PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	A transient peak of infections during onset of rheumatoid arthritis: a 10-year prospective cohort study
AUTHORS	Arleevskaya, Marina; Gabdoulkhakova, Aida; Filina, Julia; Miftakhova, Regina; Bredberg, Anders; Tsibulkin, Anatoly

VERSION 1 - REVIEW

REVIEWER	Gabriel Chodick Tel Aviv University, Israel
REVIEW RETURNED	19-Apr-2014

GENERAL COMMENTS	In this study, the authors aimed to document the infection burden in a large prospective cohort of RA patients, their relatives predisposed to RA, and healthy controls. The higher risk of infection in patients with RA when compared with a healthy control population is well-documented, and it is unclear what novel aspects the authors sought to examine in the present analysis. With a long-term follow-up with clinical comprehensive examination and relatively large study populations of RA patients and two comparison cohorts, the present study has good potential for an interesting analysis; however, the manuscript should be revised thoroughly for coherence of the written presentation, for exact definitions of study groups where patients are not used in several groups without proper time-dependent analyses, for better definitions of outcomes (e.g. survival analysis to first infection event; severity of infection, etc.), for standard statistical analysis (e.g. multivariable analysis to allow adjustment for important confounders such as age and use of DMARDs that are associated with higher likelihood of infection), for a clear attrition table (e.g. how could exclusion due to smoking, habitual alcoholism, diabetes mellitus, and leucopenia sum up to less than 5% of the total population?), for concise use of figures and tables, and for many other important methodological issues.

REVIEWER	Julia Simard Stanford
REVIEW RETURNED	09-May-2014

GENERAL COMMENTS	The introduction, which should motivate the reader and the question, needs some rewriting. For instance, without the second sentence in the introduction, the first paragraph reads more clearly. Several issues regarding the methods need to be clarified and reconsidered:

Page 4, Materials and Methods, Study population, the authors describe the population and the crossover between different categories (from relative to early RA to advanced RA) in terms of people. For the sake of this discussion, let's refer to these categories as exposures. That means that an individual can, during follow-up go from being exposed to A to B, others from B to C. It would perhaps be more clear and accurate to consider the time in this study as person-time and estimate actual infection rates (# infections/person-time) rather than proportion of those individuals with an infection (which they inaccurately call "rate")

Outcomes (infections) were identified for the non-experimental component of the study by self-reported symptoms at study visits. And "only those episodes judged by the rheumatologist to truly indicate an infection were scored" (pg 5). More information is needed here. How did the rheumatologist judge these? was objective evidence of infection necessary at that time? For the experimental subset the authors write that these were performed for a limited time of the study and hence on a fraction of the population. Other than calendar time, was there some selection mechanisms in place? Did all individuals self-reporting infection whose rheumatologists deemed their events as infection have the opportunity to be included? In essence, my concern here has to do with the validity of self-report, the definition of outcome and how rheumatologist judgement was incorporated, and whether there was sampling/selection bias for the subset undergoing the experimental procedures.

The authors look at an interesting population and observe a peak in infections earlier in the disease course in a longitudinal study with controls, early RA cases, and individuals with RA for more than 3 years. Although the question is interesting and some of the data intriguing, as it stands there are many questions about how things were defined, their validity, and so on. Perhaps simple clarification of a number of these points will do away with these concerns. I raise these points in the hope of providing constructive feedback and understanding the design more clearly.

When disease duration was calculated/estimated to define early and advanced RA, what information did the authors use? Since since diagnosis? Time since first reported symptom? Was it reported by patient? Based on a physician documented finding?

Authors may want to consider a different term for RA of longer disease duration. Some people interpret "advanced" to mean "severe"

Be cautious with your introduction and use of acronyms. For example on page 4 HC is used for healthy control in line 37 but defined in line 44.

"Healthy" relative and "Healthy" control may imply more selection than the authors intended to imply. Were these individuals actually "healthy" such as healthy blood donors? Or were they just selected and screened to not have RA? It may be sufficient to refer to them as relative controls without implying healthiness, unless they were really healthy.

Because the sampling of the experimental subset and the outcome

definition were not very clear to me, I struggled with interpreting the results in the absence of fully understanding the possible magnitude of bias and sampling. Furthermore, with data such as days of infection, it is unclear from the methods how this is a reliable/valid measure for the entire study population.

This reader would benefit from a figure depicting the timeline of data collection for the different "exposure" groups.

Some minor points:
page 5, Results, Infection quantity "In those HR with developed RA during the study the infection quantity was lower in the first examination while still a healthy relative..." The term "infection quantity" is confusing here. Can you be more descriptive?

Page 6, Infection types, second sentence "There was no infection in ..." It's difficult to know how to interpret these sentences without a real understanding of how infections were reported and who is in this denominator. During the time that the experimental subset was included? Is that the time that this represents? And who is in the denominator?

Table 1, given the cross-over across categories (HC, HR, eRA, aRA) are these categories mutually exclusive? Or are they really functions of person-time?

Table 2, supp. It is unclear, even with the footnote what column B represents. Incidence per year as in number of infections per year? Given that there are decimals, this is the mean? some other measure?

The 9 pages of supplemental writing with >100 references is excessive and needs to be more clearly introduced as supplemental material. As it stands it will be confusing to the reader how this information should be incorporated and interpreted in light of the presented study.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:
In this study, the authors aimed to document the infection burden in a large prospective cohort of RA patients, their relatives predisposed to RA, and healthy controls. The higher risk of infection in patients with RA when compared with a healthy control population is well-documented, and it is unclear what novel aspects the authors sought to examine in the present analysis. With a long-term follow-up with clinical comprehensive examination and relatively large study populations of RA patients and two comparison cohorts, the present study has good potential for an interesting analysis; however, the manuscript should be revised thoroughly for coherence of the written presentation,
The language and the presentation in general have been revised throughout all manuscript sections. More details on this comment are given in reply to the second reviewer's more specific comments.

for exact definitions of study groups where patients are not used in several groups without proper time-dependent analyses,
The study groups are described in more detail in Materials and Methods, Study population: healthy controls and healthy relatives are defined in detail, and the term advanced RA has been altered to late RA. A time-dependent, as a complement to study group-dependent, analysis of infection duration

has been included as Figure 1B.

for better definitions of outcomes (e.g. survival analysis to first infection event; severity of infection, etc.),
The outcome determined in our study is infections (as the reviewer 2 states in her para 4, first 2 words). We have no data on survival, and we have not graded the severity of infections. In reply to this comment, we now describe more clearly (in Materials and Methods, Study design) how infections were diagnosed.

for standard statistical analysis (e.g. multivariable analysis to allow adjustment for important confounders such as age and use of DMARDs that are associated with higher likelihood of infection),
The Spearman rank order correlation test has now been performed, showing that there was no correlation between age and likelihood of infections. This is stated in Results, first para, last sentence.
A more frequent use of DMARD cannot be a cause of the higher likelihood of infections observed in early-stage patients, because the eRA group (with significantly more of infections than the later-stage laRA group) infection result reflects the year preceding RA diagnosis, with a definitely lower intake of DMARD as compared with after diagnosis. DMARD as a confounding factor is now discussed at more length in Discussion, para 2, last lines, as follows:
"Furthermore, the more of anti-inflammatory medication, the less may be expected of symptoms perceived to indicate an infection. In fact, we did observe more of infections among the eRA group (in whom infections during the one-year period preceding diagnosis were scored) with certainly less of drug therapy as compared with the laRA patients with much less infections, all of whom were prescribed DMARD. Nevertheless, we find it highly unlikely that our recording of less infections in late stage patients is due to DMARD usage. Firstly, the observed size of infection reduction (from 62 days in eRA to 12 days in laRA, Figure 1A) is too large to be due to drug masking of infection signs; secondly, the highly significantly increased infection load seen at time of diagnosis in the HR subjects developing RA (Figure 1B) cannot be attributed to a decrease in DMARD intake."

for a clear attrition table (e.g. how could exclusion due to smoking, habitual alcoholism, diabetes mellitus, and leucopenia sum up to less than 5% of the total population?)
We fully understand this concern. The text now more clearly states how exclusion was done, in Materials and Methods, Study population, 7 last sentences, as follows:
"Exclusion criteria were some risk factors for infection [7]. 3.6% were excluded due to strongly immunosuppressive therapy, and another 1.3% because of smoking or habitual alcoholism (it is by tradition in this geographic area an uncommon practice for a woman to be smoker). In addition, 15-20% of the number of studied patients were excluded from participation because of a concomitant chronic disease (mainly diabetes mellitus). All patients received DMARD; 35 patients infliximab, 1 abatacept, 2 tocilizumab, 7 rituximab; 7% were on prednisolone per os averaging 10 mg daily."
We have chosen not to include a table for the exclusion criteria, in line with this reviewer's request for concise use of illustrations (see below his last comment).

for concise use of figures and tables, and for many other important methodological issues.
Figure 1 is now more concise, with both study group- and time-dependent results; the supplementary Table 2 has been deleted, because it largely overlapped with Table 1.
A number of methodological issues have been addressed, as requested by the second reviewer.

Reviewer: 2
The introduction, which should motivate the reader and the question, needs some rewriting. For instance, without the second sentence in the introduction, the first paragraph reads more clearly.
The second sentence of the introduction has been deleted.

Several issues regarding the methods need to be clarified and reconsidered:

Page 4, Materials and Methods, Study population, the authors describe the population and the crossover between different categories (from relative to early RA to advanced RA) in terms of people. For the sake of this discussion, let's refer to these categories as exposures. That means that an individual can, during follow-up go from being exposed to A to B, others from B to C. It would perhaps be more clear and accurate to consider the time in this study as person-time and estimate actual infection rates (# infections/person-time) rather than proportion of those individuals with an infection (which they inaccurately call "rate")
We appreciate very much this reviewer suggestion. A time-dependent, as a complement to study group-dependent in Figure 1A, analysis of infection duration is now included as Figure 1B. Both types of analysis reveal a transient and highly significant excess of infections during the early RA stage.

Outcomes (infections) were identified for the non-experimental component of the study by self-reported symptoms at study visits. And "only those episodes judged by the rheumatologist to truly indicate an infection were scored" (pg 5). More information is needed here. How did the rheumatologist judge these?
We now describe more clearly (in Materials and Methods, Study design, lines 7-8: "there was no formal list of criteria to be addressed during this procedure") how infections were diagnosed, making it clear that the rheumatologist based infection scoring on the patient's history in combination with written reports from practitioners visited during infection episodes.

was objective evidence of infection necessary at that time?
No, the scored infections were not limited to those present during the semiannual visits to a rheumatologist, see the above reply.

For the experimental subset the authors write that these were performed for a limited time of the study and hence on a fraction of the population. Other than calendar time, was there some selection mechanisms in place?
No.

Did all individuals self-reporting infection whose rheumatologists deemed their events as infection have the opportunity to be included?
Yes.

In essence, my concern here has to do with the validity of self-report, the definition of outcome and how rheumatologist judgement was incorporated, and whether there was sampling/selection bias for the subset undergoing the experimental procedures.
We have addressed as stated above these important concerns for validity of self-report and for how infections were scored. The clarifications are included in Materials and Methods, Study design, lines 10 and 13.

The authors look at an interesting population and observe a peak in infections earlier in the disease course in a longitudinal study with controls, early RA cases, and individuals with RA for more than 3 years. Although the question is interesting and some of the data intriguing, as it stands there are many questions about how things were defined, their validity, and so on. Perhaps simple clarification of a number of these points will do away with these concerns. I raise these points in the hope of providing constructive feedback and understanding the design more clearly.
Please find our replies point-by-point below.

When disease duration was calculated/estimated to define early and advanced RA, what information did the authors use? Since since diagnosis? Time since first reported symptom? Was it reported by

patient? Based on a physician documented finding?
We assume that the reviewer in this comment refers to the disease RA (not infection). We have addressed all the above five questions by expanding the Materials and Methods Study design section as follows: "RA was diagnosed by the coordinated decision of 3 rheumatologists using the American Rheumatology Association criteria of 1987, and from 2010 the EULAR criteria; for those patients with a diagnosis already when referred to our center, the time of diagnosis was estimated based on records from previous physicians and the patient's report. Time zero during disease progression is defined as the time point when these diagnostic criteria were met."
In addition, we have tried to make understanding of the RA duration more easy, by revising the supplementary Table 1, with information on how cross-over between the study groups occurs at specific time points.

Authors may want to consider a different term for RA of longer disease duration. Some people interpret "advanced" to mean "severe"
Throughout the text the abbreviation aRA, for advanced RA, has now been altered to laRA, for late RA.

Be cautious with your introduction and use of acronyms. For example on page 4 HC is used for healthy control in line 37 but defined in line 44.
This error example has been corrected, and also some additional ones, and we have carefully gone through the entire text for such mistakes.

"Healthy" relative and "Healthy" control may imply more selection than the authors intended to imply. Were these individuals actually "healthy" such as healthy blood donors? Or were they just selected and screened to not have RA? It may be sufficient to refer to them as relative controls without implying healthiness, unless they were really healthy.
We have kept the term Healthy control, because these subjects were free from chronic diseases, and healthy in the same sense as blood donors. This is now clearly stated in Materials and Methods, Study population, first para, last lines: "227 women with no chronic disease and with no RA among close relatives served as HC; after being found by a rheumatologist to lack signs of arthritis and to be negative in the same laboratory tests as were performed with the HR".

Because the sampling of the experimental subset and the outcome definition were not very clear to me, I struggled with interpreting the results in the absence of fully understanding the possible magnitude of bias and sampling. Furthermore, with data such as days of infection, it is unclear from the methods how this is a reliable/valid measure for the entire study population.

This reader would benefit from a figure depicting the timeline of data collection for the different "exposure" groups.
Addressing the above two paragraphs, we refer to our above replies dealing with definitions of study groups and the infection outcome, and with bias of sampling.
Regarding the timeline of data collection, we repeat the second reply to the reviewer 2:
"A time-dependent, as a complement to study group-dependent in Figure 1A, analysis of infection duration is now included as Figure 1B. Both types of analysis reveals a transient and highly significant excess of infections during the early RA stage."

Some minor points:

page 5, Results, Infection quantity "In those HR with developed RA during the study the infection quantity was lower in the first examination while still a healthy relative..." The term "infection quantity" is confusing here. Can you be more descriptive?
The term "Infection quantity" has been deleted. Accordingly, the first para of Results is now more

descriptive: "Number of days with an infection, experienced during a one-year period".

Page 6, Infection types, second sentence "There was no infection in ..." It's difficult to know how to interpret these sentences without a real understanding of how infections were reported and who is in this denominator. During the time that the experimental subset was included? Is that the time that this represents? And who is in the denominator?
A lack of infections means that no evidence of a trivial or any other type of infection was revealed.We hope that this is now made clear in Materials and Methods, Study design, describing how infections were diagnosed, over a one-year time period. Here it is stated that the rheumatologist based infection scoring on the patient's history in combination with written reports from practitioners visited during infection episodes.

Table 1, given the cross-over across categories (HC, HR, eRA, aRA) are these categories mutually exclusive? Or are they really functions of person-time?
No, the study group categories are not mutually exclusive, because there are cross-overs. This has been dealt with in detail in the above replies. The revised Figure 1B and supplemental Table 1 show the person-time function, and the extent of cross-over, respectively.

Table 2, supp. It is unclear, even with the footnote what column B represents. Incidence per year as in number of infections per year? Given that there are decimals, this is the mean? some other measure?
We agree that these data are confusing, and have chosen to delete the supplemental Table 2, which to a large extent overlaps with Table 1.

The 9 pages of supplemental writing with >100 references is excessive and needs to be more clearly introduced as supplemental material. As it stands it will be confusing to the reader how this information should be incorporated and interpreted in light of the presented study.
We agree that these pages are excessive, and have chosen to delete it. We judge them to be more suitable for a separate perspective type of report.

Hoping that you will find that we have properly addressed all the reviewers'comments, and now find our manuscript acceptable for publication.

VERSION 2 – REVIEW

REVIEWER	Gabriel Chodick
	MHS, Israel
REVIEW RETURNED	03-Aug-2014

GENERAL COMMENTS	The authors addressed most points raised. Spearman correlation is inappropriate in assessing the relationship between age and infections. Also, abbreviations in Tables and figures should be spelled out.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1
The authors addressed most points raised. Spearman correlation is inappropriate in assessing the relationship between age and infections.
In addition to Spearman correlation multiple regression analysis was performed.

Also, abbreviations in Tables and figures should be spelled out.
The spelling of the abbreviations was added to the Table and Figure letters

The authors addressed most points raised. Spearman correlation is inappropriate in assessing the relationship between age and infections.

Also, abbreviations in Tables and figures should be spelled out.