Author's response to reviews

Title: Cardiovascular disease prevalence in patients with inflammatory arthritis, diabetes mellitus and osteoarthritis: a cross-sectional study in primary care

Authors:

Markus MJ Nielen (m.nielen@nivel.nl)
Alper M van Sijl (Alper.vanSijl@vumc.nl)
Mike JL Peters (MJL.Peters@vumc.nl)
Robert A Verheij (r.verheij@nivel.nl)
Francois G Schellevis (f.schellevis@niel.nl)
Michael T Nurmohamed (MT.Nurmohamed@vumc.nl)

Version: 2 Date: 31 May 2012

Author's response to reviews: see over
May 31, 2012

Dear Sir / Madam,

Please find enclosed the revised manuscript ‘Cardiovascular disease prevalence in patients with inflammatory arthritis, diabetes mellitus and osteoarthritis: a cross-sectional study in primary care’. We would like to thank the reviewers for the valuable suggestions, which improved the quality of the manuscript. A point-by-point response to the concerns of the referees is added to this cover letter. All changes in the revised manuscript are marked blue.

We hope that you will find the manuscript suitable for publication, and look forward to hearing from you.

Yours sincerely, also on behalf of the co-authors

Mark Nielen, PhD
Reviewer: Soon H Song

Reviewer's report:
This study looked at the prevalence of cardiovascular disease among patients with inflammatory arthritis, diabetes and osteoarthritis compared to controls in a primary care setting.

Comments:
(a) The authors have shown that CVD risk in rheumatoid arthritis approaches that of diabetes mellitus. How does this study add to the authors previous work? (Major Compulsory Revisions)
In the previous published work of van Halm and Peters (references 9 and 10), a cohort of RA patients in specialized care (CARRE study) was used and compared with subjects from the Hoorn study, a population-based cohort study of glucose metabolism and other cardiovascular risk factors. In this study we used data from patients as well as controls from the same primary care cohort. This may result in a better estimation of the association between inflammatory arthritis / diabetes and CVD, since 1) using secondary care-based cohorts may lead to an overestimation of the CVD burden due to an overrepresentation of patients with more severe disease, 2) data from patients as well as controls are collected in the same manner, and 3) both patients and controls in this study are representative for the Dutch population. We added this to the introduction of the manuscript.

(b) Subjects age under 30 yrs were excluded from analysis because of lower probability of having CVD. Did the authors have any data to support this point? This can affect the overall CVD prevalence rate in this population and potentially the conclusion of this study particularly the strength of associations of each condition with CVD. (Major Compulsory Revisions)
In 2010 the Dutch Heart Foundation studied the prevalence rate of cardiovascular diseases in the Netherlands (Vaartjes I, van Dis I, Visseren FLJ, Bots ML. Hart- en vaatziekten in Nederland 2010, cijfers over leefstijl- en risicofactoren, ziekte en sterfte. Den Haag: Nederlandse Hartstichting, 2010.). They found an overall prevalence rate of 6.2 and 8.5 per 1,000 for males and females, respectively (see table below). Since the prevalence rate in patients under 30 yrs was 0.1 - 0.2 per 1,000, we do not think that this has negatively influenced the results and conclusions of our study. We added this reference to the manuscript.

(c) Practices that recorded data less than 6 months were excluded. The reason for this exclusion was not given. This can unnecessarily exclude a large number of patients and can affect the overall CVD prevalence rate in this population and the conclusion of this study similar to point (b). (Major Compulsory Revisions)
We excluded these practices, because the diagnoses of the patients are based on complaints/morbidity presented during GP consultations. When more than half of the consultations in a year are missing, the prevalence rates of the studied diseases are underestimated. This is added to the method section. We do not think that excluding these practices has influenced our conclusions, since it is not likely that the quality of the registration of a GP is associated with prevalence rates of the studied chronic diseases.

| Table: Prevalence rate (per 1,000) of CVD in the Netherlands in 2007 |
|----------------|-----------|   |
|                | Male      | Female |
| <30 years      | 0.2       | 0.1    |
| 30-34 years    | 0.3       | 0.2    |
| 35-39 years    | 0.5       | 0.3    |
| 40-44 years    | 0.8       | 0.5    |
| 45-49 years    | 1.3       | 0.9    |
| 50-54 years    | 2.4       | 1.7    |
| 55-59 years    | 4.5       | 3.4    |
| 60-64 years    | 8.1       | 6.5    |
| 65-69 years    | 15.2      | 12.8   |
| 70-74 years    | 27.5      | 24.5   |
| 75-79 years    | 47.3      | 44.4   |
| 80-84 years    | 77.0      | 76.0   |
| 85+ years      | 113.1     | 116.3  |
| Total          | 6.2       | 8.5    |
(d) Table 1 - statistical comparisons were made against controls. In addition, it would give more information if comparisons were also made for CVD risk factors between each condition studied. (Major Compulsory Revisions)

We agree with the reviewer that it would be interesting to compare the CVD risk factors between the patient groups. However, the scope of this manuscript was to associate IA, DM and OA with CVD. The CVD risk factors are only used as confounding factors. Furthermore, these analyses were not performed, since it is not possible to compare between the groups due to overlapping patients.

When performing these analyses (after excluding patients with a combination of these diseases), the (crude) prevalence rate of both hypertension and hypercholesterolemia are significantly higher in OA and DM patients compared with IA patients. The differences between DM and OA patients were also statistically significant.

(e) Do the given prevalence of diabetes, inflammatory arthritis and osteoarthritis refer to subjects with only one condition or does it also include those with mixed conditions? This needs to be clarified. It will also be helpful to give the prevalence of subjects with mixed conditions. (Major Compulsory Revisions)

The prevalence rates reported in the second sentence of the results section also include patients with multiple conditions, which is described in the sentence “Because patients can have multiple disorders, the sum of the number of patients in the four groups exceeds the total of 175,061 patients”. The percentages of patients with two conditions are shown in the table. We decided not to report the patients with all three conditions, because of the low prevalence rate (only 12 patients: 0.1 per 1000 patients). We corrected for the influence of having multiple conditions in our statistical analyses (see response on comment f) and therefore we decided not to add this to the manuscript.

(f) Table 2 - did the odds ratio refer to one individual condition (diabetes or OA or inflammatory arthritis) or did it include mixture of these conditions? If so, how can the authors conclude that inflammatory arthritis gave similar CVD risk to diabetes as this may also reflect the effect of diabetes in subjects with mixed conditions? (Major Compulsory Revisions)

In the paragraph ‘statistical analyses’ we described: “Three models were calculated: 1) a model with the three studied independent variables (presence of inflammatory arthritis, diabetes mellitus and osteoarthritis), 2) model 1 plus age and gender, and 3) model 2 plus CVD risk factors (hypertension and hypercholesterolemia). Results are presented as odds ratios with 95%-confidence intervals. Interaction terms were introduced to test for selective amplification of the CVD risk by subjects with more than one of the investigated diseases and the possible modifying effect of age”. By adding all three diseases to the model, the effects of the individual conditions are corrected for the influence of the other diseases.

(g) The authors didn’t give the breakdown of the prevalence of CVD at different age bands for diabetes, inflammatory arthritis and osteoarthritis. It would be informative to have an idea of any difference (if any) in relation to the age factor. Do subjects with inflammatory arthritis have higher burden of CVD at an earlier age compared to diabetes/osteoarthritis/controls? (Major Compulsory Revisions)

We agree with the reviewer that it would be interesting to look at prevalence rates for different age groups (see table below). We have decide not to add these rates to the manuscript because of two reasons: 1) these rates are not comparable between the three groups, because they are not corrected for all confounders, and 2) we did not found statistically significant interaction effects with age, from which it can be concluded that the association between CVD and the conditions IA, DM and OA did not differ between different age groups (see results section: “Interaction terms were added to model 3 to test for additional effects for subjects with more than one of the investigated diseases and the modifying effect of age. However, none of the interaction terms reached statistical significance (data not shown)”.

<table>
<thead>
<tr>
<th>Table: CVD prevalence rate (per 1,000) by age group</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>30-49 years</td>
</tr>
<tr>
<td>50-69 years</td>
</tr>
<tr>
<td>70+ years</td>
</tr>
</tbody>
</table>
Reviewer: Marissa Lassere

Major Compulsory Revisions
The author must respond to these before a decision on publication can be reached. For example, additional necessary experiments or controls, statistical mistakes, errors in interpretation.

This paper describes the prevalence rate of non-fatal cardiovascular in patients with inflammatory arthritis, diabetes and osteoarthritis compared to patients without these disorders in the setting of primary care. The authors have published three papers on this in 2009 (see authors publications 8, 9 and 10) Diagnoses are recorded using an electronic medical record used by GPs. A classification system is recorded when the general practitioner issues a prescription. Individuals are classified within these diagnostic groups.

The diagnosis and was recorded in 2006 and up to 2004. I'm not entirely certain what this means.
1. What is the exact period?
We used the year 2006 to define our population and used data from these patients of the period 2004-2006. By using a period of three years, we could determine a more reliable estimation of the prevalence rates, since not all patients are visiting their GP every year with complains related to the studied diseases. We clarified this in the revised manuscript.

2. How often was the diagnostic classification reviewed?
Individuals were classified within a certain patient group when the diagnosis was recorded at least once in their EMR in the period 2004-2006.

The controls would be a combination of positive and negative risk factor groups combined.
3. Is there any more information about them? Major diagnostic groups, medications used?
The controls in this study are all registered patients in the participating general practices without IA, DM and/or OA. We agree with the reviewer that the control group is a combination of individuals with positive and negative risk factors. We used this control group because they represent the 'average Dutch patient in primary care'. The main outcome in this study is CVD and therefore we have decided to only report CVD risk factors in this group in table 1.

4. How accurate is this diagnostic system. What validation has been done regarding this in general and specifically for this group of GP practices. My experience with electronic software used by GPs in Australia suggests that for some diagnoses it is very accurate and for others less so. Essentially the prescription is often a surrogate for the underlying condition that is being evaluated. There are classifications that are specifically important – the study factor classification – ie for inflammatory arthritis, diabetes, and the outcome factor i.e cardiovascular events. Both study and outcome factor must be robust for the conclusions to be valid. Study factor: Unfortunately the diagnostic code is not specific so that rheumatoid arthritis (RA) and ankylosing spondylitis (AS) two very different conditions, are combined.

Using large databases with data from EMRs of GPs has a number of disadvantages, which must be weighed against the advantages:
1) The accuracy of the ICPC coding system:
   - We agree with the reviewer that the accuracy of the ICPC coding system is a limitation of this study. In this study we used data from the Netherlands Information Network of General Practice (LINH) including 96 general practices with about 360,000 registered patients, which makes it impossible to validate diagnoses and to collect additional data in terms of time and funding. Although diagnoses are not validated by a medical specialist, the used GP diagnoses are based on the guidelines of the Dutch College of General Practitioners (NHG). Furthermore, medical specialists send a letter to the GP after the consultation of a patient with information about diagnosis and treatment. It is likely that the GP uses this information to register the diagnosis of inflammatory arthritis in the EMR of the patient. Therefore, we expect an overestimation of the number of IA patients in our primary care cohort, which might resulted in an underestimation of our results.
   - Unfortunately, it is not possible to distinguish between RA and AS, because both are registered as ICPC L88. We combined these diseases, because both diseases are characterized by chronic inflammation of the joints, which may be one of the explanation of excess CVD in both rheumatic diseases.
   - In the Netherlands, GPs have a gatekeeper role for access to specialized care and the GP is the first professional to be consulted for health problems. We expect, although the diagnoses
are not validated, that the morbidity of the studied patients are more complete compared to specialized care.

2) Completeness of data:
Data in health registries, such as LINH, are not collected in a structural way and patients are not measured periodically; we only have data from patients who visit their GP for a health problem. As a results, important confounders, such as smoking and weight, are not always registered and these variables cannot be used in the statistical analyses. Also, not all prescribed drugs in specialised care are registered in the EMRs of GPs, which makes it not possible to measure the use of DMARDs reliably.

3) Study population
Studies in specialized care often have an overrepresentation of inflammatory arthritis patients with more severe disease. Not all Inflammatory arthritis patients are under treatment of a rheumatologists, and therefore in primary care patients have a broader spectrum of disease severity, which may result in a more accurate estimate of the CVD prevalence. The results are representative for the total population of inflammatory arthritis patients. Using data from GPs also makes it possible to compare morbidity and health care utilization of inflammatory arthritis patients with other patient groups or healthy individuals. In both patients and controls, morbidity is measured with the same methods. Besides the above mentioned advantages, health registries can be used to study large groups of patients, including (combination of) diseases with low prevalence rates.

In conclusion, the use of large health registries has advantages and disadvantages. With large database including EMRs of patients in primary care, it is possible to study large groups of patients in a representative population, including control groups. Since GPs have a gatekeeper role for access to specialized care, all presented morbidity is registered with a uniform method. For more specific information about the patients’ disease activity, risk factors (smoking, overweight) and treatment, it is better to use smaller cohorts with structured follow-up measurements.

The above mentioned advantages and disadvantages are added to the manuscript

5. How was RA and AS validated. As a rheumatologist, I am aware of misclassification of these conditions. (see authors publications 9 and 10)
We agree with the reviewer that misclassification of RA and AS patients is a possibility. As described in the response on comment 4, misclassification could only have resulted in an underestimation of our results.

6. How many patients with RA were on disease modifying agents?
As described in the response on comment 4, we do not have all information about prescribed medication in specialized care.

7. Also diabetes populations differ by duration of disease and severity, control and its management (ie need for insulin or not). Is there any information on this? (see authors publications 9 and 10)
Since this is a cross-sectional design, we expect that this is a population of diabetes patients with average disease duration, severity, control and managements. Similar to the IA patients, there is no information about these parameters (see response on comment 4).

Outcome factor: Conditions such as myocardial infarction and transient ischaemic attack are not always straight forward diagnoses. These are conditions that often require adjudication within the research setting, and clearly the accuracy of the data may not be sufficient in some administrative databases. ICD-10 codes allow for the various ischaemic heart disease conditions. No information is provided regarding how the accuracy of these diagnoses was evaluated.

8. Could the authors also provide information in the manuscript regarding this? (see authors publications 8, 9 and 10)
We agree with the reviewer that it would be better to validate the CVD diagnoses. However this was not possible in this study (see response on comment 4).
9. If there is misclassification bias – is it differential or non differential? We cannot say because there is no information within the manuscript. Can the authors please comment?
We do not know whether there was misclassification in this study. As described in response on comment 4 and 5, we expect that misclassification only has resulted in an underestimation of our results.

Hypertension and hypercholesterolaemia based on blood pressure readings and blood chemistry results – there would be more GP documentation. Cholesterol could be total or LDL-cholesterol alone.

Both diagnoses are based on ICPC coded morbidity by the GP according to the guidelines. Unfortunately, we do not have the underlying blood pressure and blood chemistry results.

The statistical analysis uses multivariate logistic regression. Patients were clustered within practices presumably to look for practice differences. Patients had multiple disorders; therefore, each group exceeded the total number of patients. Table 1 shows the population characteristics. I note that the controls mean age 51 versus OA mean age of 69 is a 20 year difference. Inflammatory arthritis falls in between at 60 years. There was also a female preponderance in the study factor group. More objective data such as the use of the prescription antihypertensive agents and the use of statins shows that these are much higher in the osteoarthritis groups than in the inflammatory arthritis groups.

I would have restricted the dataset to patients with age greater than 50, as the authors did in reference 8, and provided results by decade strata, again as per reference 8.

See response on comment g) from the other reviewer.

Study Design: When determining OR the study design including the comparator/reference groups is the most important determinant of the result. If the hypothesis is that inflammatory arthritis has a greater cardiovascular risk then noninflammatory arthritis osteoarthritis should be the comparator control/reference rather than the non-specific controls (noted above). Then some other factors such as use of NSAIDS and prednisone are shared, particularly if the hypothesis is inflammatory burden rather than its management.

10. Could the authors provide this analysis. Without this I do not believe that the authors add much more to the state of knowledge over and above their publications from 2009, furthermore, this would be a better evaluation of their hypothesis.

With the used statistical methods it is not only possible to compare the ORs of the three studied diseases with controls, but also to compare between the three diseases. The presence of IA, DM and OA are added to the model as dichotomous variables, resulting in a ‘basic CVD risk’ when all these variables are zero. When using OA patients as a control group, it remains unclear whether the association of the studied diseases with CVD is higher compared with the general population. In model 3, the 95% CI for OA and IA are not overlapping, which results in a significantly higher association between IA and CVD compared with OA and CVD. Therefore, we have decided not to change the statistical analyses. Because it was not clear why we used both OA patients and controls in our analyses, we clarified this choice in more detail in the introduction of the revised manuscript.

The results are shown in table 2, model three (adjusting form some risk factors. This shows an increased odds ratio of cardiovascular events in the inflammatory arthritis and diabetes groups as compared to controls. If the authors’ intentions are to draw conclusions about inflammatory processes and risk of cardiovascular events then I would also suggest the following:

11. Please separate the results by cardiovascular events and not report just total cardiovascular events. The readers can decide how much confidence there is about the accuracy of the diagnostic codes for myocardial infarction, transient ischaemic attacks, strokes with neurological deficits etc (as the authors did in reference 8). Furthermore ankylosing spondylitis is a very different condition to rheumatoid arthritis, but these cannot be separated in their classification system.

We agree with the reviewer that it would be useful to add ORs for myocardial infarction, transient ischaemic attacks and stroke to the manuscript (see table below).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Controls</th>
<th>IA</th>
<th>DM</th>
<th>OA</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVD total</td>
<td>1</td>
<td>1.5 (1.2-1.9)*</td>
<td>1.3 (1.2-1.4)*</td>
<td>0.8 (0.7-1.0)*</td>
</tr>
<tr>
<td>MI</td>
<td>1</td>
<td>1.3 (0.9-2.0)</td>
<td>1.3 (1.1-1.5)*</td>
<td>0.8 (0.6-1.0)</td>
</tr>
<tr>
<td>TIA</td>
<td>1</td>
<td>1.4 (1.0-2.1)</td>
<td>1.1 (0.9-1.2)</td>
<td>0.9 (0.7-1.1)</td>
</tr>
<tr>
<td>CVA</td>
<td>1</td>
<td>1.5 (1.1-2.0)*</td>
<td>1.4 (1.2-1.5)*</td>
<td>0.9 (0.7-1.0)</td>
</tr>
</tbody>
</table>

* P<0.05
In the table, only the results of model 3 (corrected for all confounders) are presented. In general, these data do not change the conclusions of our manuscript. There are not many differences between the three cardiovascular diseases, although due to loss of power, not all effects are statistically significant.

This was added to the results section of the manuscript.

12. Is smoking information known therefore included in the model? The medications that are used in patients with inflammatory arthritis such as the non-steroidal anti-inflammatory agents and prednisolone are important risk factors for cardiovascular morbidity and these are not included in the analysis.

We agree with the reviewer that these risk factors are important confounders. However it was not possible to correct for these factors in the analyses (see response on comment 4). We added this to the discussion section of the revised manuscript.

13. The administrative database would also have this prescription data and so it could also be described and used to validate diagnostic codes.

Prescription data were also used in this study, but only in combination with ICPC codes: ‘In this study morbidity data and data on all prescriptions issued by the participating practices were used.’

Discretionary Revisions
For example clarifications, data that would be useful but not essential.

14. Another analysis that would add to the robustness of the results and therefore strengthen the conclusions would be to take a second snapshot to see whether the results are reproducible within these set of GP practices. I note that this appears to be the same snapshot of time as reference 8.

We agree that reproducing our data would strengthen the results of the study. However, this is not possible, because we used all patients from our registry.

If the above analyses are undertaken then we would be more confident about the results and authors conclusions regarding management in general practice, and this may have an influence on future GP management of these conditions.