NHMRC grant applications: a comparison of "track record" scores allocated by grant assessors with bibliometric analysis of publications

Marcus B Nicol, Kumara Henadeera and Linda Butler

ABSTRACT

Objectives: To investigate the correlation between the publication "track record" score of applicants for National Health and Medical Research Council (NHMRC) *project grants* and bibliometric measures of the same publication output; and to compare the publication outputs of recipients of NHMRC *program grants* with those of recipients under other NHMRC grant schemes.

Design: For a 15% random sample of 2000 and 2001 project grant applications, applicants' publication track record scores (assigned by grant assessors) were compared with bibliometric data relating to publications issued in the previous 6 years. Bibliometric measures included total publications, total citations, and citations per publication. The program grants scheme underwent a major revision in 2001 to better support broadly based collaborative research programs. For all successful 2001 and 2002 program grant applications, a citation analysis was undertaken, and the results were compared with citation data on NHMRC grant recipients from other funding schemes.

Main outcome measure: Correlation between publication track record scores and bibliometric indicators.

Results: The correlation between mean project-grant track record scores and all bibliometric indicators was poor and below statistically significant levels. Recipients of program grants had a strong citation record compared with recipients under other NHMRC funding schemes.

Conclusion: The poor correlation between track record scores and bibliometric measures for project grant applications suggests that factors other than publication history may influence the assignment of track record scores.

MJA 2007; 187: 348–352

Australia's National Health and Medical Research Council (NHMRC) provides research funding to individuals and groups through a variety of mechanisms. Until recently, the majority of researchers were funded by *project grants*, which provide support for individuals and small teams of researchers undertaking biomedical, clinical, public health or health services research. These grants are generally of 3 years' duration, and typically 20%–25% of applications are successful in obtaining funding.

In assessing project grant applications, the NHMRC uses a system of anonymous peer review, with assessors' scores providing a guide to committees in priority ranking of all applications, which effectively determines which applications are funded. One part of the assessment is the allocation of a "track record" score based on the research publication output of the project's investigators during the preceding 6 years (Box 1).

In 2001, the NHMRC initiated a revised *program grants* scheme. The scheme aims to provide support for research teams to pursue broadly based collaborative activity, and grants are typically of 5 years' duration. Inter alia, the teams are expected to contribute knowledge at a leading international level and tackle problems for which longer-term stable funding is essential. In 2001, 60% of the program grant assessment was based on the record of research achievement, with 35% of the total score relating to the applicants' publications (Box 1).

1 Assessment of grants[1]

Project grants

Project grant applications are assessed by an independent review panel consisting of three to seven expert scientists. Assessors give a rating of between 1 ("poor") and 7 ("outstanding" [ie, "in top 5% internationally"]) to each of four criteria: significance, approach, feasibility and publication track record. The track record score accounts for 25% of the assessment.

In judging applicants' track record, assessors are directed to look at the quality of publications and the standing of the journals in which the applicants published in the 6 years prior to submission of the grant application.

Program grants

The Program Grants Committee initially scores all applications against the assessment criteria, and the least competitive are removed from further consideration. Remaining applications are sent to external assessors, and applicants are interviewed by a program-grant review panel.

The criteria used to assess proposals cover three main elements, with a maximum possible score of 100: (i) record of research achievement (total score, 60, of which 35 points are for academic recognition through publications); (ii) research plan (total score, 20); and (iii) collaborative gain (total score, 20). ◆

The primary aim of our study was to examine the track record score given to applicants for project grants and to compare this with bibliometric analysis of the publications on which that assessment was based. A secondary aim was to compare the citation impact of publications from program grants with the impact of publications from other NHMRC grant schemes.

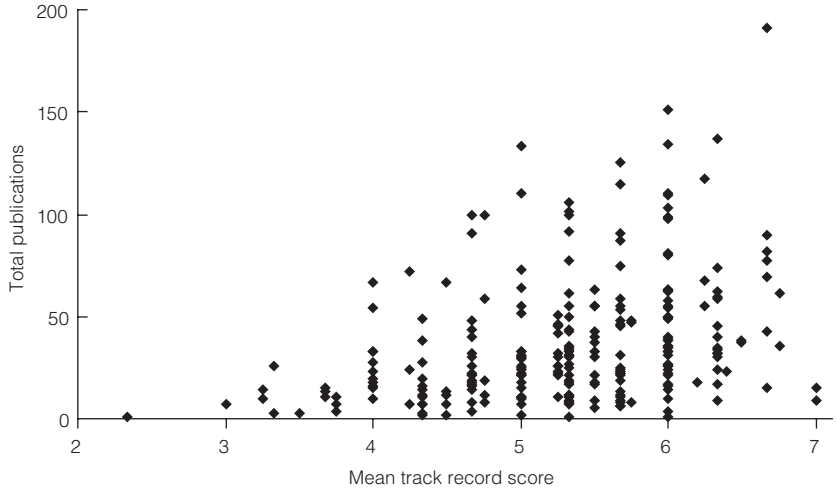
METHODS

Data sources

Bibliometric data were extracted from the Research Evaluation and Policy Project database, which contains all publications with an Australian address in the three major Thomson ISI (Institute for Scientific Information) citation indexes. The database also contains the yearly counts of citations to these publications.

Information on successful program grant applications in the 2001 and 2002 rounds of

2 Comparison of mean track record scores with total publications for project grant applications in 2001*



* Each data point represents the total publications (ie, total count of individual publications published by all authors on each grant application during the previous 6 years) for each project grant application plotted against the mean track record score given to each application by reviewers. The correlation coefficient for the relationship between the two measures was 0.375. ◆

the scheme, and on all project grant applications in 2000 and 2001, were obtained from the NHMRC. Details included the names and institutional affiliations of all investigators.

For each project grant application, we acquired information on the applicants' success and the track record scores given by each assessor (the number of assessors varied between three and seven). In addition, we collected information on the discipline panel that reviewed the application, as the NHMRC uses separate grant review panels for different scientific disciplines. We also calculated the mean, median and standard deviation of the track record scores for each application.

We did a full analysis of the program grants cohort, covering 264 investigators associated with 32 grants. However, as there were nearly 3200 project grant applications over the 2-year period, it was necessary to sample the data, as the intensive manual nature of the publication identification process made full coverage impractical. A 10% sample was selected by choosing a random number between 1 and 10, then selecting every 10th application from all 3200 applications. This sample size was later increased to 15% (see Results). The analysis covered 274 project grant applications from 2000 and 254 from 2001, involving 1340 investigators.

Identifying publications

The first step was to identify the publications that formed the basis on which assessors made their judgements. In the case of project grants, this referred to articles published by the investigators in the 6-year period preceding the grant application. For investigators listed on successful program grant applications, we restricted our publication coverage to articles published in the 5-year period 1996–2000, to make it directly comparable with a 2003 study of the publication impact of NHMRC-funded publications.[2]

The Research Evaluation and Policy Project database was initially interrogated using authors' names and institutional addresses. Where there was doubt about the relevance of a particular publication (either as a result of authors relocating to different institutions or of multiple occurrences of a common name), extra searches were performed using publication and journal titles. Duplications arising from publications being linked to one grant more than once, because of previous collaborations between investigators, were removed. The total number of publications identified was 3306 for program grants, 7435 for project grants in 2000, and 8090 for project grants in 2001.

For several reasons, we were unable to identify the publications of 93 authors: some were based overseas before the grant application; some had no ISI publications within the relevant period; and a few with common names proved impossible to identify. In a very small number of cases this resulted in a grant application being deleted from the analysis.

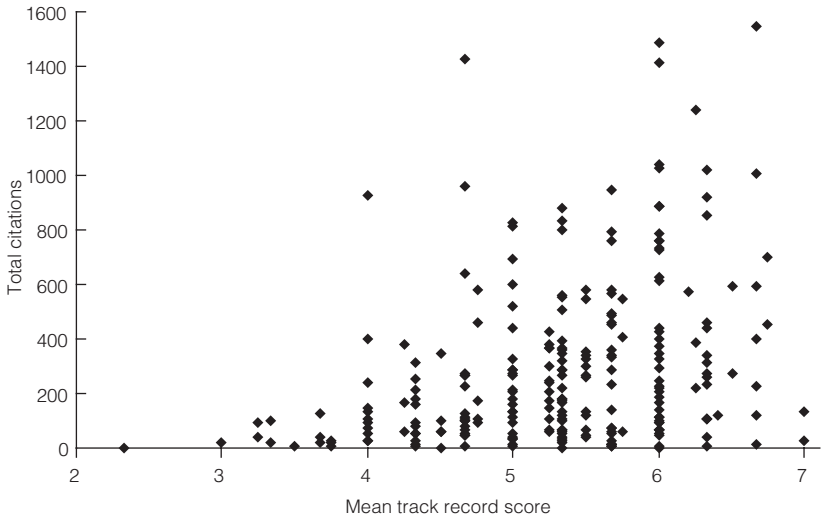
Citation analyses were undertaken on the final publication sets. For project grants, we then compared the bibliometric measures with assessors' track record scores to determine the extent of the relationship. A maximum correlation coefficient of "1" indicated a perfectly linear relationship between the two variables, while a coefficient of "0" indicated no relationship at all. In the case of program grants, the results of citation analysis were compared with data reported in the 2003 bibliometric study.[2]

Bibliometric measures — project grants

The bibliometric measures chosen for our study were those judged most closely related to the track record criteria: number of publications (productivity), quality of publications (citation impact), and "standing" of the journals in which the applicants published (journal impact). Seven measures were calculated for each grant:

- *Total publications.* This was the total number of ISI-indexed articles published by all investigators over the relevant 6-year period.
- *Total citations.* This was the total number of all citations to the applicants' articles received during the same period.
- *Total journal impact.* This is the sum of the average citation rates for all journals in which the applicants' articles appeared. The ISI journal impact factor is commonly used to assess the prestige of a journal, but it suffers from a number of methodological problems.[3] The measure we used is more robust, as it is based on a longer time frame — the same period covered by our analysis.
- *Citations per publication.* To allow for differences in the number of researchers listed on grant applications, total citations were size-adjusted by calculating an average per publication.
- *Average journal impact.* As for the previous measure, a size-adjusted figure was calculated to arrive at an average citation rate for journals in which the applicants' articles appeared.
- *Field-normalised citations per publication.* To ensure our results were not affected by field-specific citation characteristics, we calculated citation rates adjusted for the average world rate in the discipline.
- *Field-normalised average journal impact.* As for the previous measure, we calculated field-normalised journal impact data.

3 Comparison of mean track record scores with total citations for project grant applications in 2001*



* Each data point represents the total citations to all the publications published by all authors on each grant application during the previous 6 years plotted against the mean track record score assigned to each application by reviewers. The correlation coefficient for the relationship between the two measures was 0.327. ◆

Bibliometric measures — program grants

We used a single bibliometric measure — comparison of actual and expected citation rates — to summarise our analysis of program grants. This allowed us to directly compare our results with the analysis contained in the 2003 report.[2]

The *actual* citation rate is the average number of citations received by the publications being analysed. The *expected* citation rate refers to the average number of citations received by publications similar to those being assessed.

RESULTS

Project grants

The initial correlations were carried out between mean track record scores and two simple bibliometric measures — total publications and total citations. The correlations were undertaken separately for each cohort, as the publication period (and hence the citation period) differed, and we sought to remove this possible source of data "noise". The correlation coefficients for the 2000 data were 0.389 for total publications and 0.430 for total citations; the coefficients for the 2001 data were 0.375 and 0.327, respectively. Scatter plots of the 2001 data are presented in Box 2 and Box 3. These plots show that a large number of grant applications with low publication and/or citation counts had been given high track record scores (ie, > 5). These unexpected results led us to increase our initial sample from 10% to 15%, but, even with a larger sample, the results remained unchanged.

When we extended our analysis by correlating track record scores with the full range of bibliometric indicators, none produced any increase in correlation: coefficients ranged from 0.050 (for 2001 scores related to average journal impact) to 0.407 (for 2000 scores related to total journal impact).

In attempting to identify any underlying causes for the poor correlation between track record scores and bibliometric measures, we compared successful and unsuccessful grants and looked at the level of agreement between assessors (as indicated by the SD of the assessors' scores). Nearly all bibliometric variables remained weakly correlated, if at all, with the track record scores, and no correlations were statistically significant. The data from individual panels were also examined. Correlation coefficients based on four bibliometric measures for the 2001 cohort are shown in Box 4. This analysis was limited to the five panels for which robust publication counts existed.

There were considerable differences in the results across panels. High correlations were apparent for only two panels: for the immunology panel, there were strong correlations across all measures; for the endocrinology/reproduction panel, it was the publication and citation counts, unadjusted for size, that showed the strongest correlations. For the microbiology and public health panels, correlations were either extremely low or completely absent (Box 4).

We undertook further analysis to examine in detail the outliers depicted in Box 2 and Box 3. We investigated applications for which assessors had given a score of 6 or more, but for which we found < 50 publications and/or < 500 citations. We also examined applications that had been given scores of less than 5, but were above the benchmarks of 50 publications and/or 500 citations. This investigation shed little further light on the reasons for low correlations.

Program grants

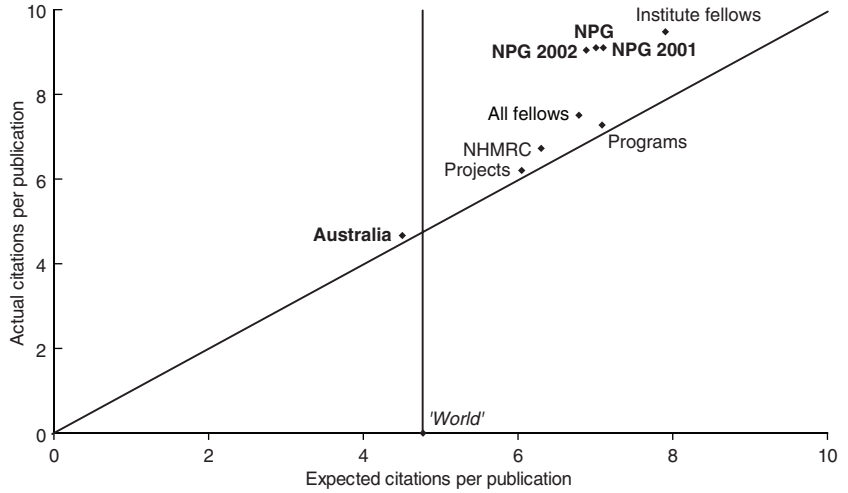
An analysis of actual and expected citation rates relating to publications by successful program grant applicants is shown in Box 5. Results for the two cohorts (2001 and 2002), shown separately and in aggregate, are compared with results for other schemes reported in 2003.[2]

The program grants scheme resulted in funding of research groups with particularly strong publication track records. They had a much higher impact than researchers in the pre-2001 program grants scheme, and outperformed all other groups, with the exception of research fellows located in block-

4 Correlation coefficients of bibliometric measures and track record scores, selected 2001 project grant review panels

Panel (discipline)	Total publications	Total citations	Citations per publication	Total journal impact
Immunology	0.593	0.772	0.627	0.713
Microbiology	0.081	0.026	0.018	0.113
Inflammation	0.305	0.508	0.358	0.361
Endocrinology/reproduction	0.686	0.608	0.334	0.466
Public health	−0.090	−0.030	0.088	−0.049
All 2001 applications	*0.375*	*0.327*	*0.110*	*0.335*

5 Relationship between actual and expected citation rates for publications by successful program grant applicants in 2001 and 2002, compared with results for other NHMRC grant schemes reported in 2003[2]*†



NHMRC = National Health and Medical Research Council.

* The "actual" citation rate is the average number of citations received by the publications being analysed. The "expected" citation rate is the average number of citations achieved by publications similar to those being assessed. The diagonal on the graph indicates the point at which the two rates converge; the vertical line indicates the world citations-per-publication rate in the medical and health sciences. A point above the diagonal and to the right of the world average indicates a citation performance above the world average in journals with high expected rates of citation.

† On the graph, "All fellows" indicates all publications from all fellowship grant recipients, including institute fellows; "Australia" indicates all Australian publications in the field; "Institute fellows" indicates all publications from NHMRC fellows in the former block-funded medical research institutes; "NHMRC" indicates all publications supported by NHMRC grants; "NPG" indicates new program grants (2001 and 2002 application rounds combined); "NPG 2001" indicates new program grants (2001 application round only); "NPG 2002" indicates new program grants (2002 application round only). ◆

funded institutes. There was modest overlap (16%) between the program-grants and institute groups, as a number of the new program grants went to researchers from the institutes.

DISCUSSION

In analysing project grants, we anticipated strong correlation between track record scores and bibliometric measures, as other studies have shown strong correlation between peer assessment and bibliometric analysis, even when the assessment took into account factors beyond the body of published research.[4,5] We expected that high track record scores would be primarily associated with grants with high publication and citation counts, but our results did not reflect this.

Studies such as those by Oppenheim[4] and Aksnes and Taxt[6] have shown much stronger correlations between bibliometric indicators and peer review rankings, with coefficients of 0.7 or better. Yet the rankings to which they were relating their measures were generally based on a much wider remit — the "quality" of the units of assessment — rather than the much more specific focus of the track record assessments we were using. As our bibliometric indicators were direct measures of the published criteria for track record scores, we expected the correlations in our study to be even stronger.

The considerable differences in results across panels may in part explain the poor level of correlation. For example, ISI citation index coverage of the publication output in the area of public health is relatively poor, and much of the output is found in other formats.[7] Weaker correlations were therefore expected for this discipline — although not the complete absence of association that we found. On the other hand, the lack of correlation in the data for the microbiology panel was unexpected and counterintuitive. As journals in this discipline are comprehensively covered by ISI indexes, bibliometric data should have correlated strongly with the scores based on the selection criteria (Box 1). Differences in ISI coverage between different grant review panels does not provide the complete answer to the poor correlations. This result raises the possibility that assessors deviated from the scoring criteria in providing track record scores.

In contrast to the perplexing outcomes of our analysis of project grants, the results for program grants were in line with our expectations. Previous studies of NHMRC-supported research[2,8] have shown that the block-funded institutes, and research fellows located in these institutes, have a citation impact well above that of other NHMRC funding schemes and other Australian research sectors. Thus, given the standing of researchers targeted by the program grants scheme and the substantial weight given to publications in the assessment criteria, we anticipated that successful applicants would have a very strong citation record. Our results confirmed this.

As the track record score is only a single component of a much larger peer review process for project grant applications, the identified lack of correlation between track record scores and bibliometric measures in project grant applications cannot be used to question the validity of the final outcomes of the application process. The assessors do appear to "get it right": the 2003 bibliometric study of NHMRC-funded research found that research projects funded by the NHMRC performed at a much higher level than those undertaken without NHMRC support, and their performance was above world and Australian benchmarks.[2]

In a study of grant proposals to the US National Science Foundation, Abrams identified two possible reasons for similar low correlations.[9] He suggested that "the ability to produce a highly-rated proposal inherently has little correlation with the ability to carry out and publish high quality research". He also suggested that the limited time scientists can devote to evaluating proposals can introduce considerable uncertainty into the process.

We have also received anecdotal evidence that, rather than rating "significance", "approach", "feasibility" and "track record" independently, assessors may judge an application as a whole, decide whether it should be funded, and give scores to each element that they believe will make it successful. With the pressure on researchers' time, and the increasing calls on them in peer review scenarios, it is not surprising that shortcuts may be taken, as Abrams suggested.

What our results appear to highlight is a lack of transparency in the process. While it would be unrealistic to expect perfect corre-

lation, it is clear that there should have been a much closer relationship between our project grant data and the scores. In contrast, the approach adopted for program grants does lead to the expected close relationship.

Perhaps now is the time to develop a more automated system of track record assessment. Why ask peers to assess track records from scratch, when there are defensible surrogates for this aspect of the grant application? Surely their scarce time is best reserved for where it is most useful, and where no alternative is possible — assessing the significance, approach and feasibility of applications. They could be relieved of the burden of assessing track record, only delving into it in the relatively few cases in which there are concerns about the automatically generated scores. Concerns about the use of such measures, raised recently in an article by Lehmann et al,[10] related not to the measures themselves, but to their potential "harmful misuse". Bibliometrics has progressed significantly in recent years, and measures are now available that are sensitive to field-specific characteristics and the concerns of researchers who are at an early stage of their careers.

ACKNOWLEDGEMENTS

Our thanks go to Roland Wise and David Porter at the NHMRC for their cheerful assistance in extracting and providing data from NHMRC records, and to Tim Brown for his advice with the statistical analysis. Certain data included here are derived from the Australian National Citation Report prepared by the Institute for Scientific Information, Philadelphia, Pa, USA. (Copyright ISI, 2000. All rights reserved.)

COMPETING INTERESTS

Marcus Nicol has been a consultant for the NHMRC for the past 4 years. The raw data collected for our article was part of a previous consultancy contract; however, the design, analysis and drafting of the article was *not* part of any paid consultancy with the NHMRC, and was done purely for academic interest.

AUTHOR DETAILS

Marcus B Nicol, BSc, MPH, PhD, AVERT (A Very Early Rehabilitation Trial) Data Manager[1]
Kumara Henadeera, PhD, Research Assistant[2]
Linda Butler, BEcon, Head[2]

1 Clinical Trials, National Stroke Research Institute, Melbourne, VIC.
2 Research Evaluation and Policy Project, Australian National University, Canberra, ACT.

Correspondence:
marcusnicol@netspace.net.au

REFERENCES

1 National Health and Medical Research Council. Record of Research Achievement (RORA) — qualitative grid. http://www.nhmrc.gov.au/funding/apply/granttype/programs/_files/rora_grid.xls (accessed Jun 2007).
2 Butler L. NHMRC-supported research: the impact of journal publication output. Canberra: National Health and Medical Research Council, 2003: 75. http://www.nhmrc.gov.au/publications/synopses/_files/butler03.pdf (accessed Jun 2007).
3 Moed H. The impact-factors debate: the ISI's uses and limits. *Nature* 2002; 415: 731-732.
4 Oppenheim C. The correlation between citation counts and the 1992 research assessment exercise ratings for British research in genetics, anatomy and archaeology. *J Doc* 1997; 53: 477-487.
5 Bornmann L, Daniel H. Reliability, fairness and predictive validity of committee peer review. *BIF Futura* 2004; 19: 7-19. http://www.bifonds.de/public/news/bornmann_e.pdf (accessed Jul 2007).
6 Aksnes DW, Taxt RE. Peer review and bibliometric indicators: a comparative study at a Norwegian university. *Res Eval* 2004; 13: 33-41.
7 Butler L, Biglia B, Henadeera K. NHMRC-supported research: the impact of journal publication output 1999–2003. Canberra: National Health and Medical Research Council, 2006. http://www.nhmrc.gov.au/publications/synopses/_files/nh75.pdf (accessed Jul 2007).
8 Butler L, Biglia B. Analysing the journal output of NHMRC research grants schemes. Canberra: National Health and Medical Research Council, 2001. http://www.nhmrc.gov.au/publications/synopses/_files/r21.pdf (accessed Jul 2007).
9 Abrams PA. The predictive ability of peer review of grant proposals: the case of ecology and the US National Science Foundation. *Soc Stud Sci* 1991; 21: 111-132.
10 Lehmann S, Jackson D, Lautrup B. Measures for measures. *Nature* 2006; 444: 1003-1004.

(Received 9 Jan 2007, accepted 19 Jun 2007) ❑