PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Impact of Smoking on Health System Costs among Cancer Patients in a retrospective cohort study in Ontario, Canada
AUTHORS	Isaranuwatchai, Wanrudee; deOliveira, Claire; Mittmann, Nicole; Evans, William; Peter, Alice; Truscott, Rebecca; Chan, Kelvin

VERSION 1 - REVIEW

REVIEWER	K. Michael Cummings Medical University of South Carolina Charleston, South Carolina, USA
REVIEW RETURNED	01-Oct-2018

GENERAL COMMENTS	This paper presents an analysis to examine how smoking status influences the average costs of cancer care among newly diagnosed cancer patients in Ontario diagnosed between April 1, 2014 and March 31, 2015 and followed for up to one year. The results show in aggregate that for most categories of health care expenditures current and recent former smokers had higher average monthly and annual health care expenditures compared to long term former and never smokers. The authors argue that the results provide evidence supporting the need for smoking cessation treatments for newly diagnosed cancer patients. While the results are novel and do show that on average newly diagnosed current and recent former smokers appear to have higher health care costs over the subsequent 12 month period even after controlling for other potential confounders, I'm not sure these results necessarily logically support the conclusion that smoking cessation programs should be integrated into cancer care treatment plans. In fact, even if the costs of cancer care were lower for smokers compared to non-smokers, given the adverse impact of continued smoking on the cancer care outcomes I would think it would be justified from a quality of care perspective to support smoking cessation as a standard for cancer care plans. The authors should make this point more directly. The authors have described their study as a population based cohort study of cancer patients. I would suggest that the paper is best described as a describe analysis of health care costs of smokers and nonsmokers following for 12-24 months after cancer diagnosis. The paper itself could be better referenced. Reference #18 is not accessible to readers and ought to be linked to a website or reader accessible pdf. Also, the authors reference general review papers and should consider utilizing the following papers which do a better job discussing the biological and clinical impacts of smoking in cancer patients, the benefits of smoking cessation in who stop

	smoking after a cancer diagnosis, as well as treatment approaches for smoking cessation in oncology centers. • Warren GW, Sobus S, Gritz ER. The Biologic and Clinical Effects of Smoking by Cancer Patients and Strategies to Implement Evidence-Based Tobacco Cessation Support. Lancet Oncol 15:e568-80, 2014. doi:10.1016/S1470-2045(14)70266-9 • Dobson Amato K, Hyland A, Reed R, Mahoney MC, Marshall J, Giovino G, Bansal-Travers M, Ochs-Balcom HM, Zevon MA, Cummings KM, Nwogu C, Singh AK, Chen H, Warren GW, Reid M. Tobacco Cessation Improves Lung Cancer Patient Survival. J Thorac Oncol 2015;10:1014-9. doi: 10.1097/JTO. • • Gritz ER, Toll BA, Warren GW. Tobacco Use in the Oncology Setting: Advancing Clinical Practice and Research. Cancer Epid Biomarkers Prev 23:3-9; 2014. doi:10.1158/1055-9965.EPI-13-0896 • Warren GW, Marshall JR, Cummings KM, Zevon MA, Reed R, Hysert P, Mahoney MC, Hyland AJ, Nwogu C, Demmy T, Dexter E, Kelly M, O'Connor RJ, Houston T, Jenkins D, Germain P, Singh AK, Epstein J, Dobson-Amato K, Reid ME. Automated Tobacco Assessment and Cessation Support for Cancer Patients. Cancer 15;120:562-9. 2014 doi:10.1002/cncr.28440 Exclusion/inclusion criteria are well defined, but it would help the reader to interpret the findings to have a better idea of how many patients were dropped from the analyses as a result of not meeting inclusion criteria. A flow diagram or table showing the total number of cancer patients identified and then excluded for various reasons would be informative. Analyses present aggregate results comparing smokers and non-smokers, but the reader is left wondering if the results hold across all cancer diagnoses, stages of disease, age groups, and oncology centers. While it is likely the results become less reliable as data are disaggregated it would be informative to provide some disaggregated data in the form of supplementary tables. For example, one question that came to mind as I was reviewing this paper was the extent smoking cessation treatments are already integrated into cancer care in Ontario oncology centers? Do some of the centers currently do a better job of providing smoking cessation than others in which case? If so, it might be informative to compare clinical and health care cost outcomes in smoking patients treated in centers with and without between those with fairly well developed smoking cessation service programs. The current analysis controls for local Health Integration Networks and data are not allow the reviewer to judge if where patients get treatment makes a differences in health care costs and clinical outcomes. Similar questions could be asked about differences in health care costs for those with early vs later stage of disease, those with different types of cancer, etc. Cancer diagnosis might be especially important to examine as a modifying factor (not just as a confounder) since smoking prevalence and health care costs are likely to differ by cancer type (i.e., higher for lung and health next cancer and lower for breast and prostate cancers). It would help to have a better idea of what adjusted clinical groups (ACGs) measures. In table 1, ACG is defined as co-morbidities. Are these co-morbidities at the time of the patient's cancer diagnosis or over the ensuing year?

	To what extent are confounders that are adjusted for in the analyses correlated? For example, one might expect that cancer stage, co-morbidities and age might be correlated. The authors have chosen to do a regression analysis to control for potential confounders. Have they considered using propensity scoring as an alternative way to test the question of how health care costs differ between smokers and nonsmokers? Since one might hope that quitting smoking might lead to lower health care costs due to improved clinical outcomes it was curious why the authors chose not to look at former smokers as a separate group rather than lumping recent former smokers (quit within 6 months of diagnosis) with current smokers and longer term former smokers with never smokers. A more informative analysis might contrast current smokers at diagnosis with recent former smokers, longer term former smokers, and never smokers. Recent former smokers might be the group who would show some clinical benefit from having stopped smoking compared to current smokers; although it may be that co-morbid conditions motivated quitting before a cancer diagnosis. It would be helpful for the authors to comment on this and perhaps explore their data more comprehensively to evaluate these questions. The authors present adjusted and unadjusted aggregate monthly costs; figure 1 (the main take home figure only presents unadjusted results). I would recommend that the authors present only adjusted results and provide unadjusted disaggregated findings for supplemental figures and/or tables. Some of the most interesting findings in this paper are summarized in table 2 showing higher rates of mortality and lower mean time from diagnosis to death in smokers compared to non-smokers. To what extent are these results due to smoking or the fact that smokers are over represented in cancers that have less successful treatment options for patients (i.e., lung and health neck vs breast and prostate)? The authors ought to highlight over a relatively short follow-up period (12-24 months) the nearly 60% higher mortality observed among smokers compared to nonsmokers. It would be useful to know if this disparity holds after stratifying by stage of diagnosis and how results compare for current smokers, recent and longer former smokers, and never smokers.

REVIEWER	Christopher Doran Central Queensland University, Australia
REVIEW RETURNED	10-Oct-2018

GENERAL COMMENTS	An interesting and well written paper. Overall I was impressed with the structure and presentation although there are several areas where improvement may be possible. The introduction is very brief and would benefit from a more concerted synthesis of empirical evidence and a clearer research objective (with subsequent hypothesis). The last paragraph of the introduction is methods and does not fit within the introduction. The explanation of person-month costs is lacking. I'm not sure what this relates to and impacts on the interpretation of results. Are health care costs pre-diagnosis considered or only post-diagnosis? A more accurate assessment would consider pre and post. I am also concerned that smokers had an overall higher severity of cancer diagnosis - could costs be estimated according to severity? Or by cancer? Does health insurance impact on costs

	or health care utilisation? Perhaps a clearer explanation of health care financing would be of benefit. The authors should provide results of statistical comparisons and perhaps subject results to sensitivity or uncertainty testing. Which result is correct? The abstract suggest that ...on average, smokers had significantly higher monthly healthcare costs ($5,091) than non-smokers ($4,847) p<0.05 The result section suggests that: total monthly health care costs were higher among smokers ($5,649 ± $7,169) than non-smokers ($4,704 ± $6,737). These differences raise concerns over validity of the results - very wide confidence intervals and hard to believe significant differences given similarity of results between groups. Discussion was good and appropriate consideration of limitations.

REVIEWER	Summer Frank-Pearce University of Oklahoma Health Sciences Center, United States
REVIEW RETURNED	26-Oct-2018

GENERAL COMMENTS	I reviewed this manuscript with a particular emphasis on the statistical methods and analyses used, as requested by the editor. This manuscript describes an exploration of the impact of smoking on health system costs among adult cancer patients who were newly diagnosed and compares the costs of cancer patients who were current smokers (or reported smoking in the last 6 months) to costs of non-smokers. Overall, the manuscript is well written and addresses an important topic. However, there are several concerns. 1) The authors should consider providing the number patients from the OCR which were excluded by the each exclusion criteria. 2) The reason(s) for excluding patients with multiple cancers should be explained. Also this exclusion should be discussed in the limitations, especially if this is a large patient group. 3) It is not clear whether the authors checked the covariates used in the adjusted model for collinearity/multicollinearity. There are some variables that might be correlated, i.e. geographical region/rurality, stage/comorbidity. This should be clarified. 4) Table 1 outlines the proportion of patients with “unknown” cancer stage which is a quarter of the population or more (depending on the smoking status). This is a large proportion of the population; however, the authors do not discuss why a cancer type would be unknown. It appears these individuals were included in the analysis. However, without more information on the cancer types that could be included in this category, it is unclear what it means to adjust for “unknown” cancer type. This should be clarified and, if necessary, discussed in the limitations. The authors should also consider performing a sensitivity analysis in which these individuals are excluded from the analysis. 5) Table 1 outlines the proportion of patients with “other” cancer site which is 29% of the population or more (depending on the smoking status), and the most commonly reported site type. However, in the results, the next most common cancers are identified as the most common type of cancer. This result needs to be more accurately reported in the results. The patients labelled “other” cancer site is a large proportion of the population; however, the authors do not discuss what types of cancers sites are categorized as “other”. It appears these individuals were included in the analysis. However, without more information on the cancer

	sites that could be included in this category, it is unclear what it means to adjust for “other” cancer site. This should be clarified and, if necessary, discussed in the limitations. The authors should also consider performing a sensitivity analysis in which these individuals are excluded from the analysis. 6) The GLM results should report the 95% confidence interval for the estimated monthly healthcare cost, which would be more informative than a p-value. 7) It is not clear why the data on smoking status was “limited to one assessment at the first ambulatory visit” as the methods describe that the information was submitted on a monthly bases to the CCO. The authors should clarify this.

VERSION 1 – AUTHOR RESPONSE

2. Comments from reviewer 1, K. Michael Cummings:

a) This paper presents an analysis to examine how smoking status influences the average costs of cancer care among newly diagnosed cancer patients in Ontario diagnosed between April 1, 2014 and March 31, 2015 and followed for up to one year. The results show in aggregate that for most categories of health care expenditures current and recent former smokers had higher average monthly and annual health care expenditures compared to long term former and never smokers. The authors argue that the results provide evidence supporting the need for smoking cessation treatments for newly diagnosed cancer patients.

While the results are novel and do show that on average newly diagnosed current and recent former smokers appear to have higher health care costs over the subsequent 12 month period even after controlling for other potential confounders, I’m not sure these results necessarily logically support the conclusion that smoking cessation programs should be integrated into cancer care treatment plans. In fact, even if the costs of cancer care were lower for smokers compared to non-smokers, given the adverse impact of continued smoking on the cancer care outcomes I would think it would be justified from a quality of care perspective to support smoking cessation as a standard for cancer care plans. The authors should make this point more directly.

Response to reviewer:

We appreciate the comments from the reviewer and the recognition that the results are novel.

We understand that the results may not directly support that smoking cessation programs should be integrated into cancer care treatment plans. We have made this point more clearly in the revised draft by stating that this work represents only one piece of evidence to further support smoking cessation programs. Please refer to the “Strengths and Limitations of this study” section (pg. 3, bullet point 1,i.e., lines 2-3), and the “Discussion” section (pg. 9, lines 32-33, and pg. 10, lines 8-12).

b) The authors have described their study as a population based cohort study of cancer patients. I would suggest that the paper is best described as a describe analysis of health care costs of smokers and nonsmokers following for 12-24 months after cancer diagnosis.

Response to reviewer:

Thank you for your comment. We have revised the manuscript accordingly. Please refer to the Methods section of the “Abstract” (pg. 2, lines 9-11), the “Introduction” section (pg.4, lines 16-18), and the “Materials and Methods” section (pg. 4, lines 26-28).

c) The paper itself could be better referenced. Reference #18 is not accessible to readers and ought to be linked to a website or reader accessible pdf. Also, the authors reference general review papers

and should consider utilizing the following papers which do a better job discussing the biological and clinical impacts of smoking in cancer patients, the benefits of smoking cessation in who stop smoking after a cancer diagnosis, as well as treatment approaches for smoking cessation in oncology centers.

 Warren GW, Sobus S, Gritz ER. The Biologic and Clinical Effects of Smoking by Cancer Patients and Strategies to Implement Evidence-Based Tobacco Cessation Support. Lancet Oncol 15:e568-80, 2014. doi:10.1016/S1470-2045(14)70266-9

 Dobson Amato K, Hyland A, Reed R, Mahoney MC, Marshall J, Giovino G, Bansal-Travers M, Ochs-Balcom HM, Zevon MA, Cummings KM, Nwogu C, Singh AK, Chen H, Warren GW, Reid M. Tobacco Cessation Improves Lung Cancer Patient Survival. J Thorac Oncol 2015;10:1014-9. doi:10.1097/JTO.

 Gritz ER, Toll BA, Warren GW. Tobacco Use in the Oncology Setting: Advancing Clinical Practice and Research. Cancer Epid Biomarkers Prev 23:3-9; 2014. doi:10.1158/1055-9965.EPI-13-0896

 Warren GW, Marshall JR, Cummings KM, Zevon MA, Reed R, Hysert P, Mahoney MC, Hyland AJ, Nwogu C, Demmy T, Dexter E, Kelly M, O'Connor RJ, Houston T, Jenkins D, Germain P, Singh AK, Epstein J, Dobson-Amato K, Reid ME. Automated Tobacco Assessment and Cessation Support for Cancer Patients. Cancer 15;120:562-9. 2014 doi:10.1002/cncr.28440

Response to reviewer:

We have changed reference 18 (now reference 22 on pg.16 after adding the 4 new suggested papers) to website reference style and added a website for it. We have also incorporated the suggested references in the updated version under the "Introduction" section (pg. 4, line 10,i.e., references 15-18). Thank you.

d) Exclusion/inclusion criteria are well defined, but it would help the reader to interpret the findings to have a better idea of how many patients were dropped from the analyses as a result of not meeting inclusion criteria. A flow diagram or table showing the total number of cancer patients identified and then excluded for various reasons would be informative.

Response to reviewer:

Thank you for the suggestion. We have added more details in terms of the number of patients excluded from the analysis as a supplementary flow chart (S1 Appendix). This has also been mentioned under the "Materials and Methods" section (pg.5, lines 1-2).

e) Analyses present aggregate results comparing smokers and non-smokers, but the reader is left wondering if the results hold across all cancer diagnoses, stages of disease, age groups, and oncology centers. While it is likely the results become less reliable as data are disaggregated it would be informative to provide some disaggregated data in the form of supplementary tables. For example, one question that came to mind as I was reviewing this paper was the extent smoking cessation treatments are already integrated into cancer care in Ontario oncology centers? Do some of the centers currently do a better job of providing smoking cessation than others in which case? If so, it might be informative to compare clinical and health care cost outcomes in smoking patients treated in centers with and without between those with fairly well developed smoking cessation service programs. The current analysis controls for local Health Integration Networks and data are not allow the reviewer to judge if where patients get treatment makes a differences in health care costs and clinical outcomes. Similar questions could be asked about differences in health care costs for those with early vs later stage of disease, those with different types of cancer, etc. Cancer diagnosis might be especially important to examine as a modifying factor (not just as a confounder) since smoking prevalence and health care costs are likely to differ by cancer type (i.e., higher for lung and health next cancer and lower for breast and prostate cancers).

Response to reviewer:

The reviewers raised many good questions which we would like to explore further in future studies. Unfortunately, we have addressed these points (e.g., the issue of small sample sizes for the less common cancers) in the “Discussion” section as limitation (pg. 10, lines 41-43). We have submitted a subsequent grant in the hopes that we could explore additional subgroup analyses and interactions.

f) It would help to have a better idea of what adjusted clinical groups (ACGs) measures. In table 1, ACG is defined as co-morbidities. Are these co-morbidities at the time of the patient’s cancer diagnosis or over the ensuing year?

Response to reviewer:

We have elaborated further and referred to additional references on this. In this study, this value was assigned at the time of the cancer diagnosis. Please refer to the “Materials and Methods “section (pg. 6, lines 27-32).

g) To what extent are confounders that are adjusted for in the analyses correlated? For example, one might expect that cancer stage, co-morbidities and age might be correlated.

Response to reviewer:

We have checked for collinearity/multicollinearity and found no evidence on this. We have added this statement in the “Materials and Methods” section (pg. 7, lines 22-23). Thank you.

h) The authors have chosen to do a regression analysis to control for potential confounders. Have they considered using propensity scoring as an alternative way to test the question of how health care costs differ between smokers and nonsmokers?

Response to reviewer:

We considered using a propensity score matching approach. Due to the current sample size, we would have lost a significant number of samples if we were to use the propensity score matching approach as we could not find a match. Today, we have larger sample size in the dataset, and we have submitted a subsequent grant in the hopes that we could explore a longer follow up with larger cohort size using a propensity score approach.

i) Since one might hope that quitting smoking might lead to lower health care costs due to improved clinical outcomes it was curious why the authors chose not to look at former smokers as a separate group rather than lumping recent former smokers (quit within 6 months of diagnosis) with current smokers and longer term former smokers with never smokers. A more informative analysis might contrast current smokers at diagnosis with recent former smokers, longer term former smokers, and never smokers. Recent former smokers might be the group who would show some clinical benefit from having stopped smoking compared to current smokers; although it may be that co-morbid conditions motivated quitting before a cancer diagnosis. It would be helpful for the authors to comment on this and perhaps explore their data more comprehensively to evaluate these questions.

Response to reviewer:

Thank you for your comments. Unfortunately, giving the data we received on smoking status, we could not separate former smokers as a separate group. We have addressed this as one of the limitations under the “Discussion” section (pg. 10, lines 22-35). Furthermore, a subject matter expert described that those that are recent quitters are much more likely to relapse so the intention is to allow them access to cessation advice if needed to prevent a relapse. Therefore, recent quitters (within 6 months) are not separated in the dataset.

j)The authors present adjusted and unadjusted aggregate monthly costs; figure 1 (the main take home figure only presents unadjusted results). I would recommend that the authors present only adjusted results and provide unadjusted disaggregated findings for supplemental figures and/or tables.

Response to reviewer:

Thank you for the suggestion. We have updated the results accordingly and moved Figure 1 to Supplemental information, i.e., S4 Appendix (pg. 9, line 4).

k) Some of the most interesting findings in this paper are summarized in table 2 showing higher rates of mortality and lower mean time from diagnosis to death in smokers compared to non-smokers. To what extent are these results due to smoking or the fact that smokers are over represented in cancers that have less successful treatment options for patients (i.e., lung and health neck vs breast and prostate)? The authors ought to highlight over a relatively short follow-up period (12-24 months) the nearly 60% higher mortality observed among smokers compared to nonsmokers. It would be useful to know if this disparity holds after stratifying by stage of diagnosis and how results compare for current smokers, recent and longer former smokers, and never smokers.

Response to reviewer:

These are great points which unfortunately we could not address (e.g., recent quitters vs. the others) based on the restriction of the dataset. We hope to explore in future research with larger sample sizes (e.g., to explore the impact by cancer type and stratify by stage of diagnosis). This study represents the first step to examine the impact of smoking based on existing available data with limitations we highlighted in the discussion (pg.10, lines 22-46). We thank the reviewers for understanding.

3. Comments from reviewer 2, Christopher Doran

An interesting and well written paper. Overall I was impressed with the structure and presentation although there are several areas where improvement may be possible.

Response to reviewer:

We would like to thank the reviewer for your valuable comments. Please see our responses and proposed revision below.

a) The introduction is very brief and would benefit from a more concerted synthesis of empirical evidence and a clearer research objective (with subsequent hypothesis). The last paragraph of the introduction is methods and does not fit within the introduction.

Response to reviewer:

We have elaborated more synthesis of empirical evidence in the discussion. The last paragraph of the introduction (pg. 4, lines 16-23) was the study objective and we have updated the text to make it clearer to the readers including the relevant hypothesis. Thank you for your comments.

b) The explanation of person-month costs is lacking. I'm not sure what this relates to and impacts on the interpretation of results. Are health care costs pre-diagnosis considered or only post-diagnosis? A more accurate assessment would consider pre and post. I am also concerned that smokers had an overall higher severity of cancer diagnosis - could costs be estimated according to severity? Or by cancer? Does health insurance impact on costs or health care utilization? Perhaps a clearer explanation of health care financing would be of benefit. The authors should provide results of statistical comparisons and perhaps subject results to sensitivity or uncertainty testing.

Response to reviewer:

Thank you for your comments. We analyzed cost as a monthly cost for each patient in order to adjust for different follow up times (as explained under the “Analysis” section on pg. 7, lines 3-5). This approach has been previously used (reference 28). Health care costs considered only post-diagnosis, and we have added further information on pg. 6, lines 4-6. We second that smokers are likely to have higher severity and have attempted to adjust for this using a validated comorbidity index (the Johns Hopkins Adjusted Clinical Group®) and cancer type (as described under the “Materials and Methods” section, pg.6, lines 20-38). In the study setting, patients are covered under publicly funded system, and thus health insurance is not likely to influence utilization – we have elaborated further in the “Materials and Methods” section on pg. 6, lines 4-7.

c) Which result is correct? The abstract suggest that ...on average, smokers had significantly higher monthly healthcare costs ($5,091) than non-smokers ($4,847) $p<0.05$. The result section suggests that: total monthly health care costs were higher among smokers ($5,649 ± $7,169) than non-smokers ($4,704 ± $6,737). These differences raise concerns over validity of the results - very wide confidence intervals and hard to believe significant differences given similarity of results between groups. Discussion was good and appropriate consideration of limitations.

Response to reviewer:

The estimates of ($5,091 and $4,847) represented the final findings from adjusted regression model. The other estimates are crude estimates of monthly costs. We have elaborated further in the manuscript under the “Results” section, pg. 9, lines 9-13.

4) Comments from reviewer 3, Summer Frank-Pearce

I reviewed this manuscript with a particular emphasis on the statistical methods and analyses used, as requested by the editor. This manuscript describes an exploration of the impact of smoking on health system costs among adult cancer patients who were newly diagnosed and compares the costs of cancer patients who were current smokers (or reported smoking in the last 6 months) to costs of non-smokers. Overall, the manuscript is well written and addresses an important topic. However, there are several concerns.

Response to reviewer:

We would like to thank the reviewer for your time and helpful input. We have responded to the points raised below.

a) The authors should consider providing the number patients from the OCR which were excluded by the each exclusion criteria.

Response to reviewer:

Thank you for the suggestion. We have added more details in terms of number of patients excluded from the analysis in a supplementary flow chart (S1 Appendix). This has also been mentioned under the “Materials and Methods” section, pg. 5, lines 1-2.

b) The reason(s) for excluding patients with multiple cancers should be explained. Also, this exclusion should be discussed in the limitations, especially if this is a large patient group.

Response to reviewer:

Thank you for your comments. We have elaborated this point further as a limitation of the study under the “Discussion” section, pg. 10, lines 41-44.

Patients with multiple cancers were excluded from the study to distinguish the impact of smoking on one type of cancer, as multiple cancers would add a layer of complexity and costs. Numerically, they were much less common. This is our first attempt to analyze the impact of smoking on costs and outcomes. Future studies will examine if multiple cancers play a role in the costs and outcomes (i.e., we have submitted a subsequent grant to explore the impact of smoking of multiple cancers).

c) It is not clear whether the authors checked the covariates used in the adjusted model for collinearity/multicollinearity. There are some variables that might be correlated, i.e. geographical region/rurality, stage/comorbidity. This should be clarified.

Response to reviewer:

We have checked for collinearity/multicollinearity and found no evidence on this. We have added this statement in the “Materials and Methods” section (pg. 7, lines 22-23). Thank you.

d) Table 1 outlines the proportion of patients with “unknown” cancer stage which is a quarter of the population or more (depending on the smoking status). This is a large proportion of the population; however, the authors do not discuss why a cancer type would be unknown. It appears these individuals were included in the analysis. However, without more information on the cancer types that could be included in this category, it is unclear what it means to adjust for “unknown” cancer type. This should be clarified and, if necessary, discussed in the limitations. The authors should also consider performing a sensitivity analysis in which these individuals are excluded from the analysis.

Response to reviewer:

We have added further clarification under both the “Materials and Methods” section (pg.6, lines 35-38) and the “Discussion” section (pg.10, lines 38-43). Data on cancer stage are limited to certain types of cancer, but the Ontario Cancer Registry system is improving the data collection of this variable including unknown tumour type, which could be beneficial for future work and analyses.

e) Table 1 outlines the proportion of patients with “other” cancer site which is 29% of the population or more (depending on the smoking status), and the most commonly reported site type. However, in the results, the next most common cancers are identified as the most common type of cancer. This result needs to be more accurately reported in the results. The patients labelled “other” cancer site is a large proportion of the population; however, the authors do not discuss what types of cancers sites are categorized as “other”. It appears these individuals were included in the analysis. However, without more information on the cancer sites that could be included in this category, it is unclear what it means to adjust for “other” cancer site. This should be clarified and, if necessary, discussed in the limitations. The authors should also consider performing a sensitivity analysis in which these individuals are excluded from the analysis.

Response to reviewer:

Thank you for your comments. We have added definition for types of cancer included in the “Other” category (please refer to the “Materials and Methods” section, pg. 6, lines 32-35). We have also added in the “Discussion” section as a limitation that due to a small sample size for each of these other types of cancers, we could not examine them separately (pg. 10, lines 41-43). As a whole (all other types of cancer combined), we could still explore the impact of smoking and hence our proposed approach.

f) The GLM results should report the 95% confidence interval for the estimated monthly healthcare cost, which would be more informative than a p-value.

Response to reviewer:

Thank you for your comments. We have added information to report the uncertainty of the findings as well (pg. 9).

g) It is not clear why the data on smoking status was “limited to one assessment at the first ambulatory visit” as the methods describe that the information was submitted on a monthly basis to the CCO. The authors should clarify this.

Response to reviewer:

This is a good question. Yes, the data are submitted on a monthly basis to CCO. However, the question was asked only one time to each patient (normally at their first ambulatory care visit or within 28 days of that visit). Therefore, we only have smoking status data for each patient at one point in time before they started the cancer treatment. We have added this clarification to the manuscript. Please refer to the “Materials and Methods” section (pg. 5, lines 4-7), and the “Discussion” section (pg. 10, lines 22-23). Thank you.

VERSION 2 – REVIEW

REVIEWER	K. Michael Cummings Medical University of South Carolina United States
REVIEW RETURNED	07-Feb-2019

GENERAL COMMENTS	Nice job responding to reviewer comments. In the discussion it would be helpful if the authors could add a comment on the need for research studies to document the specific clinical and financial benefits of smoking cessation as part of clinical care. Perhaps clinical and financial benefits could be linked to clinical trials evaluating different approaches for smoking cessation in newly diagnosed cancer patients.

REVIEWER	Summer Frank-Pearce University of Oklahoma Health Sciences Center, US
REVIEW RETURNED	20-Feb-2019

GENERAL COMMENTS	I reviewed the original submission with a particular emphasis on the statistical methods and analyses used, as requested by the editor. My review outlined several concerns, most of which were thoughtfully addressed by the authors’ revision and response letter. However, there are two points that should be addressed before this manuscript is published. 1) The terminology regarding cancer types is still confusing. The phrase “main cancer types” is used as the heading in table 2 for the cancer categories, one of which is the “other” category. However, in the results and discussion section, the authors seem to be making a distinction between the categories that include only one type of cancer (which are all categories except “other”) and the “other” category, but are still using the phrase “main cancer types” which according to the table includes the “other” category. In the discussion, the authors clearly state that the cancers included in the category “other” were less common cancers.

	Perhaps the authors should make the distinction of common and less common clear in the methods section. Then in the sentence describing the results in table 2, they could add the phrase "Among the more common cancer types" to make it clear that they are talking about cancer categories that are not "other". With the current terminology the authors appear to be misrepresenting the data by not clearly stating that the most common category is actually "other". However, this is just an issue of confusing terminology that can be remedied before publication. 2) Reporting the interquartile range of the estimated monthly healthcare cost is a valuable addition to the paper. However, it only provides information on the spread of monthly healthcare cost values. It is still important to provide the 95% confidence interval for the parameter estimate from the GLM model, an estimate of the uncertainty in the estimate of that model-based parameter. Before publication, the authors need to add this confidence interval for the adjusted model to the manuscript.

VERSION 2 – AUTHOR RESPONSE

1. Comments from reviewer 1, K. Michael Cummings:

Nice job responding to reviewer comments. In the discussion it would be helpful if the authors could add a comment on the need for research studies to document the specific clinical and financial benefits of smoking cessation as part of clinical care. Perhaps clinical and financial benefits could be linked to clinical trials evaluating different approaches for smoking cessation in newly diagnosed cancer patients.

Response to reviewer:

Thank you for this. We have included in the discussion (page. 11, lines 4-6) a suggestion that future research could consider documenting the specific clinical and financial benefits of smoking cessation as part of clinical care to allow future evaluation of smoking cessation programs.

2. Comments from reviewer 3, Summer Frank-Pearce

I reviewed the original submission with a particular emphasis on the statistical methods and analyses used, as requested by the editor. My review outlined several concerns, most of which were thoughtfully addressed by the authors' revision and response letter. However, there are two points that should be addressed before this manuscript is published.

Response to reviewer:

We would like to thank the reviewers for very helpful comments in the previous and current versions and we have updated the manuscript accordingly.

a) The terminology regarding cancer types is still confusing. The phrase "main cancer types" is used as the heading in table 2 for the cancer categories, one of which is the "other" category. However, in the results and discussion section, the authors seem to be making a distinction between the categories that include only one type of cancer (which are all categories except "other") and the "other" category, but are still using the phrase "main cancer types" which according to the table includes the "other" category. In the discussion, the authors clearly state that the cancers included in the category "other" were less common cancers. Perhaps the authors should make the distinction of common and less common clear in the methods section. Then in the sentence describing the results in table 2, they could add the phrase "Among the more common cancer types" to make it clear that

they are talking about cancer categories that are not “other”. With the current terminology the authors appear to be misrepresenting the data by not clearly stating that the most common category is actually “other”. However, this is just an issue of confusing terminology that can be remedied before publication.

Response to reviewer:

Thank you for your comment. We have updated the methods section so that there is a clear distinction between common cancers and less common cancers (page.6 lines 32-37, and page.7 lines 1-2), in the results section (page. 8, lines 5-8), and the discussion section (page. 10, lines 41-45). Table 2 (page. 8) has also been edited.

b) Reporting the interquartile range of the estimated monthly healthcare cost is a valuable addition to the paper. However, it only provides information on the spread of monthly healthcare cost values. It is still important to provide the 95% confidence interval for the parameter estimate from the GLM model, an estimate of the uncertainty in the estimate of that model-based parameter. Before publication, the authors need to add this confidence interval for the adjusted model to the manuscript.

Response to reviewer:

Thank you for your comment. We have added the 95% confidence interval to represent the uncertainty in the estimate in the results section (page 9, lines 13-14).