Article details: 2013-0087	
Title	Chronic hepatitis C associated hospitalizations: a birth cohort analysis with implications for screening
Authors	Dena L. Schanzer, MSc; Dana Paquette, PhD; Lisa M. Lix, PhD
Reviewer 1	Fiona Kouyoumdjian
Institution	University of Toronto Dalla Lana School of Public Health
General comments	Thank you for the opportunity to review this paper. This is an important paper and I think the methods and conclusions are sound. However, the writing is at times confusing, and further work is needed to clarify the presentation of information and to improve the overall flow. Here is a list of detailed comments: (please note that P=paragraph) Introduction P1-I would suggest specifying that the CHMS looks at the prevalence in the general population, and does not capture high risk groups such as IDUs, so likely underestimates true prevalence. P3- Given the general audience, it would be helpful to provide some data on the frequency of clearance (or a range) of HCV, etc. P3- I suggest that you add references for final sentence re: factors associated with progression. P4- I don't follow the logic in the sentence beginning with: "This long delay.." Methods Analysis- sentence starting with "Next rates…" does not read clearly. -Why did you use the 1970-1974 birth cohort as the reference population? Please provide this information. -line 152: I don't understand what this means: "though with the methodology used same-age rate ratios would not be sensitive to this choice of age"- can you clarify this? Results -I would suggest writing out chronic hepatitis C- associated liver disease throughout the text instead of using acronyms such as CHC-LD, which will not be familiar to a large proportion of the readership. -P1- I would be interested to know the relative size of each cohort, e.g. when you are describing the proportion of hospitalizations and fatalities in each cohort. -P1- Define "baby boom cohort" again here, i.e. specify which years of birth you are including. -P1- I think you should cut out the second part of the last sentence, i.e. "though by 2010…"- I don't see this conflicting with the first clause. Trends section- -To make it clear you are talking about birth year and not year, I think you should change "Statistically significant increases in the number of CHC-LD associated hospitalizations over the study period were detected for birth cohorts from 1950-54 to 1975-79…" to "…were detected for birth cohorts born between 1950 and 1979." -You describe the changes in numbers but not changes in numbers per cohort size. I think you should add numbers per cohort size. Rate Ratios section -The subtitle is mislabeled as "ratio ratios." -You should provide a definition of the term "morbidity burden" in the Methods section; you use this term in the Abstract and Results sections without specifying a definition. -I question whether same-age ratios provide a relative estimate of prevalence, since there may be other significant factors which affect progression to disease, access to hospitalization, etc. in various cohorts, e.g. alcohol use. -This paragraph seems to mix Methods, Results, and Interpretation, e.g. I think that "It is uncertain whether…" belongs in Interpretation, and "The estimated rate ratios at 75…" would fit better in Methods or Interpretation. Interpretation -P1- In the first sentence, you should specify significantly elevated relative to what? -P1- I think you should specify why "It is too early to fully assess the relative CHC prevalence

and disease burden for persons born after 1974…"
-P1-I don't understand the last sentence.
-P2- are you suggesting that a decrease in incident infections explains the difference between this study and prior studies? Or less diagnosis? I find this unclear, including what you mean by "the combined effects of aging." Could the methods used to generate estimates of increases in rates be a cause of the differences in your results?
-P2-I think there is an error in this sentence, which makes it hard to understand the point: "…CHC patient populations have already started to decline…" It is unclear what this sentence means.
-P3- I would specify NHANES instead of saying "an American household survey," given how well known and large this study is.
-P3- In the first sentence, I don't think you should set up the two statements as contrasting (you say "whereas"). In fact, the findings from your study on burden and the NHANES study estimates of relative prevalence are consistent (e.g. not significantly different, as per the CIs of your estimate).
-P3- I would suggest an introductory sentence to make the flow better from P2 to P3, and to clarify the focus of P3.
-P3- How do you "combine" your results with those of the CHMS? I don't understand this. I also think you need to explicitly state whether/how you would expect your burden data would to indicate prevalence; I don't think you are providing "estimates of relative prevalence," which you refer to.
-I think you miss the opportunity to talk about the strength of the data source, i.e. capturing a large proportion of all hospitalizations over several years.

Limitations
-P1- What do you mean that the "source of the HCV infection is also known to influence the all-cause mortality rate"? Do you mean the mechanism of infection?
-I think you should acknowledge the lack of knowledge re: incidence trends in each age cohort going forward.

Table 1
-I think you should cut out column for statistical significance since you provide the p values
-I would make it clear that the 95% CIs and p values refer to the average annual % change, e.g. by spreading those words across the three columns

Table 2
-I don't understand why some of the 95% CIs only have one number.
-I don't understand what the columns represent, e.g. what is the difference between 5 years and 2015? You have indicated that the data for the 5 years, 10 years, and until age 90 columns are from regression models. Where are the data for the 2015, 2020 and at age 75 columns from?
-As noted above, need to specify why 1970-1974 is the reference population.

Table 3
-How do you define the categories of relative disease burden? i.e. what does high vs. low mean? I think you need to define these categories and explain how you made them.

Figure 1
-I find it hard to match the colours in the Figure with those in the legend. You should reformat this.
-I don't think that the points of each cohort should be connected- what would those lines represent?

Figure 2
-As with Figure 1, what do lines connecting points represent? I would reconsider these lines.

Reviewer 2	Wendong Chen
Institution	University of Toronto, Toronto Health Economics and Technology Assessment Collaborative
General comments	1. Study design • It is difficult to accurately estimate chronic hepatitis C associated with hospitalization using the design in this study because there were no appropriate birth cohorts without chronic hepatitis C as control for adjustment of strong confounding effects associated with age and patient baseline comorbidities. • This study assumed that younger birth cohorts had similar hospitalization trends as older birth cohorts. Because the transmission mode of hepatitis C virus in Canada has been changed

substantially, older birth cohort likely acquired hepatitis C through blood transfusion or blood products but younger birth cohorts were mainly infected by the virus through drugs use. Thus, differences in transmission model likely make patients have different social economic status and underline comorbidities that could significantly affect disease progression and future health resource utilization.

• **The defined liver**-related disease in this study (K70 to 77, R18) might not accurately reflect the hospitalizations caused by hepatitis C related complications. For example, K70, 71, and 75 have nothing to do with chronic hepatitis C. Thus, this study likely overestimated the hospitalizations caused by CHC.

2. Analysis
• **This study only built the relationship between hospitalization and year without taking into** account any other factors that could contribute to hospitalization related to HCV. I really doubt this formula would be able to accurately predict the trend of HCV-related hospitalization. At least the study need to stratify hospitalizations by the complications caused by chronic hepatitis C. In addition, the study needs some kind evaluation approach to demonstrate the prediction model is reliable.

Author response	
	Reviewer 1: Comments to the Author Thank you for the opportunity to review this paper. This is an important paper and I think the methods and conclusions are sound. However, the writing is at times confusing, and further work is needed to clarify the presentation of information and to improve the overall flow. Here is a list of detailed comments: (please note that P=paragraph) Introduction 1. P1-I would suggest specifying that the CHMS looks at the prevalence in the general population, and does not capture high risk groups such as IDUs, so likely underestimates true prevalence. Author response: The text has been modified as suggested [A discussion of this limitation of household surveys has been added to the introduction section and interpretation section]. 2. P3- Given the general audience, it would be helpful to provide some data on the frequency of clearance (or a range) of HCV, etc. Author response: The text has been modified as suggested [Introduction: estimated at 15% by the American National Health and Nutrition Examination Survey (NHANES)]. 3. P3- I suggest that you add references for final sentence re: factors associated with progression. Author response: The additional reference has been added (repeated from above). 4. P4- **I don't follow the logic in the sentence beginning with: "This long delay.."** Author response: The referenced study by Myers and colleagues used data similar to the data used in our study to describe trends in hepatitis C hospitalization rates. Though the study was for an earlier period when the rate of increase in hepatitis C where higher, they did not look at trends by birth cohort. The sentence in the introduction section has been reworded [Introduction P4: Significant increases in liver-related hospitalization and mortality rates in Canada[6] have been attributed to HCV, though trends by birth cohort has not been described previously.] as well as the sentence referencing the same study in the interpretation section [comparison with other studies: The higher rates of increase found in the earlier studies are similar to rates of increase estimated in this study for persons aged 30-45 (Table 1), an age range that corresponds to the age of persons born in the 1950s and 1960s in the earlier studies.] Methods 5. Analysis- **sentence starting with "Next rates…" does not read clearly.**

Author response: The sentence has been corrected [now in the Appendix].

6. Why did you use the 1970-1974 birth cohort as the reference population? Please provide this information.

The number of annual hospitalizations is too small to use the either the oldest or youngest birth cohort as a reference. The 1970-74 birth cohort was chosen as the reference as a comparison with this birth cohort is an important consideration for screening recommendations and for comparability of the same-age ratios with prevalence ratios used in the development of the CDC guidelines.
The text has been modified to explain the choice [methods section in appendix, Table 2 footnote, and legend for Figure 2].

7. line 152: I don't understand what this means: "though with the methodology used same-age rate ratios would not be sensitive to this choice of age"- can you clarify this?

Author response:

The analysis sub-section has been moved to appendix 1 as suggested by the editor. The sentence has been reworded to inform the reader that the simulated projections by birth cohort are proportional as illustrated in Figure 3b.

Results
8. I would suggest writing out chronic hepatitis C- associated liver disease throughout the text instead of using acronyms such as CHC-LD, which will not be familiar to a large proportion of the readership. [Editors' note: please see comment on acceptable abbreviations in the Editors' comments.]

Author response: The abbreviation CHC-LD has been replaced with 'chronic HCV and liver disease'.

9. P1- I would be interested to know the relative size of each cohort, e.g. when you are describing the proportion of hospitalizations and fatalities in each cohort.

Author response: We have added columns specifying the number of chronic HCV and liver disease-associated hospitalizations by birth cohort in 2010/11 and the population of each birth cohort to Table 1.

10. P1- Define "baby boom cohort" again here, i.e. specify which years of birth you are including.

Author response: The baby boom cohort includes to persons born between 1945 and 1964. The sentence has been modified to include this information.

11. P1- I think you should cut out the second part of the last sentence, i.e. "though by 2010…"- I don't see this conflicting with the first clause.

Author response: The sentence has been revised. [Though numbers are smaller in younger cohorts, by 2010/11 the number of hospitalizations among the 1965-69 birth cohort had already increased to levels seen 5 years earlier in the 1960-64 birth cohort (Figure 1).]

Trends section-
12. To make it clear you are talking about birth year and not year, I think you should change "Statistically significant increases in the number of CHC-LD associated hospitalizations over the study period were detected for birth cohorts from 1950-54 to 1975-79…" to "…were detected for birth cohorts born between 1950 and 1979."

Author response: The phase 'for birth cohorts from 1950-54 to 1975-79' has been changed to 'for 5-year birth cohorts corresponding to persons born between 1950 and 1979' as suggested.

13. You describe the changes in numbers but not changes in numbers per cohort size. I think you should add numbers per cohort size.

Author response: A column specifying the number of chronic HCV and Liver Disease-associated Hospitalizations in 2010 has been added to Table 1.

Rate Ratios section
14. The subtitle is mislabeled as "ratio ratios."

Author response: The subtitle has been corrected.

15. You should provide a definition of the term "morbidity burden" in the Methods section; you use this term in the Abstract and Results sections without specifying a definition.

Author response: The term 'morbidity' has been replaced by 'hospitalization'.

16. I question whether same-age ratios provide a relative estimate of prevalence, since there may be other significant factors which affect progression to disease, access to hospitalization, etc. in various cohorts, e.g. alcohol use.

Author response: The manuscript has been revised as suggested. The description of the same-age ratios has been revised and the term 'prevalence' is used only in reference to the household surveys. We still compare this result in the interpretation section with the anti-body prevalence ratio from the US NHANES study, as this information was an important consideration for the development of a screening recommendation for Canada and comparable data is not currently available.

17. This paragraph seems to mix Methods, Results, and Interpretation, e.g. I think that "It is uncertain whether…" belongs in Interpretation, and "The estimated rate ratios at 75…" would fit better in Methods or Interpretation.

Author response: The phase 'it is uncertain' has meant to describe the same-age ratios and confidence intervals from Table 2.

The reference to Table 2 has been modified to indicate this. [same-age rate ratios and confidence intervals, Tables 2].

Interpretation
18. P1- In the first sentence, you should specify significantly elevated relative to what?

Author response: We meant to indicate that the rates were significantly elevated compared to rates for most other birth cohorts. The text has been revised as suggested.

19. P1-I think you should specify why "It is too early to fully assess the relative CHC prevalence and disease burden for persons born after 1974…"

Author response: The text has been revised. [It is too early to fully assess the relative HCV prevalence and disease burden for persons born after 1974… as the number of admissions with liver disease is still small.]

20. P1-I don't understand the last sentence.

Author response:

The sentence has been removed.

21. P2- are you suggesting that a decrease in incident infections explains the difference between this study and prior studies? Or less diagnosis? I find this unclear, including what you mean by "the combined effects of aging." Could the methods used to generate estimates of increases in rates be a cause of the differences in your results?

Author response: The higher trends of the earlier studies can be explained in part by the younger age of the birth cohorts with the highest exposure to hepatitis C. Annual increases of 15-30% are in agreement with our estimated rates of increase for persons aged 30-45 (Table 1).

The sentence has been reworded accordingly. [The higher rates of increase found in the earlier studies are similar to rates of increase estimated in this study for persons aged 30-45 (Table 1), an age range that corresponds to the age of persons born in the 1950s and 1960s in the earlier studies.]

22. P2-I think there is an error in this sentence, which makes it hard to understand the point: "...CHC patient populations have already started to decline..." It is unclear what this sentence means.

Author response: The sentence has been reworded to clarify that the number of chronic hepatitis C patients born before 1945 (aged 64+ in 2008) has already started to decline. [One difference is noted, we did not detect a statistically significant decline in the number of hospitalizations until age 80, while Zalesak and colleagues[21], found that the number of HCV patients with severe liver disease who were born before 1945 had already started to decline in 2008 (age 64+).]

23. P3- I would specify NHANES instead of saying "an American household survey," given how well known and large this study is.

Author response: The sentence has been revised as suggested.

24. P3- In the first sentence, I don't think you should set up the two statements as contrasting (you say "whereas"). In fact, the findings from your study on burden and the NHANES study estimates of relative prevalence are consistent (e.g. not significantly different, as per the CIs of your estimate).

Author response: The sentence has been revised as suggested (whereas has been removed).

25. P3- I would suggest an introductory sentence to make the flow better from P2 to P3, and to clarify the focus of P3.

Author response: An introductory sentence has been added as suggested. [As Canadian estimates of HCV prevalence are not yet available for specific birth cohorts, same-age rate ratios were calculated to compare the likely cohort effect in Canada with the American situation.]

26. P3-
i) How do you "combine" your results with those of the CHMS? I don't understand this.
ii) I also think you need to explicitly state whether/how you would expect your burden data would to indicate prevalence; I don't think you are providing "estimates of relative prevalence," which you refer to.
iii) I think you miss the opportunity to talk about the strength of the data source, i.e. capturing a large proportion of all hospitalizations over several years.

Author response:
i) The sentence has been reworded to explain that the same-age rate ratios were used to prorate the CHMS estimate.
ii) We have removed any interpretation of the same-age rate ratios as 'prevalence', though still compare the same-age ratios to the NHANES prevalence ratios.
iii) Thanks! A sentence to this effect has been added to the limitation section. [Despite the limitations, by following records of hospitalizations from an administrative database by birth cohort, this study design has provided insight into the birth cohort effects of HCV not currently available from other data sources in Canada, and model outputs appear to be in reasonable agreement with estimates from other studies.]

Limitations
27. P1- What do you mean that the "source of the HCV infection is also known to influence the all-cause mortality rate"? Do you mean the mechanism of infection?

Author response: The sentence has been reworded to state that persons who inject drugs have higher SMRs.

28. I think you should acknowledge the lack of knowledge re: incidence trends in each age cohort going forward.

Author response: Incident infections and diagnoses are included in the projections only to the extent that they were captured by historic trends. A sentence has been added to the discussion of limitations. [Cases diagnosed with HCV over the study period contributed to the historic trends if hospitalized for a liver disease, though recent discussions about the effects and treatment for HCV may increase the diagnosis of HCV at an early stage of the disease.]

Table 1
29. I think you should cut out column for statistical significance since you provide the p values

Author response: The last column has been removed.

30. I would make it clear that the 95% CIs and p values refer to the average annual % change, e.g. by spreading those words across the three columns

Author response: The column for 95%CIs has been combined with the Average Annual % Change.

Table 2
31. I don't understand why some of the 95% CIs only have one number.

Author response: We have updated the table to include the lower bound of zero.

32. I don't understand what the columns represent, e.g. what is the difference between 5 years and 2015? You have indicated that the data for the 5 years, 10 years, and until age 90 columns are from regression models. Where are the data for the 2015, 2020 and at age 75 columns from?
-As noted above, need to specify why 1970-1974 is the reference population.

Author response: Footnotes has been added for clarification and column headings have been revised.

Table 3
33. How do you define the categories of relative disease burden? i.e. what does high vs. low mean? I think you need to define these categories and explain how you made them.

Author response: Table 3 has been deleted.

Figure 1

34. I find it hard to match the colours in the Figure with those in the legend. You should reformat this.
Author response: The graphs and legend has been reformatted for both Figure 1 and 2a, 2b.

35. I don't think that the points of each cohort should be connected- what would those lines represent?
* The editors found these last 2 points to be of lesser consequence

Author response:

The legend has been revised to indicate that "The resulting curve is the number of admissions by age for one birth cohort. The cohort curves indicate the trend in hospitalization at the age specified on the x-axis, and show an increase from age 25 to 55, and decrease from age 80 on".

Figure 2
-As with Figure 1, what do lines connecting points represent? I would reconsider these lines.

Author response: The figure has been removed as the values are represented in Table 1.

Reviewer 2:

Comments to the Author
1. Study design
a. It is difficult to accurately estimate chronic hepatitis C associated with hospitalization using the design in this study because there were no appropriate birth cohorts without chronic hepatitis C as control for adjustment of strong confounding effects associated with age and patient baseline comorbidities.

Author response: We have revised the manuscript for statements that could be interpreted to suggest that this study design would estimate the true burden attributable to hepatitis C. Rather we used the number of hospital admissions for patients with a diagnosis of liver disease and chronic hepatitis C as an indicator of the burden – an indicator that is available directly from an existing data base and one that could be used to monitor progress in reducing the significant burden of hepatitis C without the collection of additional data.

b. This study assumed that younger birth cohorts had similar hospitalization trends as older birth cohorts. Because the transmission mode of hepatitis C virus in Canada has been changed substantially, older birth cohort likely acquired hepatitis C through blood transfusion or blood products but younger birth cohorts were mainly infected by the virus through drugs use. Thus, differences in transmission model likely make patients have different social economic status and underline comorbidities that could significantly affect disease progression and future health resource utilization.

Author response: The discussion of the limitations has been revised to address this point more fully.

c. The defined liver-related disease in this study (K70 to 77, R18) might not accurately reflect the hospitalizations caused by hepatitis C related complications. For example, K70, 71, and 75 have nothing to do with chronic hepatitis C. Thus, this study likely overestimated the hospitalizations caused by CHC.

Author response: K70 (alcoholic liver disease), K71 (Toxic liver disease) and K75 (Other inflammatory liver disease that excludes hepatitis) are not directly linked to hepatitis C. We reviewed all records for the use of these diagnostic codes in combination with chronic hepatitis C. K71 and K75 rarely occurred in combination with B18.2 (chronic hepatitis C), though K70 accounted for approximately 25% of hospital admissions associated with both chronic hepatitis C and liver disease. As persons diagnosed with chronic hepatitis C are advised to reduce or eliminate alcohol consumption, it may be useful to explore this relationship further.

No attempt has been made in this study to attribute only a portion of the B18.2, K70 admissions to chronic hepatitis C, or to claim that all chronic hepatitis C – liver disease admissions were caused by HCV. We used the word 'associated' to simply mean that the diagnosis was recorded in the electronic discharge record. We also did not account for under reporting of HCV in the patient's discharge records, which was apparent through the linkage of scrambled patient ids in the CIHI database.

The text has been modified to comment on the association with alcoholic and discuss the possibility other misclassifications. While the topic of data quality is an important consideration, it is out of scope of this manuscript. The discussion of the limitations has been revised to address this point more fully.

2. Analysis
• This study only built the relationship between hospitalization and year without taking into account any other factors that could contribute to hospitalization related to HCV. I really doubt this formula would be able to accurately predict the trend of HCV-related hospitalization. At least the study needs to stratify hospitalizations by the complications caused by chronic hepatitis C. In addition, the study needs some kind evaluation approach to demonstrate the prediction model is reliable.

Author response: With only 7 years of data we looked at the accuracy of a one year ahead prediction, though this is not sufficient validation and we did not report this work

Various revisions to the limitation sub-section have been made. For example, we acknowledge that projections are hypothetical as better treatments should reduce hospitalizations to levels below our status quo estimate [limitations sub-section].