PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Sex-Specific Temporal Trends in Ambulatory Heart Failure Incidence, Mortality and Hospitalization in Ontario, Canada from 1994-2013 – a Population-Based Cohort Study
AUTHORS	Sun, Louise; Mielniczuk, Lisa; Liu, Peter; Beanlands, Rob; Chih, Sharon; Davies, Ross; Coutinho, Thais; Lee, Douglas; Austin, Peter; Bader Eddeen, Anan; Tu, Jack

VERSION 1 – REVIEW

REVIEWER	Song Li University of Washington, Seattle, WA, USA
REVIEW RETURNED	07-Sep-2020

GENERAL COMMENTS	In the present manuscript, Sun et al. analyzed the sex-specific temporal trends in incidence, mortality, and heart failure (HF) hospitalizations in newly diagnosed ambulatory HF patients in a universal healthcare system that allows for complete capture of all patients in the most populous Canadian province. The authors found that over a 20-year period, age-standardized mortality and hospitalization rates declined more for men than for women while the age-standardized incidence declined similarly in both men and women. The authors advocated for further research into the potential causes of the above differences. First of all, I applaud the authors for a very well written and organized manuscript. For the most part, the language is clear, succinct, and a pleasure to read. This manuscript is also investigating a very fundamental topic in HF, and the ambulatory and sex-specific analyses have not been well published before, further contributing to the significance and originality of this manuscript. The research methods are standard and sound. There are some major technical issues with the tables and figures that I listed below but they should be easily fixable. While it is unfortunate that this study was not able to examine the subtypes of HF or the management of HF in females versus males to explain the cause of observed sex-specific differences, the authors acknowledged these limitations appropriately and advocated for further research. My comments on specific issues are below: 1. The authors included adults >= 40 years of age. Is there a reason for not including younger patients? There are important subtypes of HF that are more prevalent in young patients such as peripartum cardiomyopathy that may be missed by excluding them.

2. The authors used 2 outpatient billing claims to identify HF. While this method has been validated at a single point in time, is it possible that potential changes in billing codes or practices in the 20-year period could have led to an artificial change in calculated incidence?

3. The method of calculating age-standardized rates could be further elaborated. For example, what age brackets did the authors use in the calculations?

4. Rather than testing the statistical significance of the differences between the historical and the modern cohorts for many comparisons, why not test for a statistically significant trend over the four 5-year periods, or better yet, over the 20 years using a simple linear regression?

5. Not able to see Figure 4 or eFigure 2 in the PDF file.

6. In Table 1, in the Women column under the Modern Cohort, the numbers and percentages below the "Uncomplicated HTN" line appear shifted and out-of-place.

7. In addition to the authors' conclusions that male outcomes improved more than female outcomes in general, I recommend also mentioning that males started with higher mortality and hospitalization rates and now those are similar to females.

8. For Figure 1, I recommend setting the lower bound of the y-axis to zero in order to avoid over-representing the magnitude of the difference between the two sexes.

9. In the Results section, Figure 3 and 4 were referenced in the text before Figure 2.

REVIEWER	Simonetta Scalvini ICS Maugeri SpA Benefit Society IRCCS (Institute of Research and Care) ITALY
REVIEW RETURNED	15-Sep-2020

| GENERAL COMMENTS | The paper well describes the work of the authors; there are some minor comment that the authors have to answer:
1- "During the historical period (1994-1998), a total of 47,676 (0.36%) incident HF cases were identified in women and 43,907 (0.36%) in men. During the modern era (2009-2013), 42,746 (0.24%) incident cases were identified in women and 47,961 (0.29%) in men". The authors reported to use an algorithm reported in reference 9 in which they conclude that " cases of CHF can be identified with a high degree of accuracy from administrative data with an algorithm of one hospitalization or ambulatory record followed by a second record from other source within one year" It will be better to know in this case how patients have ambulatory record only or ambulatory and hospitalization in the following year. Describe the casuistic in this sense.
2- both in the results in the abstract and in the main manuscript, there is the description:
"Age-standardized one-year mortality rates (AMR) also declined in both sexes but the magnitude of reduction was greater in men than in women. Men had higher AMR than women at most time points |

	prior to 2009 (Figure 2). Specifically, the female AMR per 1000 was 104.4 (95% CI, 90.5-119.8) in 1994 and 84.8 (75.2-95.3) in 2013, representing a 19% reduction. Conversely, male AMR per 1000 was 123.0 (110.6-136.5) in 1994 and 83.0 (75.3-91.2) in 2013, representing a 33% reduction (eTable 3)" In my opinion this part is quite difficult to understand for the readers because we are not used to "per 1000" but we use %; it will be better to change in this way if it is possible. In other case the authors use % and so it is better to define better the different use of % and per 1000. 3- Some reference are too old and need to be reduce and they need to find more recent papers of the same arguments

REVIEWER	Julie A. Murphy, PharmD, FASHP, FCCP, BCPS University of Toledo College of Pharmacy and Pharmaceutical Sciences United States
REVIEW RETURNED	27-Sep-2020

GENERAL COMMENTS	Methods 1. Design a. Strengths: i. 20 years of data ii. Validated algorithm iii. Acknowledged limitation of databases lacking physiologic and laboratory measures b. Page 5, line 12: change reimbursed to reimburses c. Page 5, lines 13-14: necessary? Consider deleting. d. Page 5, line 17: change 2 to two. e. Page 6, lines 6-7: was the secondary outcome also within one year? Based on the statistical analysis, it would appear so, but maybe this could be clarified here. f. Page 6, lines 11-16: it is mentioned that "other co-morbidities" were identified within 5 years. Is this the same for hypertension, asthma, COPD, and diabetes? This could be clarified. Results 1. Page 9, lines 4-8: the information provided here does not match Table 1 a. The sentence specifically mentions the modern cohort, when in fact for all the cohorts women were more likely to be older, have lower income status, have HTN, hypothyroidism, depression. b. Similarly, for all the cohorts, men were more likely to have MI, and diabetes. c. According to the table, women were not more likely to be frail or have dementia, but the text indicates this. d. Anemia is mentioned in the text, but not in Table 1. e. Text states peripheral arterial disease, but the table says peripheral vascular disease. Be consistent. f. Table indicates alcohol abuse was more common in women in the modern cohort. The opposite is stated in the text. 2. Page 10, lines 2-3 (and Figure 4a and 4b): the stacked Kaplan-Meier curves did not come across in my copy, so I could not assess them. 3. Page 10, line 18 (and table 2): the data for male sex was not provided in Table 2. 4. Page 11, line 15 (and eFigure 2a): the stacked Kaplan-Meier curves did not come across in my copy, so I could not assess them.

REVIEWER 1

Reviewer Name: Song Li

Institution and Country: University of Washington, Seattle, WA, USA

Please state any competing interests or state 'None declared': None declared

In the present manuscript, Sun et al. analyzed the sex-specific temporal trends in incidence, mortality, and heart failure (HF) hospitalizations in newly diagnosed ambulatory HF patients in a universal healthcare system that allows for complete capture of all patients in the most populous Canadian province. The authors found that over a 20-year period, age-standardized mortality and hospitalization rates declined more for men than for women while the age-standardized incidence declined similarly in both men and women. The authors advocated for further research into the potential causes of the above differences.

First of all, I applaud the authors for a very well written and organized manuscript. For the most part, the language is clear, succinct, and a pleasure to read. This manuscript is also investigating a very fundamental topic in HF, and the ambulatory and sex-specific analyses have not been well published before, further contributing to the significance and originality of this manuscript. The research methods are standard and sound. There are some major technical issues with the tables and figures that I listed below but they should be easily fixable. While it is unfortunate that this study was not able to examine the subtypes of HF or the management of HF in females versus males to explain the cause of observed sex-specific differences, the authors acknowledged these limitations appropriately and advocated for further research.

RESPONSE: We thank the reviewer for her support of our manuscript.

My comments on specific issues are below:

1.   The authors included adults >= 40 years of age. Is there a reason for not including younger patients? There are important subtypes of HF that are more prevalent in young patients such as peripartum cardiomyopathy that may be missed by excluding them.

RESPONSE/REVISION: Thank you for raising this important point. Our age inclusion was limited by the algorithm used for identifying HF patients, which was validated only for patients over the age of

40.   We have now acknowledged this in the Limitations on page 15 line 13: "Secondly, our algorithm

for ascertainment of HF is validated in patients who are 40 years of age and older, thus limiting the generalizability of our findings."

2. The authors used 2 outpatient billing claims to identify HF. While this method has been validated at a single point in time, is it possible that potential changes in billing codes or practices in the 20-year period could have led to an artificial change in calculated incidence?

RESPONSE/REVISION: Thank you for the astute comment. The code used for HF billing in Ontario has stayed the same over the 20-year study period (ICD-8 code 428).

3. The method of calculating age-standardized rates could be further elaborated. For example, what age brackets did the authors use in the calculations?

RESPONSE/REVISION: Thank you for the opportunity to clarify our methodology. We used Canadian patients >=40 years of age as the standardizing population, corresponding to our age inclusion criterion. This was explained on page 7 line 11: "These rates were directly standardized by age using the 1991 Canadian population aged ≥ 40 years as the reference population."

4. Rather than testing the statistical significance of the differences between the historical and the modern cohorts for many comparisons, why not test for a statistically significant trend over the four 5-year periods, or better yet, over the 20 years using a simple linear regression?
RESPONSE/REVISION: Thank you for this suggestion. We have used linear regression with year as the independent variable to evaluate temporal trends.

Methods, page 7 line 13: "We used linear regression with fiscal year as the independent variable to assess for temporal changes in HF incidence and outcomes in women and men across the 20-year period." Results: page 8 line 17: "Although the incidence of HF declined in both sexes over the 20-year period (linear regression slope, -0.031; p<0.0001 in women and -0.025; p<0.0001 in men), it remained higher in men than in women (Figure 1)."

Page 9 line 4: "…demonstrate an improvement in male survival (linear regression slope, -0.020; p<0.0001) but relatively little change in female survival over time (linear regression slope, -0.010; p=0.001)."

Page 10 line 13: "HF hospitalization rates declined in men (linear regression slope, -0.010; p=0.0002) but remained unchanged in women (linear regression slope, -0.005; p=0.11) during the 20-year period (eFigure 1 and eTable 4)."

5. Not able to see Figure 4 or eFigure 2 in the PDF file.

RESPONSE/REVISION: Thank you for pointing these out. We have submitted reformatted versions of these figures.

6. In Table 1, in the Women column under the Modern Cohort, the numbers and percentages below the "Uncomplicated HTN" line appear shifted and out-of-place.

RESPONSE/REVISION: Thank you for pointing this out. This has been reformatted.

7. In addition to the authors' conclusions that male outcomes improved more than female outcomes in general, I recommend also mentioning that males started with higher mortality and hospitalization rates and now those are similar to females.

RESPONSE/REVISION: Thank you for this suggestion. We have added "Specifically, mortality and

hospitalization rates were higher in men than women at the start of the study period and were similar between sexes towards the end of this period." On page 15 line 6.

8.   For Figure 1, I recommend setting the lower bound of the y-axis to zero in order to avoid over-representing the magnitude of the difference between the two sexes.

RESPONSE/REVISION: We have done as the reviewer has asked.

9.   In the Results section, Figure 3 and 4 were referenced in the text before Figure 2. RESPONSE/REVISION: Thank you for pointing this out. We have renumbered the figures.

****

REVIEWER 2

Reviewer Name: Simonetta Scalvini

Institution and Country: ICS Maugeri SpA Benefit Society IRCCS (Institute of Research and Care),

ITALY

Please state any competing interests or state 'None declared': None declared

The paper well describes the work of the authors; there are some minor comment that the authors have to answer:

1- "During the historical period (1994-1998), a total of 47,676 (0.36%) incident HF cases were identified in women and 43,907 (0.36%) in men. During the modern era (2009-2013), 42,746 (0.24%) incident cases were identified in women and 47,961 (0.29%) in men". The authors reported to use an algorithm reported in reference 9 in which they conclude that " cases of CHF can be identified with a high degree of accuracy from administrative data with an algorithm of one hospitalization or ambulatory record followed by a second record from other source within one year" It will be better to know in this case how patients have ambulatory record only or ambulatory and hospitalization in the following year. Describe the casuistic in this sense.

RESPONSE/REVISION: Thank you for the opportunity to clarify our methods. The HF algorithm described by Schultz et al in Chronic Dis Inj Can. 2013;33:160-6 indeed uses either 2 ambulatory records, or one ambulatory record + one hospitalization record. As we included only patients with an ambulatory diagnosis of HF, we used the definition of 2 ambulatory records in the year prior to HF diagnosis: "…based on two outpatient billing claims for HF within one year." on page 5, line 19. The same method was used in an earlier publication by our group (reference 19) to identify ambulatory HF patients.

2- both in the results in the abstract and in the main manuscript, there is the description: "Age-standardized one-year mortality rates (AMR) also declined in both sexes but the magnitude of

reduction was greater in men than in women. Men had higher AMR than women at most time points prior to 2009 (Figure 2). Specifically, the female AMR per 1000 was 104.4 (95% CI, 90.5-119.8) in 1994 and 84.8 (75.2-95.3) in 2013, representing a 19% reduction. Conversely, male AMR per 1000 was 123.0 (110.6-136.5) in 1994 and 83.0 (75.3-91.2) in 2013, representing a 33% reduction (eTable 3)"

In my opinion this part is quite difficult to understand for the readers because we are not used to "per

1000" but we use %; it will be better to change in this way if it is possible. In other case the authors use % and so it is better to define better the different use of % and per 1000. RESPONSE/REVISION: We have revised to % as you have suggested: "Specifically, female AMR at one-year was 10.4% (95% CI, 9.1-12.0) in 1994 and 8.5% (7.5-9.5) in 2013, and male AMR at one-year was 12.3% (11.1-13.7) in 1994 and 8.3% (7.5-9.1) in 2013. Conversely, age standardized HF hospitalization rates declined in men (11.4% [10.1-12.9] in 1994 and 9.1% [8.2-10.1] in 2013) but remained unchanged in women (9.7% [8.3-11.3] in 1994 and 9.8% [8.6-11.0] in 2013)." in the abstract, and have made similar revisions in the Results section of the manuscript.

3- Some reference are too old and need to be reduce and they need to find more recent papers of the same arguments

RESPONSE/REVISION: Thank you for this suggestion. We have replaced the reference to women being under-represented in HF clinical trials with a newer reference from European Heart Journal 2019. We were unfortunately unable to replace the papers in reference to validated algorithms in administrative databases, as the references used were the most recent ones.

We thank the reviewer for her support of our manuscript.

****

REVIEWER 3

Reviewer Name: Julie A. Murphy, PharmD, FASHP, FCCP, BCPS

Institution and Country: University of Toledo College of Pharmacy and Pharmaceutical Sciences, United States

Please state any competing interests or state 'None declared': None declared

Methods

1.   Design

a. Strengths:

i.   20 years of data

ii.   Validated algorithm

iii. Acknowledged limitation of databases lacking physiologic and laboratory measures
RESPONSE: We thank the reviewer for her appreciation of our manuscript.

b.   Page 5, line 12: change reimbursed to reimburses

RESPONSE/REVISION: Thank you for pointing this out. We have revised accordingly.

c.   Page 5, lines 13-14: necessary? Consider deleting.

RESPONSE/REVISION: We have revised to "Patients and the public were not involved in the design and conception of this study. However, the results will be publicly disseminated. " at the Editor's request.

d.   Page 5, line 17: change 2 to two.

RESPONSE/REVISION: Thank you for pointing this out. We have revised accordingly.

e.   Page 6, lines 6-7: was the secondary outcome also within one year? Based on the statistical analysis, it would appear so, but maybe this could be clarified here.

RESPONSE/REVISION: Thank you for the opportunity to clarify our methods. We have revised to "Secondary outcome was HF hospitalization within one year of HF diagnosis, which was ascertained using the Discharge Abstract Database."

f.   Page 6, lines 11-16: it is mentioned that "other co-morbidities" were identified within 5 years. Is this the same for hypertension, asthma, COPD, and diabetes? This could be clarified.
RESPONSE/REVISION: Thank you for the opportunity to clarify. All comorbidities were identified within 5 years. We have revised to: "We identified hypertension10, asthma 15, chronic obstructive pulmonary disease (COPD) 16 and diabetes mellitus 12 using validated algorithms applied on patient encounters within five years of HF diagnosis."

Results

1.   Page 9, lines 4-8: the information provided here does not match Table 1

a.   The sentence specifically mentions the modern cohort, when in fact for all the cohorts women were more likely to be older, have lower income status, have HTN, hypothyroidism, depression. RESPONSE/REVISION: Thank you for pointing this out. We have revised to "Throughout the study period, women with HF were more likely to be older, more frail, of lower income status, have co-morbid conditions such as hypertension, hypothyroidism, anemia, dementia and depression, but were less likely to have myocardial infarction (MI), peripheral arterial disease, diabetes and alcohol abuse compared to men (Table 1)."

b.   Similarly, for all the cohorts, men were more likely to have MI, and diabetes. RESPONSE/REVISION: Thank you for pointing this out. Please see our revision in response to your previous comment.

c.   According to the table, women were not more likely to be frail or have dementia, but the text indicates this.

8

RESPONSE/REVISION: Table 1 indicates that women were more likely to be frail than men in all time cohorts (9.0% vs. 6.1%, p<0.001; 13.0% vs. 10.4%, p<0.001; 24.8% vs. 16.4%, p<0.001; 26.3% vs. 17.5%, p<0.001; respectively), as well as to have dementia (4.1% vs. 2.7%, p<0.001; 4.6% vs.

3.0%, p<0.001; 4.4% vs. 2.6%, p<0.001; 4.2% vs. 2.5%, p<0.001; respectively).

d.   Anemia is mentioned in the text, but not in Table 1.

RESPONSE/REVISION: Thank you for pointing this out. We have added anemia in Table 1.

e.   Text states peripheral arterial disease, but the table says peripheral vascular disease. Be consistent. RESPONSE/REVISION: Thank you for pointing this out. We have revised to "PAD (peripheral arterial disease)" in Table 1.

f.   Table indicates alcohol abuse was more common in women in the modern cohort. The opposite is stated in the text.

RESPONSE/REVISION: Thank you for the astute observation. The position of the "women" column

in the modern cohort had been shifted by one space. We have corrected this and have marked our corrections in red. Women were in fact less likely to abuse alcohol than men in the modern cohort.

2.   Page 10, lines 2-3 (and Figure 4a and 4b): the stacked Kaplan-Meier curves did not come across in my copy, so I could not assess them.

RESPONSE/REVISION: Thank you for pointing this out. We have reformatted the figures.

3.   Page 10, line 18 (and table 2): the data for male sex was not provided in Table 2. RESPONSE/REVISION: Male sex was used as the reference category in the time-to-event model. Therefore, the HR (95% CI) represented the hazard of death in women relative to men.

4.   Page 11, line 15 (and eFigure 2a): the stacked Kaplan-Meier curves did not come across in my copy, so I could not assess them.

RESPONSE/REVISION: Thank you for pointing this out. We have reformatted the efigures.