Reviewer's report

Title: A method for measuring productivity and encouraging the practice of research in hospitals: development and feasibility

Version: 2
Date: 20 May 2015

Reviewer: Iain Moppett

Reviewer's report:

Please number your comments and divide them into

The authors have piloted an array of metrics to attempt to provide some quantification of research activity. They found that it seemed to have some discriminant validity (specialized medicine outperformed other areas). This is an increasingly important topic as funders (direct and indirect) seek value for money in those they support. The authors are to be commended for trying to integrate various metrics of time and output.

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.

Essentially this is a paper describing the development of a measurement tool. It therefore falls into the domain of psychometrics – does the tool measure reliably some underlying construct that is important? I fully appreciate that with any new tool there is a balance between publishing early to let others see what is happening and use it but without extensive validation, and conversely producing a vast amount of robust internal and external validation but at the risk of never publishing anything. At the moment I feel the authors have erred slightly too far towards early publication. I would suggest that the authors take the principles of psychometrics / tool validation and apply (or at least discuss their application) to their tool. This might usefully be objective three.

The effect of this will be that the authors will include some (new) information about reliability, construct validity, face validity, discrimination etc. Some of this they already have but it is implied rather than up front. For instance, they have shown that their tool highlights specialized medicine as being research heavy – is this supported by other methods of assessment? (The question of whether this research bias is appropriate is for debate elsewhere).

The authors describe use of the JCR as a source for normalized impact factor. I have just been onto JCR and can’t find this; Scopus does something along these lines. So at present I am unable to replicate the study method.

How much intra-individual variation is there? Grants and papers are episodic so
any single year may be good or bad for any single individual.

The authors describe a literature search. Can they please provide details.

Minor Essential Revisions

The author can be trusted to make these. For example, missing labels on figures, the wrong use of a term, spelling mistakes.

Line 189: ‘…indicators for each HCP and data analysis was about 9 weeks/person…’ There were 896 individuals – did it really take 8000 weeks to do this study?

Discretionary Revisions

These are recommendations for improvement which the author can choose to ignore. For example clarifications, data that would be useful but not essential.

Level of interest: An article whose findings are important to those with closely related research interests

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

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests’ below