Author's response to reviews

Title: Data-driven identification of co-morbidities associated with rheumatoid arthritis in a large US health plan claims database

Authors:

Hans Petri (hans.petri@roche.com)
Debra Maldonato (debra.maldonato@roche.com)
Noah J Robinson (jamie.robinson@roche.com)

Version: 3 **Date:** 6 September 2010

Author's response to reviews: see over

Response to comments of reviewers of second version of of the manuscript 'Data-driven identification of co-morbidities associated with rheumatoid arthritis..' Hans Petri 31-8-2010
(also on behalf of co-authors)

Reviewer Dr Vlad:

The reviewer had no further comments and stated that the authers have adequately addressed the issues raised.

Reviewer Dr Treharne:

(reply to remaining issues with questions 3 and 7):

Q3
How were data MarketScan de-identified. Are patients aware their data are included in MarketScan database; are they able their data to be removed from the database?

i Patients from the various health plans are not specifically informed about inclusion of their data in MarketScan; however when joining the health plan they sign a waiver allowing their anonymous data to be used for research and administrative purposes.
ii Patients have legal right to ask their data are removed from MarketScan. Individual is to be notified in case of any breach of confidentiality.

An overview of de-identifaction is provided in reference 11 of the text [Adamson et al]. The steps went beyond HIPAA requirements for limited-use datasets; for example three-digit ZIP codes for employees and providers are used in the the databases, instead more detailed ones.

The statistical analysis by a third party to verify that they met HIPAA requirements for fully de-identified data sets was carried out by Dr Dan Barth-Jones of Columbia University.

Q7
Comment referring to lack of internal validation of using relative risks as a tool to quickly identify comorbidities:

We have followed the suggestion of other reviewer Dr Agarwal (Q4) to use a different control group and to do analysis again, comparing RRs and ORs with first original analysis (information in next section).

Reviewer Dr Agarwal:

(replying to remarks/questions 1-5)

1. Issues with quality of data and external validity.

We acknowledge that the data have limitations, because of the nature of claims with their administrative (payment) focus, instead of a clinical or scientific one. These limitations have been described in the last paragraphs of the manuscript. We believe that if these limitations are kept in mind the data can be used for the purpose descibed: quickly establishing a co-morbidity profile of

patients with a condition of interest in a very large population.

2. Type of study: case-control or cohort design.

MarketScan is a longitudinal database and selection of data was by disease status (RA/ Control condition), hence the cohort approach. RRs for a large number of events (outcomes) have been calculated for a one-year follow-up period. 'Cases' in the text refer to the RA group, not to outcome/co-morbidity (of which there are many).

3. Relative Risk versus Odds Ratio

We have added calculations of Odds Ratio together with the Relative Risk values that were provided earlier in Table 2. The comparison has been uploaded as Additional file 2. Differences between ORs and RRs are small.

4. Choice of controls

The comparator group consists of patients with at least two claims for dermatitis/eczema. This diagnosis was chosen as it is a (relatively minor) chronic condition. Advantage should be that such a control group should have similar likelihood to have comorbidities detected as RA group (both RA and comparator groups have contact with health care).
As requested, we have added an analysis of an additional (random) control group (uploaded as Additional file 3), with an overview of rank-ordered RRs and ORs.
When comparing with the results for the dermatitis/eczema control group (Table 2, Additional file2) largely the same comorbities end up in the top-30 list. RRs/ORs are however higher, which could be explained by the contrast beween a disease (rheumatoid arhtritis) and a random comparator group, with less contact with health care compared to the other control group (dermatitis/eczema).

Descriptions of the Additional files 2 and 3 have been added/tracked (only changes in the body of the manuscript)