (Congestive Heart Failure readmission rate, First visit in first trimester, Smoking cessation counselling for asthmatics, Blood pressure measurement, Re-measurement of blood pressure for those with high blood pressure, Initial laboratory investigations for hypertension and Hospitalisation for ambulatory care sensitive conditions

2.4. Other public, private and empirical approaches

Pan American Health Organisation (PAHO)

PAHO has a Basic Health Indicator Data Base which is a multidimensional query tool that offers a collection of 108 indicators from 1995 to 2005. The system presents data and indicators on: demography, socioeconomic, mortality by cause indicators, morbidity and risk factors and access, resources and health services coverage. Selected indicators can be disaggregated into age groups, sex and/or urban/rural region. Generated tables can be exported and printed as required. The report stated that the data presented was updated annually with the latest country information. (http://www.paho.org/English/DD/AIS/cp_840.htm)

Health care indicator measures to participate in global ICT community

Certain health care indicators were used to measure the above in Botswana by (Maitlamo-National Policy for ICT Development, 2004) to see the progress in ICT policy development in comparison with. The indicators used were: Life Expectancy at Birth, Infant Mortality Rate, Adult Prevalence of HIV/AIDS, Health Expenditure per Capita and Physicians per 1,000 Inhabitants. The countries that were included in the comparative analysis were: Canada, Estonia, Malaysia, Mauritius, Namibia, South Africa and Trinidad and Tobago.
Empirical studies on health care inequalities

Research on health care inequalities is extensive and dates back to the 1940s of William Farr’s works. For a long time the research was based on mortality and survival data. The need to include other factors such as socioeconomic and health care factors as an additional dimension seem to be a more recent one and there a large number of studies, reviews and books on this subject. Interestingly, some even see this area of study as a subject of social justice and covering all that is far beyond the scope of this paper. However, some major directives taken to search for and measure health inequalities from an investigation conducted thus far in this research are outlined herein. As stated in a review essay titled “Health inequalities; still making policy in a fog?” by (Martin, 2006), the authors argue the policy perspectives as the “Achilles heel of health inequalities literature”. The reviews finds policy makers and health decision markers blaming each other as “politically naïve people” and “policy making in a fog” respectively.

In (Munga & Mæstad, 2009) qualitative data is analysed to understand the distribution of health status across a given population and to see how it changes as a result of any policy intervention. The paper presented an approach to evaluate qualitative data from self-reported health status, a record of people’s own perception of their health status. The approach is based on a partial inequality ordering and a second partial ordering defined to indicate a more “spread out” and the “overall health level rises” respectively. The approach is found to be a useful way to overcome the straightforward use of qualitative data as traditional Lorenz curve methods could be only be used with quantitative data.

Pradhan, Sahn, & Young (2003) using the height of children under 36 months and very complicated formulae produced a new index called “world height inequality”. The study, even though did not use any income, GDP or any other socioeconomic factors, found these factors as the main cause and also concluded that the within country variation as the source for most inequality rather that of the difference between countries. However, in the results there was a dramatic difference between the income and health inequalities, the difference being particularly more with the use nutrition indicators producing a concave function. This is due to the fact even modest improvements in a country’s welfare could compress the population’s distribution significantly as weight is seen to have a genetic upper bond whereas income was not as stated in the paper.

3. HEALTH CARE DATA ANALYSIS USING K-MEANS CLUSTERING

The section details the data and results of k-means clustering being invested herein to develop indicators to inform decision/ policy making management. The following are the data used in the analysis;

1. Total population (000s)
2. GNI per Capita
3. Expected years of schooling
4. Adult literacy rate (% of population ages 15+)
5. Average annual population growth rate (%)
6. Age dependency ratio (dependents as a proportion of working-age population)
7. Total fertility rate (births per woman)
8. Adolescent fertility rate (births per 1,000 women ages 15-19)
9. Contraceptive prevalence rate (% of women ages 15-49), any method
10. Life expectancy at birth (years)
11. Infant mortality rate (per 1,000 live births)
12. Under-5 mortality rate (per 1,000)
13. Maternal mortality ratio (per 100,000 live births), modelled estimates
15. Child immunization rate, measles (% of ages 12-23 months)
16. Child immunization rate, DPT3 (% of ages 12-23 months)
17. Births attended by skilled health staff (% of total)
18. Physicians (per 1,000 people)
19. Hospital beds (per 1,000 people)
20. Tuberculosis treatment success rate (% of registered cases)
21. Health expenditure, total (% of GDP)
22. Health expenditure, public (% of GDP)
23. Health expenditure, public (% of total health expenditure)
24. Health expenditure per capita ($) 
25. Prevalence of HIV, total (% of population ages 15-49)
26. Tuberculosis incidence (per 100,000 people)
(Source: The World Bank)

4. DISCUSSION AND CONCLUSIONS

The 208 countries analysed in the paper are clustered into 12 clusters and the cluster profiles (Figure 1) are studied to see the health care status within and among countries. The difference seen in the k-means cluster profiles are considerably stark as stated by WHO reports and other studies discussed earlier in the paper and they cannot be ignored even though the difference could not be confirmed by any statistical methods as stated by Rosling (2007). Furthermore, statistical significance tests are for validating sample data and not for trends and patterns within data sets such as analysed herein.

175
**Figure 1.** k-means country cluster profiles generated with 25 variables relating to health care and related socioeconomic factors. In the clustering four clusters consist of single countries. They are cluster 2 with Japan, 4 with India, 11 with USA and 12 China. As pointed out in the many WHO reports and other studies the difference between countries are stark. For example, the Health expenditure per capita ($) is 6096.2 in the US (cluster 11) and 30.4, the lowest of all in India (cluster 5). Cluster 9 countries (in italics) show 42, the highest Age dependency ratio (dependents as a proportion of working-age population). These countries as well show the worst figures for most of the variables analysed from the HNP Group Data (http://go.worldbank.org/VRLLAR6BG0 (last accessed 2006)).

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Japan</th>
<th>India</th>
<th>USA</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>June</td>
<td>India</td>
<td>USA</td>
<td>China</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C 1: Angola, Burkina Faso, Cambodia, Cameroon, Chile, Cote d’ivoire, Cuba, Ecuador, Ghana, Guatemala, Iraq, Kazakhstan, Korea D, Madagascar, Malawi, Malaysia, Mali, Morocco, Mozambique, Nepal, Niger, Peru, Romania, Saudi Arabia, Senegal, Sri Lanka, Syrian Arab Rep, Uganda, Uzbekistan, Venezuela RB, Yemen Rep, Zambia, Zimbabwe / C 2: Japan / C 3: Argentina, Canada, Colombia, Congo Dem, Kenya, Korea Rep, Myanmar, Poland, South Africa, Spain, Sudan, Tanzania, Ukraine / C 4: American Samoa, Andorra, Aruba, Austria, Australia, Austria, Belgium, Bermuda, Cayman Islands, Channel Islands, Denmark, Faeroe Islands, Finland, Greenland, Guam, Hong Kong, Iceland, Ireland, Isle of man, Kuwait, Liechtenstein, Luxembourg, Mayotte, Monaco, N Mariana Islands, Netherlands, Netherlands Antilles, New Zealand, Norway, San Marino, Singapore, Sweden, Switzerland, UAE, Virgin Islands / C 5: India / C 6: Brazil, Indonesia / C 7: Germany, Mexico / C 8: Egypt, Ethiopia, France, Iran, Italy, Philippines, Thailand, Turkey, Ukraine / C 9: Antigua Barbuda, Armenia, Azerbaijan, Baham, Bahrain, Barados, Batarus, Belize, Benin, Bhutan, Bolivia, Bosnia Herzegovina, Bosn, Brunei Darussalam, Bulgaria, Brundi, Cape Verde, Central Africa, Chad, Comoros, Congo, Costa Rica, Croatia, Cyprus, Czech, Djibouti, Dominica, Dominican R, El Salvador, Equatorial Guinea, Eritrea, Estonia, Fiji, French polynesia, Gabon, Gambia, Geogia, Greece, Grenada, Guinea, Guinea Bissau, Guyana, Haiti, Honduras, Hungary, Israel, Jamaica, Jordan, Kiribati, KyrgyzReo, Lao PDR, Latina, Lebanon, Lesotho, Liberia, Libya, Lithuania, Macao china, Macedonia FYR, Maldives, Malta, Marshall Islands, Mauritania, Mauritius, Micronesia Fed Sts, Moldova, Mongolia, Namibia, New Caledonia, Nicaragua, Oman, Palau, Panama, Papua New Guinea, Paraguay, Portugal, Puerto Rico, Qatar, Rwanda, Samoa, Sao Tome, Serbia Montenegro, Seychells, Sierra Leone, Slovak Rep, Slovenia, Solomon Islands, Somalia, St KittsNevis, St Lucia, St Vincent, Suriname, Swaziland, Tajikistan, Timor Leste, Togo, Tonga, Trinidad, Tunisia, Turkmenistan, Uruguay, Vanuatu, West Bani Gaza / C 10: Nigeria, Pakistan, Russian Fed / C 11: USA / C12: China.
A systematic review by Derose (2007) found it difficult to draw firm conclusions on the effects of social capital on health care access as there was a lack of congruence in how social capital was measured and interpreted, and a general inconsistency in findings. The review was based on 21 abstracts that met the examining criteria, some measure of social capital and its effects on health care access, out of a total of 2,396. The data used in this paper was formulated in 2005 for analysing progress in MDG target and does not include any indicators later designed in the OECD report discussed in section 2 hence further research is in progress including time series data on preventive health, such as changes in diet, physical activity, and tobacco use as this is considered to have dramatically increased in rich and poor counties, the risks of chronic disease such as heart disease, stroke, cancer and diabetes contributing for more than half of deaths (35 million out of 53 million) each year (World Health Organisation, 2007). Also, research is underway to look for trends within countries based on a study produced success in New Zealand (Signal, et al., 2007)

REFERENCES
Rosling, H. (2007). No more boring data: TEDTalks www.youtube.com/watch?v=hVimVzgtD6w