ELSEVIER

Journal of Clinical Epidemiology ■ (2015) ■

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

A randomized trial of fellowships for early career researchers finds a high reliability in funding decisions

Philip Clarke[a], Danielle Herbert[b,c], Nick Graves[b,c], Adrian G. Barnett[b,c,*]

[a]*Melbourne School of Population and Global Health, The University of Melbourne, 207 Bouverie Street, Melbourne, Victoria, 3010, Australia*
[b]*School of Public Health and Social Work, Queensland University of Technology, Victoria Park Road, Brisbane, Queensland, 4059, Australia*
[c]*Institute of Health and Biomedical Innovation, Queensland University of Technology, 60 Musk Avenue, Kelvin Grove, Brisbane, Queensland 4059, Australia*

Accepted 24 April 2015; Published online xxxx

Abstract

Objectives: Funding for early career researchers in Australia's largest medical research funding scheme is determined by a competitive peer-review process using a panel of four reviewers. The purpose of this experiment was to appraise the reliability of funding by duplicating applications that were considered by separate grant review panels.

Study Design and Methods: Sixty duplicate applications were considered by two independent grant review panels that were awarding funding for Australia's National Health and Medical Research Council. Panel members were blinded to which applications were included in the experiment and to whether it was the original or duplicate application. Scores were compared across panels using Bland−Altman plots to determine measures of agreement, including whether agreement would have impacted on actual funding.

Results: Twenty-three percent of the applicants were funded by both panels and 60 percent were not funded by both, giving an overall agreement of 83 percent [95% confidence interval (CI): 73%, 92%]. The chance-adjusted agreement was 0.75 (95% CI: 0.58, 0.92).

Conclusion: There was a comparatively high level of agreement when compared with other types of funding schemes. Further experimental research could be used to determine if this higher agreement is due to nature of the application, the composition of the assessment panel, or the characteristics of the applicants.

Keywords: Funding; Peer review; Reliability; Research; Fellowship; Career

1. Introduction

Research funding is most commonly allocated through a competitive process in which applicants are assessed by a panel of their peers using predefined criteria. Surprisingly, the reliability of grant applicant assessment processes has only rarely been examined. In 1977, the US National Science Foundation rereviewed 150 proposals using a second independent peer-review panel and found a 24−30% disagreement in funding outcomes [1]. A Canadian study of 248 proposals submitted to two major funding agencies with similar peer-review processes found a 27% disagreement in funding [2]. In 2009, the Academy of Finland randomly assigned peer reviewers to two panels assessing the same 65 proposals and found a 31−35% disagreement [3].

A review as part of the Cochrane collaboration examining studies of the allocation of grant funding concluded that: "experimental studies assessing the effects of grant giving peer review on importance, relevance, usefulness, soundness of methods, soundness of ethics, completeness and accuracy of funded research are urgently needed." [4] To date, all experimental studies assessing reliability of peer-review panels in allocating research grants have examined the allocation of project funding and not people funding such as awards and fellowships.

The purpose of this study was to assess the interpanel reliability in assessing applicants for fellowship support. It uses an experimental design of applications for an Australian scheme for Early Career Fellowships, administered by Australia's National Health and Medical Research Council (NHMRC) [5]. The experiment involved duplicating a randomly selected subset of applications that were

Conflict of interest: None.

Funding: The work was funded by the National Health and Medical Research Council (NHMRC) (Project Grant 1023735). The NHMRC contributed to the study design and performed the data collection. The NHMRC had no role in the analysis, interpretation of the data, writing of the article, or decision to submit for publication.

* Corresponding author. Tel.: +61731386010; fax: +61731386030.
E-mail address: a.barnett@qut.edu.au (A.G. Barnett).

http://dx.doi.org/10.1016/j.jclinepi.2015.04.010
0895-4356/

What is new?

- Independent panels of peer reviewers judging what applicants deserved fellowship funding had a relatively high level of agreement.
- Billions of dollars are awarded through competitive peer review every year, but there are only a handful of studies examining the reliability of decisions.
- The greater reliability of fellowship compared with project funding gives support for awarding more research dollars to people instead of projects.

assessed by two independent review panels. The panel scores were compared to determine the reliability of the scoring and the agreement in the awarding of funding.

2. Materials and methods

The NHMRC is the main government funding body for health and medical research in Australia [6]. In addition to funding for research projects, the NHMRC also funds individual researchers including early career postdoctoral researchers who can apply up to 2 years after the awarding of their PhDs [5]. In 2013, the scheme funded 123 (24%) of the 523 applications received at total cost of $21 million (Australian dollars). Applicants are assessed for funding by multiple panels, each consisting of four senior Australian health and medical researchers with a background in either biomedical, clinical, public health, or health services.

Eight independent panels assess 30−40 applications that are split by field. All applications are independently scored by all panel reviewers on three domains: (1) personal achievement that includes information on undergraduate track record, a supervisor's report, and information on prizes and conference presentations, and this contributes 35 percent to the score; (2) project that involves a research plan and description of the institutional environment, 25 percent; (3) research output in terms of quality and quantity including publications, patents, and influence on policy and research funding, 40 percent [7]. Next, the secretariat of NHMRC calculates an average score that determines a ranking of applicants from highest to lowest, and this is then reviewed in a teleconference of all panel members. The NHMRC rules require that panel members sign a confidentiality agreement and separate teleconferences were held; hence, the decision of each panel should be independent.

This study was undertaken with the assistance of the secretariat of the NHMRC. In 2013, all applicants for the Early Career Fellowships were invited to participate in an experiment to test the reliability of the assessment procedure. Applicants provided their consent to participate using the online submission system. The NHMRC secretariat randomly selected 60 applications, and these were duplicated so they could be assessed by two different review panels: one assesses the original and the other the duplicate. Sixty applications were the maximum number considered feasible by the secretariat both for internal NHMRC administration and to not over burden panel members. The only alteration to the duplicate was to alter the applicant ID number, and this was assigned to a separate review panel. Each fellowship panel reviewed around five experimental applications, with the rest being applications that were not part of the experiment. Panel members, who also consented to participate in the experiment, were blinded to which applications were included in the experiment. The study was approved by the Queensland University of Technology (QUT) Human Research Ethics Committee (number 1200000547).

At the end of the experiment, the NHMRC secretariat provided modified final combined scores of both panels, which involved them adding fixed amount to every score that was not disclosed to the research team, to all experimental scores supplied for analysis. This was done to prevent applicants identifying their own scores. The only applicant-specific information provided by the NHMRC to the research team was the applicants' area of research that was one of: Basic Science, Clinical Medicine and Science, Public Health, or Health Services.

The agreement between the two panel scores for the original and duplicate applications was analyzed using a Bland−Altman plot and 95% limits of agreement. This provided a graphical method to plot the difference scores from each panel against the mean for each subject. We also used the percent overall agreement in funding and 95% bootstrap confidence limits. We adjusted for chance agreement using Gwet's statistic, for which 0 means no agreement and 1 means perfect agreement [8]. However, we note that chance is an acknowledged part of the funding process [9], and so, the most useful measure of agreement includes chance.

To compare our results with the previous agreement in funding from studies of project funding, we used a Bayesian approach to estimate the overall agreement probability from previous studies examining agreement. We assumed that the observed number of agreements in each study had a binomial distribution with a common overall probability, which we modeled using a beta distribution (using a noninformative beta(1,1) prior). We tested the hypothesis that the mean agreement from our study was equal to the overall probability from previous studies.

3. Results

Of the 60 applicants involved in the experiment 29 were from Basic Science (48 percent), 17 (28 percent) from Public Health, 9 from Clinical Medicine and Science (15 percent), and 5 from Health Services (8 percent).