

Breast Cancer and Income Loss in Denmark: Heterogeneous Outcomes and Longitudinal Effects

Corresponding Author: Ms Emily Johnson

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The study by Johnson et al. under review presents an observational study that aims to quantify the financial losses experienced by breast cancer patients in Denmark using matched controls, and with the inclusion of a difference-in-difference analysis. Compared to previous studies addressing a similar research question and study population, this work stands out for its methodological robustness and, notably, it extends the follow-up period beyond that of previous research, accounting the long-term financial implications for patients, including the timing of when the income losses seem to dissipate. Moreover, by addressing the heterogeneity within the population, the authors have been able to identify subgroups whose income appears to be disproportionately affected by the cancer diagnosis. These methodological strengths, and the depth of its findings, make this study a likely important contribution to characterising the sustained economic impact of breast cancer in Denmark.

The conclusions of the study are original in their context but do not introduce entirely novel findings. Rather, they reinforce and validate the outcomes of prior research by employing more robust analytical techniques, thereby improving the reliability and generalizability of the results. The study's findings have potential applications in the wider field of health economics, especially for economic evaluations that adopt societal perspectives and would involve a quantification of lost economic potential due to breast cancer in Denmark. That said, the generalizability of the results to other contexts is limited by the specific social and economic protection measures that are available in Denmark, making it unlikely that the estimates could be used elsewhere.

The methodology employed in the study is both detailed and transparent. The authors provide a clear explanation of the data used, matching techniques, and difference-in-difference framework. Further explanation on what data had been made available through Statistics Denmark, especially in relation to analyses that the authors mention that they could not carry out, would be appreciated, even as part of supplementary material. The methods chosen to address the research question are suited to the study objectives and have been applied appropriately, demonstrating a sound understanding of the analytical tools required for this type of observational research. The use of fixed effects is well-justified given the study context, as it effectively accounts for unobserved time-invariant heterogeneity between the matched groups and isolates the effect of breast cancer diagnosis on financial outcomes.

The manuscript does not present critical flaws that would prevent its publication, but suggestions on clarifying some assumptions and/or limitations are presented below:

1. While acknowledged in the limitations, the authors do not provide a rationale for their decision not to account for cancer stage or subtype in their analysis, leaving it unclear whether this omission was due to a lack of available data or a deliberate methodological choice. Accounting for cancer stage and subtype is likely to be critical, as these factors significantly influence treatment intensity, duration, and associated financial costs. For example, patients with advanced-stage or aggressive subtypes often face higher out-of-pocket expenses due to extended treatment courses, additional supportive care needs, and potential loss of income over a likely prolonged period. The absence of these variables could limit the study's ability to fully capture the heterogeneity of financial impacts experienced by patients across the disease spectrum. Providing an explanation for this—whether due to data constraints or other considerations—would improve the transparency and comprehensiveness of the study. Moreover, if data on stage and subtype were indeed unavailable, discussing potential implications and suggesting directions for future research could strengthen the study's contribution to the literature.

2. Additionally, and related to the previous point, the inclusion criteria specify “no history of previous breast cancer in the past 5 years.” While this approach may have been designed to avoid double-counting events or including patients with multiple diagnoses, it introduces a potential concern. This criterion is likely to exclude cases of early-stage breast cancer (with some potentially diagnosed before the study period) that recur within five years and progress to advanced or metastatic stages (with the second diagnosis within the study period). Such patients are likely to experience significant financial losses due to the compounded challenges of recurrent cancer and the intensive treatments often required. This criterion, while understandable in its intent to streamline the analysis and focus on first diagnoses, raises questions about the rationale behind the five-year cutoff versus a more stringent “no prior history of breast cancer” criterion. Clarification or further discussion from the authors on why they opted for the five-year threshold would address this potential limitation (even if this 5-year cut-off is standard in the literature).

3. The assumption and process described on Line 404 (and 131), wherein a deceased individual’s income is set to zero if death happens before the retirement age of 65, appears inconsistent with the stated perspective of the analysis. While this approach technically aligns with the comparison framework—since the matched controls are still alive and earning income—it raises questions about its relevance to the patient perspective, which is ostensibly the focus of this part of the analysis. From a societal perspective, setting income to zero in the case of death could appropriately reflect productivity losses due to cancer mortality. However, within the framework of personal financial losses, this assumption seems less appropriate, as it does not account for the fact that a deceased individual is no longer personally experiencing the financial impacts of income loss.

Moreover, while the authors acknowledge that the same process was not applied to household income due to the need for more nuanced and stronger assumptions (line 411), this distinction further complicates the interpretation. Categorizing a deceased individual’s income as a personal financial loss arguably requires even stronger assumptions than applying the same logic to household income. In the latter case, the household survives the deceased and directly experiences the economic fallout, making the loss tangible and measurable.

The discussion on Lines 245–248, which reframes this issue as lost economic potential, aligns better with a societal perspective, emphasizing the broader impact of cancer mortality on economic productivity. However, figures derived from this framing should ideally account for welfare substitution effects, as these could partially mitigate the societal cost of income losses due to death. Clarifying these distinctions and providing a more cohesive rationale for these methodological choices would improve the study’s conceptual consistency.

4. The authors note that no analysis has been conducted to decompose the income effect from the substitution effect, which would be instrumental in isolating the pure financial impact of breast cancer on income losses from the mitigating effects of welfare. This type of analysis would untangle the direct economic burden of the disease from the support mechanisms provided by welfare systems, offering a clearer understanding of the unmitigated financial challenges faced by patients. However, it is not specified whether this omission is due to a lack of access to the necessary data or because it falls outside the original scope of the research.

Providing a clear statement regarding why this analysis was not undertaken would help readers better understand the boundaries of the research. If the limitation is due to data constraints, the authors could highlight the types of data that would be required to perform such analysis, offering guidance for future research efforts. Conversely, if the decision reflects a deliberate narrowing of scope, the authors could explain how this aligns with their study objectives. In either case, explicitly expanding on this limitation would be beneficial to the study’s clarity and transparency.

Furthermore, should access to relevant data be available, conducting an analysis to quantify the pure income losses would add significant value to the research. It would provide a more precise estimation of the economic burden attributable to breast cancer itself, independent of the cushioning effects of welfare systems, and could offer more targeted insights for policymakers aiming to address these financial challenges.

5. On Line 106, the authors mention computational limitations but do not specify what these limitations entail or the rationale behind choosing a single representative year of incidence for their analysis. Clarifying the nature of these computational constraints—whether related to processing power or statistical challenges—would provide greater transparency and better representation for the selected population. Additionally, the choice of a single representative year of incidence is not explained, leaving readers to question why this particular year was selected over others. Was it chosen for reasons related to data completeness, or to reflect a typical year in terms of healthcare utilization and economic conditions?

The conclusions and data interpretation presented in the study are robust, valid, and reliable, reflecting the authors’ rigorous application of methodologies and analytical processes. The study’s use of matched controls and a difference-in-difference framework ensures that observed financial impacts are meaningfully attributed to the breast cancer diagnosis rather than confounding factors. Moreover, the overall transparency in reporting methods and results reinforces the credibility of findings, allowing for consistent interpretation by readers. The conclusions are supported by a thorough analysis of both aggregate trends and subgroup differences and demonstrates a nuanced understanding of the data. By integrating appropriate statistical techniques and addressing potential biases where possible, the authors provide a substantiated interpretation of the financial burdens associated with breast cancer. Overall, these factors make the study a reliable source for informing policy and guiding potential future research.

The abstract is well-written and accessible. The abstract, introduction, and conclusion sections are appropriately structured, presenting a logical progression of the study’s purpose, approach, and implications. However, while the abstract provides a valuable overview, it would benefit from the inclusion of more specific numeric estimates rather than using the term “modest” to describe the financial losses observed.

Nothing of concern in terms of inflammatory material, or Springer Nature’s commitment to diversity, equity, and inclusion.

(Remarks on code availability)

none

Reviewer #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

Thank you for the opportunity to review this interesting manuscript. It is a matched cohort study on the estimation of some indirect costs of breast cancer. The authors analysed a set of linked individual-level data on Danish female breast cancer patients and estimated the causal effect of breast cancer diagnosis on personal and household income in the course of a 10-year period.

The effect was higher on personal than household income, particularly among the non-retired population. Personal income loss peaks at 5 years after diagnosis and remains significant at 10 years after diagnosis. Household income loss peaks at 3 years and recovers by 10 years post-diagnosis. Accounting for mortality, magnifies the effect on income loss. The population groups most affected are individuals in school at the time of diagnosis and those in worse health. Interestingly, personal income of retirees is overall higher than the controls.

This study is original in the design, analytical methods and wealth of data. It also has the advantage of looking at a longer period after diagnosis than similar studies, which helps to identify when income loss peaks, if it dissipates and when. These findings can help other countries with similar wealth and resources as Denmark, to design welfare policies to support breast cancer survivors. However, as authors have partly acknowledged and partly adjusted for with the DiD analysis, prognosis of breast cancer might have changed within those 10-20 years for reasons which could be further discussed in the Discussion section. Could they perhaps add a few points in the penultimate paragraph, explaining some of these reasons in the specific context of Denmark ?

I think authors should also acknowledge that although personal income losses increase with higher income quartile, this is perhaps a reflection of the overall higher income in those groups than in the more deprived. More deprived patients may already be getting support from the welfare system in which case the loss in income will be less than in the more affluent. In the same context, it has to be acknowledged that loss in personal or household income is part of the indirect costs of cancer. There is also the Wider Societal Impact that includes unpaid activities such as child-care, domestic work and volunteering, all of which may affect different population groups differently, widening or reducing inequalities between age-groups, income quartiles, education groups etc.

The manuscript has good flow between the different parts, and conclusions are cautious. I would however, suggest a few changes to just improve clarity and flow. The analysis has different levels and is complex, so I would suggest that the authors give more specific sub-titles in the Results section oriented to what they describe. Also, the analysis on the non-retired population could be a different section.

In addition, in the first paragraph of the Results section, authors list the sub-groups with few individuals. I think, instead of the proportions, the reader would be more interested in a summary of which those sub-groups were as numbers can be found in the Supplementary files.

A "cosmetic" suggestion: could you please add the years on the horizontal axis in Figure 5?

Overall, this is a well thought out study with rigour in methods and attention to detail. The findings are interesting and original but I would recommend a more comprehensive Discussion that puts everything in context. I would also recommend that the Methods should be reviewed by a statistician with expertise in the causal framework.

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)

This paper studies breast cancer development and corresponding income loss in Denmark. The paper has many strengths, such as the valuable and interesting data source, the methodological approach, and the clear writing. The topic is well trod, the authors make a point of identifying how their particular approach is better than others (a testament to the data's quality), but ultimately the perception of the contribution of the paper is how important one treats the income loss from breast cancer. I imagine the authors could do a similar analysis for any cancer, maybe they intend to do a complete analysis of all income lost to cancer in the future. Breast cancer specifically is heavily studied and one of the most survivable cancers (despite the authors' claim it has a high mortality rate), so while the paper is quite good the topic might be unexciting for some. To me that's not a failing of the paper, but I am writing it here in the comments to the authors for their information; I find this paper interesting but many in the general audience might be exhausted from the attention breast cancer commanded in the past 20

years.

The paper runs both a matching exercise and an event study to determine the amount of money a person/household loses upon a diagnosis of breast cancer. The results are presented succinctly and the loss seems to be moderate for the population, but the authors identify some specific groups for whom the loss of income is more substantial. I will organize my comments into major and minor:

Major comments.

1. The result has no context. There is no real declaration of whether the losses are large or small, only larger or smaller than other estimates. For example, the first estimate of income loss per person is 3636 Euros over the observation period. Over ten years, is 363.3 Euros a lot of money? What does that mean in Denmark, how about compared to other sources of income loss from the literature? What is the equivalent to this loss? Without this context, the reader could think that less than one Euro a day is not that bad and potentially testament to how great the Danish system is at mitigating the consequences of cancer. The authors suggest that their results show exactly this, but when they make that statement it's unclear if they are referring to the overall estimate being small or they are referring to Denmark having a lot of social programs which jointly, somehow, support people with cancer. They also make the point that most people who get breast cancer are retirement age and on income supports already, it's basically unclear whether these results are important for society.

2. Incomes are not really comparable. Denmark's system of illness supports and retirement benefits are identified in the paper, which complicates the idea of "lost income" among retirees. So the lost income from getting breast cancer while on a series of guaranteed sources of income is not obviously important, those people are already supported. The people who are earning wages, however, who are treated as a subgroup by the authors, are the people who immediately spring to mind when using the term "lost income due to breast cancer". I would think it is common to assume the authors are talking about people who were otherwise earning an income in the labour market, rather than people collecting a guaranteed income. This group has a larger estimate of income lost than the general population (which includes a lot of the older adults) and death in this group is more consequential from an income and years of life lost perspective. My advice would be to put this group upfront as the main group for the analysis. Their lost income is true value gone, they were creating something of value in the labour market, and that's the value of breast cancer prevention, rather than someone dying and no longer receiving government benefits (which is savings from a government perspective.) This point is why statements like the ones in para starting p7 line 249 are not evident, it is not clear from this analysis that mitigation of the consequences of breast cancer is coming from government policies.

3. The discussion of vulnerable groups is interesting but underdeveloped. The authors mention students when they speak of "vulnerable groups" on p8 para starting line 276. Breast cancer is not homogeneous, so I imagine that the versions that affect young women to have different consequences medically. Thus this discussion is important, are there other vulnerable groups worth mentioning? If not, then the general term "vulnerable groups" should be replaced with something more specific, which ties into my next point.

4. General terms when specific are better abound. The authors use the term "health" to describe adjusting for a comorbidities index, and that will be poorly received by people who like to point to more general definitions of health than "diagnosed with one or more of a list of diseases." Further, the issues the authors mention, like depression from a diagnosis, are not included in the Charlson index so there are consequential health conditions that are unadjusted for by this "adjustment for health". I would recommend using the specific term.

Other examples are health care use not being the same as a hospital admission, the term "baseline" being unclear (is it entry into the cohort?), and what disposable income is in Denmark vis-à-vis mandatory deductions that might not exist in other countries.

Minor comments.

5. The authors set income = 0 for people who die at one point, I cannot tell if that's a temporary thing or they only did so at one point to emphasize the importance of death. I think that if they use an expected age of death (75, or similar) then this method could identify losses from breast cancer among working people. If it's a temporary thing then simply adding "for this analysis only" or similar wording would clear that up.

6. It is not obvious that the household income is adjusted for household size until the final section, which makes the household income < personal income facts confusing. Including "adjusted for household size" at the first mention of household income solves this problem.

7. Figures need more descriptive titles. E.g., Figure 2 is completely reliant on the reader knowing what the CACE is in the text, which is inefficient. I recommend trying to make the figures stand alone with titles and enough description in the subtitle.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Thank you for taking into account my comments. I am satisfied that the paper is now suitable for publication.

(Remarks on code availability)

Accurately reflects the analysis

Reviewer #3

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Code is well structured and appropriately commented. Results are reproducible, and a README file with an appropriate level of instructions is present.

Reviewer #4

(Remarks to the Author)

Thank you for addressing the comments provided. The manuscript is improved.

(Remarks on code availability)

Reviewer #5

(Remarks to the Author)

The authors have addressed all my comments effectively.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Dear reviewers,

We are very appreciative of the thoughtful and well articulated comments on our manuscript. Below, we provide point by point responses to each reviewer along with quotes from the manuscript and supplement of the changes which we have implemented. Comments and quotes are provided in red text.

Reviewer #1 (Remarks to the Author):

The study by Johnson et al. under review presents an observational study that aims to quantify the financial losses experienced by breast cancer patients in Denmark using matched controls, and with the inclusion of a difference-in-difference analysis. Compared to previous studies addressing a similar research question and study population, this work stands out for its methodological robustness and, notably, it extends the follow-up period beyond that of previous research, accounting the long-term financial implications for patients, including the timing of when the income losses seem to dissipate. Moreover, by addressing the heterogeneity within the population, the authors have been able to identify subgroups whose income appears to be disproportionately affected by the cancer diagnosis. These methodological strengths, and the depth of its findings, make this study a likely important contribution to characterising the sustained economic impact of breast cancer in Denmark.

Thank you for your summary and kind words.

The conclusions of the study are original in their context but do not introduce entirely novel findings. Rather, they reinforce and validate the outcomes of prior research by employing more robust analytical techniques, thereby improving the reliability and generalizability of the results. The study's findings have potential applications in the wider field of health economics, especially for economic evaluations that adopt societal perspectives and would involve a quantification of lost economic potential due to breast cancer in Denmark. That said, the generalizability of the results to other contexts is limited by the specific social and economic protection measures that are available in Denmark, making it unlikely that the estimates could be used elsewhere.

We agree about the primary contributions of this work and the uniqueness of the Danish context in terms of the generalizability of results. We also agree that some elements of our analysis – particularly our analytical perspective and our methodology – do have potential applications in the health economics field more widely. For this reason, we have clarified this in the concluding paragraph of the Discussion (page 10):

Our comprehensive methodology captures heterogeneous effects across the population and adds rigor to the literature on the economic consequences of breast cancer, and we suggest that similar approaches which control for confounding and follow a longer period should be adopted in other country-level cost of illness studies where data allows.

The methodology employed in the study is both detailed and transparent. The authors provide a clear explanation of the data used, matching techniques, and difference-in-difference framework. Further explanation on what data had been made available through Statistics Denmark, especially in relation to analyses that the authors mention

that they could not carry out, would be appreciated, even as part of supplementary material. The methods chosen to address the research question are suited to the study objectives and have been applied appropriately, demonstrating a sound understanding of the analytical tools required for this type of observational research. The use of fixed effects is well-justified given the study context, as it effectively accounts for unobserved time-invariant heterogeneity between the matched groups and isolates the effect of breast cancer diagnosis on financial outcomes.

Thank you for this detailed consideration of our data and methods. We have added additional information under Supplementary Material Table 1 concerning the contents of each register, including the breast cancer register:

Brief descriptions of each register:

Danish breast cancer group register: clinical information collected on December 31st of each year for every positive confirmed case of breast cancer within the year. Data includes information on the type and stage of tumor, the tumor size, and any procedures and treatments performed. Importantly, this data only includes the most recent characteristics of the cancer and the treatment performed, so any information about the disease at the first diagnosis is not possible to infer.

Danish national population register: demographic information for each resident of Denmark, collected on January 1st of each year. Data includes individual age, sex, relationship status, family identifier, household and residence information and immigration status.

Danish income statistics register: income information derived for tax purposes, collected on December 31st of each year. Data includes the personal and household disposable income levels, the primary source of individual and household income, and other income-related variables.

Integrated student register: educational attainment information collected on December 31st of each year. Data includes information about educational attainment within the data year as well as highest level of educational attainment which corresponds to ISCED categories.

National inpatient register: all inpatient and outpatient hospitalizations, including diagnosis, procedure, and length of stay information for each hospitalization. Data is collected on December 31st of each calendar year.

We have also clarified some of the important limitations of data availability in the Methods section (page 11):

We include all cases of breast cancer regardless of stage and subtype at diagnosis, as the register data does not distinguish these cancer characteristics at diagnosis from those recorded as the cancer progresses.

The manuscript does not present critical flaws that would prevent its publication, but suggestions on clarifying some assumptions and/or limitations are presented below:

1. While acknowledged in the limitations, the authors do not provide a rationale for their decision not to account for cancer stage or subtype in their analysis, leaving it unclear whether this omission was due to a lack of available data or a deliberate methodological

choice. Accounting for cancer stage and subtype is likely to be critical, as these factors significantly influence treatment intensity, duration, and associated financial costs. For example, patients with advanced-stage or aggressive subtypes often face higher out-of-pocket expenses due to extended treatment courses, additional supportive care needs, and potential loss of income over a likely prolonged period. The absence of these variables could limit the study's ability to fully capture the heterogeneity of financial impacts experienced by patients across the disease spectrum. Providing an explanation for this—whether due to data constraints or other considerations—would improve the transparency and comprehensiveness of the study. Moreover, if data on stage and subtype were indeed unavailable, discussing potential implications and suggesting directions for future research could strengthen the study's contribution to the literature.

We agree that cancer subtype and stage are an important factor in the personal experiences of individuals with breast cancer. We do not stratify our analysis by subtype, stage, or treatment partly because of limits to data availability, and partly because they were not in scope for this piece of work. To improve transparency on this topic, we have included additional detail in the methods and supplement to describe the data available from the Danish Breast Cancer Register and why stage and subtype cannot be identified (see quoted changes above).

We have also clarified in the Discussion section (page 9) that this is an important limitation which deserves expansion upon in future studies:

Our results do not account for changes in the underlying constitution of the breast cancer population over this time period, nor do they stratify by the type or stage of breast cancer at diagnosis due to limitations in the available data... Future work is necessary to understand how disease prognosis and treatment outcomes determine income loss patterns.

2. Additionally, and related to the previous point, the inclusion criteria specify “no history of previous breast cancer in the past 5 years.” While this approach may have been designed to avoid double-counting events or including patients with multiple diagnoses, it introduces a potential concern. This criterion is likely to exclude cases of early-stage breast cancer (with some potentially diagnosed before the study period) that recur within five years and progress to advanced or metastatic stages (with the second diagnosis within the study period). Such patients are likely to experience significant financial losses due to the compounded challenges of recurrent cancer and the intensive treatments often required.

This criterion, while understandable in its intent to streamline the analysis and focus on first diagnoses, raises questions about the rationale behind the five-year cutoff versus a more stringent “no prior history of breast cancer” criterion. Clarification or further discussion from the authors on why they opted for the five-year threshold would address this potential limitation (even if this 5-year cut-off is standard in the literature).

This is a key consideration. Our reasoning for a 5-year washout period was twofold: first, we wanted to ensure consistency across years in the length of the exclusion window. This prevents bias in the oldest years of analysis where only 5 years of lookback are available, as compared to a diagnosis in more recent calendar years where we can confirm incidence over 15 or 20 years. Second, we thought it appropriate to align our definition of incident breast cancer with the distinction made by clinical

guidelines between a recurrent case of cancer (within <5 years) and a new incident case (after 5 years cancer free). With our current design, we would measure any income loss which is exacerbated by recurrence, within the 10-year follow up period. We clarify the justification for our 5-year window in the Methods section (page 11):

In all models, the exposed breast cancer group is defined as all females in Denmark with an incident breast cancer diagnosis between 2000-2015 and no history of breast cancer for at least 5 years prior to date of diagnosis. A 5-year washout window was chosen in order to maintain consistency in the length of washout for each year of data, and to align with clinical guidelines which reduce screening for recurrence after 5 years.

3. The assumption and process described on Line 404 (and 131), wherein a deceased individual's income is set to zero if death happens before the retirement age of 65, appears inconsistent with the stated perspective of the analysis. While this approach technically aligns with the comparison framework—since the matched controls are still alive and earning income—it raises questions about its relevance to the patient perspective, which is ostensibly the focus of this part of the analysis. From a societal perspective, setting income to zero in the case of death could appropriately reflect productivity losses due to cancer mortality. However, within the framework of personal financial losses, this assumption seems less appropriate, as it does not account for the fact that a deceased individual is no longer personally experiencing the financial impacts of income loss.

Moreover, while the authors acknowledge that the same process was not applied to household income due to the need for more nuanced and stronger assumptions (line 411), this distinction further complicates the interpretation. Categorizing a deceased individual's income as a personal financial loss arguably requires even stronger assumptions than applying the same logic to household income. In the latter case, the household survives the deceased and directly experiences the economic fallout, making the loss tangible and measurable.

The discussion on Lines 245–248, which reframes this issue as lost economic potential, aligns better with a societal perspective, emphasizing the broader impact of cancer mortality on economic productivity. However, figures derived from this framing should ideally account for welfare substitution effects, as these could partially mitigate the societal cost of income losses due to death. Clarifying these distinctions and providing a more cohesive rationale for these methodological choices would improve the study's conceptual consistency.

This is an insightful comment. The question of measuring premature mortality in this project was one that we considered deeply, and we believe that there are many possible approaches which make different assumptions. Nonetheless, the difference in mortality between the breast cancer and control population was something that we thought deserved discussion given its impact on the results, especially over 10 years and across subpopulations. In response to your suggestions, we have added Section 2 to the Supplement which describes our considerations for the mortality analysis in more detail, including acknowledging the shift in perspective and the reason that we could not apply the same methods to household equivalized income:

Section 2: Mortality and income loss: methods and interpretation of secondary analysis

Differential mortality rates over time between the control and exposed populations, and within exposed subpopulations, have implications for the interpretation of cumulative income loss. Survival in each year is required to observe impacts on income in that year, but this skews the population distribution overtime to the population who is able to survive.

In our primary analysis, we condition on survival so that income loss measured in each year is not adjusted for the mortality rate in each match group. As a secondary analysis, we aim to illustrate the effect of premature mortality alongside income losses in the survivors, to demonstrate the theoretical value lost by early death at the population level. To do so, we impute years after death within the follow-up period as years of zero income, as long as the death occurs before age 65 which is the retirement age in Denmark. Deaths of any cause for the control or exposed population are included. This was only possible for personal income loss, as household income losses would require identification of all income earners who constitute the household and could not be applied to equivalized household income since this measure is adjusted for the size of the household in each year.

Unlike our primary results, this secondary analysis does not conform to the perspective of individual experienced income loss, and there are other methods which could be more consistent with this perspective, but which require other assumptions. The income losses reported in this secondary analysis are intended as an exploratory analysis to illuminate the outsized impact of mortality, but a more rigorous comparison of mortality adjustment methods is warranted to draw conclusions about the non-mortality counterfactual.

We have also clarified the language in the lines you mention (page 7) to draw more appropriate conclusions from our analysis.

Although these economic losses are modest, it is important to recognize the outsized effect that premature mortality has on the population with breast cancer and it is likely that the most economically vulnerable are also those at highest risk of early death, skewing income loss results towards the survivors.

4. The authors note that no analysis has been conducted to decompose the income effect from the substitution effect, which would be instrumental in isolating the pure financial impact of breast cancer on income losses from the mitigating effects of welfare. This type of analysis would untangle the direct economic burden of the disease from the support mechanisms provided by welfare systems, offering a clearer understanding of the unmitigated financial challenges faced by patients. However, it is not specified whether this omission is due to a lack of access to the necessary data or because it falls outside the original scope of the research.

Providing a clear statement regarding why this analysis was not undertaken would help readers better understand the boundaries of the research. If the limitation is due to data constraints, the authors could highlight the types of data that would be required to perform such analysis, offering guidance for future research efforts. Conversely, if the decision reflects a deliberate narrowing of scope, the authors could explain how this aligns with their study objectives. In either case, explicitly expanding on this limitation would be beneficial to the study's clarity and transparency.

Furthermore, should access to relevant data be available, conducting an analysis to quantify the pure income losses would add significant value to the research. It would provide a more precise estimation of the economic burden attributable to breast cancer itself, independent of the cushioning effects of welfare systems, and could offer more targeted insights for policymakers aiming to address these financial challenges.

We agree that it is highly relevant to understand the independent occurrences of wage loss following a diagnosis and the uptake of welfare payments to mitigate income loss. Though out of scope for this paper, we are working now on a second project which measures welfare payouts following a breast cancer diagnosis stratified by the type of payout (sickness leave, long-term disability, etc.). Unfortunately, separate measures of wage income were unavailable to us for this analysis. We have clarified the limitations of scope for this analysis in the Methods section (page 11):

Economic burden is sourced from the Danish Income Statistics Registry, and we collect disposable personal income and disposable equivalized household income which adjusts total household earnings for the household size [18]. Disposable income was chosen as an outcome measure to reflect the ultimate financial experience of individuals with breast cancer, and as stratifications by income source deserve their own body of work they were considered out of scope for this analysis.

We have also added clearer suggestions for future work in the Discussion section (page 9-10):

... disaggregation between sources of income into wages and welfare benefits, though relevant to the broader topic of economic burden, was beyond the scope of this analysis due to data limitations. Consideration of all types of disposable income best reflects our perspective which is concerned with individual experience of income loss. However, estimation of direct wage loss and separate measurement of the mitigating effects of welfare payment would be highly relevant information for policymakers.

5. On Line 106, the authors mention computational limitations but do not specify what these limitations entail or the rationale behind choosing a single representative year of incidence for their analysis. Clarifying the nature of these computational constraints—whether related to processing power or statistical challenges—would provide greater transparency and better representation for the selected population. Additionally, the choice of a single representative year of incidence is not explained, leaving readers to question why this particular year was selected over others. Was it chosen for reasons related to data completeness, or to reflect a typical year in terms of healthcare utilization and economic conditions?

Thank you for this point. We have clarified our decision about using the single year for descriptive characteristics in the Results section (page 3) and also clarified why the specific single year was chosen:

Descriptives are shown for the most recent match year of analysis to avoid reporting duplicate counting, as individuals are eligible to act as controls for every year of the data that they qualify. The population characteristics shown for the exposed and control groups do not vary strongly year over year.

The conclusions and data interpretation presented in the study are robust, valid, and reliable, reflecting the authors' rigorous application of methodologies and analytical processes. The study's use of matched controls and a difference-in-difference framework ensures that observed financial impacts are meaningfully attributed to the breast cancer diagnosis rather than confounding factors. Moreover, the overall transparency in reporting methods and results reinforces the credibility of findings, allowing for consistent interpretation by readers. The conclusions are supported by a thorough analysis of both aggregate trends and subgroup differences and demonstrates a nuanced understanding of the data. By integrating appropriate statistical techniques and addressing potential biases where possible, the authors provide a substantiated interpretation of the financial burdens associated with breast cancer. Overall, these factors make the study a reliable source for informing policy and guiding potential future research.

The abstract is well-written and accessible. The abstract, introduction, and conclusion sections are appropriately structured, presenting a logical progression of the study's purpose, approach, and implications. However, while the abstract provides a valuable overview, it would benefit from the inclusion of more specific numeric estimates rather than using the term "modest" to describe the financial losses observed.

Nothing of concern in terms of inflammatory material, or Springer Nature's commitment to diversity, equity, and inclusion.

Thank you for your feedback and thoughtful review of our conclusions, interpretation, and abstract. We have included numeric estimates in the Abstract to improve the specificity of the summarized results:

Our study finds personal income loss averaging €7138 over 10 years in the non-retired population.

Reviewer #1 (Remarks on code availability):

none

Reviewer #3 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review

and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Thank you for your thorough consideration and comments.

Reviewer #4 (Remarks to the Author):

Thank you for the opportunity to review this interesting manuscript. It is a matched cohort study on the estimation of some indirect costs of breast cancer. The authors analysed a set of linked individual-level data on Danish female breast cancer patients and estimated the causal effect of breast cancer diagnosis on personal and household income in the course of a 10-year period.

The effect was higher on personal than household income, particularly among the non-retired population. Personal income loss peaks at 5 years after diagnosis and remains significant at 10 years after diagnosis. Household income loss peaks at 3 years and recovers by 10 years post-diagnosis. Accounting for mortality, magnifies the effect on income loss. The population groups most affected are individuals in school at the time of diagnosis and those in worse health. Interestingly, personal income of retirees is overall higher than the controls.

This study is original in the design, analytical methods and wealth of data. It also has the advantage of looking at a longer period after diagnosis than similar studies, which helps to identify when income loss peaks, if it dissipates and when. These findings can help other countries with similar wealth and resources as Denmark, to design welfare policies to support breast cancer survivors. However, as authors have partly acknowledged and partly adjusted for with the DiD analysis, prognosis of breast cancer might have changed within those 10-20 years for reasons which could be further discussed in the Discussion section. Could they perhaps add a few points in the penultimate paragraph, explaining some of these reasons in the specific context of Denmark?

Thank you for your thorough comments and for this suggestion. We have updated the penultimate paragraph of the Discussion with more specific reference to the Danish context of cancer treatment guidelines:

Treatment guidelines for breast cancer have become more targeted and personalized over the study period, with aims to improve quality of life and reduce mortality¹⁸. In principle, treatments which improve quality of life could have a mitigating impact on potential economic losses, though our calendar year-specific results indicate that income losses at the aggregate do not vary substantially over time as treatment guidelines were modified. Future work is necessary to understand how disease prognosis and treatment outcomes determine income loss patterns.

I think authors should also acknowledge that although personal income losses increase with higher income quartile, this is perhaps a reflection of the overall higher income in those groups than in the more deprived. More deprived patients may already be getting support from the welfare system in which case the loss in income will be less than in the more affluent.

We appreciate your suggestion to acknowledge the effects of loss relative to the income level of the individual, and we agree that it is important to point out that a smaller loss for a lower-income group may still cause more harm than a loss for someone with more economic resources. We've made the following revision to our subgroup loss section in the Discussion:

... although income losses are greatest in absolute terms for the highest income quartiles, relative losses are in the range of 1-2% for all income quartiles except the lowest one, where substitution of wages by welfare benefits is most significant. The implication of this relative loss is that while high-income individuals may have the greatest income available to lose, smaller losses in the rest of the income distribution will have an equally meaningful impact on personal budget in the short term, and wealth accumulation in the long-term.

We've additionally made clearer the impact of welfare substitution on the lowest income quartile in the Discussion:

Income for the lowest income quantile cohort increased a small amount following breast cancer diagnosis, which represents an effective transition of this population from social security benefits to disability benefits or other welfare protections. Nonetheless, individuals on fixed income before diagnosis including those in school, those on leave and the unemployed population still experience losses compared to individuals with the same economic situation but without a breast cancer diagnosis, indicating the potential of long-term accumulation of income effects which are not fully mitigated by the Danish welfare system.

In the same context, it has to be acknowledged that loss in personal or household income is part of the indirect costs of cancer. There is also the Wider Societal Impact that includes unpaid activities such as child-care, domestic work and volunteering, all of which may affect different population groups differently, widening or reducing inequalities between age-groups, income quartiles, education groups etc.

Thank you for this suggestion. We have added the following to our second paragraph in the Discussion to put income loss in context of the broader impact of breast cancer:

Additionally, disposable income losses are only one indicator of the broader social and economic impact of breast cancer. Unpaid labor including caregiving and domestic work, and additional economic measures including wealth accumulation are not represented by disposable income loss, but are likely affected by a breast cancer diagnosis with different trajectories over time and across sub-populations.

The manuscript has good flow between the different parts, and conclusions are cautious. I would however, suggest a few changes to just improve clarity and flow. The analysis has different levels and is complex, so I would suggest that the authors give more specific sub-titles in the Results section oriented to what they describe. Also, the analysis on the non-retired population could be a different section.

We appreciate this recommendation. We now include the following subsections in the Results:

- Data and sample size
- Aggregate income losses
- Subgroup variation in income loss

Event study difference-in-difference
Model specification and robustness

We have also decided to restructure the primary results to focus on the non-retired population, where income losses are the greatest since retirees receive a fixed income. The revisions to the third paragraph in the Results are as follows:

When exact matching on all variables described above and censoring at death, the total ACE of breast cancer on personal income for non-retirees is €7138 (95% CI €6777 – €7498) in personal income loss over 10 years, representing 1.50% (1.40% – 1.55%) of the income of the control population over this time period. Losses for the entire population including retirees totaled €3636 (€3392 – €3881) over the period, which reflects the protection of retirees by the fixed pension payouts which are agnostic to disease status.

In addition, in the first paragraph of the Results section, authors list the sub-groups with few individuals. I think, instead of the proportions, the reader would be more interested in a summary of which those sub-groups were as numbers can be found in the Supplementary files.

Thank you for this suggestion. We cannot easily include a table of this type because of the number of total match groups: a table with one row for each match group with a small count of exposed individuals would be over 4500 rows. Most of the small groups occur because of rare combinations of variables – for example, individuals aged 60-64 in school. This also makes it hard to summarize at aggregate levels, as it is only the unique combination of multiple variables which makes the groups so small. We've included more clarifying language in this paragraph, in lieu of a table:

Small groups were most often defined by rare combinations of match variables, such as individuals unemployed or on leave (but not disabled) with high education and multiple comorbidities.

A “cosmetic” suggestion: could you please add the years on the horizontal axis in Figure 5?

We have made this cosmetic suggestion for further clarity in Figure 5.

Overall, this is a well thought out study with rigour in methods and attention to detail. The findings are interesting and original but I would recommend a more comprehensive Discussion that puts everything in context. I would also recommend that the Methods should be reviewed by a statistician with expertise in the causal framework.

Thank you for your review and your comments. In addition to the specific points of feedback you provide for our Discussion, we have a few additional changes which put our results into context and make the Discussion more comprehensive.

First, we include direct comparisons of the scale of income loss in the Danish context:

Both [household and personal income loss] represent a small proportion of the average household budget in Denmark, and both income loss measures recover within 10 years of diagnosis. Although these economic losses are modest, it is important to recognize the outsized effect that premature mortality has on the population with breast cancer and it is likely that the most economically vulnerable are also those at highest risk of early death, skewing income loss results towards the survivors.

Additionally, we have expanded our discussion of vulnerable subpopulations to include those in poor baseline health:

... individuals with higher hospital utilization before diagnosis and those with a higher CCI have more profound and persistent income effects than those in better health at baseline. These increasing losses reflect the financial challenges which can be brought on by multimorbidity and a gap in protection for those who are already in poor health.

Reviewer #5 (Remarks to the Author):

This paper studies breast cancer development and corresponding income loss in Denmark. The paper has many strengths, such as the valuable and interesting data source, the methodological approach, and the clear writing. The topic is well trod, the authors make a point of identifying how their particular approach is better than others (a testament to the data's quality), but ultimately the perception of the contribution of the paper is how important one treats the income loss from breast cancer. I imagine the authors could do a similar analysis for any cancer, maybe they intend to do a complete analysis of all income lost to cancer in the future. Breast cancer specifically is heavily studied and one of the most survivable cancers (despite the authors' claim it has a high mortality rate), so while the paper is quite good the topic might be unexciting for some. To me that's not a failing of the paper, but I am writing it here in the comments to the authors for their information; I find this paper interesting but many in the general audience might be exhausted from the attention breast cancer commanded in the past 20 years.

The paper runs both a matching exercise and an event study to determine the amount of money a person/household loses upon a diagnosis of breast cancer. The results are presented succinctly and the loss seems to be moderate for the population, but the authors identify some specific groups for whom the loss of income is more substantial.

Thank you for this overview and your perspective on the contributions of this article. We agree with your points about the ultimate contribution of this paper, and of the repeat emphasis on breast cancer in the literature. The amount of evidence on income loss due to breast cancer was a part of the motivation for us to write this paper: it presented an opportunity to illuminate the importance of methods construction in the resulting outcomes measurement and the potential threats of confounding which must be considered. We are also pursuing a comparison of income losses across diseases which have less attention in the literature using these methods in ongoing work.

Major comments.

1. The result has no context. There is no real declaration of whether the losses are large or small, only larger or smaller than other estimates. For example, the first estimate of income loss per person is 3636 Euros over the observation period. Over ten years, is 363.3 Euros a lot of money? What does that mean in Denmark, how about compared to other sources of income loss from the literature? What is the equivalent to this loss? Without this context, the reader could think that less than one Euro a day is not that bad and potentially testament to how great the Danish system is at mitigating the

consequences of cancer. The authors suggest that their results show exactly this, but when they make that statement it's unclear if they are referring to the overall estimate being small or they are referring to Denmark having a lot of social programs which jointly, somehow, support people with cancer. They also make the point that most people who get breast cancer are retirement age and on income supports already, it's basically unclear whether these results are important for society.

This is a very helpful piece of feedback. We recognize the importance of contextualizing the results, especially because they are modest in magnitude at the average level across the population. Though income losses are relatively small, the heterogeneity in effects across subpopulations despite the same sickness, disability and welfare provisions being equally available is something we believe is of relevance to policymakers.

We have modified the primary results section to add more context to the total income losses and have ensured that each income loss reported in absolute amounts in euros is paired with a relative loss which is a percent of the comparator control's income. An example of this change is shown here (page 5):

Individuals with postgraduate education experienced the highest absolute loss, at €20,072 (€17,735 – €22,408), or 2.60% (2.30% – 2.91%) of income in that period. Both health measures had a nonmonotonic relationship with income losses: individuals with breast cancer and a CCI of 3 have total losses of €22,945 (€19,160 – €26,731) or 6.50% (5.45% - 7.61%), while lower CCI scores have smaller losses, as low as €4245 (€2673 – €5817) or 1.00% (0.63% - 1.37%) for a CCI of 2. Individuals who have had 6-10 inpatient encounters in the third and second year before diagnosis lost €9985 (€9083 – €10,886) or 2.20% (2.00% - 2.40%), while those with no inpatient encounters lost €6408 (€5768 – €7047) or 1.30% (1.17% - 1.43%). Conversely, those with a CCI greater than 4 or with 11 or more inpatient encounters experience no measured effect on income. Breast cancer patients aged 40 to 45 have the highest income losses at any age at €11,732 (€10,585 – €12,879) or 2.10% (1.86% - 2.26%) and income losses are at a similar magnitude at prime working ages between 30 and 59. Individuals aged 18-29 have no significant income loss due to small sample size, and those age 60-64 has losses as low as €2685 (€1878 – €3492) or 0.70% (0.49% - 0.92%) of control income.

Additionally, we have clarified the interpretation of the magnitude of our results in the Results section (page 4):

This equates to approximately 8% of the annual budget for housing and utilities in Denmark, or half of the budget for restaurants and hotel accommodations¹⁷.

Finally, we have added more interpretation to the Discussion which put income losses in context (page 7):

Income losses at the household are approximately a quarter of the magnitude of personal income losses and less than half of one percent of the average non-retired household income of the control over 10 years. Both outcomes represent a small proportion of the average household budget in Denmark, and both income loss measures recover within 10 years of diagnosis.

2. Incomes are not really comparable. Denmark's system of illness supports and retirement benefits are identified in the paper, which complicates the idea of "lost income" among retirees. So the lost income from getting breast cancer while on a series of guaranteed sources of income is not obviously important, those people are already supported. The people who are earning wages, however, who are treated as a subgroup by the authors, are the people who immediately spring to mind when using the term "lost income due to breast cancer". I would think it is common to assume the authors are talking about people who were otherwise earning an income in the labour market, rather than people collecting a guaranteed income. This group has a larger estimate of income lost than the general population (which includes a lot of the older adults) and death in this group is more consequential from an income and years of life lost perspective. My advice would be to put this group upfront as the main group for the analysis. Their lost income is true value gone, they were creating something of value in the labour market, and that's the value of breast cancer prevention, rather than someone dying and no longer receiving government benefits (which is savings from a government perspective.) This point is why statements like the ones in para starting p7 line 249 are not evident, it is not clear from this analysis that mitigation of the consequences of breast cancer is coming from government policies.

This is an insightful comment. It is true that average income effects over the entire exposed population hide the mitigating impact of welfare benefits and the fact that retirees, who make up a large share of the breast cancer population, do not lose pension due to ill health. We have adopted your advice to report income losses on the non-retired population as our primary result, and to communicate upfront that the retired population does not experience loss (page 4):

When exact matching on all variables described above and censoring at death, the total ACE of breast cancer on personal income for non-retirees is €7138 (95% CI €6777 – €7498) in personal income loss over 10 years, representing 1.50% (1.40% – 1.55%) of the income of the control population over this time period. Losses for the entire population including retirees totaled €3636 (€3392 – €3881) over the period, which reflects the protection of retirees by the fixed pension payouts which are agnostic to disease status.

We have also made necessary updates in the Discussion (page 8) which highlights both the value of the Danish welfare system in mitigating economic harm but also its limitation in fully protecting the population:

In the years following diagnosis, individuals with breast cancer are more likely to be receiving income from sick leave, disability pension, and social assistance than their peers, indicating that these state functions are being employed by those affected by breast cancer which would consequently buffer the effects of lost wages. Nonetheless, individuals on fixed income before diagnosis including those in school, those on leave and the unemployed population still experience losses compared to individuals with the same economic situation but without a breast cancer diagnosis, indicating the potential of long-term accumulation of income effects which are not fully mitigated by the Danish welfare system.

3. The discussion of vulnerable groups is interesting but underdeveloped. The authors mention students when they speak of “vulnerable groups” on p8 para starting line 276. Breast cancer is not homogeneous, so I imagine that the versions that affect young women to have different consequences medically. Thus this discussion is important, are there other vulnerable groups worth mentioning? If not, then the general term “vulnerable groups” should be replaced with something more specific, which ties into my next point.

We appreciate your call for expansion upon the topic of vulnerable groups in our analysis, especially since estimation of heterogeneous income losses is one of the most important contributions of the paper in our opinion. We have added wording to the Discussion (page 8) which reflects the disproportionate loss for those with high healthcare utilization and CCI index, and the relative magnitude of loss for lower income quartiles:

Additionally, individuals with higher hospital utilization before diagnosis and those with a higher CCI have more profound and persistent income effects than those in better health at baseline. These increasing losses reflect the financial challenges which can be brought on by multimorbidity and a gap in protection for those who are already in poor health. Furthermore, although income losses are greatest in absolute terms for the highest income quartiles, relative losses are in the range of 1-2% for all income quartiles except the lowest one, where substitution of wages by welfare benefits is most significant. The implication of this relative loss is that while high-income individuals may have the greatest income available to lose, smaller losses in the rest of the income distribution will have an equally meaningful impact on personal budget in the short term, and wealth accumulation in the long-term.

4. General terms when specific are better abound. The authors use to term “health” to describe adjusting for a comorbidities index, and that will be poorly received by people who like to point to more general definitions of health than “diagnosed with one or more of a list of diseases.” Further, the issues the authors mention, like depression from a diagnosis, are not included in the Charlson index so there are consequential health conditions that are unadjusted for by this “adjustment for health”. I would recommend using the specific term. Other examples are health care use not being the same as a hospital admission, the term “baseline” being unclear (is it entry into the cohort?), and what disposable income is in Denmark vis-à-vis mandatory deductions that might not exist in other countries.

Thank you for your call for greater specificity and precision in our language. We have updated all of the mentioned vague descriptors with more precise wording throughout the manuscript. Some examples of our updates are included here:

- (Page 3) We employ a matching approach to adjust for confounders beyond the basics (such as age, baseline income, education), and expand to include employment status, household size, Charlson Comorbidity Index (CCI), total hospital utilization, and relationship status, allowing us to estimate average causal effects on the exposed (ACE) for personal and household disposable income over a 10-year period post-diagnosis.
- (Page 5) Within the non-retired population, additional stratifications highlight differences in outcomes across subgroups. Figure 2 shows loss of personal income by age,

educational attainment, income quintile, hospital utilization and CCI at cohort entry for the nonretired population.

- (Page 9) Including income quantile, household size, and relationship status as matching variables at cohort entry in addition to age and education is crucial to mitigate confounding produced by socio-demographic characteristics. Excluding hospital utilization and CCI produces personal income losses on the non-retired population which are 5% lower than our final results.

Minor comments.

5. The authors set income = 0 for people who die at one point, I cannot tell if that's a temporary thing or they only did so at one point to emphasize the importance of death. I think that if they use an expected age of death (75, or similar) then this method could identify losses from breast cancer among working people. If it's a temporary thing then simply adding "for this analysis only" or similar wording would clear that up.

Thank you for this note. We have reworked our language about the analysis including years after death as zero income.

- (Page 4) Our secondary analysis accounting for mortality by setting years after death but before age 65 or end of follow-up to zero income produces larger economic gaps between the exposed and unexposed.
- (Page 12) Because early death is still a loss of potential future income, we perform a secondary analysis where we consider years after death as years with no personal income.

We've also added a new section in the supplement which describes our analytical approach to estimate the hypothetical income impacts of mortality compared to the income loss of survivors.

Section 2: Mortality and income loss: methods and interpretation of secondary analysis

Differential mortality rates over time between the control and exposed populations, and within exposed subpopulations, have implications for the interpretation of cumulative income loss. Survival in each year is required to observe impacts on income in that year, but this skews the population distribution overtime to the population who is able to survive.

In our primary analysis, we condition on survival so that income loss measured in each year is not adjusted for the mortality rate in each match group. As a secondary analysis, we aim to illustrate the effect of premature mortality alongside income losses in the survivors, to demonstrate the theoretical value lost by early death at the population level. To do so, we impute years after death within the follow-up period as years of 0 income, as long as the death occurs before age 65 which is the retirement age in Denmark. Deaths of any cause for the control or exposed population are included. This was only possible for personal income loss, as household income losses would require identification of all income earners who constitute the household and could not be applied to equivalized

household income since this measure is adjusted for the size of the household in each year.

Unlike our primary results, this secondary analysis does not conform to the perspective of individual experienced income loss, and there are other methods which could be more consistent with this perspective, but which require other assumptions. The income losses reported in this secondary analysis are intended as an exploratory analysis to illuminate the outsized impact of mortality, but a more rigorous comparison of mortality adjustment methods is warranted to draw conclusions about the non-mortality counterfactual.

6. It is not obvious that the household income is adjusted for household size until the final section, which makes the household income < personal income facts confusing. Including “adjusted for household size” at the first mention of household income solves this problem.

We now clarify the interpretation of household income in the first paragraph of the results and the last paragraph of the introduction (Page 3).

- Our study addresses these limitations by examining the long-term impacts of breast cancer on personal and household equivalized disposable income across the Danish population from 2000 to 2018, including an analysis of heterogeneity across sub-populations.
- The median personal income at the index was €42,128 (€27,131 - €59,460), and the median household equivalized income (which controls for household size) was €33,765 (€25,270 - €44,296).

7. Figures need more descriptive titles. E.g., Figure 2 is completely reliant on the reader knowing what the CACE is in the text, which is inefficient. I recommend trying to make the figures stand alone with titles and enough description in the subtitle.

We have updated the figure captions for each figure and improved the legend and axis labels in some of the figures where necessary. An example of these modifications for the figure caption is given here (page 18):

Figure 1: Average causal effect on the exposed in personal and household equivalized disposable income loss for the non-retired population. Relative income loss is shown as a proportion of control income over each year and absolute losses are adjusted to 2023 euros. Personal and household effects are censored at mortality, so each year of income loss reflects the effects only on the surviving population at that year.