Author's response to reviews

Title: The Burden of Disease and Injury in the United States 1996

Authors:

Catherine M Michaud (cmichaud@hsph.harvard.edu)
Mathew T McKenna (mtm1@cdc.gov)
Stephen Begg (s.begg@sph.uq.edu.au)
Majid Ezzati (mezzati@hsph.harvard.edu)
Maria T Bulzacchelli (mbulzacc@hsph.edu)
Jessica Gaber Kreiser (jessicagaber@hotmail.com)
Niels Tomijima (nmt2102@columbia.edu)
Meghna Majmudar (mmajmud@yahoo.com)
Shahul Ebrahim (sbe2@cdc.gov)
Joshua A Salomon (jsalomon@hsph.harvard.edu)
Christopher JL Murray (christopher_murray@harvard.edu)

Version: 2 Date: 25 May 2006

Author's response to reviews: see over
We would like to sincerely the two reviewers for their helpful comments and suggestions. We took these comments very seriously and provided replies based on our understanding of these criticisms. We will be very happy to provide further clarifications if it is necessary.

Reviewer # 1 – Debbie Bradshaw

While the study reflects a substantial technical effort, it is written as a monograph and not in a typical article format. Furthermore, sections C1 and C2 of the methods reads like a manual on how to undertake a BODI with few specifics or details about this study. The report has an abrupt ending with very little discussion about the implications of the findings and lacks a synthesis of the findings of this study with other work on health and disparities in the USA.

In my view the article needs revision to article format including reference to literature reviewing what is already known about the burden of disease in the USA and the racial disparities in health.

1. The revised article should present selective results by making greater use of the facility to reference supporting documents – and still provide a full description of the methodology and a comprehensive presentation of results. Figures 8 and 9 are a particularly useful display of a large amount of detail on the race difference in the burden of disease.

Reply - The paper submitted in a report format was extensively revised to an article format. Specifically:

1. Sections C1 and C2 of methods have been deleted, and replaced in the text by the following: “The conceptual and computational details of how these parameters were estimated for individual conditions in the GBD have been presented elsewhere (Murray and Lopez, 1996). A summary overview of GBD methods is provided in Annex 1 for easy reference.

2. The discussion section has been rewritten and includes a discussion of the implications of findings and a synthesis of findings of the study with other work on health and disparities in the USA.
Methodological issues

2. There is a need to comment on the completeness of registration of deaths in the USA. While it is quite likely that most deaths are registered – there are suggestions that it differs by race groups? A rapid investigation of this topic indicates that misclassification of race on death certificates together with differential undercount in the census by race group does affect the estimates of mortality rates by race – see the NCHS Vital and Health Statistics Series 2 No 28 September 1999 by Rosenberg et al. This work would suggest that the low rates observed for Asians in the BODI may be a result of mis-classification on the death certificates. It would be important to assess the impact of this on the results of this study.

Reply - This comment is important and we added the following in the methods section:

“The inclusion of race in the analysis posed particular challenges for minority populations. Death rates are derived from two independent data sets: the number of deaths in the numerator comes from the detailed mortality file, and population numbers in the denominator from the census. There is no discrepancy in reporting of race in both data sets for Whites and Blacks, but race misclassification was found to be problematic for Asians and American Indians. Self-reporting of race in the census tended to be higher, particularly for American Indians, than was the attribution of race (by a third party) on death certificates – which will yield an underestimate of death and YLL rates (Rosenberg et al., 1999). However, we did not correct for race misclassification in Native Americans and Asians because evaluations showed that there was no systematic bias, as discrepancies in race reporting varied from year to year. Therefore observed race differentials in YLL rates may have been slightly overestimated “.

3. It is difficult to follow the rationale for adjustment to the number of IHD deaths. The clear correlation between the proportion of cardiovascular deaths (excluding stroke) that are IHD with the proportion ‘garbage cardiovascular codes’ demonstrates the need to reclassify some of the ‘garbage cardiovascular codes’ to IHD. Table 6 shows the coefficients for multiple linear regressions of the proportion of cardiovascular deaths that are IHD regressed on the lung cancer death rates and the cardiovascular garbage code death rates. It is unclear why the R2 for this multiple regression should be used as the proportion of the garbage category that should be reallocated to IHD. Why not use the regression equation to estimate the proportion of IHD based on the garbage code rate? In the extreme case where the R2 is 100%, it would mean that together, the lung cancer rate and the garbage code rate would perfectly predict the proportion of cardiovascular that are IHD. It is not clear to me why that would mean that all the garbage code death should be reclassified as IHD?
Reply – We revised this section and provided a more detailed explanation of the rationale and methodology of correction for underreporting of IHD in the US:

The most problematic aspect of cause of death coding pertains to coding of ischemic heart disease (IHD) (ICD-9 codes 410-414), which is one of the leading causes of premature mortality. The wide variations that exist in IHD reported mortality rates shed doubts about the validity of cross-national comparisons. The underlying causes of that might explain these variations were further explored in the context of the GBD and were convincingly attributed to variations across countries in coding practices. Physicians may use several ICD-9 codes that are actually due to IHD when they assign the cause of death. These include heart failure (428), ventricular dysrhythmias (427.1, 427.4, 427.5), general atherosclerosis (440.9) and ill-defined descriptions and complications of heart disease (429.0, 429.1, 429.2 and 429.9). IHD deaths may be assigned to these ill-defined cardiovascular codes, or “garbage codes” because of insufficient clinical information at the time of death, local medical diagnostic practices or simply by error. The statistical approach developed to correct for likely undercoding resulting from different coding practices in the GBD included a two step procedure which includes an ordinary least squares (OLS) regression equation predicting the proportion of cardiovascular death for each age group assigned to ill-defined codes as a function of the proportion of deaths assigned to IHD, and the correction of proportions for each country within set constraints, based on the assumption that the cluster of countries where ill-defined coding was low defined the standard coding practices (Murray and Lopez, 1996a, pp.131-133).

An exploration of cardiovascular death coding in the United States showed important differences in coding practices between states. Indeed, the proportion of all cardiovascular deaths (minus stroke) coded to cardiovascular “garbage” codes ranged from 14% in New Mexico to 37% in Alabama and New Jersey (Table 4.) Figure 1 illustrates the enormous variation across US States in coding practices with respect to these ill-defined cardiovascular codes. For each state, the fraction of cardiovascular deaths (excluding stroke) that are assigned to ICD –9 codes 410-414 is shown on the y-axis. On the x-axis the fraction of cardiovascular deaths (excluding stroke) that are assigned to the ill-defined cardiovascular codes is measured. The strong negative relation between IHD mortality and that from ill-defined cardiovascular codes strongly supports the suggestion that the quality of CVD death certification varies substantially across states, as it does across countries in the world.

This preliminary analysis confirmed the need to correct for under-registration of IHD in the US. To estimate the fraction of IHD deaths assigned to ill-defined cardiovascular codes, the regression equation applied in the GBD was revised. Age and sex specific lung cancer death rates were added to the model. Lung cancer mortality rates measure the cumulative effects of tobacco exposure as a risk factor for IHD (Peto et al., 1992).

The regression model for the US included age and sex specific lung cancer death rates, and ill-defined CVD rates for Blacks and Whites. These regression equations predicted the proportion of ill-defined CVD death by age and sex for Whites and Blacks.
4. Related to this, the plausibility of the coefficients in Table 6 warrant some comment. Is there any reason why the coefficients in younger ages differ from those in older ages and why the coefficients for white females should differ from other groups.

**Reply** - The finding that the extent of miscoding increased in older age groups is consistent with GBD regression results, which notes that “R-square increases with increasing age, [is] providing further evidence that ill-defined codes are indeed being used for IHD which is more common in older ages” (Murray and Lopez, 1996a, p.). Differences in coding practices observed by race as well as gender warrants further exploration (Table 5). A recent study concluded that “the greater presence of medical knowledge at the time of death, reflected by place of death and cardiologist per capita, reduces the use of ill-defined cardiovascular clusters. Racial and gender effects on coronary heart disease (CHD) assignment may reflect disparities in access to care and quality of care” (Murray et al., 2006).

5. It is not clear why the authors choose to compare the rankings with the OECD countries and not the DALY rates? A comparison of the rates may be useful in showing much higher or lower rates in the USA.

**Reply** – We added the following in the text: “International comparisons may pertain to the difference of in the magnitude of YLL (expressed as YLL rates), or to differences in the distribution of major causes of YLL. The latter was the focus of interest of our study, and therefore we examined differences in rankings of major causes of YLL (rather than YLL rates) between the United States and other comparable countries.”

6. A discretionary issue is whether there is any advantage to comparing age standardized rates rather than crude DALY rates?

**Minor revisions**

7. *Y. Popova is listed twice in author list.*

We revised the list of authors.

We added

1. Matthew T. McKenna, Centers for Disease Control who was instrumental at all stages of the study and contributed substantially to data analysis as well as to the write up of the report;
2. Shahul H. Ebrahim, Centers for Disease Control, who contributed the analysis of congenital birth defects.

We do no longer include Y.Popova, JA Kurichety as authors because we were unable to contact them. Their contribution is acknowledged.
The reason why neither of them was included in the report submitted to PHM was that the report has not been cleared by CDC, which we wrongly assumed precluded them to be listed as co-authors. In fact this is not the case if we include the following disclaimer “The views and opinions in this report represent those of the authors and not the Centers for Disease Control and Prevention.”

8. **Table 1 does not add any additional information and can be omitted.**

   **Reply** - Table 1 has been deleted

9. **Table 2 - It would be helpful to the reader to add the coefficient of variation to the table to back up the point made in the paragraph on page 8.**

   Now Table 4. Disability weights from person trade-off exercise conducted in Atlanta compared to composite scores from other exercises conducted at various international sites

   **Reply** – The coefficient of variation calculations have been added to the table per the reviewer’s suggestion.

10. **Table 3 – A bit more description of the data presented in the table should be given in the text.** Now Table 5. Pearson’s correlation coefficients for median disability weights for each exercise based on 19 conditions common to all person trade-off exercises

   **Reply** – The section on the comparison of median disability weight values between groups has been expanded and now states: “The coefficients of variation (C.V. = standard deviation/point estimate) were much smaller for conditions associated with more severe disability such as quadriplegia and severe dementia. However, despite variation between individuals within these groups, the correlation between groups for the disability weight values was very strong. Nineteen conditions were included in every exercise. For each group the median disability weight value for each of these 19 conditions was calculated. The correlations between pairs of groups for the 19 disability weight values were very high (Table 5, Range of Pearson’s correlation coefficients = 0.82-0.99).

11. **Page 9 – The ICD codes should be fully referenced.**

   **Reply** – The reference has been added

12. **Page 11 – There should be a reference to some of the work showing the wide range in
IHD mortality and questioning the validity of cross national comparisons.

**Reply** – The main reference is the Global Burden of Disease (pp. 131-133) and has been added. Also see reply provided above to point 2, first paragraph.

13. Table 10. Column headings in right hand side of table show “DALYs” but should be “YLLs”

**Reply** – Table 10 (Table 9 in the revised version) – The headings were changed from “DALYs” to “YLL”

14. Table 14. The title should indicate what burden measure is presented – YLDs or DALYs?

**Reply** – Table 14 has been dropped in the revised version

15. Page 17. The first 2 paragraphs should form part of the Methods section and not be part of Results. The next paragraph is repetitious and could be left out. The Results should start with Section A1 on page 19.

**Reply** – The two paragraphs were deleted and are now included in the methods

16. Page 42. There is a blank space between sentences in line 4.

**Reply** – The Discussion section has been re-written.

17. Page 44. The comparison of ranks cannot reflect difference in “magnitude” – only in the ranking.

**Reply** – Yes!

18. Page 45. I would disagree that a BODI is the first step to look at inequalities – there are many simpler indicators that reflect the disparities eg mortality rates?

**Reply** – The revised discussion takes took this point into consideration.

**Annex I**

19. The details given in the Annex are in very broad terms about what data was used and how it was adjusted. For example, it is not always give clear details on whether the data were analyzed by race and if not, how the estimates for each race group were obtained. It would also be useful to include tables showing the estimates derived – for example, the incidence by age and sex and duration.

**Reply** – The issue of whether race specific estimates were derived for all conditions is addressed in the methods section of the paper. This section states: “YLL were computed
by age, sex and race based on the detailed mortality datafile for 1996 (NCHS, 1996). YLD estimates were based on epidemiological data for the United States for 72 conditions that appear to account for at least 90 percent of the DALY total for this country. Main data sources for epidemiologic assessments in the United States were National population surveys (NHIS, NHANES, NLAES), disease registers, epidemiological studies and the National Hospital Discharge Database. A detailed presentation of analytic methods, data sources, and datasets used to develop the estimates for major causes of diseases and injuries are presented in Annex 2. If data on race and gender specific subgroups were too sparse to derive reasonable epidemiological parameters for particular conditions, YLD were estimated for these subgroups using YLD to YLL ratios for the overall population. For the remaining 26 conditions, YLD were calculated using YLD to YLL ratios from Established Market Economies (EME) countries in the GBD, applied to US specific estimates of YLL.”

Annex 2 provides more detailed explanations of the race, sex and age specific computations. A comprehensive presentation of all the epidemiologic estimates has been added as supplementary materials (Annex 5: Epidemiological estimates)

20. Page 3 ”I general” should read ”In general”; Page 41 “Motor” not “Motore

Reply – The text has been edited as requested

21. It is difficult to read the caption in Figure 1

Reply - Unfortunately this cannot be changed

Powerpoint Annexure

22. The Annex should be referenced in the main text.

Reply – Proper reference has been added page 14 – first paragraph of results section.

Reviewer # 2 – Theo Vos

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

This manuscript has obviously been written as a project report rather than a journal paper:

1. It has no abstract

Reply – The abstract was provided but has been revised as follows:
2. It repeats much of what is written elsewhere (e.g. the basic GBD methods)
   The methods section on YLD/dismod gives a textbook-like introduction but fails to
   inform the reader what was done for the US study

**Reply** – The method section has been extensively edited. It now focuses entirely on US
specific methodological adjustments to standard GBD methods. The general overview of
GBD methods has been included as Annex 1, for easy reference.

3. In the results section quite a bit of space is reserved to present differences between
deadth, YLL and YLD patterns which no longer have any ‘novelty value’ for a journal
   paper

**Reply** –

4. Intros to results sections repeat methods

**Reply** – intros to results sections that repeated methods have been removed in the revised
version.

5. Discussion is a rehash of results and does not critically examine the methods nor does
it place findings of this study against previous findings

**Reply** – the discussion has been re-written in the light of these important comments.

6. In its current form not acceptable as a journal article. The authors should consider a
major rewrite of the material as a journal article concentrating on what was particular
to the US study (e.g. the DW validation, the CVD redistribution algorithm) and
potentially making e.g. the ethnic comparisons its main focus. I have no problem
having a relatively large amount of basic methods and results added as an annex.

**Reply** – The article has been extensively revised and is now presented as a journal article
as was requested by both reviewers.

*Other issues:*

7. Comparisons between ethnic groups and with other OECD results are mostly made
for rankings; age-standardized death/YLL/YLD and DALY rates would be more
informative as are presented in fig17, fig 20-23 for selected conditions and for total
DALY rates between ethnic groups which curiously only appear in the discussion
section

**Reply** – This comment has been addressed in the reply to reviewer #1 (point3).
A discussion of figures 20-23 has been added in the results section.
8. The CVD redistribution methods are not fully explained: Are age, sex and race also included as independent variables?

Reply - Variables for lung cancer and garbage codes variables are age, race and sex specific rates.

9. Why are lung cancer rates (a proxy for smoking? and if so, why?) and garbage code ‘rates’ (or proportions?) included as independent variables? I have my doubts about the inclusion of the rate/proportion of garbage codes as this is highly correlated with the % coded to IHD, the dependent variable

Reply - These questions have been addressed in the revised methods section (see replies to reviewer #1 (points 2 and 3).

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

10. Following the GBD hierarchy, perinatal conditions should not appear in a ranking list of diseases (i.e. comparison ought to be with all cancer rather than individual cancers)

Reply – The authors agree with the reviewer that the choice of presenting major cause groups (e.g. cancer), or individual conditions (e.g. lung cancer) in the rankings of diseases is not consistent in the Burden of Disease literature. In this manuscript the authors made these choices based on the overall size of the burden associated with individual conditions, and the etiologic and pathophysiologic heterogeneity associated with each of the conditions within a major cause grouping. For example, Ischemic Heart Disease and Cerebrovascular Disease are listed separately because of the overall large burden associated with each of these conditions, and the different organ systems effected – even thought the underlying vascular mechanisms that precipitate pathology are similar (i.e. atherosclerosis). However, Homicide and Violence are listed in the rankings because the underlying mechanism for most of the burden resulting from these two different injury conditions in the U.S. is the same (i.e. use of firearms). In the case of perinatal conditions, most of the burden arises from conditions precipitated by premature birth – specifically, low birth weight and perinatal asphyxia. Therefore, the major cause grouping for these conditions was presented. However, in the case of cancer, the populations effected, and etiology for conditions such as breast cancer and lung cancer are very different, and each of these conditions has a relatively large burden. Therefore, they are presented separately.

11. There are quite some examples of sloppy editing/presentation, e.g.
   - death (should be deaths) on p18 (line 5 following heading A1) and p19 (line 14)
   - p36 (line 9) causes (should be cause)
   - figure 2, Y-axis legends should not include (thousands)
Figure 18 has legend as well as labels
Figure 21: sloppy position of legends
Page 41 neuropsychiatrique?

Reply - The revised article has been corrected as needed.