PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Retrospective Review of Tertiary and Neurosyphilis Cases in Alberta, 1973 to 2017
AUTHORS	Landry, Takaaki; Smyczek, Petra; Cooper, Ryan; Gratrix, Jennifer; Bertholet, Lindsay; Read, Ron; Romanowski, Barbara; Singh , AE

VERSION 1 - REVIEW

REVIEWER	Michael Marks London School of Hygiene and Tropical Medicine, UK
REVIEW RETURNED	04-Sep-2018

GENERAL COMMENTS	This is an interesting paper. In particular showing the correlation between infectious cases of syphilis and the resurgence of NS is particularly valuable. I have a few comments of which the major one is about the categorisation used especially for Late NS. I am not sure the distinction that the authors use between early/late neurosyphilis makes complete sense - I think it would make more sense to separate out 1) Early NS (defined as per authors) 2) Late but not Tertiary NS (i.e NS > 1year not Tabes/GPI) 3) Late Tertiary NS (Tabes & GPI) This is because I think many clinicians take late NS to == Tertiary NS; and then when reading this the authors report that late NS went up in the context of syphilis outbreaks - clearly here (I presume as the data is not shown) the authors mean late but not tertiary NS because, given the time it takes to develop Tertiary NS, Tertiary NS can not be the thing driving increases in NS at the time of an outbreak. As part of separating them out I would then present what proportion of "Late NS" was Late non-tertiary and what what Late-Tertiary, and look at patterns of these separately against time. I think a table showing Manifestations against early vs late NS would complement the material presented in the text of the results. The methods should explain how the authors defined an outbreak - what % increase over how long etc.

REVIEWER	Dennis Cordato Department of Neurophysiology, Liverpool Hospital, Sydney, Australia
REVIEW RETURNED	25-Sep-2018

GENERAL COMMENTS	This is an excellent, well designed and excellently written article that is worthy of publication. The authors present an interesting and relevant analysis of tertiary and neurosyphilis cases in Alberta and the context of these cases in relation to 3 outbreaks of syphilis during the study period. There are two minor typos - line 47/48 page 5 references 1,8 need to both be upper case and line 37, page 18, sentence doesn't need to begin with 'As well'

REVIEWER	Min Liu Department of Epidemiology and Biostatistics, School of Public Health, Peking University
REVIEW RETURNED	13-Feb-2019

GENERAL COMMENTS	This manuscript has its merits, as the author stated:“An important strength of our study was the consistent reporting of all cases with positive syphilis serology over the 44 year period by laboratories as well as active follow up of all cases by the provincial STI program.” However, there was some defects. The source of the data was not clear and the data quality is not described. This study did not use any statistical methods and the results were not very reliable.

REVIEWER	Yuzo Arima Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Japan
REVIEW RETURNED	19-Feb-2019

GENERAL COMMENTS	In the original article, “Retrospective Review of Tertiary and Neurosyphilis Cases in Alberta, 1973 to 2017”, Landry et al. provide a retrospective description of the reported tertiary and neurosyphilis cases in Alberta, Canada, based on the provincial surveillance data, from 1973-2017. Descriptive epidemiologic and clinical information are provided, with stratified presentation of symptomatic vs.asymptomatic and early vs. late stages of neurosyphilis. Key characteristics assessed were demographics, route of transmission (same sex vs. heterosexual), HIV status, clinical manifestation, and treatment method. Distributions were compared between early vs late neurosyhilis cases and also between asymptomatic and symptomatic cases, stratified by early vs. late stage. As much of syphilis surveillance focuses on primary and secondary stage syphilis because they represent recent infection (and most infectious stages) and hence are useful for assessing (proxy) incidence trends (and less prone to surveillance bias due to changes in testing activity of asymptomatic cases), it is true that there are limited data on other stages of syphilis. In addition, neurosyphilis represents a serious clinical outcome, and understanding what kind of individuals with what kind of

	manifestations are being detected are important for the clinical and public health sectors. Thus, the reviewer agrees that these descriptive data are informative and important to share. I have several comments which I describe in detail below. Abstract: - Objectives: Since most people use “prevalence” to mean point prevalence, I would avoid the term, since it is not a cross-sectional point prevalence. While it could be seen as a period prevalence measure of prevalent cases notified over the entire period, the results are not presented in this manner (the denominator would then be a mid-interval population, for instance), so it is safer to avoid this term (the authors can then also avoid mixing the terms rates with prevalence). I think it would be more accurate to report it as notifications or notification rate, as these are based on notified surveillance data (and are presented as such in Figures 1 and 2). - Methods should probably explicitly state that distributions were compared between early vs late neurosyhilis cases and asymptomatic vs. symptomatic cases (stratified by early vs. late stage), if this were the a priori planned assessment. - In the results, perhaps simply report as something like, “Relative to late neurosyphilis cases, early neurosyphilis cases were more likely to be …” I say this because the results are presented in a binary manner, and comparing the two groups, if one group were younger, the other group is necessarily older, if one group had more same sex partners, that means the other group had relatively less, etc. So rephrasing this sentence would make it more succinct. Background - Line 30-32: since syphilis is described as a notifiable disease in Alberta, it would be good to explicitly state as notification rate (rather than simply “rates”). - See comment regarding prevalence above. Methods - It would be useful for the readers if the reason for the specific dates selected for the assessment period are explicitly stated. The authors state that the serological testing method changed in September 2017 so I am assuming that was why the cutoff was in 2017 (although not sure why March was selected) but I could not guess why 1973 was selected as the start year. - I have reservations regarding the linear trend line (and statistical testing for it). Please see below under the Results section. - Were the p-values two-sided? It would be good to be explicit. Results - Does the statement “defined as an increase in cases of two standard deviations above the baseline” refer to the data presented in Figure 1? And for all syphilis notifications or infectious syphilis notifications? The authors should be clear what the “cases” refer to here. - I have reservations regarding Figure 2. First, the raw data are a lot more informative and as in the upper left figure, the linear trend line poorly fits the actual trend. Also in the lower right figure, even though the p-value is significant, the linear trend line is not representing the actual cyclical pattern (as the authors state, there

	were 3 outbreaks during the period). Also, while the authors state that “significant rises were seen during the outbreak periods (outbreak #2, p<0.001; Figure 2)”, it is clear from the lower left figure for “outbreak #2” that this trend line is being strongly influenced by the large increase in the third outbreak (and, such a statement would result from comparing the notification rate during the outbreak periods vs. the non-outbreak periods, which the linear trend line is not doing). Thus, I would not include this linear trend line (I think the raw data alone are fine; a moving average would be better than the linear trend line as it would smooth the fluctuations but still not lose so much information). I believe it would be more informative to plot the absolute number of cases along with the notification rate, given values on the y-axis (e.g. peak of just over 0.04 per 100,000)…the authors state that the current population is 4.3 million, and even if the population were smaller in the 1980s, there were only 254 cases over a period of more than four decades (a simple average would be <6 cases per year), so these are very small numbers when assessed annually; showing the absolute numbers will be useful for the reader for interpretation. - Last paragraph: “Asymptomatic late neurosyphilis cases were less likely to be treated with penicillin G (44.7%; n=21) as compared to late neurosyphilis (90.0%; n=63; p<0.001)”. I think the second group refers to the “symptomatic” late nuerosyhilis cases (from the table results). Discussion - The various study periods listed were a bit confusing (Abstract, Methods and Results (Table 2): 1973 to 2017 vs. Results (in text) and Discussion: 1975 to 2016). They should be consistent. - Regarding the long gap period, the authors state possible reasons: “a well-established and sustained prevention and control program for STIs in the province, emergence of HIV and the mass education that occurred during this time period”. Were there declines in other STIs during this period? If these are the hypothesized reasons, it seems likely that other STIs would also have declined/remained low as the behaviors/policies would act on all STIs (although a common bias such as reduction in STI surveillance activities could also explain such similar trends)… - I agree regarding the statement that an increase in neurosyhilis would increase concomitantly with an increase in infectious syphilis, under usual assumptions. Just as with congenital syphilis, while rare, if the denominator of infected persons increase, certain outcomes would be expected to increase in proportion. - Considering the time lag, that early neruosyhilis increased during the outbreak periods agrees with what is known about the natural history of syphilis and provides assurance that the surveillance system was detecting trends that would be expected were there a true increase in syphilis incidence. Unless I am missing something, the delayed rise in late neurosyphilis would also seem to be in agreement with what would be expected, since there would be a relative temporal delay in the onset. That is, I am not sure if it would be attributed to a surveillance artefact such as “heightened awareness and increased testing due to public health announcements” (congenital rubella syndrome is also known to have a delayed increase following an increase in rubella cases, since there is a temporal lag from infection in the mother to birth of the infant).

	- More frequent testing for syphilis among HIV-positive persons has been reported. That early neurosyphilis was associated with HIV positivity, along with the finding that asymptomatic status was also associated with HIV positivity among early neurosyphilis cases, appear to agree with more intensive testing among HIV positive individuals—confirming such practice with clinicians in the field could be valuable as it could increase the plausibility of the results. - It might be useful to conduct some sensitivity analysis by decade or outbreak periods to see if the associations observed for the entire period were consistent over time (in surveillance, "person" (or "place") information can change over time, and it would be assuring to know that the distributions observed in Table 2 are not modified over time). As the numbers would get small, I would not run any statistical tests but look at the directionality of the results and see if they are qualitatively similar. - Considering the relevance of these findings for practice, perhaps there could be a little discussion on also focusing on those who are being detected at a late stage (i.e. older, born outside of Canada, heterosexual); these individuals were also more likely to be reported at a symptomatic state (Table 2), and may be an important group for outreach (or for clinicians to be aware of). Strengths and Limitations - It seems a little contradictory to state that an important strength was the consistent reporting of all cases with positive serology while a limitation was changes in data collection practices over time (which is common in surveillance). Even if there were consistent reporting, factors upstream of reporting (e.g. testing policies/practices, healthcare accessibility/access behaviors) can affect notification trends. It would thus be important to be clear and specific with the language here regarding the strengths and limitations. - The retrospective application of the current case definition to all cases is also listed as a strength and a limitation ("possibility of inaccurate classification of cases"). Do the authors have concerns that such misclassification would have been differential for any of the variables assessed (that is, for instance, do the authors think older cases could have been more likely to be inaccurately classified than younger cases? If so, that type of differential bias could pose a concern regarding the findings…)?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Michael Marks, London School of Hygiene and Tropical Medicine, UK	
I have a few comments of which the major one is about the categorisation used especially for Late NS. I am not sure the distinction that the authors use between early/late neurosyphilis makes complete sense - I think it would make more sense to separate out 1) Early NS (defined as per authors) 2) Late but not Tertiary NS (i.e NS > 1year not Tabes/GPI)	The separation of cases by early and late are based on the provincial and national surveillance case definitions for our jurisdiction. The inclusion of a manifestations table, as suggested in the next comment, will address the proportion of cases with ataxia and cognitive impairment. Unfortunately, insufficient detail on the neurological findings was available for late

3) Late Tertiary NS (Tabes & GPI) This is because I think many clinicians take late NS to == Tertiary NS; and then when reading this the authors report that late NS went up in the context of syphilis outbreaks - clearly here (I presume as the data is not shown) the authors mean late but not tertiary NS because, given the time it takes to develop Tertiary NS, Tertiary NS cannot be the thing driving increases in NS at the time of an outbreak. As part of separating them out I would then present what proportion of "Late NS" was Late non-tertiary and what Late-Tertiary, and look at patterns of these separately against time.	cases with ataxia to determine if they met diagnostic criteria for tabes dorsalis or for late cases with dementia to determine if they met diagnostic criteria for general paresis of the insane; this sentence has been added to the Results section.
I think a table showing Manifestations against early vs late NS would complement the material presented in the text of the results.	Additional table created with manifestations.
The methods should explain how the authors defined an outbreak - what % increase over how long etc.	Text has been added to the methods section to define the term outbreak.
Reviewer 2: Dennis Cordato, Department of Neurophysiology, Liverpool Hospital, Sydney, Australia	
There are two minor typos - line 47/48 page 5 references 1,8 need to both be upper case and line 37, page 18, sentence doesn't need to begin with 'As well'.	Both references have been superscripted and ‘as well’ has been removed from the sentence.
Reviewer 3: Min Lui, Department of Epidemiology and Biostatistics, School of Public Health, Peking University	
The source of the data was not clear and the data quality is not described.	Text was added to the methods describing the provincial medical records that are kept for syphilis cases. In addition, the data quality was added as a limitation to the discussion
This study did not use any statistical methods and the results were not very reliable.	Statistical methods were refined according to suggestions from Reviewer 4.
Reviewer 4: Yuzo Arima, Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Japan	
Abstract:-Objectives: Since most people use “prevalence” to mean point prevalence, I would avoid the term, since it is not a cross-sectional point prevalence. While it could be seen as a period prevalence measure of prevalent cases notified over the entire period, the results are not presented in this manner (the denominator would then be a mid-interval population, for instance), so it is safer to avoid this term (the authors can then also avoid mixing the terms rates with prevalence). I think it would be more accurate to report it as notifications or notification rate, as these are based on notified	The term prevalence has been replaced with notification rate as suggested.

surveillance data (and are presented as such in Figures 1 and 2).	
Abstract - Methods should probably explicitly state that distributions were compared between early vs late neurosyphilis cases and asymptomatic vs. symptomatic cases (stratified by early vs. late stage), if this were the a priori planned assessment.	Text has been added to the methods section as suggested.
Abstract- In the results, perhaps simply report as something like, “Relative to late neurosyphilis cases, early neurosyphilis cases were more likely to be …” I say this because the results are presented in a binary manner, and comparing the two groups, if one group were younger, the other group is necessarily older, if one group had more same sex partners, that means the other group had relatively less, etc. So rephrasing this sentence would make it more succinct.	Text has been added to the methods section as suggested.
Background - Line 30-32: since syphilis is described as a notifiable disease in Alberta, it would be good to explicitly state as notification rate (rather than simply “rates”). - See comment regarding prevalence above.	The text has been modified as suggested for rates pertaining to Alberta data throughout the manuscript.
Methods - It would be useful for the readers if the reason for the specific dates selected for the assessment period are explicitly stated. The authors state that the serological testing method changed in September 2017 so I am assuming that was why the cutoff was in 2017 (although not sure why March was selected) but I could not guess why 1973 was selected as the start year.	Reasons for data collection period has been added. A correction was made to the change in serological testing method to September 2007.
Methods - I have reservations regarding the linear trend line (and statistical testing for it). Please see below under the Results section.	The linear trend line has been removed as recommended below.
Methods - Were the p-values two-sided? It would be good to be explicit.	Text has been added to indicate p-values were two-sided.
Results - Does the statement “defined as an increase in cases of two standard deviations above the baseline” refer to the data presented in Figure 1? And for all syphilis notifications or infectious syphilis notifications? The authors should be clear what the “cases” refer to here.	The outbreak definition only applies to infectious syphilis. A note has been added to Figure 1 for more clarity.

Results -I have reservations regarding Figure 2. First, the raw data are a lot more informative and as in the upper left figure, the linear trend line poorly fits the actual trend. Also in the lower right figure, even though the p-value is significant, the linear trend line is not representing the actual cyclical pattern (as the authors state, there were 3 outbreaks during the period). Also, while the authors state that "significant rises were seen during the outbreak periods (outbreak #2, p<0.001; Figure 2)", it is clear from the lower left figure for "outbreak #2" that this trend line is being strongly influenced by the large increase in the third outbreak (and, such a statement would result from comparing the notification rate during the outbreak periods vs. the non-outbreak periods, which the linear trend line is not doing). Thus, I would not include this linear trend line (I think the raw data alone are fine; a moving average would be better than the linear trend line as it would smooth the fluctuations but still not lose so much information). I believe it would be more informative to plot the absolute number of cases along with the notification rate, given values on the y-axis (e.g. peak of just over 0.04 per 100,000)…the authors state that the current population is 4.3 million, and even if the population were smaller in the 1980s, there were only 254 cases over a period of more than four decades (a simple average would be <6 cases per year), so these are very small numbers when assessed annually; showing the absolute numbers will be useful for the reader for interpretation.	Figure 2 has been redone to plot absolute number of cases and notification rate over time using a dual axis.
Results - Last paragraph: "Asymptomatic late neurosyphilis cases were less likely to be treated with penicillin G (44.7%; n=21) as compared to late neurosyphilis (90.0%; n=63;p<0.001)". I think the second group refers to the "symptomatic" late neurosyphilis cases (from the table results).	Thank you, symptomatic has been added.
Discussion - The various study periods listed were a bit confusing (Abstract, Methods and Results (Table 2): 1973 to 2017 vs. Results (in text) and Discussion: 1975 to 2016). They should be consistent.	Data was available for the chart review of tertiary and neurosyphilis beginning in 1973 to March 2017; however, surveillance data with staging of all syphilis cases was only available from 1975 to 2016 (data presented in Figure 2).

Discussion -Regarding the long gap period, the authors' state possible reasons: "a well-established and sustained prevention and control program for STIs in the province, emergence of HIV and the mass education that occurred during this time period". Were there declines in other STIs during this period? If these are the hypothesized reasons, it seems likely that other STIs would also have declined/remained low as the behaviors/policies would act on all STIs (although a common bias such as reduction in STI surveillance activities could also explain such similar trends)…	Rates of gonorrhea and chlamydia also declined during this gap, until 1998, when notification rates began to climb to present conditions. A sentence summarizing this has been added to the discussion.
Discussion I agree regarding the statement that an increase in neurosyhilis would increase concomitantly with an increase in infectious syphilis, under usual assumptions. Just as with congenital syphilis, while rare, if the denominator of infected persons increase, certain outcomes would be expected to increase in proportion. Considering the time lag, that early neruosyhilis increased during the outbreak periods agrees with what is known about the natural history of syphilis and provides assurance that the surveillance system was detecting trends that would be expected were there a true increase in syphilis incidence. Unless I am missing something, the delayed rise in late neurosyphilis would also seem to be in agreement with what would be expected, since there would be a relative temporal delay in the onset. That is, I am not sure if it would be attributed to a surveillance artefact such as "heightened awareness and increased testing due to public health announcements" (congenital rubella syndrome is also known to have a delayed increase following an increase in rubella cases, since there is a temporal lag from infection in the mother to birth of the infant).	A statement has been added to the discussion together with a reference (Golden, 2003) to account for the identification of late neurosyphilis cases as soon as two years post-infection.
Discussion More frequent testing for syphilis among HIV-positive persons has been reported. That early neurosyphilis was associated with HIV positivity, along with the finding that asymptomatic status was also associated with HIV positivity among early neurosyphilis cases, appear to agree with more intensive testing among HIV positive individuals—confirming such practice with	A sentence and reference have been added to the discussion to indicate that more frequent testing for syphilis in HIV positive persons likely occurred as per DHHS guidelines (cited reference).

clinicians in the field could be valuable as it could increase the plausibility of the results.	
Discussion It might be useful to conduct some sensitivity analysis by decade or outbreak periods to see if the associations observed for the entire period were consistent over time (in surveillance, "person" (or "place") information can change over time, and it would be assuring to know that the distributions observed in Table 2 are not modified over time). As the numbers would get small, I would not run any statistical tests but look at the directionality of the results and see if they are qualitatively similar.	The suggested sensitivity analysis was completed and reported in the results section.
Discussion Considering the relevance of these findings for practice, perhaps there could be a little discussion on also focusing on those who are being detected at a late stage (i.e. older, born outside of Canada, heterosexual); these individuals were also more likely to be reported at a symptomatic state (Table 2), and may be an important group for outreach (or for clinicians to be aware of).	A sentence has been added to the discussion to highlight this concern
Strengths and Limitations - It seems a little contradictory to state that an important strength was the consistent reporting of all cases with positive serology while a limitation was changes in data collection practices over time (which is common in surveillance). Even if there were consistent reporting, factors upstream of reporting (e.g. testing policies/practices, healthcare accessibility/access behaviors) can affect notification trends. It would thus be important to be clear and specific with the language here regarding the strengths and limitations.	Text has been modified from data collection practices to testing policies and practices, as well as changes in social norms.
Strengths and Limitations -The retrospective application of the current case definition to all cases is also listed as a strength and a limitation ("possibility of inaccurate classification of cases"). Do the authors have concerns that such misclassification would have been differential for any of the variables assessed (that is, for instance, do the authors think older cases could have been more likely to be inaccurately classified than younger cases? If so, that type of differential bias could pose a concern regarding the findings…)?	Although the data quality improved over time, the authors (AS and PS) who reviewed the cases, felt that that sufficient information was available to accurately classify the majority of cases. In only one case did the 2 reviewers feel that insufficient information was available to classify the case into early versus late neruosyphilis; this case was classified as unknown duration.

VERSION 2 – REVIEW

REVIEWER	Michael Marks London School of Hygiene & Tropical Medicine
REVIEW RETURNED	19-Mar-2019

GENERAL COMMENTS	I am satisfied that the authors have responded to the comments

REVIEWER	Yuzo Arima Infectious Disease Surveillance Center, National Institute of Infectious Diseases, Japan
REVIEW RETURNED	05-Apr-2019

GENERAL COMMENTS	The reviewer appreciates the revisions to the original article by Landry et al., "Retrospective Review of Tertiary and Neurosyphilis Cases in Alberta, 1973 to 2017". The reviewer agrees that these descriptive data are informative and important to share to the public health and medical community. I only have a few minor comments/suggestions, listed below. Abstract: - This is a minor point, but I had suggested in the first review that there is probably no need to repeat the same point twice regarding the early vs. late neurosyphilis cases (as stated in the Results line 187-189 in revised version). What I meant (my apologies if it was not clear) was to simply state as, "Relative to late neurosyphilis cases, early neurosyphilis cases were more likely to be younger, Caucasian, born in Canada, HIV positive and reporting same sex partners." There is no need to repeat the same findings as "while late neurosyphilis cases were more likely to be older, born outside of Canada and less likely to report same sex partners". Since you are comparing two groups (early vs. late neurosyphilis), if one group is younger, the other group is necessarily older, if one group had more same sex partners, that means the other group had relatively fewer, etc. It is confusing to the reader when the same information is presented twice. Methods - Regarding my suggested optional sensitivity analysis, I am afraid what I meant did not come across. Since this an assessment over such a long time period, my suggestion was to check to see if the reported key findings (e.g. that early neurosyphilis cases were more likely to be younger, Caucasian, born in Canada, HIV positive and reporting same sex partners than late neurosyphilis cases) were true over different time periods, such as prior to 2000 (outbreak 1) and in the 2000s (outbreaks 2 and 3). Sometimes aggregate summaries can dilute or hide important differences, so stratification can be a useful tool to see if the overall summary is true when disaggregated (for instance, was reporting same sex partners associated with early neurosyphilis prior to 2000 as well as in the 2000s?). This was an optional suggestion and I leave it to the authors and the editor to decide whether to consider this assessment or to simply leave it out. Results

	- Thank you for clarifying the definition regarding an outbreak, that they are restricted to infectious syphilis cases; I would suggest that the definition of baseline also be included, per standard surveillance practice—I say this because baseline values can depend on the number of years used for the calculation, whether it is based on all weeks of the year or for the same calendar week, etc. Excuse me if I am missing something, but I also cannot seem to see the note that the authors refer to ("a note has been added to Figure 1 for more clarity."). - Thank you for considering my suggestion and reconstructing Figure 2; the situation is a lot easier to comprehend now. It would be good to have the Figure 2 title also updated (unchanged, with "linear trend line"). Results and Discussion - Please see my comment above regarding the optional sensitivity analysis I had suggested (that the overall key findings were true over different time periods). The reviewer did not mean to ask about the distribution of a variable among cases (e.g. that cases reporting same sex partners consistently made up the highest proportion of cases, as stated in the Results), but rather if the key reported findings were true regardless of time (e.g. cases reporting same sex partners were associated with early neurosyphilis cases, whether prior to 2000 or afterwards). I was simply suggesting some disaggregated assessment since the study covered more than 40 years of time; it would be reassuring to see that the key findings from the entire period were also true when stratified over time periods (e.g. pre 2000 period and the 2000s period, just as an example). This was an optional suggestion and I leave it to the authors and the editor to decide whether to consider this assessment or to simply leave it out.

VERSION 2 – AUTHOR RESPONSE

Response to Reviewer 4.

The reviewer appreciates the revisions to the original article by Landry et al., "Retrospective

Review of Tertiary and Neurosyphilis Cases in Alberta, 1973 to 2017". The reviewer agrees that these descriptive data are informative and important to share to the public health and medical community. I only have a few minor comments/suggestions, listed below.

Abstract:

- This is a minor point, but I had suggested in the first review that there is probably no need to repeat the same point twice regarding the early vs. late neurosyphilis cases (as stated in the Results line 187-189 in revised version). What I meant (my apologies if it was not clear) was to simply state as, "Relative to late neurosyphilis cases, early neurosyphilis cases were more likely to be younger, Caucasian, born in Canada, HIV positive and reporting same sex partners." There is no need to repeat the same findings as "while late neurosyphilis cases were more likely to be older, born outside of Canada and less likely to report same sex partners". Since you are comparing two groups (early vs. late neurosyphilis), if one group is younger, the other group is necessarily older, if one group had

more same sex partners, that means the other group had relatively fewer, etc. It is confusing to the reader when the same information is presented twice.

Thank you for the suggestion, the text “while late neurosyphilis cases were more likely to be older, born outside of Canada and less likely to report same sex partners” has been deleted in the abstract.

Methods

- Regarding my suggested optional sensitivity analysis, I am afraid what I meant did not come across. Since this an assessment over such a long time period, my suggestion was to check to see if the reported key findings (e.g. that early neurosyphilis cases were more likely to be younger, Caucasian, born in Canada, HIV positive and reporting same sex partners than late neurosyphilis cases) were true over different time periods, such as prior to 2000 (outbreak 1) and in the 2000s (outbreaks 2 and 3). Sometimes aggregate summaries can dilute or hide important differences, so stratification can be a useful tool to see if the overall summary is true when disaggregated (for instance, was reporting same sex partners associated with early neurosyphilis prior to 2000 as well as in the 2000s?). This was an optional suggestion and I leave it to the authors and the editor to decide whether to consider this assessment or to simply leave it out.

Thank you for the further clarification. The previous edits have been removed from the manuscript.

A sensitivity analysis was completed disaggregating the cases by prior to 2000 and 2000+; however due to the small number (n=19) of early neurosyphilis prior to 2000, the authors chose to exclude early neurosyphilis. Several factors impacted the late neurosyphilis cases during the 2000+ period including the change to reverse sequence syphilis screening in 2007 and two infectious syphilis outbreaks leading to additional diagnoses of late neurosyphilis, which we have outlined in the paper as major events. Therefore, we found the time periods to be distinct and expected, with only age as a consistent univariate result across the two time periods.

We have included a sentence in the methods to explain our consideration of the sensitivity analysis.

Results

- Thank you for clarifying the definition regarding an outbreak, that they are restricted to infectious syphilis cases; I would suggest that the definition of baseline also be included, per standard surveillance practice—I say this because baseline values can depend on the number of years used for the calculation, whether it is based on all weeks of the year or for the same calendar week, etc.

Text has been added to the methods to define baseline as the previous 5 year quarterly average.

Excuse me if I am missing something, but I also cannot seem to see the note that the authors refer to ("a note has been added to Figure 1 for more clarity.").

On the bottom of Figure 1 a note reads: "Outbreaks pertain to infectious syphilis".

- Thank you for considering my suggestion and reconstructing Figure 2; the situation is a lot easier to comprehend now. It would be good to have the Figure 2 title also updated (unchanged, with "linear trend line").

Thank you, the title of the figure has been updated in the manuscript text.

Results and Discussion

- Please see my comment above regarding the optional sensitivity analysis I had suggested (that the overall key findings were true over different time periods). The reviewer did not mean to ask about the distribution of a variable among cases (e.g. that cases reporting same sex partners consistently made up the highest proportion of cases, as stated in the Results), but rather if the key reported findings were true regardless of time (e.g. cases reporting same sex partners were associated with early neurosyphilis cases, whether prior to 2000 or afterwards). I was simply suggesting some disaggregated assessment since the study covered more than 40 years of time; it would be reassuring to see that the key findings from the entire period were also true when stratified over time periods (e.g. pre 2000 period and the 2000s period, just as an example). This was an optional suggestion and I leave it to the authors and the editor to decide whether to consider this assessment or to simply leave it out

Thank you for the additional details, please see above for our reply.