

Reviewer's report

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

Version: 2 **Date:** 23 April 2010

Reviewer: Gareth John Treharne

Reviewer's report:

With reference to the original numbering of my comments:

3. Can the authors explain how and when the linked data were de-identified and explicitly clarify whether or not patients:

- i. are aware their data are included in the MarketScan database;
- ii. had any ability to request their data be removed from the MarketScan database.

Can they also cite who carried out the "statistical analysis by a third party to verify that they met HIPAA requirements for fully de-identified data sets." (p. 6).

7. My original comment referred to the lack of internal validation of using relative risks as "a tool to quickly identify comorbidities" (p. 11) rather than the external validation of the findings in comparison to other studies. Internal validation would be established by comparing other assessments of comorbidity in the same sample, and is ideally required before conclusions about the robustness of a tool are made.

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.