PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Clinical and Epidemiological Indicators and Spatial Analysis of Leprosy Cases in Patients Under 15 Years Old in an Endemic Area of Northeast Brazil: an ecological and time series study
AUTHORS	Santos, Márcio; Santos, Allan; Barreto, Aline; Souza, Mariana; Goes, Marco; Barreto Alves, Jose Antonio; Barreto, Ikaro; Silva, José-Rodrigo; Oliveira, Daniela; Araújo, Karina; Duthie, Malcolm; Jesus, Amélia

VERSION 1 - REVIEW

REVIEWER	Mathys Vanessa PhD Sciensano, Belgium
REVIEW RETURNED	17-May-2018

GENERAL COMMENTS	The topic is interesting. The study was well conducted. The manuscript is clear and well written.

REVIEWER	Lintang Dian Saraswati Diponegoro University, Indonesia
REVIEW RETURNED	20-Aug-2018

GENERAL COMMENTS	-for each variable, give sources of data and details of methods of measurement - Explain how missing data were addressed - Describe any efforts to address potential sources of bias

REVIEWER	Amal Mitra Jackson State University, USA
REVIEW RETURNED	24-Nov-2018

GENERAL COMMENTS	Figure 1 and Figure 2 are too small to review the data. Please submit revised versions of the figures.

REVIEWER	Susanna Cramb Cancer Council Queensland, Australia
REVIEW RETURNED	16-Dec-2018

GENERAL COMMENTS	The paper covers an interesting and important topic. The analyses conducted are relatively simple, but nonetheless useful. However,

	greater clarity in the aims, methods and interpretation of (some) results would be very helpful. 1. Please state the aims of the paper more clearly. They are stated in lines 90-92, but there is no mention of time series analysis, yet line 97 refers to conducting 'an ecological and time series' study. I would suggest breaking the aims into 3 parts: 1) Clinical and epidemiological indicators 2) Spatial analysis 3) Temporal analysis (Please feel free to further modify this – for instance, you might want to separate into 4 areas. Whichever way they are ordered, keeping the methods and results sub-sections in the same order would be easier to follow.) 2. Please also clearly state the details of the data obtained – in lines 104-105 you talk about analyzing leprosy data for those under 15 years old, yet your results present details on all ages as well as children. A clear statement is needed about the years, variables (including ages), and (preferably) any details on the completeness/coverage of data over the time period (e.g. whether it is consistent). Leprosy is obviously a notifiable disease – is this a legislative requirement? If so, how are details of cases collected? 3. Likewise, in lines 108 to 114 the collected variables are each mentioned, yet no location information is included here. Please clearly state is the household address was collected and made available to investigators, or if only an area was provided to investigators for the location. 4. More details on the municipalities would also be helpful for those less familiar with Brazil – e.g. provide the median population over the 75 areas for a specific year, and the population range. 5. More details on the kernel intensity estimator, including the bandwidth, are needed. Were any sensitivity tests conducted on the choice of bandwidth, etc? If so, please specify. 6. Brief details on the assigning of local Moran's I values to the quadrants and high/high etc would be helpful. Given the usefulness of the Moran scatterplot for identifying outliers, perhaps consider providing the scatterplot as a figure also? 7. What goodness of fit checks were performed on the kernel intensity estimator? 8. Would suggest using the same capitalisation when describing variables in equations as is in the equation. E.g. equation 1 has lower case y, yet in the text defining these terms, (lines 136-137) uppercase Y is used instead. Please keep these identical, and likewise for equation 2. 9. In the section Statistical Analysis (starting at line 159): Please provide a brief justification for performing descriptive analyses and testing for differences rather than using statistical regression. Also, please provide references for the tests used. 10. Would recommend using a joinpoint analysis for the time series analysis (free software can be used to run these, available at: https://surveillance.cancer.gov/joinpoint/). This allows for trends to

	change over time. Currently you are fitting a straight line across all time points, and this does not reflect the variation which you mention in the text (lines 182-184). Please also make it clear when reporting rates whether you are using crude or age-standardised rates (even ages 0-14 can be presented as age-standardised rates). 11. Figure 2A: Instead of reporting proportions of MB and PB, consider showing as stacked bar charts based on the number of PB and MB cases in each year. This would give an indication of the proportion, but is much more informative. 12. The results in Figure 3 (Moran's I versus kernel maps) seem to identify different areas. Some areas identified as high/high in Moran's I are identified as low in the kernel maps. Please discuss in the paper potential reasons as to why these conflict (i.e. red in Moran's I and green or blue in kernel). 13. Normally data and study methods would have some limitations associated with them, yet I could not find any discussion on the limitations of this study. The summary strengths and limitations mentions there can be under-notification of cases and missing data, yet this is not mentioned within the manuscript. Please add this in. Minor comments 14. Line 62: There is a full stop in the middle of a number (213.899) that should probably be a comma. 15. Line 150: mentions Y_i – this subscript should be j. 16. Lines 181-182: Please mention the units for the 0.76 and/or 1.2. 17. Line 202: Expect "PM" should be "PB" 18. Some of the methods are rather unclear, probably due to grammatical issues. Would particularly flag the following statements for further refining: lines 126-7: "…indicating cases agglomeration in a spatial distribution and a continuous surface from point data" line 132: "…the contiguity of autocorrelation." line 163: "…to analyze the normal distribution of data." line 166: "…tendency analyze by linear regression," 19. Please define all abbreviations used in a table underneath it.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name: Mathys Vanessa PhD

Please leave your comments for the authors below

Q1 - The topic is interesting.

The study was well conducted.

The manuscript is clear and well written.

Answer 1: We thank the reviewer for these comments.

Reviewer: 2

Reviewer Name: Lintang Dian Saraswati

Please leave your comments for the authors below

Q1- for each variable, give sources of data and details of methods of measurement

Answer 1: The source of all data from this study was the leprosy cases and the information of each individual case notified by the health centers of the municipalities to the SINAN (Information System on Notifiable Diseases) from the State of Sergipe, Brazil. SINAN is an important database of the Secretariat of Health of all States of Brazil, to report information about sociodemographic, clinical features and the address of notifiable diseases, such as leprosy. We clarified this information in the material and methods section (page 6).

Q2- Explain how missing data were addressed

Answer 2: In Brazil we have a specific normative that is an obligation to notify several diseases to SINAN, and leprosy is one of them. SINAN is an important database of the Secretariat of Health of all States of Brazil, to report information about sociodemographic, clinical features and the address of each diagnosed case. This source of data can present under notification because Leprosy is an asymptomatic disease and the active search would be important to detect more cases, but all diagnosed cases are reported to SINAN. In fact, one of the findings of this study is that the cases diagnosed by the exam of contacts are less severe than the patients that were forwarded from a primary clinic to a leprosy reference center. The limitation mentioned about missing data is not very important in the SINAN database in the case of the disease prevalence, but the complete information about the cases follow-up, such as degree of neurological disability at the end of treatment, leprosy reactions and treatment details, because it is a secondary database that depends on other doctors or nurses from the health care centers to fulfill the information. Despite this, those data reported high endemicity of leprosy cases in patients under 15 years old, and this study did not focus on patients follow-up. We have pointed that in the ‘Strengths and limitations’ (page 3) and in the discussion section of manuscript (page 17).

Q3 - Describe any efforts to address potential sources of bias

Answer 3: We did not do any efforts to address data bias, because the limitation mentioned about missing data is not very important in the SINAN database in the case of the disease prevalence, but the complete information about the cases follow-up, such as degree of neurological disability at the end of treatment, leprosy reactions and treatment details, and this study do not focus on these problems. We have addressed this point in the discussion section of manuscript (page 17).

Reviewer: 3

Reviewer Name: Amal Mitra

Please leave your comments for the authors below

Q1 - Figure 1 and Figure 2 are too small to review the data. Please submit revised versions of the figures.

Answer 1: We will resubmit the manuscript with better figures resolution and appropriate sizes.

Reviewer: 4

Reviewer Name: Susanna Cramb

Please leave your comments for the authors below

The paper covers an interesting and important topic. The analyses conducted are relatively simple, but nonetheless useful. However, greater clarity in the aims, methods and interpretation of (some) results would be very helpful.

1. Please state the aims of the paper more clearly. They are stated in lines 90-92, but there is no mention of time series analysis, yet line 97 refers to conducting 'an ecological and time series' study. I would suggest breaking the aims into 3 parts:

1) Clinical and epidemiological indicators

2) Spatial analysis

3) Temporal analysis

(Please feel free to further modify this – for instance, you might want to separate into 4 areas. Whichever way they are ordered, keeping the methods and results sub-sections in the same order would be easier to follow.)

Answer 1: We thank the reviewer for that important suggestion. We rewrote the objectives in the introduction section and organized it according to the methods and results sections (page 05).

2. Please also clearly state the details of the data obtained – in lines 104-105 you talk about analyzing leprosy data for those under 15 years old, yet your results present details on all ages as well as children. A clear statement is needed about the years, variables (including ages), and (preferably) any details on the completeness/coverage of data over the time period (e.g. whether it is consistent). Leprosy is obviously a notifiable disease – is this a legislative requirement? If so, how are details of cases collected?

Answer 2: The main objective of this study was to analyze leprosy data in those patients under 15 years old. However, we also compare those data with data in all ages and with the occurrence of physical disability, with a purpose to visualize the total prevalence together with the prevalence in children. Moreover, leprosy is a notifiable disease in Brazil, as a legislative requirement. All leprosy cases have to be notified by the SINAN, including information about social and demographic features of patients and their clinical forms. We have now included that information in the material and methods section (page 6).

3. Likewise, in lines 108 to 114 the collected variables are each mentioned, yet no location information is included here. Please clearly state is the household address was collected and made available to investigators, or if only an area was provided to investigators for the location.

Answer 3: The SINAN database has the patient address. We have now included that information in the material and methods section (page 6).

4. More details on the municipalities would also be helpful for those less familiar with Brazil – e.g. provide the median population over the 75 areas for a specific year, and the population range.

Answer 4: Sergipe is located on the coast of Northeast Brazil. The State has 75 municipalities and the capital is Aracaju. It has a population of 2,068,017 inhabitants. The median population per county was 27,573.56 in 2015. We have now included that information in the material and methods section (page 6).

5. More details on the kernel intensity estimator, including the bandwidth, are needed. Were any sensitivity tests conducted on the choice of bandwidth, etc? If so, please specify.

Answer 5: The kernel technique was applied to identify the intensity of the distribution of leprosy cases in the state of Sergipe. This is an appropriate data interpolation and smoothing technique for application in point location data. The point distribution was transformed into a smoothed surface and presented as a continuous map, representing different levels of case intensity. The amount of smoothing, that is, the width of the radius of influence was defined as 3,000 meters, since this value generated an adequate representation of the distribution of cases of leprosy in the municipalities, minimizing the overlapping bias or the occurrence of sub distribution patterns smoothed. We now included these information in the method section (pages 7 and 8).

6. Brief details on the assigning of local Moran's I values to the quadrants and high/high etc would be helpful. Given the usefulness of the Moran scatterplot for identifying outliers, perhaps consider providing the scatterplot as a figure also?

Answer 6: Spatial autocorrelation between leprosy rates was used to investigate whether the spatial distribution of the disease occurs randomly or follows some pattern of occurrence in space. The Moran Map was used to indicate the clusters and their relationship with the neighbors. This analysis verifies the existence of spatial dependence and risk patterns: Q1 (high/high) and Q2 (Low/Low), which indicate municipalities with similar values between their neighbors and Q3 (high/low) and Q4 (low/high) for municipalities with different values between their neighbors and no spatial association. A spatial proximity matrix obtained by the contiguity criterion was adopted. The level of significance was 5% and the Moran Global Index (I) varying between -1 and +1, representing the spatial autocorrelation of leprosy detection rate in the geographic space analyzed to identify spatial clusters and risk areas. Values between 0 and +1 indicate positive spatial autocorrelation (Q1 and Q2) and between -1 and 0 negative spatial autocorrelation (Q3 and Q4). We now included these information in the method section (page 9).

7. What goodness of fit checks were performed on the kernel intensity estimator?

Answer 7: We did not check the goodness of fit for Kernel analyzes. The amount of smoothing, that is, the width of the radius of influence was defined as 3,000 meters, since this value generated an adequate representation of the distribution of cases of leprosy in the municipalities, minimizing the overlapping bias or the occurrence of sub distribution patterns smoothed.

8. Would suggest using the same capitalisation when describing variables in equations as is in the equation. E.g. equation 1 has lower case y, yet in the text defining these terms, (lines 136-137) uppercase Y is used instead. Please keep these identical, and likewise for equation 2.

Answer 8: We corrected that in the text (pages 8 and 9).

9. In the section Statistical Analysis (starting at line 159): Please provide a brief justification for performing descriptive analyses and testing for differences rather than using statistical regression. Also, please provide references for the tests used.

Answer 9: We used now the Joinpoint program (version 4.0.4 - Surveillance Research, National Cancer Institute, USA) to evaluate the temporal trend for annual incidence of leprosy cases instead statistical regression, as required. We included this following information in the Material and Methods section, as required (pages 9 and 10).

10. Would recommend using a joinpoint analysis for the time series analysis (free software can be used to run these, available at: https://nam01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fsurveillance.cancer.gov%2Fjoinpoint%2F&data=02%7C01%7C%7C206331fefdb642aada5708d670c479f4%7C84df9e7fe9f640afb435aaaaaaaaaaaa%7C1%7C0%7C636820386489633050&sdata=OtJXSCy%2FlzvTyfCBPqRugCpLQ6Jq5h1z28e7oUW2%2FGU%3D&reserved=0). This allows for trends to change over time. Currently you are fitting a straight line across all time points, and this does not reflect the variation which you mention in the text (lines 182-184). Please also make it clear when reporting rates whether you are using crude or age-standardised rates (even ages 0-14 can be presented as age-standardised rates).

Answer 10: Thank you for your suggestion. We now analyzed the time series data using joinpoint software. We included this following information in the Material and Methods section, as required (page 10)

“In order to enable trend analysis, annual incidence of leprosy was calculated as dependent variables (y), and the years of the study period as the independent variables (x). Initially, trend analysis was performed with the Joinpoint program, version 4.0.4 (Surveillance Research, National Cancer Institute, USA). This program estimates the annual percentage change (APC) of a segmented linear regression (jointpont regression) and identifies inflection points. Each inflection point reflects changes in the increase or decline of leprosy rates. Poisson regression is used to determine the number of segments required to adequately explain the relationship between two variables. We considered the points of trend change that presented p-value < 0.05.”

11. Figure 2A: Instead of reporting proportions of MB and PB, consider showing as stacked bar charts based on the number of PB and MB cases in each year. This would give an indication of the proportion, but is much more informative.

Answer 11: We thank the reviewer for this suggestion. We changed figure 2A, as required.

12. The results in Figure 3 (Moran's I versus kernel maps) seem to identify different areas. Some areas identified as high/high in Moran's I are identified as low in the kernel maps. Please discuss in the paper potential reasons as to why these conflict (i.e. red in Moran's I and green or blue in kernel).

Answer 12: The maps present certain disagreements regarding the occurrence of leprosy cases in the state of Sergipe because they use distinct techniques of spatial analysis. The Kernel estimator produces a continuous surface, with densities calculated at all locations, based on total number of cases and no considering the geographical boundaries of the municipalities. The Kernel technique presents greater advantages to the quick visualization of areas that deserve attention, besides not being affected by political-administrative division, while the Moran technique, constructs maps considering the political-administrative divisions of the state and the clusters are based on the number of cases by the municipalities population rates. We have included this information in the discussion section (Page 16 and 17).

13. Normally data and study methods would have some limitations associated with them, yet I could not find any discussion on the limitations of this study. The summary strengths and limitations mentions there can be under-notification of cases and missing data, yet this is not mentioned within the manuscript. Please add this in.

Answer 13: We have now included that information in the Discussion section: page 17.

Minor comments

14. Line 62: There is a full stop in the middle of a number (213.899) that should probably be a comma.

Answer 14: We corrected that.

15. Line 150: mentions Y_i – this subscript should be j.

Answer 15: We corrected that.

16. Lines 181-182: Please mention the units for the 0.76 and/or 1.2.

Answer 16: We included that in the text.

17. Line 202: Expect "PM" should be "PB"

Answer 17: We corrected that..

18. Some of the methods are rather unclear, probably due to grammatical issues. Would particularly flag the following statements for further refining:

lines 126-7: "…indicating cases agglomeration in a spatial distribution and a continuous surface from point data"

line 132: "…the contiguity of autocorrelation."

line 163: "…to analyze the normal distribution of data."

line 166: "…tendency analyze by linear regression,"

Answer 18: We corrected these sentences in the manuscript.

19. Please define all abbreviations used in a table underneath it.

Answer 19: We have included all abbreviations in the tables.

VERSION 2 – REVIEW

REVIEWER	Susanna Cramb Cancer Council Queensland, Australia
REVIEW RETURNED	23-Feb-2019

GENERAL COMMENTS	The authors have been very thorough in addressing my comments, and there are only a few of their changes I would like to comment on. 1. In respect to my original point 8, regarding the symbols used in the equations: I can see you have made modifications, however, some were missed (e.g. the capital omega is used in the text (line 174), while it is lowercase omega in the equation (line 170), while others simply do not match (e.g. gamma is used in the equation on lines 154-156, yet the text mentions y on line 158. Would it be possible to simply copy the symbol used in the equation into the text where it is referred to, so that both the text and equations have exactly the same symbol, with exactly the same formatting? 2. Line 174 does define Z_i twice – should one of those be Z_j? 3. Thank you for using Joinpoint for the trend analyses. Could you also briefly state the constraints placed on the model, as this can influence the results. The key ones to mention are the minimum number of data points between a joinpoint and either end of the data series and the minimum number of data points between joinpoints, as well as the maximum number of joinpoints allowed.

	4. Lines 222-224: I would recommend ensuring the trends do follow your description of trends in the text though. It is fine to say incidence rates fluctuated due to low numbers, but rather than saying that rates increased then decreased, just report what joinpoint reports for trends. Note that by altering the number of minimum data points between joinpoints/end of data series, the trends can change. 5. You have specified "annual incidence", and sometimes "annual incidence rate", but I still could not find a statement as to whether the rates were age-standardised or crude. Please briefly insert details on the rate calculation, and if age-standardised, also details of the standard population used. (Please note that for the general population particularly, when examining changes over time, or comparing between areas, often using age-standardised rates are helpful to see how things are changing once removing the influence of population growth and ageing – but it really depends on your aims.) 6. Figure 2A is much better now. I noticed you did not create a stacked bar chart (which would have MB on top of PB, and so enable seeing the total number of leprosy cases at a glance), but it is fine if you prefer this version. Please modify the y-axis to remove the "(%)" as this is rather confusing. 7. Line 229-230: This mention of "no significant reduction" is confusing as it is a non-significant increasing trend – please modify. Minor: Line 205 – "le rates" should probably be “leprosy rates” Lines 370-371: Should this be "…number of cases divided by the municipalities…"?

VERSION 2 – AUTHOR RESPONSE

1. In respect to my original point 8, regarding the symbols used in the equations: I can see you have made modifications, however, some were missed (e.g. the capital omega is used in the text (line 174), while it is lowercase omega in the equation (line 170), while others simply do not match (e.g. gamma is used in the equation on lines 154-156, yet the text mentions y on line 158. Would it be possible to simply copy the symbol used in the equation into the text where it is referred to, so that both the text and equations have exactly the same symbol, with exactly the same formatting?

Answer 1: We thank the careful revision of the reviewer. We corrected all the symbols, as suggested.

2. Line 174 does define Z_i twice – should one of those be Z_j?

Answer 2: We corrected that in the text.

3. Thank you for using Joinpoint for the trend analyses. Could you also briefly state the constraints placed on the model, as this can influence the results. The key ones to mention are the minimum number of data points between a joinpoint and either end of the data series and the minimum number of data points between joinpoints, as well as the maximum number of joinpoints allowed.

Answer 3: The number of inflections used in the analysis was the result of models defined by the program itself, in order to allow the best representation of the trend, with the lowest number of inflection points. We applied the Grid Seacrh method. The minimum number of observations from a joinpoint to either end of the data were 3 and the minimum number of observations between two joinpoints were 4. The result showed growth (positive APC values), reduction (APC negative values) or maintenance (APC value equal to zero) of the trend throughout the historical series analyzed (2006-2014). A statistically significant trend, other than zero, was considered when the p-value was less than 0.05. We included this information in the new version of the manuscript.

We also included in the manuscript a more detailed description “The joinpoint regression provided the adjustment of a series of lines as well as their inflection points on a logarithmic scale by means of the annual trend test. To obtain the adjustment based on the best line of each analyzed segment, the Monte Carlo permutation method was used as a test of significance. From the definition of the follow-ups, the annual percentage change (APC) and the average annual percentage change (AAPC), with their respective 95% confidence intervals, were estimated and tested. If the occurrence of an inflection point with inverted direction was verified, the study periods were analyzed separately.”

4. Lines 222-224: I would recommend ensuring the trends do follow your description of trends in the text though. It is fine to say incidence rates fluctuated due to low numbers, but rather than saying that rates increased then decreased, just report what joinpoint reports for trends. Note that by altering the number of minimum data points between joinpoints/end of data series, the trends can change.

Answer 4: We have now rewrite that sentence in the text for: “The incidence of leprosy in children under 15 years has declined from 6.29 cases per 100,000 inhabitants in 2002, to 3.78 in 2015, confirmed by Joinpoint regression analyzes (APC = -5.3 and p-value < 0.05; Figure 1A and B).”

5. You have specified "annual incidence", and sometimes "annual incidence rate", but I still could not find a statement as to whether the rates were age-standardised or crude. Please briefly insert details on the rate calculation, and if age-standardised, also details of the standard population used. (Please note that for the general population particularly, when examining changes over time, or comparing between areas, often using age-standardised rates are helpful to see how things are changing once removing the influence of population growth and ageing – but it really depends on your aims.)

Answer 5: In fact, we calculated annual incidence rates for the general population, and for patients younger than 15 years the rates were age-standardised, and the standard population used was the population under 15 years, according to the data from the Brazilian Institute of Geography and Statistics (IBGE). We included this information in the manuscript (pg 9, lines 190 to 193).

6. Figure 2A is much better now. I noticed you did not create a stacked bar chart (which would have MB on top of PB, and so enable seeing the total number of leprosy cases at a glance), but it is fine if you prefer this version. Please modify the y-axis to remove the "(%)" as this is rather confusing.

Answer 6: Sorry if we forgot to modify the Y-axis legend. We modified that now.

7. Line 229-230: This mention of "no significant reduction" is confusing as it is a non-significant increasing trend – please modify.

Answer 7: We modified that in the text.

Minor:

Line 205 – "le rates" should probably be “leprosy rates”

Answer: We modified that in the text.

Lines 370-371: Should this be "…number of cases divided by the municipalities…"?

Answer: We modified that in the text.

VERSION 3 - REVIEW

REVIEWER	Susanna Cramb Cancer Council Queensland, Australia
REVIEW RETURNED	06-Apr-2019

GENERAL COMMENTS	The authors have done an excellent job in revising their paper, and only two minor modifications are suggested. 1. The additional information included in the manuscript for joinpoint in the methods still does not mention the model constraints, as I suggested. Could you please include the details you mention in your response to me, specifically: "The minimum number of observations from a joinpoint to either end of the data was 3 and the minimum number of observations between two joinpoints was 4." Could you also mention the maximum number of joinpoints allowed (this should be determined by the above specifications and the number of data points in your dataset, but it can be helpful to specify exactly what was used so that people know it wasn’t set at a lower number.) 2. The annual incidence rates were calculated for the general population – could this be changed to "the crude annual incidence rates were calculated for the general population", just to make the methods completely clear?

VERSION 3 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 4

Reviewer Name

Susanna Cramb

Institution and Country

Cancer Council Queensland, Australia

Please state any competing interests or state ‘None declared’:

None declared

Please leave your comments for the authors below The authors have done an excellent job in revising their paper, and only two minor modifications are suggested.

1. The additional information included in the manuscript for joinpoint in the methods still does not mention the model constraints, as I suggested. Could you please include the details you mention in your response to me, specifically: "The minimum number of observations from a joinpoint to either end of the data was 3 and the minimum number of observations between two joinpoints was 4." Could you also mention the maximum number of joinpoints allowed (this should be determined by the above specifications and the number of data points in your dataset, but it can be helpful to specify exactly what was used so that people know it wasn’t set at a lower number).

Answer 1: We thank you for your contribution that will make our manuscript clearer. As you suggested, we included in the manuscript, methods section, a more detailed description of the joinpoint analyze. “All of the models were run under the same specifications. The minimum number of observations from a joinpoint to either end of the data was 3 and the minimum number of observations between two joinpoints was 4. The maximum number of joinpoints allowed was 2.”, Page 10, lines 209 – 212.

2. The annual incidence rates were calculated for the general population – could this be changed to "the crude annual incidence rates were calculated for the general population", just to make the methods completely clear?

Answer 2: We have now rewritten that sentence in the text, as you suggested. Page 09, Line 189.