**Author’s response to reviews**

**Title:** Evaluation of the Mortality Registry in Ecuador (2001 - 2013) - Social and Geographical Inequalities in Completeness and Quality

**Authors:**

Alejandro Andrés Peralta Chiriboga (aperalta@aspb.cat)

Joan Benach (joan.benach@upf.edu)

Carme Borrell (cborrell@aspb.cat)

Verónica Espinel (vespinel@aspb.cat)

Lucinda Cash-Gibson (lucinda.cashgibson01@estudiant.upf.edu)

Bernardo L. Queiroz (blanza@gmail.com)

Marc Marí-Dell’Olmo (mmari@aspb.cat)

**Version:** 1  **Date:** 09 Sep 2018

**Author’s response to reviews:**

We would like to thank the reviewers for their thoughtful comments and suggestions. We believe that the proposed changes have improved the quality and rigor of the current study. We have put a lot of effort in responding to your suggestions and answers. It was hard work, but we believe it was worth it.

Reviewer #1: Date: 12 July 2018

Reviewer's report:

The manuscript "Evaluation of the Mortality Registry in Ecuador" is an article of potential importance to policy-makers and researchers in this field. The authors apply best-practices methods to both estimate completeness and quantify a number of indicators of quality. Most importantly, the authors describe the results by province, sex and age to explore inequalities within Ecuador. Every section is well-organized and easy to understand.
This manuscript is of excellent quality and professionalism and I congratulate the authors. I offer some minor suggestions which I would consider important to the publication of the manuscript and some discretionary revisions which I believe would improve the interpretation of the manuscript, but are not critical. I have no major compulsory revisions.

Major Compulsory Revisions

None

Minor Essential Revisions

1. The application of SEG, GGB and hybrid SEG/GGB is the accepted approach to death distribution methods, however there is insufficient information to ensure that these methods have been applied correctly. I suggest a supplementary methodological appendix which describes the DDM approaches applied using text and figures to assure the audience of the rigor of this work.

We agree with your comment and have included a short supplementary appendix describing the Death Distribution Methods (DDM) used in the study. This clarifies the data and methods used, their assumptions and limitations, the selection of the appropriate age trims and of the summary measure (harmonic mean).

2. The authors do an excellent job of describing the results by geography, sex and age. Some of the results could also be presented in terms of changes over time. I suggest the authors also describe the temporal trends, perhaps even the time trends by geography, in order to inform the reader of expectations for the future.

We agree with this necessity to show how completeness and quality have evolved during the study period. Unfortunately, one of the limitations of death distribution methods is their (lack of) timeliness, as they depend on the information of two censuses to be developed. For an assessment of evolution we will have to wait until the 2020 census that is being planned in the country. For quality indicators this could be done. We have added two supplementary figures showing the evolution of the percentages of garbage codes for Ecuador and for each province for women (supplementary figure 1) and men (supplementary figure 2). The following changes have been made in the text to reflect this new analysis:
Methods - the following sentence was added to the text:

“The annual percentages of garbage codes were also obtained to describe the temporal evolution in each study area.”

Results - the following sentences were added to the text:

“The temporal evolution of garbage coding in all the study areas can be seen in Supplementary figures 1 and 2. When we analyse the temporal evolution of garbage coding, we can observe that there has been a steady descent in garbage codes in Ecuador in the study period; but that the situation in individual provinces is heterogeneous. Provinces located in the Amazonic, northern Coastal and in the central Andean regions show consistently high garbage code percentages through the study period. Other provinces show important declines in garbage code percentages. For example, provinces in the southern Coastal region (Guayas/Santa Elena, Los Rios and El Oro) and Azuay in the Andean region show this trend.”

Discussion – When discussing the geographical patterns of completeness and quality indicators, we added a brief comment on the temporal evolution of garbage coding.

This visual description of the temporal evolution of garbage coding allowed us to see several interesting things: 1) In the country as a whole, a slow but constant descending trend can be observed. 2) The evolution of garbage coding among provinces is heterogeneous. Important progress is observed in some provinces (Guayas/Santa Elena, Azuay and los Rios for example). Other provinces have consistently high garbage code percentages through all the study period (Esmeraldas, Bolivar and Chimborazo for example). Most of the provinces in the Amazonic region show consistently high percentages of garbage codes with an irregular temporal pattern. The irregular temporal pattern observed in Galápagos reflects its small population and number of yearly deaths.

3. The authors note the use of a harmonic mean of the three DDM methods as a summary. While this is acceptable, it should be noted in the Methods section before being mentioned in the Results section.
The following sentence was added to the part of the Methods section describing completeness:

“For each study area and for all Ecuador (for both men and women), we present the completeness estimates for each DDM method and the harmonic mean of the three methods.”

Also, in the supplementary appendix, we describe why we decided to use this measure.

4. As the focus of the manuscript is related to equity, it is important for the authors to provide more discussion surrounding the relationship between CRVS performance and ethnic minorities and socioeconomic indicators. The authors duly note that the patterns of poverty and marginalized ethnicities coincide with the patterns of CRVS performance, but offer no explanations or hypotheses why. Could it be access to services? Poor-functioning health systems? Differences in health-seeking behavior? Differences in epidemiology? Citations to any other studies on the subject would also be expected.

Following the recommendation of other reviewer, we performed an analysis comparing the percentage of deaths not coded by MDs n each province, with the percentage of garbage codes in each province in two ways: 1) visually, through the creation of scatterplots form men and women that allowed us to examine the type of possible association and how it changes by geographical region (Supplementary Figure 3); and 2) estimating the Pearson’s correlation coefficient and its 95% CI. We believe that the poor access to services and the poor quality of them can explain much of the observed inequalities in the CRVS. In the analyses, we can clearly observe how the quality of the mortality registry worsens when deaths are not coded by MDs (possibly because of lack of access to healthcare services); but it is also noticeable that a good percentage of the deaths coded by MDs are still garbage codes (reflecting lack of adequate training for example). The following changes have been introduced in the text’s methods, results, and discussion sections to introduce this new analysis:

Methods - the following paragraph have been added:

“As the quality of death registration can be affected by the professional background of the people registering the deaths (8), the association between the percentage of garbage codes and the percentage of deaths not coded by a medical doctor (MD) was explored visually through the use of scatterplots and quantitatively through the estimation of Pearson’s correlation coefficients and their 95% confidence intervals.”
Results - the following sentences have been added:

“The percentage of deaths not registered by an MD ranged from 1.5% (in Pichincha/Santo Domingo) to 50.17% (in Napo) for men; and from 1.03% (in Pichincha/Santo Domingo) to 53.12% (in Napo) for women. In all areas and in both men and women the percentages of garbage codes were higher than the percentage of deaths not coded by an MD. Supplementary Figure 3 shows visually how the percentage of deaths not coded by an MD relate to the percentage of garbage codes in the study areas. It can be noticed that many provinces in the Amazonic region and one province in the northern Coastal region (Esmeraldas) have high percentages of deaths not coded by an MD and high percentages of garbage codes. Also, some provinces in the Andean region (and Manabí in the coastal region) have high percentages of garbage codes, even though their percentages of deaths not coded by an MD is lower. Pearson’s correlation was 0.82 (95%CI 0.61 to 0.92) for men and 0.84 (95%CI 0.65 to 0.93) for women.”

Discussion - the following lines have been added:

“The correlation between the professional background of the people who certify deaths and the quality of the mortality registry has been observed in other countries in the region; finding that mostly rural areas have less deaths certified by medical doctors and lower quality of the mortality data (1). We have found a similar pattern in Ecuador, especially in the provinces in the Amazonic region and Esmeraldas in the Coastal region. Populations in provinces with disperse populations in rural areas could have greater difficulties finding healthcare personnel to certify deaths. One thing that can be noticed is that there is still an important percentage of deaths certified by medical doctors and assigned to garbage codes, especially in many provinces in the Andean region. This means that work is needed not only to improve access to healthcare professionals, but also to improve the quality of death certification by medical doctors. “

5. Because the authors intend this work to be applicable for policy makers to "focus on completeness, quality or both", more discussion is necessary on the provinces which paradoxically have the lowest completeness but also the highest quality and vice versa. Are there unique local systems in place? What conditions might explain this?

We agree with your comment. Most of the provinces with high completeness estimates and low quality are located in the central Andean region. We believe the particular characteristics of the vital statistics system in those areas should be studied more in depth. The following sentences have been added to the discussion section to highlight this point:
“Attention should be brought also to the provinces that paradoxically have high quality and low completeness or high completeness and low quality of the mortality data. Galapagos has a low percentage of garbage codes (high quality) and also a low completes estimate (low quality). We believe that this could be explained by the low population (and consequent low number of deaths) and relatively high migration rates in the Province that could make DDM methods to perform poorly (completeness being underestimated). On the other side, many provinces in the central Andean region show high mortality completeness estimates and low-quality data. We believe that the particular characteristics of these areas should be studied more in depth in order to explain this phenomenon.”

Discretionary Revisions

1. The authors note that a period of re-drawing of province borders occurred, but do not return to this topic. Is there any evidence that this administrative shift had an effect on the health system, and by extension the CRVS system? I suggest the authors note either in the Results of Discussion section whether the "new" provinces appear to be high or low in completeness or quality.

During the study period, Santo Domingo and Santa Elena provinces were created. Santo Domingo was part of Pichincha and Santa Elena part of Guayas. Pichincha and Guayas are the most populated provinces in the country and have much more medical services. Santa Elena and Santo Domingo are less populated and more rural. We believe that, if analyses could have been made separately for the 4 areas, Pichincha and Guayas would have greater completeness and quality and Santa Elena and Santo Domingo much lower completeness and quality indicators. This was stated more clearly in the methods section.

2. An additional limitation worth noting in the discussion section might be construct validity. Garbage codes, age/sex missingness and internal consistency are just proxy indicators for the overall quality of the system and should probably be acknowledged as such. Other important quality issues such as classification accuracy are near-impossible to measure without primary data collection, but are certainly other aspects of quality that are not captured here.

We included this in the limitations paragraph in the discussion section.
Reviewer #2:

The authors in this paper propose to answer the question of whether there are inequalities between subnational administrative-geographical divisions in Ecuador in quality, completeness and internal consistency of death certification/registration. Ecuador's Civil registration and vital statistics system will no doubt benefit of a more detailed analysis of mortality registration completeness at the subnational level. Such estimations will allow targeting interventions for improvement, given an estimated country completeness of death registration of around 78.3 for the period 2010-2015 (OPS. 158 Sesión del Comité Ejecutivo. CE158/INF/8. Plan De Acción Regional para el Fortalecimiento de las Estadísticas Vitales y de Salud: Informe Final. 8 de Abril del 2016. Anexo). This is also of interest to a more general audience in low and middle-income countries wishing to learn on methods to estimate completeness of mortality registration.

The authors need to clarify the following issues in the methodology:

Completeness estimates:

1. The authors refer to Murray et al's (Murray CJL, Rajaratnam JK, Marcus J, Laakso T, Lopez AD (2010) What Can We Conclude from Death Registration? Improved Methods for Evaluating Completeness. PLoS Med 7(4): e1000262. doi:10.1371/journal.pmed.1000262) article for the methodology used in the Death Distribution Methods. This research clearly points to best age groups to use for each DDM and recommends that: "if an estimate of completeness must be obtained solely from the optimal DDMs at time period, then the median result across these three best age trims may be used". Peralta et al used the three DDMs: GGB, SEG and GGB-SEG to estimate completeness. They provide a summary estimate based on the harmonic mean of the three estimates and provide in table 1 for each province an age interval. In the methods section they say that "...for each method used, DDM automatically chooses the age interval that best fits the models and minimizes the residuals...". The authors need to explain why they chose to use the harmonic mean and how they arrived at the single age interval in each province.
Thank you for your comment. We think it is important to clarify these points better. Initially, we estimated the completeness estimates using the age trims recommended in the paper by Murray et al. (2010); we then used the automatic age trim feature of the DDM package and compared both. Results obtained from the R DDM package automatic selection feature seemed to be more consistent; while results from the methodology recommended by Murray et al. did not seem to relate to migratory patterns (one of its potential advantages was to reduce bias caused by migration). For that reasons, we decided to keep the results obtained using the automatic age trims of the DDM package. Regarding the harmonic mean, we decided to use it because the use of the median (of three values), consistently selected results from the SEG family. To obtain a better measure of central tendency (that actually summarized the three methods used), we picked the harmonic mean. Harmonic means are not as affected by extreme values as other means and are useful when we work with rates. Following the recommendation by other reviewer, we decided to include a supplementary appendix explaining the DDM methods used with greater detail. In there we explain the selection of appropriate age trims and why we use the harmonic mean in the following way:

“4. Strategies to minimize bias IN DDMs

To try to minimize bias induced by non-compliance with the assumptions of DDMs, researchers have adopted various strategies, many times subjectively. One of the main strategies is to use an age range that is more compliant with the assumptions of the methods. To minimize the effect of migration, some authors have preferred to use age ranges considered to be affected less by migration (over 60 years old for example). Other authors have used simulated data or comparisons of methods to real data in places in which complete registration is presumed in order to determine the appropriate age ranges to be used (1,2,4). Hill, You and Choi (2009) conclude that the best strategy is to use GGB-SEG with an age range of 5 to 65 years old. On the other hand, Murray et al. (2010) conclude the most appropriate strategy is to combine the three types of DDMs using the following age trims: 55 to 80 years old for SEG, 40 to 70 years old for GGB, and 50 to 70 years old for GGBSEG. Other approaches to select the appropriate age trims are based on graphical analyses for each DDM method. The DDM R package presents a solution that automates the choice of age groups for users (5). The main idea is to provide an estimator that will give the best solution from all possible combinations of age ranges. The estimator implemented by the DDM R package is the Root Mean Square Error (RMSE). The Root Mean Square Error, also called the root mean square deviation, is a measure of the difference between values predicted by a model and the values actually observed. These individual differences are also called residuals, and the RMSE serves to aggregate them into a single measure of predictive power. For each individual area, the age range that minimizes the RMSE is taken as the age range for the GGB coverage estimate, and this criterion is designed to mimic the more traditional eyeball fitting used for the GGB method. In other words, the automatic solution is to determine death registration coverage based on the age range that produces the smallest residual difference between observed and estimated death rates and their adjustment line (5)…”
7. Analysis

The R DDM package, developed by Tim Riffe, Everton Lima and Bernardo Queiroz was used to estimate completeness using the 3 DDM methods. The R DDM package manual describes the methods used as follows (5):

GGB:

“The method is based on finding a best-fitting linear relationship between two modeled parameters (right term and left term), but the fit, and resulting coverage estimate, depend on exactly which age range is taken. This function either finds a nice age range for you automatically, or you can specify an exact vector of ages.”

SEG:

“The method estimates age-specific degrees of coverage. The age pattern of these is assumed to be noisy, so we take the arithmetic mean over some range of ages. One may either specify a particular age-range, or let the age range be determined automatically. If the age-range is found automatically, this is done using the method developed for the generalized growth-balance method. Part of this method relies on a prior value for remaining life expectancy in the open age group. By default, this is estimated using a standard reference to the Coale-Demeny West model life table, although the user may also supply a value.”

GGB-SEG:

“The method estimates age-specific degrees of coverage. The age pattern of these is assumed to be noisy, so we take the arithmetic mean over some range of ages. One may either specify a particular age-range, or let the age range be determined automatically. If the age-range is found automatically, this is done using the method developed for the generalized growth-balance method. Part of this method relies on a prior value for remaining life expectancy in the open age group. By default, this is estimated using a standard reference to the Coale-Demeny West model life table, although the user may also supply a value. The difference between this method and SEG is that here we adjust census 1 part way through processing, based on some calculations similar to GGB.”
Two different approaches were performed initially: 1) Estimate completeness using the age ranges automatically chosen by the R DDM package; and 2) Estimate completeness estimates using the specific age ranges for each method as described in the Murray et al. 2010 paper. Results from the three DDM methods with both age range selection strategies were obtained for men and women for all Ecuador and for the 22 study areas. Following Murray et al. (2010) recommendation, a summary measure of the three DDM methods was obtained. In that study, the authors used the median as summary measure. The problem with that approach is that the median of three numbers will always be one of the three values. In our case that meant selecting either the SEG or the GGBSEG methods most of the time. For that reason we decided to use the harmonic mean, as it is a central tendency measure not affected by extreme values and appropriate for rates.

The following table compares the completeness results obtained using the automatic age range selection with the ones obtained using the specific age ranges:

Table 1: Comparison of DDM estimates using automatic age trims of the R DDM package and age trims recommended by Murray et al. (2010)

(See table in Appendix 1)

From the table shown above, we can see that for both sexes, completeness estimates were higher using the automatic selection of age ranges in most of the study areas. When applying DDM methods for the whole country, both strategies obtained similar results (mean difference of 3.43% for women and 5.51% for men). The difference of the harmonic means was much greater at the provincial level; ranging from -0.35% (Pichincha / Santo Domingo) to 23.54% (Esmeraldas) in women; and from -9.40% (Galápagos) to 26.90% (Esmeraldas) in men. We decided to keep the results obtained through the automatic age range selection for two reasons:

1) These estimates give more conservative estimates of completeness, closer to the ones observed in previous literature for Ecuador.
2) One of the potential benefits of using the specific age ranges approach was to reduce the bias associated with migration. We compared the changes in DDM estimates with the provincial percentages of net migration in 2001 and 2010 reported by Royuela and Ordóñez (2018) (7). The change seen between the two approaches did not seem to be related to the migratory patterns in Ecuador."
2. Murray et al found that the uncertainty around completeness of death registries estimates could be up to +20%. Peralta et al make no effort to estimate uncertainty bounds and overlook the fact of the possibly very important effect of age misreporting. De Lima et al, also referred to in this paper (de Lima y Queiroz Evolution of the deaths registry system in Brazil: associations with changes in the mortality profile, under-registration of death counts, and ill-defined causes of death. Cad. Saúde Pública, Rio de Janeiro, 30(8):1721-1730, ago, 2014) warn about the limitations of the application of DDMs for small areas where the assumptions inherent to these methods do not hold. The uncertainty issue is clearly important when comparing provinces in Ecuador as the difference in population size and number of deaths in the period between provinces is important, and some of the provinces have very small populations and number of deaths in the period. As the objective is to identify provinces needing strengthening of death registration or using completeness estimates to adjust for death rates and other mortality analysis, these estimates might be misleading.

We agree with you in the sense that the uncertainty of completeness estimates is an important issue. In the supplementary appendix that describes DDM methods used in the study we describe this limitation with more detail. We also compare the completeness results obtained using the automatic age trims selected by the R DDM package and the results obtained using the specific age trims recommended by Murray et al. (2010). We hoped this could serve as a small sensitivity analysis; and, in fact, it reveals that in some areas results can have great variability depending on the methods used.

The idea of this study was to give an overall view of the mortality registry in Ecuador. It is part of a thesis project that has as one of its final objectives to create a mortality Atlas for Ecuador. The next study under creation will estimate SMRs using Bayesian hierarchical models that include completeness estimates for each area. Similar approaches have been used by other authors. For example, Schmertmann and Gonzaga (2018) use Bayesian models to estimate age-specific mortality in small areas in Brazil (2). One of the advantages of these models is that they take the completeness estimates for each area and use information from the neighbour areas to estimate credible intervals for the completeness and SMR estimates. These types of analyses are complex and the research team thinks their application in the context of Ecuador requires a separate study detailing the specific methods used. As the limitation of uncertainty is still present, we mentioned it in the discussion and in the methodological appendix.
The following changes have been introduced in the text’s discussion section to explain this limitation better:

Discussion - the following sentences have been added:

“Other important limitation of DDM methods is uncertainty. Murray et al. (2010) reported that “…the uncertainty around relative completeness of registration is likely to be at least +/-20% of the estimated level, and perhaps considerably more…” (11). In supplementary text 1, we compare completeness estimates presented in the article with estimates obtained using the age groups recommended in the article by Murray and colleagues. We found that, at a national level, results were quite similar between the two approaches (3.43% difference between harmonic mean of the three methods in women and 5.51% in men). Nevertheless, results showed a much greater variability in the provincial estimates. All this means that the completeness estimates presented should be taken as rough estimates, helpful to orientate public policy and have a sense of the geographical pattern of completeness in the country; but not as definitive and precise estimates. Greater work is needed to refine methods to estimate completeness at subnational levels and in countries were data is not completely compliant with the assumptions of DDM methods.”

Quality:

1. Unusable causes of death during the period analysed: 2000-2013 by province could have been affected by the nature of the person certifying as explained in the paper, death certificates may be completed by health personnel (not clear if it is possible to identify if medical doctor or other), civil registrars and police authorities. It would help to interpret these results if for each province the proportion of death certificates issued by medical doctors and others could be known. It would be no surprise if provinces with ethnic and disperse rural population had less access to medical staff to certify deaths. This is important as the interventions needed to improve quality in the different provinces will vary according to what are the causes. These differences might also help explain the results of internal consistency obtained.
We agree with the comment. The percentage of deaths registered by MDs or by others at a provincial level could be estimated from the annual national mortality databases. We analysed the association between the % of deaths registered by non-MDs and the percentage of garbage codes at a provincial level in two ways: 1) visually, through the creation of scatterplots form men and women that allowed us to examine the type of possible association and how it changes by geographical region (Supplementary Figure 3); and 2) estimating the Pearson’s correlation coefficient and its 95% CI. The following changes have been introduced in the text’s methods, results, and discussion sections to introduce this new analysis:

Methods - the following paragraph have been added:

“As the quality of death registration can be affected by the professional background of the people registering the deaths (8), the association between the percentage of garbage codes and the percentage of deaths not coded by a medical doctor (MD) was explored visually through the use of scatterplots and quantitatively through the estimation of Pearson’s correlation coefficients and their 95% confidence intervals.”

Results - the following sentences have been added:

“The percentage of deaths not registered by an MD ranged from 1.5% (in Pichincha/Santo Domingo) to 50.17% (in Napo) for men; and from 1.03% (in Pichincha/Santo Domingo) to 53.12% (in Napo) for women. In all areas and in both men and women the percentages of garbage codes were higher than the percentage of deaths not coded by an MD. Supplementary Figure 3 shows visually how the percentage of deaths not coded by an MD relate to the percentage of garbage codes in the study areas. It can be noticed that many provinces in the Amazonic region and one province in the northern Coastal region (Esmeraldas) have high percentages of deaths not coded by an MD and high percentages of garbage codes. Also, some provinces in the Andean region (and Manabí in the coastal region) have high percentages of garbage codes, even thou their percentages of deaths not coded by an MD is lower. Pearson’s correlation was 0.82 (95%CI 0.61 to 0.92) for men and 0.84 (95%CI 0.65 to 0.93) for women.”
Discussion - the following lines have been added:

“The correlation between the professional background of the people who certify deaths and the quality of the mortality registry has been observed in other countries in the region; finding that mostly rural areas have less deaths certified by medical doctors and lower quality of the mortality data (1). We have found a similar pattern in Ecuador, especially in the provinces in the Amazonic region and Esmeraldas in the Coastal region. Populations in provinces with disperse populations in rural areas could have greater difficulties finding healthcare personnel to certify deaths. One thing that can be noticed is that there is still an important percentage of deaths certified by medical doctors and assigned to garbage codes, especially in many provinces in the Andean region. This means that work is needed not only to improve access to healthcare professionals, but also to improve the quality of death certification by medical doctors. “

Reviewer #3:

This paper provides a much-needed analysis of national and subnational death registration and cause of death data in Ecuador.

1. A primary issue with the paper is regarding the death distribution methods used, their suitability for subnational analysis and the estimates of registration completeness produced. There should be more prominent acknowledgement of specific limitations of death distribution methods, including their lack of timeliness (an issue in Ecuador given their last 2 censuses were in 2001 and 2010) and issues with their accuracy as identified in the Murray et al 2010 paper (reference 11). In the discussion, the authors acknowledge issues regarding internal migration in Ecuador, however the reader does not get a sense of the extent of this. Can the authors include some evidence of the extent of inter-provincial migration in Ecuador?

A methodological appendix describing the DDM methods used was included. In there we acknowledge the specific limitations of the methods and their applications. Regarding internal migration in Ecuador, we added one relevant citation (Royuela and Ordoñez, 2018) and the following text to the manuscript:
Discussion -

“Royuela & Ordoñez (2018) describe the internal migration patterns in the Provinces of Ecuador, using data from the national censuses since 1982. They describe important differences between provinces. For example, in 2010, percentages of net migration ranged from –8.7 (Bolivar) to 10.9 (Galápagos).”

2. There is also no mention of other methods to estimate completeness of registration, including capture-recapture methods and a more recently published empirical method (http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0197047). These should be mentioned in the Background.

We agree with the comment. Other methods have been included in the methods section as follows:

“DDMs are one of several types of methods to estimate mortality completeness that also include capture-recapture methods, comparison of mortality with other estimates, and a newly developed empirical method that models key drivers of mortality in a population to estimate completeness (11).”

3. There is also no mention made of completeness of registration using estimates of total deaths according to the Global Burden of Disease and UN Population Division. Both sources estimate registration completeness of approximately 80-85% in Ecuador, significantly different to that estimated in this paper. Can the authors discuss their findings in the context of these other estimates (including the data sources the GBD and UN used), which indicate the uncertainty of completeness estimation in Ecuador.
We have included this in the discussion section as follows:

“Nevertheless, our results are different from the ones in other studies. For example, the UN population division estimates completeness to be a little higher (between 70 and 79%), and the GBD 2016 project estimates a mortality completeness over 80% for the study period (18,19). The differences found between the different sources can be explained by the use of diverse methods to estimate completeness and the high uncertainty of the methods. For example, the GGB project uses the three DDM methods as recommended by Murray et al. (2010) to estimate adult mortality. They then combine the results from the DDM methods with estimations of under-5 completeness for several time periods. When necessary, they use sibling histories too. Even after all this complex process, they still recognize that estimates may change between GBD revisions (20). Certainly, much work is needed to improve completeness estimation methods worldwide.”

4. It would be insightful if there were a description of death registration processes in Ecuador, and the implications of these findings for improving these processes. What happens when a person dies in a facility or in the community? What onus on people to register deaths? How are data transferred to the central authority? How is the cause of death ascertained? How can medical certification be improved?

A brief description of the death registration process is included in the methods section as follows:

“In Ecuador, by law, the declaration and registration of all deaths is mandatory. The mortality registry contains all death certificates completed in the country by health professionals or (in their absence): 1) police or civil authorities or 2) civil registration officials. Death certificates include sociodemographic information (such as age and sex), home and death addresses, and causes of death. The National Institute of Statistics and Censuses (INEC) uses the information in the certificates to determine the underlying causes of death. For the entire study period, cause of death information was coded using the International statistical classification of diseases and related health problems - 10th Revision (ICD10).”

We certainly believe a wider description of the system is needed, but no more information could be found from official sources or scientific literature.
We believe that the biggest problems with death registration completeness and quality are derived from lack of access to health services (completeness and quality) and poor training of professionals coding deaths (quality). We believe that the process of universalization of healthcare happening in the country since 2008 could be improving access to healthcare and thus improve quality and completeness of mortality. Moreover, professionals coding deaths certainly require more training to improve the quality of the registry.

5. The cause of death analysis should reference each of these publications on assessing registration data quality:

http://www.getinthepicture.org/sites/default/files/resources/Mortality%20statistics%20a%20tool%20to%20improve%20un


Both references have been included in the methods section.

6. The garbage code categories used are different from that of Naghavi et al 2010 - what is the justification for this?

Naghavi et al. (2010) describe the algorithms for redistributing deaths coded with garbage codes to 56 target causes of death. The authors obtained a garbage code list that redistributed to those causes of death. The global burden of disease projects use a wider categorization of garbage codes and target causes of death (235 in GBD 2010, 240 in GBD 2013 and 264 in GBD 2016 for example). Consequently, the lists of garbage codes have changed between the GBDs. For that reason, we compare that garbage code list with the Global Burden of Disease (GBD) 2013 and 2015 cause lists. This process allowed us to obtain a consistent list of GBD codes and garbage codes in which all deaths with non-garbage codes were assigned to a GBD code.

7. What are the specific causes of higher garbage codes in ages 80+ - is it senility for example?
We analysed the garbage codes in that age group and found that the principal garbage codes in that age group, for both men and women, were heart failure, senility and primary hypertension. We believe the three of them reflect the difficulties of correctly diagnosing and coding causes of death in people with multiple chronic diseases. The following text was added in the discussion section to clarify this point:

“The age pattern of garbage codes was similar to that observed in other contexts, and could be explained by the difficulty of determining the cause of death accurately in older people (with multiple comorbidities) (26). For example, we analysed the principal garbage codes in people over 80 years old and found that heart failure, senility and essential hypertension accounted for approximately 40% of the garbage codes in that age group. Deaths coded as heart failure (ICD10 code I50) accounted for 17.21% of the garbage codes in women and 16.30% in men; senility (ICD10 code R54) accounted for 15.34% of garbage codes in women and 13.55% in men; and essential (primary) hypertension (ICD10 code I10) accounted for 10.54% of garbage codes in women and 9.76% in men.”

8. Were other thresholds of garbage code acceptability considered (other than the national level), to enable international comparison? Naghavi et al. 2010 states that garbage codes are 22% globally. This suggests the threshold used is too high.

At an initial moment, the global percentage was considered. Nevertheless, we dropped that idea when noticing that none of the provinces had garbage code percentages bellow that threshold for both sexes. That meant all areas would have been categorized as deficient. As the idea was to guide public policies in order to reduce inequalities, we decided to use a threshold based on the national percentage.

Other comments:

1. In the objectives there is mention of social inequalities by age and sex, however these are more appropriately described as demographic differences.

Our research group works under a social determinants of health framework. Under this framework, gender and age can be important factors in the creation of social inequalities in health, acting as axes of inequality. Sex as a variable can operate through gender relations and patterns; and relate to health inequalities. For that reason, we decided to keep mentioning social inequalities by age and sex.


2. I suggest only using one decimal place.

We have reviewed this item with the research team and decided to leave the two decimal points, as tables and graphs are homogeneous and understandable throughout the text.

3. Error in this sentence: "In both men and women, only one area was classified as acceptable (Pichincha/Santo Domingo for women and Azuay for women)" - last word should be "men"

The error was corrected. Thank you.