London School of Hygiene & Tropical Medicine
(University of London)

Keppel Street, London WC1E 7HT
Direct Dial: 020-7927-2680
Fax: 020-7580-6897 Email*: sinead.langan@lshtm.ac.uk*

Faculty of Epidemiology & Population Health
Department of Non-communicable Disease Epidemiology

6 March 2018

Dear Dr Godlee,

Many thanks for considering our manuscript, and for the helpful comments made by your reviewers. We have reviewed the manuscript thoroughly and addressed the issues raised. We believe that this has significantly improved the paper. Please find below our detailed responses to all the individual comments.

The BMJ's manuscript committee meeting

0.1 First, please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below.

Please see responses to reviewer comments below.

0.2 Subjects were matched for age within a 15-year time frame. This seems excessively wide. Why was this age gap chosen?

The statistical models use age as the underlying timescale, therefore we have adjusted closely for the confounding effects of age in all analyses. We have made this clearer in the revised manuscript.

We agree that the 15-year matching interval is wider than in most matched cohort studies. Finer age-matching led to a large loss of atopic eczema patients due to a lack of any potential matches. Specifically, matching within 1, 5 and 10 years led to a loss of 8.2%, 7.1% and 6.6% of atopic eczema patients respectively. Although the participants are therefore not matched closely on age.

We have acknowledged this in the Discussion of the manuscript, as follows:

"Patients were matched within 15-year age intervals as finer age-matching led to a significant loss of atopic eczema patients, who did not have any eligible matches. However, age was accounted for as the underlying timescale in all analyses, closely adjusting for the confounding effects of age."

0.3 The statistical analysis of the data excluded a rather large percentage (20%) of subjects with missing values. The potential bias is acknowledged in the Discussion. Why was this rather simplistic method used to deal with missing values? Why, for example, was a multiple imputation procedure not used?

Many thanks for raising this issue. Missingness is largely driven by BMI, for which 19.2% of our cohort have missing data. Other covariates with missing data were smoking (5.3% missing) and Index of Multiple Deprivation (0.1% missing). We did consider using multiple imputation, however this method assumes data are missing at random (MAR); that is, there may be systematic differences between the observed and unobserved data, but these differences can be explained by other variables[1]. This assumption is unlikely to hold true for BMI and smoking variables, as the BMI and smoking values themselves may determine missingness[2]; for example obese patients may be more likely to have their BMI recorded. We therefore chose not to carry out multiple imputation because we believed that the data were likely to be missing not at random (MNAR).

Instead, we conducted a complete case analysis. Complete case analysis is valid under the missing not at random assumption if missingness (or being a complete case) is not related to the outcome, conditional on the covariates.

Since BMI and smoking only enter the analysis in the mediation model, it is possible to fit the unadjusted and adjusted models in the 99.9% of the cohort who have complete data on the variables in these models (i.e. even in the presence of missing BMI and/or smoking data). These results are presented in Table R1, below. Comparing these results with the equivalent models in Table 3 of the manuscript (which additionally require complete BMI and smoking data) it can be seen that the estimated HRs are very similar, suggesting that conducting a complete case analysis in the presence of missing BMI and smoking data does not result in appreciable bias.

We have made this clearer in the Methods (Statistical analysis) section, as follows:

"Patients with missing BMI, smoking or IMD data were excluded. These data are likely to be missing not at random (as missingness is likely to depend on the actual values). Multiple imputation would therefore not be appropriate. Complete case analysis is valid where the missingness is independent of each of the cardiovascular outcomes, conditional on the model covariates.[1]"

We have clarified the nature of the missingness in the Results section, as follows:

"…(19.2% missing BMI, 5.3% missing smoking and 0.1% missing IMD)…"

In the Discussion, we had previously written:

"Exclusion of patients without complete data on the analysis variables reduced the sample by nearly 20% which may also have introduced some bias."

This has been revised to:

"Exclusion of patients without complete data on BMI, smoking and IMD reduced the sample by nearly 20%. However, we believe that there are no factors leading to missingness which would independently affect the CVD outcomes in question, therefore using a complete case analysis was valid."

0.4 The estimated absolute cardiovascular rates should be presented. This is important to enable a careful consideration of the impact of the estimated risk estimates, given that many of the relative risks are rather low.

We agree with the reviewer that consideration of absolute incidence rates of the CVD outcomes in each exposure group can be useful in assessing the impact of the atopic eczema. Unfortunately, calculation of the incidence rates of the CVD outcomes is not straightforward in this dataset. Matching means that while we can reliably estimate the incidence rate of the CVD outcomes in the atopic eczema exposed patients (since our sample includes all such patients from the population), we cannot reliably estimate the incidence rate of the CVD outcomes among people without atopic eczema (since such individuals in our sample are not necessarily representative of the population). Therefore, we have estimated the incidence rates of the CVD outcomes in the atopic eczema exposed patients using the data in our sample and then estimated the incidence rates of the CVD outcomes among people without

atopic eczema by multiplying these by our corresponding estimated hazard ratio (after having first inverted it so that it compares unexposed to exposed).

These additional results are presented in Table R2, below. They confirm the increased absolute risk in CVD outcomes among people with atopic eczema.

We have included these results within the manuscript as follows.

Methods (Statistical analysis section):

"Incidence rates of each cardiovascular outcome in the atopic eczema exposed patients were estimated using the data in our sample. Incidence rates of each cardiovascular outcome among people without atopic eczema (which cannot be reliably estimated from the sample due to the matching) were then estimated by multiplying the incidence rate in the atopic eczema exposed patients by our corresponding estimated HR (after having first inverted it so that it compares unexposed to exposed)."

Results section (interpreting the attributable risks we have provided in response to point 3.7 rather than the exposure group-specific incidence rates themselves):

"Estimated attributable risks confirm the increased incidence rates of cardiovascular outcomes among atopic eczema patients (Table 4). Attributable risks were greatest for heart failure (40 per 100,000; 99% CI 22, 57) and atrial fibrillation (37; 15, 55)."

Discussion:

"Estimated attributable risks confirm the increased incidence rates of cardiovascular outcomes among atopic eczema patients"

0.5 The median follow-up period was only just over 5 years. Is this a long enough follow-up period, particularly as results for cardiovascular mortality changed in the sensitivity analysis using subjects with at least 5-years follow-up?

Although the median follow-up period was 5.08 years, a quarter of individuals were followed up for at least 9.75 years (Table 2 in the manuscript) and the maximum duration of follow-up was 17.25 years. The follow-up time is a feature of the dataset we have used: we used all the follow-up data available to us. We had sufficient follow-up periods across different ages, times since diagnosis and calendar times to obtain adequately powered estimates of the association between atopic eczema and each outcome of interest. With the current 5 year median follow-up, we were able to detect important effects of atopic eczema on CVD risk. Therefore, we do not believe the median 5 year follow-up was a substantial limitation.

The sensitivity analysis restricting to those with at least 5 years of follow-up was to explore any potential bias caused by atopic eczema patients with short follow-up periods being more likely to have either none or all of their follow-up with active atopic eczema. This restriction therefore only affected one of our analyses (the atopic eczema activity analyses in Figure 3). As pointed out by the reviewer, we observed an increased risk of cardiovascular death among patients with "never active" atopic eczema. However, this finding may be explained by poor capture of activity data – as Reviewer 1 points out, some patients may have active atopic eczema but not adhere to treatment, thus being misclassified as never active.

We have added this last point to the Discussion of the manuscript to improve clarity on this issue, as follows:

"This finding may be explained by poor capture of activity data: some patients may have active atopic eczema but not adhere to treatment, thus being misclassified as never active."

0.6 There are confounders that are not measured. Diet, exercise and medication use are all pretty important for cardiovascular outcomes.

Yes, we acknowledge that diet, exercise and medication use could be important confounders.

Although the Clinical Practice Research Datalink (CPRD) is a rich data source, it does not capture all lifestyle variables and, unfortunately data on diet and exercise are simply not available. However, we have controlled for BMI, smoking and severe alcohol use, which will capture diet and exercise to some extent and, in further analyses, we have controlled for hyperlipidaemia and hypertension, which also relate to these exposures. Indeed, the ability to control for these life style factors was better than in most previous studies.

Regarding medications which may be associated with both atopic eczema and cardiovascular disease, it is true that anti-hypertensives may be potential confounders. However, as we have adjusted for hyperlipidaemia and hypertension, this will already have captured to some degree use of medications. We have additionally adjusted for use of high-dose oral corticosteroids, as discussed in response to point 0.12 below.

0.7 You "matched" for age within 15 years but in fact there are more elderly people in the eczema cohort (eg 12.4% vs 9.0% over 70) and fewer young people (18.1% vs 20.6% 30to39) except in the 18-19 age group who probably don't contribute many events anyway.

Please see our response to point 0.2, which raised a more general concern around the extent of age matching.

All our modelling used age as the underlying timescale (rather than just relying on the matching to control for age), implicitly adjusting closely for the confounding effects of age. Thus, any imbalances within matched sets would not cause bias in our estimated associations.

We do agree that within this matched cohort it is possible to get overall age imbalances. This is due to post-matching exclusions, which may result in imbalances even if the exclusions are non-differential between the exposure groups. If a particular type of individual were more likely to have missing data (and hence be excluded from the study) then even if this affected the two exposure groups to the same extent it would result in a greater exclusion of unexposed individuals from the analysis (and hence a perceived imbalance in the factor being considered) because within each matched set it is possible for up to 4 unexposed individuals to be excluded and it still remain valid (i.e. containing both exposed and unexposed individuals), whereas as soon as the sole exposed individual is excluded then the matched set becomes invalid (and all 5 unexposed individuals are excluded). It is important to be clear that only valid matched sets (with broad matching for age) were analysed and that individuals in the analysis sample were representative of those from the eligible cohort.

0.8 The completeness and accuracy of codes in CPRD is not great for some things such as alcohol and BMI. You should report this and discuss for all covariates.

Data completeness for smoking and BMI are discussed in point 0.3 above. We have discussed the accuracy of BMI, smoking and alcohol use below.

BMI

We followed a published algorithm with regard to data cleaning to obtain BMI records.[2] Research on BMI data in CPRD showed that when BMI at specific time points was assigned based on the most recent record; calendar–year-specific mean BMI statistics underestimated equivalent Health Survey for England statistics by 0.75–1.1 kg/m^2. Therefore, although BMI appears to be underestimated in CPRD, and thus the BMI category possibly misclassified, the absolute error appears relatively minimal.

Smoking
A CPRD study assessing the validity of smoking records in CPRD found that the prevalence estimates for current smoking and non-smoking in CPRD were similar to those from nationally representative surveys, but former smoking may be under-recorded.[3] However, we do not believe that missingness would be related to either the outcomes or our exposure, and therefore this would not result in bias.

Alcohol use
We agree that the accuracy of alcohol use data may be poor in CPRD. We therefore chose to define severe alcohol use at the first of: i) "high" alcohol use code or ii) treatment for severe alcohol use recorded in CPRD, as this is likely to be more reliable.

0.9 The treatment of missing data is poor. To simple exclude 20% in this day and age and not make any attempt at including via imputation or other methods is verging on unacceptable.

We agree that ignoring this issue would be unacceptable; however, although we did not describe this clearly in our manuscript, we have given this issue a great deal of consideration. Please see our response to point 0.3 where we have addressed concerns about missing BMI and smoking data.

0.10 The fully adjusted associations are pretty small and many of them don't reach statistical significance (even though you have used 99% CIs).

We agree that some of the fully adjusted associations are small. However, in terms of absolute effects even these smaller associations correspond to important differences in the absolute rates in the two exposure groups (see Table R2, below, and our responses to points 0.4 and 3.7). Clearly, using 99% confidence intervals rather than 95% confidence intervals may have limited our ability to detect statistically significant findings. However, we believe this approach was warranted as we were undertaking multiple analyses.

We would like to highlight that many of the fully adjusted associations do suggest evidence of an important association; there is very good evidence of an association between eczema and five out of the seven outcomes assessed (see Table 3). Given the very high prevalence of atopic eczema in the general population, we believe these are critical findings.

0.11 Table numbering has gone astray.

Thank you for informing us of this, and apologies for the error. We have now checked and revised our references to the tables throughout the manuscript.

0.12 Another editor said this was an interesting research question, and while not novel, surely the largest study until now. He saw no info on medication, and steroid use might be an important variable.

We are pleased to hear that this research question was considered interesting. Regarding medication use, as mentioned in response to point 0.6, we believe adjusting for hypertension and hyperlipidaemia will have accounted for any effects of antihypertensive drugs.

Regarding corticosteroid use, we agree with the reviewer that oral corticosteroid use may be an important confounder and we thank them for raising this issue. This issue will partly have been addressed by adjusting for time-updated asthma, but we agree that it warrants further exploration. We have therefore conducted an additional analysis, to be added as an additional sensitivity analysis in the manuscript, which included time-

updated exposure to high-dose oral corticosteroids as an additional covariate in the mediation model only. The results are presented in a Table R3, below. n = 47,062 (1.8%) of the 2,636,006 successfully matched individuals eligible for cohort entry had one or more periods of high dose oral corticosteroid treatment. Oral corticosteroid use was found to be positively associated with all the CVD outcomes with the exception of coronary revascularisation (conditional on the other variables in the model; results not shown). However, comparison of the results in Table R3 with those in Table 2 of the manuscript show that this additional adjustment had no meaningful impact on the findings of the study.

The manuscript has been updated as follows.

Methods (Covariates) section:

"Patients were defined as exposed to high-dose (≥20mg/day) oral corticosteroids for the duration of their prescription and 3 months following the end of the prescription."

Methods (Statistical analyses) section (Table 1):

"The primary analysis (mediation model only) was repeated with additional adjustment for time-updated high dose corticosteroid use"

"To examine whether the omission of this covariate in the primary analysis may have introduced bias"

Results section:

"Results adjusted for time-updated high dose corticosteroid use were consistent with the main analysis (Table S15), similarly suggesting limited bias from omission of this covariate."

Supplementary Material (Methods S1):

"Daily dose of oral corticosteroid was calculated as follow: numeric daily dose (NDD) x dose per tablet. Where NDD was missing, a "hot-deck" style imputation method was adopted, which replaced missing data with comparable data from the same set. An extra binary variable for quantity of tablets per prescription was created, categorising quantity about the median number (42) into low and high. If a patient had any other record with the same quantity and dose per tablet, the median NDD among those records was used where NDD was missing. If a patient had no recorded NDD but had any other record of the same dose per tablet and quantity as a binary variable, the median NDD among those records was used. If a patient did not have a recorded NDD or quantity, but had records for the same dose per tablet, then the median NDD among those records was used. If there was no record of NDD, dose per tablet or quantity, but there were other patients in the dataset in the same 5-year age band, of the same gender, with the same dose per tablet and quantity, the median NDD for those records was used. Finally, if none of the above were possible, patients in the dataset in the same 5-year age band, of the same gender, with the same dose per tablet and quantity as a binary variable, the median NDD among these records was used."

Supplementary Material (Tables):

Table R3, below, has been added as Table S15.

Regarding other atopic eczema drugs, ciclosporin may increase the risk of cardiovascular disease, as it is known to be associated with hypertension, hyperlipidaemia and renal impairment.[4] Regarding methotrexate, some of the existing literature suggests that this drug might have cardioprotective effects in psoriasis patients.[5] We therefore conducted separate sensitivity analyses removing all patients ever exposed to ciclosporin (n = 3196 (0.1%) of the 2,636,006 successfully matched individuals eligible for cohort entry) and all patients ever exposed to methotrexate (n = 20,506 (0.8%) of the 2,636,006 successfully matched individuals eligible for cohort entry). The

results are presented in a Tables R4 and R5, below. Comparison of these results with those in Table 2 of the manuscript shows that these exclusions had no meaningful impact on the findings of the study. These results have not been included in the manuscript or supplementary material, but we would be happy to include them if the reviewers deem it necessary

0.13 Another editor said that these statements don't seem supported by data and should be revised and toned down: "The results support targeted screening and focus on primary prevention strategies to reduce cardiovascular disease among patients with severe or predominantly active atopic eczema."

We have revised this as follows:

"If these results are robustly replicated, it would support targeted screening and focus on primary prevention strategies to reduce cardiovascular disease among patients with severe or predominantly active atopic eczema."

0.14 He also said that you need to emphasise that the effect size is small, and the effect size should be mentioned in the abstract.

As mentioned in our response to point 0.10, we agree that some of the fully adjusted associations are small. However, in terms of absolute effects even these smaller associations correspond to important differences in the absolute rates in the two exposure groups (see Table R2, below, and our response to point 0.4).

0.15 Another editor was supportive given that the study addresses a relevant research question.

We thank the reviewer for this kind comment.

0.16 Another editor was in favour but said the title should not announce the findings.

We have revised the title as follows:

"Severe and predominantly active atopic eczema in adulthood and long-term cardiovascular disease risk: a UK population-based cohort study, 1998–2015"

Reviewer 1

1.1 I understand that a requirement for all manuscripts submitted to The BMJ is a statement regarding patient involvement - I could not see this, so please add. As this was a non-interventional study, the authors did not have the opportunity to involve the patients whose data they used in the study. Nevertheless, it would be useful to know if they discussed the study design and the chosen outcomes with a group of patients with atopic eczema (and if not, why not).

The design, conduct and initial results of this study have been overseen by Dr Sinéad Langan's Wellcome Senior Clinical Fellowship steering committee, which includes lay representation. We have now added a statement to this affect, and acknowledged a patient who contributed greatly to these discussions.

We have added the following statement in the Methods section under a new subheading of "Patient involvement":

“The research questions, design, conduct and initial results and interpretation of the findings of this study have been overseen by Dr Sinéad Langan’s Wellcome Senior Clinical Fellowship steering committee, which includes lay representation.”

We have added the following in the Acknowledgements:

“We thank Amanda Roberts for her important input as the lay member of Dr Sinéad Langan’s Wellcome Senior Clinical Fellowship steering committee.”

1.2 In the Discussion, the authors state that "it is not possible to disentangle the effects of therapy and severity, but this was not an objective of the present paper." However, this is an extremely important distinction for patients. I feel strongly that The BMJ has engaged patient reviewers like myself in order to encourage authors not to make such statements without really considering the impact on patients. Many patients with atopic eczema remain distrustful of long-term topical steroid and/or immunosuppressant use. Would it have been possible to disentangle the effects of therapy from the effects of disease when assessing CV risk with an altered study design? If so, please discuss why it was not done. If it was not possible, some discussion of this impact that this has on the usefulness of the study, particularly the practical impact for PCPs and their patients, would be good.

We thank the reviewer for this insightful comment. We entirely agree that more complete explanations for our observed effects, i.e. the extent to which the increased risks are attributable to underlying disease or to treatments, are crucial research questions. However, our study was undertaken to establish and quantify the risks of cardiovascular disease among people with atopic eczema, and we completed this major aim. Future work to disentangle the effects of drugs and disease severity would require a different study design using different data sources.

We have revised the Discussion section of the manuscript as follows:

"It is not possible to disentangle the effects of therapy and severity, due to the nature of routinely collected data sources and observational settings whereby those with more severe disease are given specific therapies, and where those therapies might be continued for long periods of time.”

We hope to address this issue in the near future with international colleagues working as part of the TREatment of ATopic eczema (TREAT) Registry Taskforce as we understand that this research question is of paramount importance to patients and physicians.

1.3 Another point I wanted to raise was a potential limitation that has not been covered in the Discussion. In my experience, patients who suffer long-term from eczema often do not adhere to therapy, or have episodes of adherence alternating with non-adherence. This means that many patients in the study who actually have active disease may not be classified as such, because they may not have had a CPRD/HES record within the last 12 months. This potentially could confound the results.

We agree with the reviewer that adults with long-term atopic eczema may become disenfranchised and may not attend their GP, despite the presence of active disease, as they may perceive that current care is suboptimal. We believe that this observation is very valid. This sort of exposure misclassification would bias the results towards the null, underestimating the magnitude of the associations. In particular, this issue could at least partially explain the observed increased risk of cardiovascular death among patients with “never active” atopic eczema.

We have revised the Discussion section of the manuscript as follows:

“This finding may be explained by poor capture of activity data: some patients may have active atopic eczema but not adhere to treatment, thus being misclassified as never active.”

1.4 The median follow up seems quite low to me, given the time period of the study and the fact that eczema is a long-term disease. Could the authors comment?

Please see our response to point 0.5.

1.5 In the supplementary tables, I think it is clearer to describe the groups as With atopic eczema/Without atopic eczema as in Table 2, rather than as Exposed/Unexposed. Either way, I think the tables should be consistent in this regard.

We agree and have now revised these descriptions.

Reviewer 2

2.1 Page 10, second paragraph: I did not quite understand the sentence: Any patients with an atopic eczema diagnosis...., could you please clarify why they were put into the group of unexposed individuals

Prior to an individual's atopic eczema diagnosis, that individual does not have atopic eczema. They are therefore "unexposed" to atopic eczema. As such, they are theoretically eligible to be in the pool of "unexposed" patients up until their atopic eczema diagnosis. This is standard practice in database studies and an important technique that ensures bias is not introduced.

This statement has been revised, as follows:

"Any patients with an atopic eczema diagnosis were included in the pool of eligible unexposed patients up until the date of their atopic eczema diagnosis; prior to their atopic eczema diagnosis, these patients were not considered to have atopic eczema and were therefore eligible to contribute to unexposed person time. Patients with an atopic eczema diagnosis who did not go on to meet the full definition of atopic eczema (at least one diagnosis code and two treatment codes on separate dates) were also in the pool of eligible unexposed patients up until the date of their atopic eczema diagnosis code. Removing these patients from the pool of eligible unexposed at the point of diagnosis rather than allowing them to remain until they met the full validated definition of atopic eczema ensured greater certainty that the pool of unexposed patients did not have atopic eczema."

In addition, we have undertaken analyses varying our definition of atopic eczema and our definition of the unexposed pool and observed similar findings, suggesting that this not a major issue.

2.2 Page 11, second paragraph the classification of patients into moderate or severe disease. You can have a severe disease and still use only topical treatment. If the patients would have been seen and classified by a dermatologist it would have been more reliable, but I understand the difficulty when the population base is so large. This is still a limitation of the study which could me mentioned in the discussion.

We agree with this statement and have now added a statement about misclassification of severity to the discussion. As only 3% of individuals with atopic eczema in the UK ever see a dermatologist, we believe that requiring assessment of severity by a dermatologist would answer a very different question.

The Discussion section of the manuscript has been revised as follows:

"Similarly to activity of atopic eczema, misclassification of disease severity is possible, for example if patients with severe disease used only topical treatment. However, such misclassification would bias our result towards the null, underestimating the magnitude of the associations."

Reviewer 3

3.1 I suggest a statistical reviewer to check that Cox regression stratified for matched set is a good method to take matching into account, and that the simple methods used for missing data are adequate.

We thank the reviewer for this important comment. We are very fortunate in our team to have statistical expertise. Richard Silverwood is Assistant Professor of Medical Statistics. In addition, we have consulted with Krishnan Bhaskaran, Associate Professor of Medical Statistics, to ensure that our analyses of a matched cohort are appropriate.

We have now added this detail to the Acknowledgements as follows:

"[We thank] Krishnan Bhaskaran for his important input in advising on the analysis of our matched cohort study."

3.2 I do not agree with the sentence (page 10, lines 30-4)" Any patients with and atopic eczema diagnosis….this ensured greater certainty that the pool of unexposed patients did not have atopic eczema. " I think that including atopic eczema patients before their diagnosis in the unexposed pool does not ensure that this pool does not have eczema patients. It might lead to some differential misclassification (patients having eczema among the unexposed group) that would lead to a result biased towards the null effect. This does not compromise the study, but I think that authors should consider modifying this sentence.

Another reviewer raised a similar comment (see point 2.1). There were some patients with an atopic eczema diagnostic code, but no eczema treatments, therefore they did not meet our atopic eczema definition. However, these patients could only contribute person time up until their atopic eczema diagnosis, to ensure the pooled of unexposed patients did not have any evidence of atopic eczema.

We have revised the statement as follows:

"Any patients with an atopic eczema diagnosis were included in the pool of eligible unexposed patients up until the date of their atopic eczema diagnosis; prior to their atopic eczema diagnosis, these patients were not considered to have atopic eczema and were therefore eligible to contribute to unexposed person time. Patients with an atopic eczema diagnosis who did not go on to meet the full definition of atopic eczema (at least one diagnosis code and two treatment codes on separate dates) were also in the pool of eligible unexposed patients up until the date of their atopic eczema diagnosis code. Removing these patients from the pool of eligible unexposed at the point of diagnosis rather than allowing them to remain until they met the full validated definition of atopic eczema ensured greater certainty that the pool of unexposed patients did not have atopic eczema."

As discussed above, we have also undertaken sensitivity analyses varying our definition of atopic eczema.

3.3 Both the diagnosis definition and the measures of severity are probably associated with more frequent interaction with health providers, that might also lead to a higher probability of being diagnosed with the outcome. This might be more relevant for "softer" outcomes (atrial fibrillation?) and should be discussed.

We do not believe that ascertainment bias is driving our results, as most of the study outcomes would require medical intervention and therefore presentation in primary or secondary care. (Indeed, this is a major strength of

the study.) An exception to this may be atrial fibrillation, which may be diagnosed more often in frequent attenders; ascertainment bias may therefore explain part of the observed association with eczema and atrial fibrillation.

We have now added a statement to the Discussion section of the manuscript as follows:

"The study outcomes would require medical intervention and therefore presentation in primary or secondary care, meaning that ascertainment bias is unlikely to be a problem. One exception may be asymptomatic atrial fibrillation, though this is only a minor proportion of the total patients with atrial fibrillation so any ascertainment bias is likely to be limited."

3.4 I suggest adding in table 2 a category of patients without classic risk factors for CVD (DM, hypertension, hyperlipidemia, smoking, overweight-obese). This are the patients that might specially benefit from the results of this paper.

Although this would be an interesting area of research, the present study does not allow us to examine whether or not exposure to atopic eczema among individuals without classic risk factors for CVD is a particular risk factor for CVD outcomes. To properly address this question would require a different analysis approach (in particular, either using a specific categorisation for the absence of classic risk factors for CVD or allowing for interactions between each classic risk factor for CVD and eczema). This is beyond the scope of the present study.

3.5 Figure 2 and 3 seem to show a dose-effect response between eczema severity and outcome, and less clear for activity. This might be included in the models (assuming a linear relationship) and be further discussed, as it strengthens the argument of the paper.

While there does appear to be a dose-response relationship between atopic eczema severity and several of the outcomes, this does not necessarily appear to be linear for many – and certainly not for any of the relationships with atopic eczema activity. Therefore, we believe that imposing a linear trend here, even if only for the purpose of a statistical test, would not be a useful addition to the manuscript, which currently allows the nature of each of these relationships to speak for itself through graphical display (Figs. 2 and 3).

3.6 I suggest that authors clearly indicate how many of the atopic dermatitis patients do not have classic risk factors for CVD. This would be the patients that might benefit for additional counselling (exercise and diet prescription, increased screening for risk factors, etc.)

As in point 3.4 above, this would be an interesting area of research, but one that is beyond the scope of the present study. The present study does not allow us to examine whether or not atopic eczema patients without classic risk factors for CVD would benefit from specific interventions, so the added value in singling out these individuals for specific attention within the manuscript is not clear.

3.7 I also suggest that authors should discuss the expected impact of their findings. Assuming a cause-effect relationship, what are the attributable risk and population attributable risks of atopic dermatitis?

We thank the reviewer for the suggestion of further discussion of the impact of our findings. We refer the reviewer to our response to point 0.4, above, for details of our approach to estimating exposure group-specific incidence rates, which are presented in Table R2, below. We then calculated the attributable risk as the difference between these exposure group-specific incidence rates. The population attributable risk was estimated as $P(HR - 1) / (1 + P(HR - 1))$ where P, the prevalence of atopic eczema, is assumed to be 10%[6] and HR is the estimated hazard ratio comparing atopic eczema exposed patients to people without atopic eczema.

These additional results are also presented in Table R2, below. They confirm the increased absolute risk in CVD outcomes among people with atopic eczema and suggest that the population attributable risks may be as high as 2.4% (unstable angina).

We have included these results within the manuscript as follows.

Methods (Statistical analysis section):

"Attributable risks were calculated as the difference between these exposure group-specific incidence rates. The population attributable risk of each cardiovascular outcomes was estimated using the estimated HR and assuming the prevalence of atopic eczema to be 10%.[6]"

Results section:

"Estimated attributable risks confirm the increased incidence rates of cardiovascular outcomes among atopic eczema patients (Table 4). Attributable risks were greatest for heart failure (40 per 100,000; 99% CI 22, 57) and atrial fibrillation (37; 15, 55). The greatest population attributable risks were estimated for unstable angina (2.4%; 99% CI 1.1, 3.9)" and heart failure (1.9; 1.0, 2.9)."

Discussion:

"Estimated attributable risks confirm the increased incidence rates of cardiovascular outcomes among atopic eczema patients, with population attributable risks of 2% or more for some outcomes."

Reviewer 4

4.1 Using use of immunosuppressants to (partially) define severe atopic eczema may be slightly problematic, as these drugs could represent confounders. This is acknowledged by the authors on page 18. Were they controlled for in the analyses?

We did not control for immunosuppressant use in the analysis. As well as being used to (partially) define eczema and severe eczema, these drugs may be on the causal pathway between eczema and our cardiovascular outcomes. It would therefore have been inappropriate for us to include them in our adjusted model. We could instead have included them in our mediation model, but mediation by immunosuppressant use was not of specific interest in the present study.

As per our response to point 0.12, we have conducted additional analyses in which we have considered high-dose oral corticosteroid use as a time-varying confounder in the mediation model and (separately) removed all patients ever exposed to ciclosporin or methotrexate. Comparison of these results (Tables R3, R4 and R5, below) with those in Table 2 of the manuscript shows that this additional adjustment and these exclusions had no meaningful impact on the findings of the study.

4.2 It may be helpful to tease out the use of immunosuppressants. This could be done in several ways, by either controlling for immunosuppressant use as a covariate, analyzing those with severe eczema excluding those on immunosuppressants, or including patients on immunosuppressants who do not have eczema (i.e. those using these drugs for other indications). Another question would be duration of therapy on immunosuppressants, as extensive data have suggested a dose-response relationship for important clinical outcomes such as cancer risk.

We agree that the exact role of immunosuppressants in cardiovascular disease risk is very interesting; however, it is beyond the scope of this study. It was not specified in our aims prior to the analysis and would be better

investigated in a different study design. We would therefore be reluctant to explore the effects of timing and duration of immunosuppressant use among atopic eczema and other patient groups in this study. However, we agree it is an interesting research question and one we shall consider in future.

4.3 Table 1 could possibly be changed to a supplemental table, unless required per BMJ style.

If the editors would prefer this table to be in the Supplementary Material, we are happy to oblige. However, many of the sensitivity analyses are very important for the credibility of our findings; therefore, we would prefer this table stayed in the main paper if the editors agree.

4.4 Line 26: Why was the age matching performed so broadly (+/-15 years)?

Please see our response to point 0.2.

We hope that these responses address all the reviewers' comments. Please let us know if further clarifications are necessary. We look forward to hearing from you.

Yours Sincerely,

Richard Silverwood on behalf of co-authors

References

1. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat Med* 2010;29(28):2920-31. doi: 10.1002/sim.3944
2. Bhaskaran K, Forbes HJ, Douglas I, et al. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). *BMJ Open* 2013;3(9) doi: 10.1136/bmjopen-2013-003389
3. Booth HP, Prevost AT, Gulliford MC. Validity of smoking prevalence estimates from primary care electronic health records compared with national population survey data for England, 2007 to 2011. *Pharmacoepidemiology and Drug Safety* 2013;22(12):1357-61. doi: 10.1002/pds.3537
4. Williams H, Bigby M, Diepgen T, et al. Evidence-Based Dermatology: John Wiley & Sons 2009
5. Armstrong AW, Brezinski EA, Follansbee MR, et al. Effects of biologic agents and other disease-modifying antirheumatic drugs on cardiovascular outcomes in psoriasis and psoriatic arthritis: a systematic review. *Curr Pharm Des* 2014;20(4):500-12.
6. Silverberg JI. Public Health Burden and Epidemiology of Atopic Dermatitis. *Dermatol Clin* 2017;35(3):283-89. doi: 10.1016/j.det.2017.02.002 [published Online First: 2017/06/05]

Table R1. Association between atopic eczema and cardiovascular outcomes. Fitted to patients with complete data for all variables included in the models and from valid matched sets[1]. n = 2,629,754 (2,541,626 unique patients).

				Hazard ratio and 99% confidence interval[2]			
	n	P-Y at risk	Events	Unadjusted		Adjusted	
Primary outcomes							
Myocardial infarction							
Unexposed	2,163,466	12,221,132	20,135	1.00	(ref)	1.00	(ref)
Exposed	466,288	2,902,664	5,959	1.13	1.09, 1.18	1.10	1.02, 1.19
Unstable angina							
Unexposed	2,163,466	12,256,061	8,004	1.00	(ref)	1.00	(ref)
Exposed	466,288	2,911,792	2,578	1.26	1.18, 1.34	1.30	1.16, 1.46
Heart failure							
Unexposed	2,163,466	12,233,399	21,085	1.00	(ref)	1.00	(ref)
Exposed	466,288	2,903,117	7,179	1.22	1.18, 1.28	1.21	1.12, 1.31
Atrial fibrillation							
Unexposed	2,163,466	12,167,585	33,303	1.00	(ref)	1.00	(ref)
Exposed	466,288	2,884,386	10,570	1.15	1.11, 1.19	1.14	1.07, 1.21
Stroke							
Unexposed	2,163,466	12,218,326	26,463	1.00	(ref)	1.00	(ref)
Exposed	466,288	2,901,358	7,994	1.09	1.05, 1.13	1.09	1.02, 1.17
Cardiovascular death							
Unexposed	2,163,466	12,294,754	41,450	1.00	(ref)	1.00	(ref)
Exposed	466,288	2,924,334	12,852	1.04	1.01, 1.08	0.95	0.90, 1.01
Secondary outcome							
Coronary revascularisation							
Unexposed	2,163,466	12,219,310	17,883	1.00	(ref)	1.00	(ref)
Exposed	466,288	2,901,657	5,184	1.15	1.10, 1.21	1.18	1.09, 1.28

[1]Matched sets including one exposed patient and at least one unexposed patient.
[2]Estimated hazard ratios from Cox regression with current age as underlying timescale, stratified by matched set (matched on age at cohort entry, gender, date at cohort entry and practice).
Unadjusted: No adjustment.

Adjusted: Adjusted for current calendar period (1997-1999, 2000-2004, 2005-2009, 2010-2015), time since diagnosis (0-4, 5-9, 10-14, 15-19, 20+ years), IMD at cohort entry, and time-varying asthma.

Table R2. Absolute incidence rates, incidence rate differences (attributable risks) and population attributable risks of cardiovascular outcomes.

	Estimated incidence rate per 100,000 person-years in atopic eczema exposed patients	Hazard ratio and 99% confidence interval[1]		Inverse hazard ratio and 99% confidence interval[2]		Estimated incidence rate and 99% CI per 100,000 person-years in people without atopic eczema[3]		Estimated incidence rate difference (attributable risk) and 99% CI per 100,000 person-years[4]		Estimated population attributable risk (%) and 99% confidence interval[5]	
Primary outcomes											
Myocardial infarction	205	1.06	0.98, 1.15	0.94	0.87, 1.02	193	178, 209	12	-4, 27	0.6	-0.2, 1.5
Unstable angina	89	1.25	1.11, 1.41	0.80	0.71, 0.90	71	63, 80	18	9, 26	2.4	1.1, 3.9
Heart failure	248	1.19	1.10, 1.30	0.84	0.77, 0.91	208	191, 226	40	22, 57	1.9	1.0, 2.9
Atrial fibrillation	366	1.11	1.04, 1.18	0.90	0.85, 0.96	329	311, 351	37	15, 55	1.1	0.4, 1.8
Stroke	276	1.10	1.02, 1.19	0.91	0.84, 0.98	251	232, 270	25	6, 44	1.0	0.2, 1.9
Cardiovascular death	440	0.98	0.92, 1.06	1.02	0.94, 1.09	449	414, 480	-9	-40, 26	-0.2	-0.8, 0.6
Secondary outcome											
Coronary revascularisation	179	1.14	1.05, 1.24	0.88	0.81, 0.95	158	145, 170	21	9, 34	1.4	0.5, 2.3

[1]Comparing atopic eczema exposed patients to people without atopic eczema. Estimated hazard ratios from Cox regression with current age as underlying timescale, stratified by matched set (matched on age at cohort entry, gender, date at cohort entry and practice) comparing atopic eczema exposed to atopic eczema unexposed. Adjusted for current calendar period (1997-1999, 2000-2004, 2005-2009, 2010-2015), time since diagnosis (0-4, 5-9, 10-14, 15-19, 20+ years), IMD at cohort entry, and time-varying asthma.

[2]Comparing people without atopic eczema to atopic eczema exposed patients.

[3]Calculated by multiplying the estimated incidence rate in atopic eczema exposed patients by the inverse estimated hazard ratio and 99% CI.

[4]Calculated by subtracting the estimated incidence rate and 99% CI in people without atopic eczema from the estimated incidence rate in atopic eczema exposed patients.

[5]Estimated as P(HR - 1) **/** (1 + P(HR - 1)) where P, the prevalence of atopic eczema, is assumed to be 10%[6] and HR is the estimated hazard ratio comparing atopic eczema exposed patients to people without atopic eczema.

Table R3. Association between atopic eczema and cardiovascular outcomes, including adjustment for high dose oral corticosteroid use. Fitted to patients with complete data for all variables included in the models and from valid matched sets1. n = 1,915,916 (1,842,759 unique patients).

	n	P-Y at risk	Events	Hazard ratio and 99% confidence interval[2] Mediation model	
Primary outcomes					
Myocardial infarction					
Unexposed	1,528,477	9,361,522	17,178	1.00	(ref)
Exposed	387,439	2,569,214	5,561	1.03	0.95, 1.12
Unstable angina					
Unexposed	1,528,477	9,392,370	7,059	1.00	(ref)
Exposed	387,439	2,578,165	2,460	1.16	1.03, 1.32
Heart failure					
Unexposed	1,528,477	9,375,383	16,983	1.00	(ref)
Exposed	387,439	2,570,412	6,441	1.16	1.06, 1.27
Atrial fibrillation					
Unexposed	1,528,477	9,316,331	28,571	1.00	(ref)
Exposed	387,439	2,552,311	9,892	1.07	1.00, 1.14
Stroke					
Unexposed	1,528,477	9,361,252	21,387	1.00	(ref)
Exposed	387,439	2,568,749	7,149	1.08	1.00, 1.16
Cardiovascular death					
Unexposed	1,528,477	9,427,420	30,116	1.00	(ref)
Exposed	387,439	2,590,305	10,813	0.96	0.89, 1.03
Secondary outcome					
Coronary revascularisation					
Unexposed	1,528,477	9,358,381	16,195	1.00	(ref)

Exposed	2,567,93				
	387,439	2	5,056	1.08	0.99, 1.18

[1]Matched sets including one exposed patient and at least one unexposed patient.
[2]Estimated hazard ratios from Cox regression with current age as underlying timescale, stratified by matched set (matched on age at cohort entry, gender, date at cohort entry and practice).
Unadjusted: No adjustment.
Adjusted: Adjusted for current calendar period (1997-1999, 2000-2004, 2005-2009, 2010-2015), time since diagnosis (0-4, 5-9, 10-14, 15-19, 20+ years), IMD at cohort entry, and time-varying asthma.
Mediation model: Adjusted for current calendar period (1997-1999, 2000-2004, 2005-2009, 2010-2015), time since diagnosis (0-4, 5-9, 10-14, 15-19, 20+ years), IMD at cohort entry, time-varying asthma, BMI and smoking at cohort entry, and time-varying hyperlipidaemia, hypertension, depression, anxiety, diabetes, severe alcohol use **and high dose oral corticosteroid use**.

Table R4. Association between atopic eczema and cardiovascular outcomes **having excluded all patients ever exposed to ciclosporin**. Fitted to patients with complete data for all variables included in the models and from valid matched sets[1]. n = 1,908,782 (1,836,009 unique patients).

				Hazard ratio and 99% confidence interval[2]					
	n	P-Y at risk	Events	Unadjusted		Adjusted		Mediation model	
Primary outcomes									
Myocardial infarction									
Unexposed	1,522,440	9,318,632	17,075	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	386,342	2,560,201	5,532	1.10	1.05, 1.15	1.06	0.98, 1.15	1.04	0.96, 1.13
Unstable angina									
Unexposed	1,522,440	9,349,311	7,023	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	386,342	2,569,124	2,448	1.22	1.14, 1.31	1.24	1.10, 1.40	1.16	1.02, 1.31
Heart failure									
Unexposed	1,522,440	9,332,437	16,889	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	386,342	2,561,409	6,407	1.21	1.16, 1.26	1.19	1.09, 1.29	1.17	1.08, 1.28
Atrial fibrillation									
Unexposed	1,522,440	9,273,644	28,428	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	386,342	2,543,370	9,849	1.12	1.08, 1.16	1.11	1.04, 1.18	1.07	1.00, 1.15
Stroke									

Unexposed	1,522,440	9,318,293	21,279	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	386,342	2,559,714	7,125	1.07	1.03, 1.12	1.10	1.02, 1.19	1.08	1.00, 1.16
Cardiovascular death									
Unexposed	1,522,440	9,384,126	29,954	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	386,342	2,581,189	10,759	1.07	1.03, 1.11	0.99	0.92, 1.06	0.96	0.89, 1.03
Secondary outcome									
Coronary revascularisation									
Unexposed	1,522,440	9,315,553	16,090	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	386,342	2,558,927	5,026	1.11	1.06, 1.17	1.14	1.05, 1.24	1.08	0.99, 1.18

[1]Matched sets including one exposed patient and at least one unexposed patient.

[2]Estimated hazard ratios from Cox regression with current age as underlying timescale, stratified by matched set (matched on age at cohort entry, gender, date at cohort entry and practice).

Unadjusted: No adjustment.

Adjusted: Adjusted for current calendar period (1997-1999, 2000-2004, 2005-2009, 2010-2015), time since diagnosis (0-4, 5-9, 10-14, 15-19, 20+ years), IMD at cohort entry, and time-varying asthma.

Mediation model: Adjusted additionally for BMI and smoking at cohort entry, and time-varying hyperlipidaemia, hypertension, depression, anxiety, diabetes and severe alcohol use.

Table R5. Association between atopic eczema and cardiovascular outcomes **having excluded all patients ever exposed to methotrexate**. Fitted to patients with complete data for all variables included in the models and from valid matched sets[1]. n = 1,877,760 (1,806,905 unique patients).

				Hazard ratio and 99% confidence interval[2]					
	n	P-Y at risk	Events	Unadjusted		Adjusted		Mediation model	
Primary outcomes									
Myocardial infarction									
Unexposed	1,495,325	9,118,224	16,528	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	382,435	2,527,663	5,397	1.09	1.05, 1.14	1.06	0.97, 1.15	1.04	0.95, 1.13
Unstable angina									
Unexposed	1,495,325	9,147,745	6,805	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	382,435	2,536,131	2,406	1.23	1.15, 1.31	1.26	1.12, 1.43	1.18	1.04, 1.34
Heart failure									
Unexposed	1,495,325	9,131,610	16,337	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	382,435	2,528,742	6,275	1.21	1.16, 1.27	1.20	1.10, 1.30	1.18	1.08, 1.29
Atrial fibrillation									
Unexposed	1,495,325	9,074,399	27,596	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	382,435	2,510,919	9,666	1.12	1.08, 1.16	1.11	1.03, 1.18	1.07	1.00, 1.15
Stroke									

Unexposed	1,495,325	9,117,838	20,658	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	382,435	2,526,916	7,023	1.08	1.03, 1.12	1.11	1.03, 1.20	1.09	1.01, 1.18
Cardiovascular death									
Unexposed	1,495,325	9,181,591	29,117	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	382,435	2,547,994	10,608	1.07	1.04, 1.11	0.99	0.93, 1.07	0.97	0.90, 1.04
Secondary outcome									
Coronary revascularisation									
Unexposed	1,495,325	9,115,157	15,591	1.00	(ref)	1.00	(ref)	1.00	(ref)
Exposed	382,435	2,526,191	4,926	1.12	1.07, 1.17	1.14	1.04, 1.24	1.08	0.99, 1.18

[1]Matched sets including one exposed patient and at least one unexposed patient.

[2]Estimated hazard ratios from Cox regression with current age as underlying timescale, stratified by matched set (matched on age at cohort entry, gender, date at cohort entry and practice).

Unadjusted: No adjustment.

Adjusted: Adjusted for current calendar period (1997-1999, 2000-2004, 2005-2009, 2010-2015), time since diagnosis (0-4, 5-9, 10-14, 15-19, 20+ years), IMD at cohort entry, and time-varying asthma.

Mediation model: Adjusted additionally for BMI and smoking at cohort entry, and time-varying hyperlipidaemia, hypertension, depression, anxiety, diabetes and severe alcohol use.