

Reviewer's report

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

Version: 1 **Date:** 9 March 2010

Reviewer: Gareth John Treharne

Reviewer's report:

Major Compulsory Revisions

1. The factual claims made in the first paragraph of the Background are not referenced; for example, the claim that "in practice often adverse events are reported that were not foreseen." is not provided with an empirical foundation.
2. No coverage of existing studies of comorbidities of rheumatoid arthritis and the 'control' condition, eczema/dermatitis, is provided in the Background.
3. The way in which participants consented for their data to be used is not described. Specifically, is consent to be in the MarketScan database run on an opt-in or an opt-out basis? Were participants required to be in the database in order to receive insurance coverage for their treatment?
4. The fact that the data were not originally collated for research purposes leads to a number of flaws in the data, including the inclusion of overlapping codes (e.g., elevated sedimentation rate) as well as the absence of important data, notably treatment information, ethnicity, income etc.
5. Given the large sample, I would expect to see a boot-strapping approach taken to ensure reliability of the analyses.
6. The Discussion overlooks the likelihood that current clinical knowledge about comorbidities of rheumatoid arthritis or eczema/dermatitis is likely to drive the specific comorbidities that clinicians test for and deem it worth recording.
7. The Discussion and Conclusion refer to rank ordering of relative risk of comorbidities as a 'tool' to quickly identify comorbidities, but no comparison assessment of comorbidity prevalence was included (or referenced) hence the reliability and validity of this 'tool' is uncertain.

Minor Essential Revisions

8. The sentence "It is usually a challenge to quantify the association of less common but serious events with the target disease." in the Background section of the Abstract is unclear.
9. The phrase "Adverse effect to medicine/biological" in the Results section of the Abstract is unclear.
10. The meaning of the abbreviation FDA is not provided (p. 4).
11. The meaning of the terms 'capitated', 'Medicare' and 'Medicaid' are not

explained (p. 5).

Level of interest: An article of limited interest

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests.