Research Evaluation, 28(4), 2019, 313–325
doi: 10.1093/reseval/rvz019
Advance Access Publication Date: 5 August 2019
Article

Measuring scientific impact of fisheries and aquaculture research-for-development projects in South East Asia and the Pacific

Kiros Hiruy[1,2,*], Ray Murphy[2], Tom Lewis[2], William T. White[3] and Steven W. Purcell [4]

[1]Centre for Social Impact (CSI) Swinburne, Faculty of Business and Law, Swinburne University of Technology, PO Box 218 Hawthorn, VIC 3122, Australia, [2]RDS Partners, Hobart, TAS, Australia, [3]CSIRO Oceans & Atmosphere, Hobart, TAS, Australia and [4]National Marine Science Centre, Southern Cross University, Coffs Harbour, NSW, Australia

*Corresponding author. Email: khiruy@swin.edu.au; khiruy@gmail.com.

Abstract

Research-for-development (R4D), as a mode of foreign aid, is a practical way to support developing countries. However, few studies have assessed the scientific impacts of R4D projects. Here, we applied an integrated assessment approach to evaluate the scientific impact of research projects commissioned by the Australian Centre for International Agricultural Research's Fisheries Program in the Asia-Pacific region between 2000 and 2012. We use publication metrics and expert panel reviews to examine trends in scientific outputs among 73 fisheries and aquaculture research projects. 'Among projects, there was considerable variation in publication types and outputs (including academic papers, magazine articles, policy reports, books, and book chapters), and projects produced, on average, 10 publications at a funded-dollar cost of AU$41,000 per publication. Bilateral and multilateral projects tended to produce more refereed journal articles and citations per funded dollar than single-country projects, and publication outputs were poor for certain countries. With the exception of fisheries projects, larger (more highly funded) projects did not produce more journal articles per funded dollar than smaller projects. Project duration had no significant effect on the number of refereed publications, citation rates, or total citations per funded dollar across projects. Aquaculture projects had greater publication impact, per funded dollar than fisheries projects. Beyond the bibliometric measures, qualitative findings indicated that scientific impact was strongly influenced by motivations of project leaders and their institutions. We identified disparate impact performance among industries and countries. These findings could inform future overseas aid investments, policies and strategies. This study offers useful benchmarks for gauging scientific impacts of R4D programs and shows the value of using approaches that go beyond bibliometric measures.

Key words: research for development; research evaluation; scientific impact; bibliometrics; Southeast Asia; Pacific Islands.

1. Introduction

Research-for-Development (R4D) in developing countries constitutes a significant portion of foreign aid budgets of developed nations and philanthropic organizations (White 1992). This is mainly due to the assumption that scientific research can contribute to economic growth, human capital development, the production of new products and technologies, and provide evidence for policy and practice in developing countries (Pearson et al. 2012; DFID 2014; Confraria, Mira and Wang 2017). However, the impact of foreign aid can be highly contextual (McGillivray et al. 2006). Thus, attention has turned to critically evaluating the outcomes from R4D organizations and identifying strategies that better direct foreign aid investment in research (Herdt 2012; Confraria, Mira and Wang 2017). Strategies by funding organizations on the best duration and size (i.e. funding quantum) of projects would benefit from evaluations across projects within industry types, yet few such comparative analyses exist.

R4D impacts can take the form of (1) academic or scientific impact: knowledge and capacity building; and (2) development impact: policy and product development, social, economic, and livelihood impacts (DFID 2014). Gauging the impact of government-funded research is integral to justifying foreign-aid funding and adapting funding strategies to maximize impact. While most studies are concerned with evaluating the economic impact of projects, scientific impact is more loosely defined (Godin and Doré 2004; Bollen et al. 2009). Nonetheless, there is increasing global interest in assessing the scientific impact of publicly funded research, i.e. the outcome of the knowledge produced as a result of research in technological innovation, and its impact on society, the economy and the environment (Godin and Doré 2004; Hicks et al. 2004; Campbell et al. 2010; Pearson et al. 2012; Bertocchi et al. 2015; Wilsdon et al. 2015; Bartlett 2016; Bozeman and Youtie 2017).

Hence, an impetus for this study was to understand trends in the scientific impact of publicly funded research conducted in developing countries. This is important for two reasons. On one hand funding countries seek to account for public funds and understand the academic and development impacts of foreign aid. On the other, as the implications for development can be profound, developing countries are becoming increasingly interested in understanding the contribution of R4D in developing their scientific capabilities (Mazzoleni and Nelson 2005). Here, we report the scientific impact of the Australian Centre for International Agricultural Research (ACIAR) Fisheries Program in the Asia-Pacific region between 2000 and 2012 through bibliometrics measures and assess capacity building impacts of the program through qualitative research. Our analysis does not include broader societal development impacts.

2. Publication evaluation

A principal issue in evaluating 'scientific impact' has been the lack of a widely accepted definition of the term itself. Bollen et al. (2009) concluded that scientific impact is a multi-dimensional construct that cannot be adequately understood or measured by any single indicator, and called for further analysis and discussion on the utility of various metrics of scientific impact. Most scientific research is intended to support, advance, or achieve goals that are extrinsic to science itself (Sarewitz and Pielke 2007). Hence, when evaluating the scientific impact of large research programs integrated assessment models that incorporate both quantitative (e.g. bibliometrics) and qualitative (e.g. peer review) assessment methods are more useful than assessments based either on bibliometrics or peer review alone (Campbell et al. 2010; Wilsdon et al. 2015; Bozeman and Youtie 2017).

Bibliometric indicators are traditionally considered as objective, reliable, and cost-effective measures of published research outputs, and are complementary to informed peer review (Braun, Glänzel Schubert 2006; Bollen et al. 2009; Campbell et al. 2010; Bertocchi et al. 2015; Waltman 2016). Citation rate of published works is one of the most commonly used proxy measures of scientific impact (Braun, Glänzel Schubert 2006; Wouters et al. 2015). While a large number of citation impact indicators exist, most are derivations of a few basic indicators, namely the total and average number of citations of the publications of a research unit (Waltman 2016). However, indicators based on counting highly cited publications are regarded as more robust than those based on total or average citation counts (Wouters et al. 2015). This is because citation distributions tend to be highly skewed, and therefore the total or the average number of citations of a set of publications may be strongly influenced by one or a few highly cited publications ('outliers') (Waltman 2016).

The three most important multidisciplinary databases to derive bibliometric data are Web of Science (WoS), Scopus, and Google Scholar (GS) (Waltman 2016). Both Scopus and WoS cover the mainstream English international scientific journals and are comparable in terms of citation data. Coverage of the natural sciences fields in GS was previously considered poor compared to WoS and Scopus, but that appears no longer to be the case (Wouters et al. 2015; Waltman 2016). GS has the advantage of counting citations beyond ISI indexed journals, such as from international non-English journals and published reports and theses, and so can be considered to represent scientific impact both within and beyond academia (Harzing and Van der Wal 2008). However, GS has less quality control over citation coverage (Wouters et al. 2015; Waltman 2016) and requires more effort to collate bibliometric data than WoS or Scopus (Meho and Yang 2007).

Altmetrics, or 'webometrics', are alternative metrics to measure scientific impact from website usage data. Altmetrics have arisen from the online platforms on which scientific literature is now published, accessed and communicated through social media. Website usage rates provide a prediction of future citation rates (Brody, Harnad and Carr 2006). However, Costas, Zehedi and Wouters (2015) found the correlations weak and surmised that altmetrics do not reflect the same kind of impact as citations. The consensus is that altmetrics cannot presently be used for formal research assessment (Wouters and Costas 2012; Thelwall et al. 2013).

Assessment of the scientific impact of research or academic institutions using bibliometric indicators would usually occur by retrieving papers in which the name of the institution is found in the authors' addresses. However, R4D organizations that fund research would not usually be found in the authors' affiliations on published papers. Therefore, a publication database has to be manually constructed for each project or researcher who received funding from that R4D organization, as practised by Campbell et al. (2010) for a Canadian research funding organization.

While bibliometric indicators are accepted measures for evaluating the output and outcome of a research program, they can be complemented by experts' opinions or peer review (Seglen 1997; Bertocchi et al. 2015). Seglen (1997), for example, prefers scientific impact to be evaluated by a panel of experts with an in-depth understanding of the particular field of research, arguing that bibliometric indicators are temporally biased in the way citation data is collected and calculated. On the other hand, Bertocchi et al. (2015) note that the assessments of scientific impact should not be based on peer review alone, citing potential 'conflicts of interest' and the challenge that peer review poses in establishing a uniform yardstick across disciplines or research areas. The assessment of the scientific impact of an organization by an appointed panel of experts involves making qualitative and quantitative assessments according to established rules (Seglen 1997: 1). However, this necessitates appointing experts with sufficient level of expertise and experience.

3. Measuring the scientific impact of R4D: a case study

R4D organizations strive to inform solutions for socio-economic problems and industry development through scientific research that could influence practice and policy. A formal assessment of the impact of projects funded by R4D organizations is necessary for accountability of the use of public funds, for benchmarking performance, to improve the design of future projects and to set performance improvement targets. The term R4D is used here to describe organizations that catalyze development in developing countries through brokering and funding research partnerships between scientists and their institutions.

The ACIAR operates as an R4D organization under Australia's Foreign Aid Program. ACIAR's Fisheries Research Program, which covers aquaculture and capture fisheries, hereafter called 'the Program', aims to improve the productivity and sustainability of fisheries and aquatic farming systems in partner countries and Australia through international research partnerships. Key goals are to improve food security, improve livelihoods, and better manage natural resources. The geographic focus of the Program is Southeast Asia (Indonesia, Timor Leste, the Philippines, and Mekong countries), Pacific Island countries and Papua New Guinea (PNG). Projects within the Program are facilitated and funded by ACIAR. Research studies of varying duration and funding amount are coordinated through a variety of commissioned Australian research institutions (including universities, state departments, and other research providers) or International Agricultural Research Centres (IARCs).

In the current study, we undertook a retrospective evaluation of the scientific impact of the Program from 2000 to 2012. This assessment complements an economics-focused evaluation of ACIAR projects published previously (Farmer et al. 2013). Our primary objectives were to establish scientific impact benchmarks and to identify trends for improving the design of future projects, and in doing so, identify lessons learned to guide future project/program design, as distinct from auditing past performance. One of our guiding research questions was whether the duration or funding quantum to projects improved the efficiency of project outputs and measures of scientific impact—in essence—do longer, or larger projects give better 'bang for the buck' than shorter or smaller projects?

While the debate continues to underlie the term 'scientific impact' (Bollen et al. 2009), for the purposes of this article, we have adopted it to mean 'impact on the scientific community'. It comprises both the 'publication output and impact' (i.e. uptake of the Program's research by the scientific community) and 'academic impact' (i.e. formal scientific capacity development arising from the Program). We used an integrated assessment approach for impact assessment (Bryman 2008; Patton 2015). This approach integrates results from publication metrics, expert judgement, document analysis, surveys, interviews, and case studies. This study describes the evaluation of scientific impact and presents approaches to measuring the scientific impact of R4D programs using the Program's projects in Southeast Asia and the Pacific countries as a case study.

4. Data and methods

4.1 Data collection methods

4.1.1 *Publication database*

Our analysis focused on the Program projects that commenced and concluded between the beginning of 2000 and the end of 2012. The census date for the inclusion of publication and citation data was 17 March 2016. We considered this as a reasonable citation window to allow sufficient time for most of the relevant data and findings to have been published and made available for our inclusion in the publication and citation analysis (see Waltman 2016). In total, 90 research projects across southeast Asia and the Pacific countries fell within this timeframe. However, our publication and citation analysis accounted for only 73 of the 90 projects[1] for which we were able to obtain data (see Supplementary Table 1). Collectively these 73 projects had received AU$30.4 million in ACIAR funding (See Supplementary Table 1). The number of projects per funding range for these projects is presented in Table 1. The remaining 17 projects with no data were similar to the 73 projects that produced publications, in terms of industry sector and country of research. However, the majority (13 out of 17 projects), were small, short-term research and development activities having an average project duration of slightly above 1 year and average funding of $67,000. These projects were excluded from the analysis as the majority were scoping studies with no forecasted publication outputs.

The Program provided the initial publication database. To verify all types of publications and identify any additional publications for each project, we contacted respective research project leaders. The publication database was further populated using appropriate online search engines for the publications provided. Projects were categorized by funding amount, project leaders, industry sector, commissioned agency, and country/region. Project publications were identified as either a book, chapter in a scholarly book, an article in a scholarly peer-reviewed journal, conference publication (peer-reviewed or non-peer-reviewed), final report, technical report, internal report, magazine article, policy paper, or database.

Both the Elsevier's Scopus database and Thomson Reuters' WoS database were then searched for all scholarly peer-reviewed journal articles listed in the publications database. For articles indexed in the WoS, we recorded the 2014 5-year impact factor of the journal in which the article was published.

To account for the capacity building outputs of the Program, we extracted information on project capacity building activities from project reports. These included educational tours and visits, mentoring activities, workshops and seminars, masterclasses, and fellowships and scholarships. The data were confirmed through interviews and survey with project leaders.

4.1.2 *Surveys, interviews and case studies*

In addition to creating a publications database, we conducted surveys, interviews and case studies to confirm the publication record of projects and to collect further information about scientific impact. Thirty-nine research project leaders, representing 60% of the project leaders between 2000 and the end of 2012, completed a questionnaire survey.

Table 1. Number of projects per project funding range

Project funding range (AU$)	Number of projects
1,000–49,999	9
50,000–99,999	10
100,000–199,999	13
200,000–499,999	20
500,000–999,999	14
1,000,000–1,999,999	7

We further contacted project leaders[2] of the 73 projects for interviews. Of these, 15 project leaders representing 35 of the 73 projects[3] who were willing and available for interviews were interviewed. Project leaders were asked about their view on the scientific and policy impact of their projects, the publication decision-making process and motivation to publish, and barriers within the publication process. We also interviewed 15 partner-country collaborating-institution leaders and researchers from five case study countries, namely: Indonesia, Lao PDR, PNG, Solomon Islands, and Vietnam. These interviews sought to understand the academic/scientific impact: capacity building and knowledge production impacts of the projects from the perspective of the partner country scientists and leaders. Also, five case studies of the Program projects across the five selected countries were carried out. These case studies were selected based on project diversity and representativeness. The case studies provided insights into the academic impact of the Program projects.

4.2 Peer review—the Expert Panel

An Expert Panel was appointed to review the Program's scientific impact. The Expert Panel was made up of three well-published and regarded experts (from Australia, the UK, and Canada) with extensive experience in managing R4D organizations with relevant national and international aquaculture and fisheries research programs. The Expert Panel was involved in assessing both the methodology—whether the publication metrics chosen to measure the publication output and impact of the Program were appropriate for this purpose—and the results obtained.

4.3 Data analysis

To assess the publication output and academic impact of the Program, we used standard publication metrics. These include publication output, citations per publication output, and a field-weighted citation impact (FWCI). Quantitative data analyzes were complemented by, and triangulated with, a qualitative analysis of the survey data and semi-structured interviews and findings of the Expert Panel (Patton 2015). All qualitative data were manually coded by one researcher for consistency. Narrative and deliberative analyzes were used to extract emerging themes related to knowledge production and capacity building (Bryman 2008).

We sought to compare the results of this study with other comparable research and development programs such as the Canadian International Development Agency, and the UK Department for International Development but could not find comparable data.

For each project, we calculated the average number of refereed journal articles per AU$100,000 funded. We also measured the 'normalized citation rate' (NCR)—citations attracted by projects per year per funded dollar (Waltman 2016). The NCR was calculated as the number of citations across each project's total refereed publications, normalized to the recency of projects relative to the census date of data collection. The citation data included only the citations of Scopus-listed journal articles. Reported start date of each project was usually when contracts were signed, and the research activities usually commenced one year afterwards, so 1 year was subtracted from the number of years from project start dates for all projects for consistency. The NCR per AU$100,000 funded to each project was calculated as:

$$\mathrm{NCR} = \frac{\sum C_n}{Y-1} \div \frac{F}{100{,}000}$$

where,

C_n = Total citations of all (n) refereed publications for a project at the census date

Y = Years from the reported project start date to the census date

F = Total project funding

Projects were considered Bi/Multilateral if the research was conducted in more than one country. The number of Scopus citations across all refereed journal articles published from a project was counted only for projects producing refereed journal articles listed in Scopus.

Average citation rates of refereed journal articles produced by projects reflect the 'quality' or 'impact' of the research within the scientific community. Average citation rates were calculated across refereed publications within each project by dividing the total citations at the census date by the number of years since each article was published.

Relationships between the duration of R4D projects and impact metrics and the total funding for R4D projects and impact metrics were tested using non-linear regression modelling. The modelling was performed on data from Aquaculture, Fisheries, and Shared (Aquaculture/Fisheries) projects separately to avoid potential bias and error due to potential pre-existing relationships among the various types of projects. To avoid overfitting and spurious relationships, we tested the best fitting curves against one-parameter and two-parameter equation forms using DataFit software. The best curve by r^2 value was chosen. Only statistically significant relationships were displayed in graphs. Confidence intervals to statistically significant relationships were fitted using SigmaPlot software.

5. Results

5.1 Publication metrics

5.1.1 *Overall publication outputs of the ACIAR Fisheries Program*

The 73 projects included in this study produced 734 publications during the reporting period, equating to an average output of ten publications per project. The most frequent type of publication (28%, n = 208) was scholarly peer-reviewed journal articles (Figure 1). The next most common type was final reports (n = 110; 15%) and non-peer-reviewed conference papers (n = 86; 12%). The number of final reports (110) is more than the number of projects as some projects had several 'mini-projects' for which separate 'final reports' were submitted. 'Other' publications are outputs such as newsletters, brochures, bulletins, technical notes, or maps. With overall funding for the 73 projects amounting to AU$30.4 million, the 734 publications were produced at an average investment of about AU$41,000 per publication. Some of the publication types could be classed as non-scientific, such as policy papers and magazine articles. Many others, such as conference publications, theses, final reports and book chapters, could present scientific data and results such that we could not logically make a simple distinction between publications that were scientific or not.

The number of publications produced per project was highly variable (Figure 2). More than 90% of the publications came from half of the projects (n = 37), with half of all publications coming from eight projects. One project, 'Improved hatchery and growout technology for marine finfish in the Asia-Pacific region' (FIS/2002/077), had the greatest output with 127 total publications (17% of

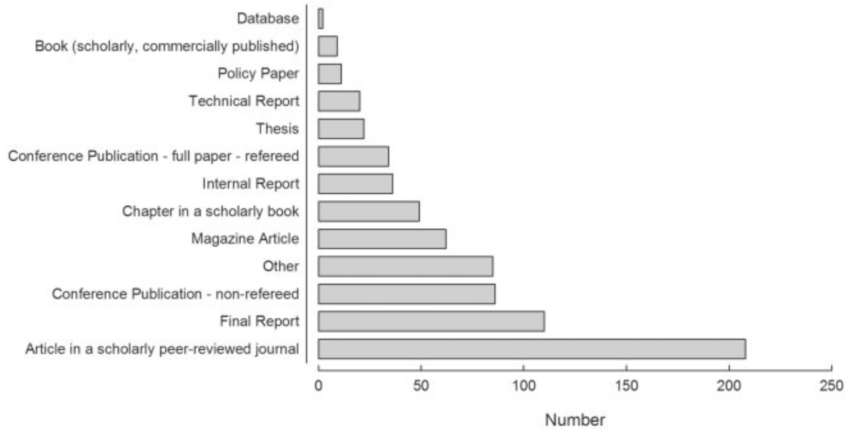

Figure 1. Publication output of the Program by publication type.

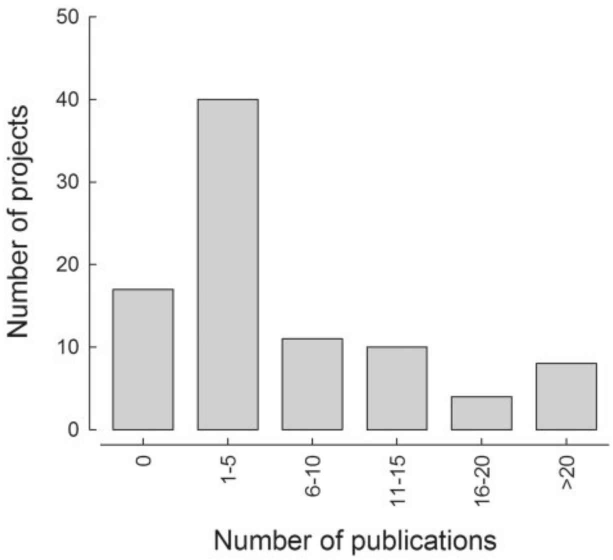

Figure 2. Publication output distribution by project.

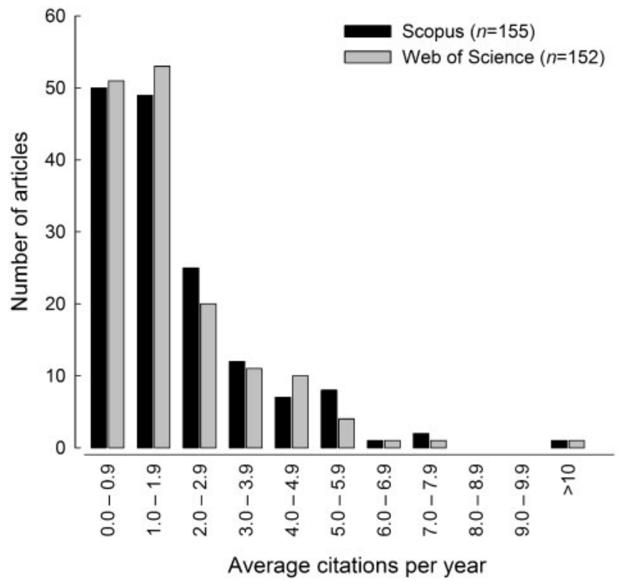

Figure 3. Number of scholarly peer-reviewed journal articles by average citation by year.

all publications). These were: magazine articles (n = 33), chapters in a scholarly book (n = 29), conference publications—non-peer-reviewed (n = 24), journal articles (n = 17), and other publications (n = 24).

5.1.2 *Peer-reviewed journal article outputs*

The 73 projects for which publication data were available produced 208 articles in scholarly peer-reviewed journals during the reporting period. This equates to an average output of 2.8 journal articles per project, representing an average project funding of $146,000 per peer-reviewed scientific paper. In comparing this cost against other projects, one must bear in mind that many of the ACIAR projects employ researchers with the project funds, so this value does not simply reflect the direct costs of research. As in the case of total publication output for the Program, there was an uneven distribution of journal articles per project. Half of the projects (n = 37) produced 98% of the journal articles (204 out of 208) during the census period. Nearly 45% of the projects (n = 33) had not published any journal articles by the census date. Eight projects (10%) produced 56% of the total journal articles (116 of 208 journal articles) for the Program across the review period.

5.1.3 *Citations per publication (journal articles)*

The number of journal articles with citation information in Elsevier's Scopus database (155) was marginally higher than Thomson Reuters' WoS database (152). Around two-thirds of the journal articles were reported to have been cited less than twice per year by both Scopus (64%) and the WoS (68%) (Figure 3). Scopus reported 31 journal articles (from 14 projects) as being cited on average between three and 12 times per year. Out of those 31 articles, 12 were co-authored by one or more partner-country researcher, and partner-country researchers were the lead authors on seven articles.

5.1.4 *Field-weighted citation impact (journal articles)*

FWCI takes into account the differences in research across disciplines and compares the number of citations of publications with

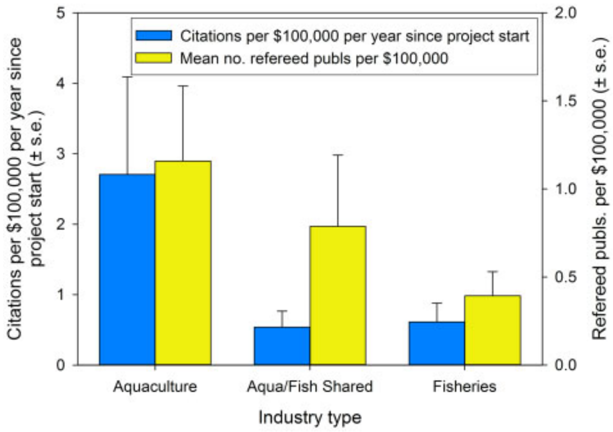

Figure 4. Mean number of Scopus citations per AU$100,000 funded to projects per year since projects started (blue bars, left-hand axis scale) and mean number of refereed journal articles per AU$100,000 funded to projects (yellow or grey bars, right-hand axis scale), with project data grouped by industry type. Error bars are standard errors of the mean. Sample sizes (n): Aquaculture = 41, Aquaculture/Fisheries Shared = 7, Fisheries = 25.

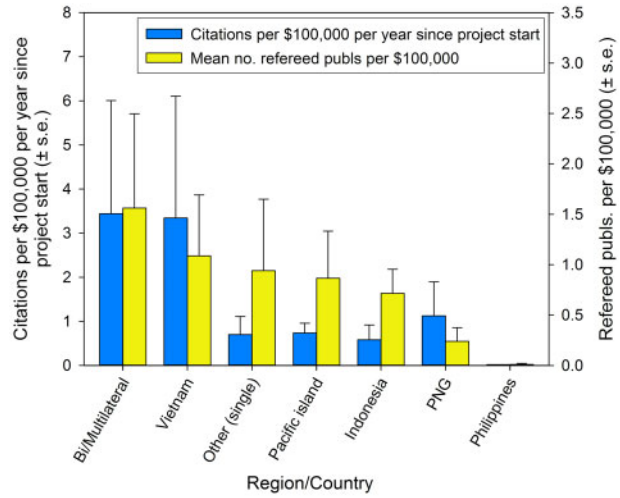

Figure 5. Mean number of Scopus citations per AU$100,000 funded to projects per year since projects started (blue or black bars, left-hand axis scale) and mean number of refereed journal articles per AU$100,000 funded to projects (yellow bars, right-hand axis scale), with project data grouped by country/region. Error bars are standard errors of the mean. Number of projects (n): Bi/Multilateral = 17, Indonesia = 15, Philippines = 6, Vietnam = 8, PNG = 11, Pacific Islands = 12, and Other (single) = 4. Countries in the "Other" category were Sri Lanka, India, and Cambodia.

the average of other similar publications indexed in the Scopus database (Colledge 2017). A score of 1.0 indicates that the publications have been cited at world average for similar publications, while greater than 1.0 indicates that the publications have been cited more than would be expected based on the world average for similar publications (Colledge 2017). FWCI metrics is useful to benchmark and provide a useful way to evaluate researchers' citation performance. Over three-quarters of the journal articles had FWCI of less than 1.0, or the article was not listed in the Scopus database at all. The journal article with the highest FWCI was a paper titled 'Vulnerability and resilience of remote rural communities to shocks and global changes: Empirical analysis from Solomon Islands' published in the journal *Global Environmental Change* (Schwarz et al. 2011). This article was an output of the project 'Improving resilience and adaptive capacity of fisheries-dependent communities in the Solomon Islands' (FIS/2007/116). The paper had FWCI of 5.3 and the most citations per year (12 citations per year) of any journal article in the publications database.

5.1.5 *Scopus citations and refereed publications by industry type*

Overall, the average number of journal articles published per AU$100,000 of ACIAR investment was much higher for Aquaculture projects than for either capture fisheries, termed 'Fisheries' hereafter, or Shared projects. Likewise, the average number of all Scopus citations per $100,000 of funding per year since project start was higher for Aquaculture (2.7) compared to Fisheries (0.6) and Shared projects (0.5) (Figure 4). However, as noted previously, project publication and citation data in this review are skewed by a number of relatively small projects. In particular, two relatively small Aquaculture projects (Projects FIS/2002/083 and FIS/2006/009, both in the AU$50,000–$99,000 funding range) dominated the publications and citations per AU$100,000 per year since project start for the Program overall. This is reflected in the large standard errors for Aquaculture projects.

5.1.6 *Scopus citations and refereed publications by country*

The number of refereed articles and corresponding citation rates per funded dollar were higher for bilateral and multilateral projects than those conducted in a single country (Figure 5). Moreover, three of the top five projects with a high number of citations were bi/multilateral. A few outlying projects had a strong influence on this significant variation. In particular, the small bilateral project FIS/2002/083 was a notable outlier to the rest of the Program projects, producing 16.4 journal articles and 31.5 citations per $100,000 per year since the project started compared to a mean of 1.6 journal articles and 3.4 citations per $100,000 per year since projects started across all bi/multilateral projects. Projects conducted solely in Vietnam also tended to yield relatively high numbers of refereed articles and citations per funded dollar. However, the data for both metrics were particularly variable for both bi/multilateral projects and projects in Vietnam, as indicated by large error bars (Figure 5).

There was little difference between the numbers of refereed journal article citations per funded dollar per year since project start for projects conducted solely in Indonesia, PNG, Pacific Islands, and Other single-country projects (blue or black bars, Figure 5). Projects conducted solely in PNG or the Philippines tended to produce fewer refereed journal article publications per funded dollar, compared to the other regions/countries (yellow or grey bars, Figure 5). Despite the lower average productivity, it is worth noting that some multi-country (bi/multilateral) projects involving these other countries did produce many refereed articles and had relatively high citation rates.

5.1.7 *Publication and citation numbers as a function of funding*

Project publication and citation data were skewed by the relatively large number of journal articles produced from a few outstanding projects. The number of refereed journal articles produced by projects per funded dollar increased significantly for Fisheries projects funded by more than $600,000 ($F_{1, 23} = 7.4$, $P = 0.012$) (Figure 6a). However, the level of funding to projects did not significantly influence the number of refereed articles produced by either Aquaculture projects or Shared projects ($P > 0.05$).

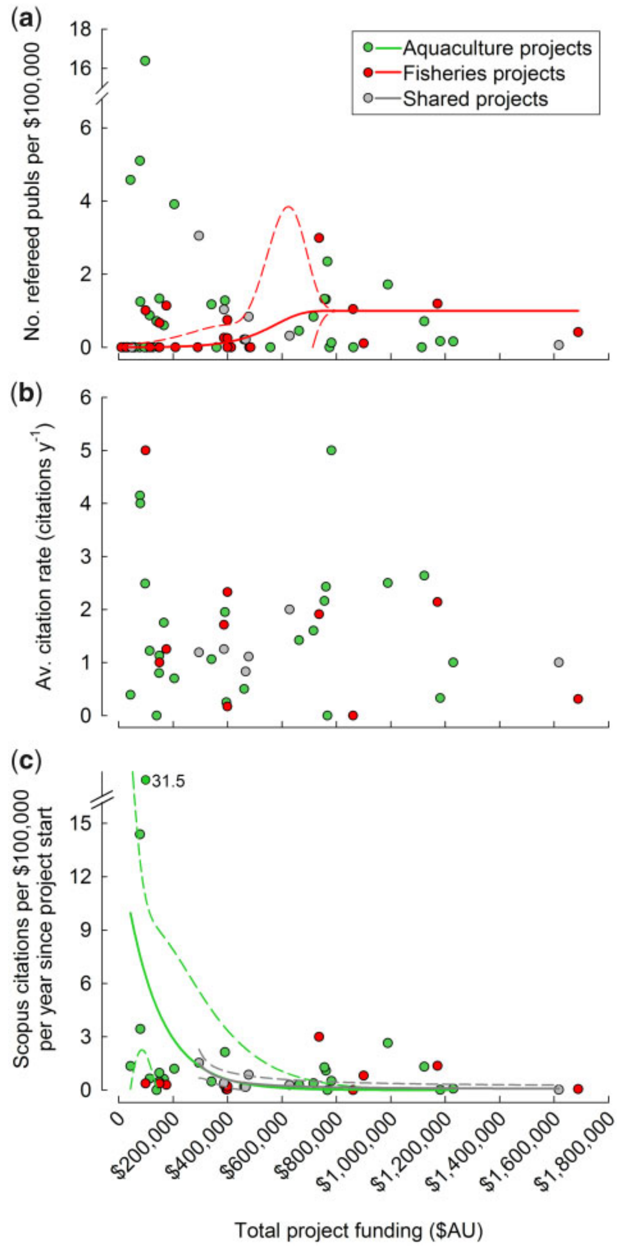

Figure 6. Scatter plots of total project funding versus scientific impact metrics: (a) number of refereed journal publications per AU$100,000 funded, (b) average citation rate of publications produced by projects (average number of Scopus citations per year), and (c) Scopus citations per AU$100,000 funded per year since projects started until the census date. Only the statistically significant relationships for any project type are shown. Solid lines are the fitted non-linear relationships; dashed lines are 95% confidence intervals. Line colours or the different shades correspond to the colour coding by project type given in the legend (green or grey = Aquaculture; red or black = Fisheries; light grey= Shared); refer to the electronic version of the article for colours.

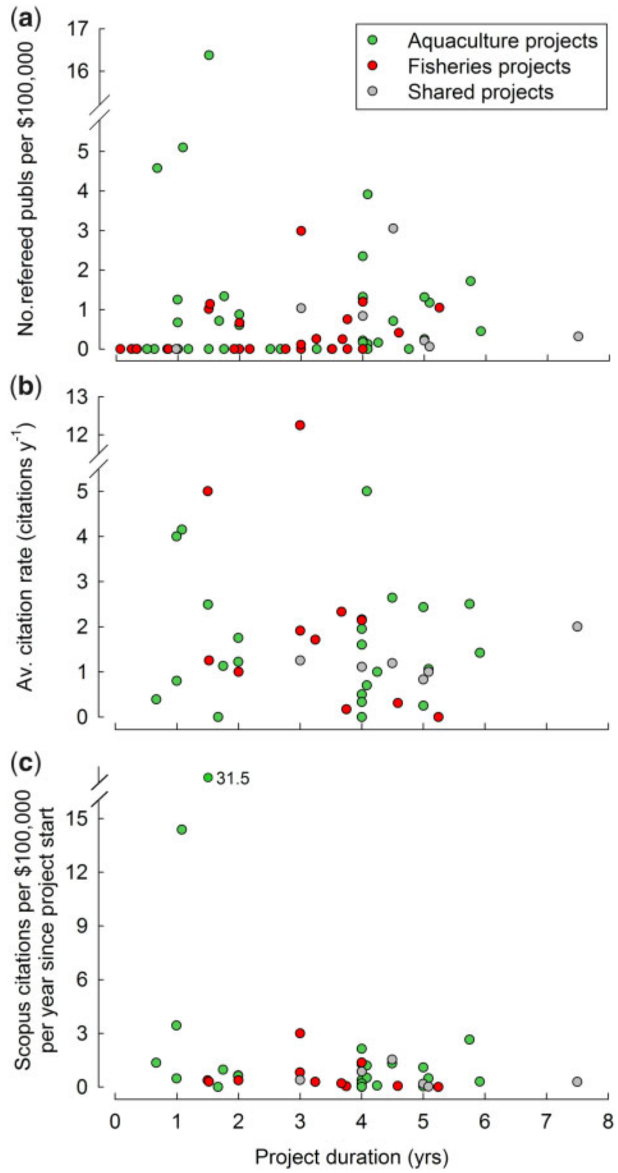

Figure 7. Scatter plots of project duration vs scientific impact metrics: (a) number of refereed journal publications per AU$100,000 funded, (b) average citation rate of publications produced by projects (average number of Scopus citations per year), and (c) Scopus citations per AU$100,000 funded per year since projects started until the census date. Fitted relationships are not shown, as none were statistically significant for any project type.

Our analysis indicated that average citation rates of refereed journal articles were not related to the level of project funding for neither Aquaculture, Fisheries, nor Shared projects ($P > 0.05$ for all three industry types) (Figure 6b). That is, projects with greater amounts of funding did not tend to produce refereed journal articles with greater or lesser impact than projects with modest funding.

The number of citations of all refereed journal articles produced by projects per funded dollar, normalized by years since projects started, was significantly higher for projects with slighter funding for both Aquaculture ($F_{1,\ 22}$=4.9, $P = 0.04$) and Shared projects ($F_{1,\ 4}$=13.9, $P = 0.02$), but not for Fisheries projects ($P = 0.20$) (Figure 6c). That is, Aquaculture and Shared projects with greater funding tended to produce articles that, on average, had lower 'impact' than projects with less funding. Notably, the (best) fitted curves were closely overlapping for these two industry types, showing high consistency in the relationship across the fitted values.

5.1.8 *Publication and citation numbers as a function of project duration*

Analysis of publication and Scopus citation numbers for Fisheries, Aquaculture, and Shared projects as a function of project duration (Figure 7a–c) revealed no significant relationships ($P > 0.05$ for all industry types on all three panels). Simply put, the duration of a project did not tend to significantly affect either the number of refereed publications per funded dollar, the average citation rate of refereed publications produced by projects, or the number of citations of all refereed articles produced by projects per funded dollar per year after project activities started.

5.2 Scientific impact—knowledge production and capacity building in partner countries

Inconsistent reporting of research capacity building activities and impacts in partner countries among projects limited our capacity to fully account for the scientific (research) impact of the Program (Mullen, Gray and Meyer 2015). However, a synthesis of reported outputs from Program projects (Table 2) shows that, beyond publishing scientific outputs, projects invested considerable resources in capacity building activities. Many projects organized educational tours and visits to Australia or other countries, institutionalized mentoring processes, short training, and postgraduate programs.

The majority of the workshops were noted as having been structured as 'train-the-trainer' activities, to enable participants to provide subsequent training to others. At least 215 different workshops were mentioned in the project documents, with over 2,500 participants recorded. However, only about a quarter of reports noted the actual number of people who attend the events, suggesting that the actual number of participants may well be several thousand more. The most common outcomes of the training and workshops were considered to be (1) increased capability of project members and their institutions in research planning and execution; and (2) increased publication output, particularly the number of papers authored by local staff.

With regards to short courses and postgraduate studies, at least 29 people attended master classes, with the majority in the field of aquaculture. Twenty-eight people undertook project-supported Masters-level courses, and six undertook PhD degrees. All these students attended Australian Universities, inferring a corresponding increase in the pool of highly trained scientists with whom Australian scientists could collaborate. For every Australian researcher involved in the Program research projects, there were at least three partner country researchers involved.

All project leaders and partner-country collaborators interviewed during this study agreed that the Program contributed significantly to the scientific capacity development of partner-country researchers. Project leaders reported that the training and capacity development activities contributed to the improvement in experimental design and data analysis skills of researchers in partner countries, thereby positively influencing research output. Several project leaders also mentioned their time investments in mentoring partner country scientists and government officers through collaborative writing of scientific publications and reports. These forms of capacity building resulted in the increased international exposure of the work of partner country scientists. As one scientist from Indonesia put it, '. . . the Australians have helped us to become good scientists. That is very important, we work on the research and learn how to become a good scientist'.

The case studies also indicated significant capacity-building impact. For example, the capacity building of staff at the Research Institute for Aquaculture No. 1 (RIA 1) supported the establishment of a state-of-the-art oyster hatchery research facility in Cat Ba, Vietnam in 2006. The capacity building of RIA 1 Vietnamese scientists included the training of seven aquaculture officers in Australian hatcheries, 25 undergraduate students and three MSc students were involved directly in the Program. In addition, three staff members received John Allwright Fellowships[4] to undertake PhD studies in Australia. The increased capabilities at RIA 1 through these capacity building activities has significantly enhanced the capacity of RIA 1 staff to pursue independent research and contribute to industry development.

In Indonesia, the ACIAR Program has been working with the Institute for Mariculture Research and Development (IMRAD) through several projects and has contributed significantly to the development of 'grassroots' hatchery technology for milkfish and grouper. ACIAR projects contributed to the development of the capability of IMRAD as a research institute and in developing the capacity of its scientists. More than 100 hatchery operators and researchers from 22 countries received training. Indonesian project staff attest that the projects improved the capacity of partner country scientists to plan and conduct research, and improved scientists' confidence and capability to present at international conferences, which in turn, improved staff publication rates.

The Program also provided capacity building outcomes for Australian researchers and Australian research institutions (Commissioned Agencies) by providing significant, varied and ongoing staff and student training, development, and publication opportunities. Nearly all the Australian researchers interviewed considered the Program to have played a major role in their career development and teaching practices. Australian researchers also reported that participation in these projects provided opportunities to learn about and be trained in areas that are of international significance, improving the broader research capacity of Australian science. From the perspectives of researchers involved, the Program clearly contributed to building and retaining high calibre researchers and scientific capability in Australia.

5.3 Factors influencing the publication agenda and project outputs

Qualitative analysis of the interview and survey data highlighted that three main factors influenced the publication agenda and outputs and involvement in the publication process: (1) project leaders' leadership and motivational ability; (2) institutional culture and drive of participating institutions (both Australian and partner agencies); and (3) individual drive of project leaders and partner agency researchers.

5.3.1 *Project leaders' leadership and motivational ability*

Individual project leaders were the primary drivers influencing both the publication agenda and outputs and the motivation to publish within project teams. In particular, the ability of a Project Leader to harness team members' aspirations and motivate the team was considered as an important factor in determining publication output. Project leaders who were keen to engage their team members in publishing had higher publication outputs. Project leaders considered that encouraging partner agency researchers to publish was an essential element of their role.

Table 2. Reported formal capacity development activities

Main activity	No. of people involved	Specific activities and processes	Who was involved	Notes
Visits (educational tours and visits to Australia and other countries)	120	Project initiation and planning, knowledge sharing, postgrad studies, staff exchange, placement, on-farm training, manuscript preparation, conferences, and workshops	Partner country researchers, officers, farmers; managers, staff, trainees, and government officials	Country visit's duration can be in the order of weeks or months
Mentoring	Not recorded	Information exchange through collaborations, scientific support, joint field trips, structured learning opportunities, organized task forces, institutional 'twinning', planning, and websites (platforms)	Mainly partner country researchers, officers, farmers, managers, staff, and trainees	Although extensive formal and informal mentoring was noted in projects documents, numbers were not reported
Workshops and seminars	2,500	Various—too many to list	Partner country researchers, officers, farmers, managers, staff, trainees, and government officials.	Of the 215 workshops/seminars cited, only a quarter recorded numbers of participants.
Masterclass (training)	29	Aquaculture Nutrition Master Class, Aquafeed Extrusion Masterclass	Partner country researchers, technical officers	
John Dillon Fellowship	10	Various	Researchers and research managers	
John Allwright Fellowship (MSc and PhD)	28 MScs, 6 PhDs, 6 others	Various	Partner country researchers and research managers	

5.3.2 *Institutional culture*

A secondary driver of publication agenda and output was the publication culture of the Project Leader's employers. Institutional culture reportedly influenced the team members' drive to publish. Project leaders reported that there was significant variance between partner agencies regarding the perceived importance of publications and that this was largely influenced by institutional culture and related factors. However, it was noted that this variance was more pronounced at the start of projects and diminished as projects progressed (i.e. both the perceived importance of publications and the drive to publish increased over time). Substantial in-kind support by partner agencies and project partners also was reported to bolster publication success.

5.3.3 *Personal motivation of project leaders and partner agency researchers*

Although the data demonstrated that the above two factors were the main factors that played a significant role in publication agenda and output, the individual drive of project leaders and partner agency researchers was also important. Project leaders reported that publication records were highly regarded as evidence of capability and capacity of partner agency researchers.

In addition to the above three factors, other, secondary, factors that were also reported to influence publication decision making and outputs included:

- The Program's flexibility in project planning, review and adaptive project management;
- Individual researchers' and organizational relationships and credibility with a specific journal;
- Previous publication records for partner agency researchers; and
- Evolution within the global publication environment, including the growing reach of open access journals.

6. Discussion

6.1 Measuring and benchmarking the scientific impact of an R4D organization

Our bibliometric analysis provides a point of departure for benchmarking R4D organizations in the absence of comparable data for other similar R4D organizations. The Expert Panel engaged for this project considered that, overall, the reported publication output and citation rate were relatively low, with a corresponding relatively high average investment per publication. However, caution is needed when comparing the number of 'all publications' across projects because publication value/impact will vary among publication types and the same research can be used in different types of publications. The Panel noted that it is important to consider the operating environment from which publications arise when comparing publication and citation data. The publications in this case study arise from projects in developing countries with facilities and field conditions incomparable to those of developmental research in developed countries. The Panel noted that ACIAR is both a development agency and a research organization and that measuring scientific impact of the Program based on its publication outputs alone was insufficient.

Our findings showed that data on publication outputs and citation rates were skewed because of a few outstanding projects or publications. Waltman (2016) noted that this effect, driven by a few relatively highly cited publications, is commonplace. Sub-field differences also influence citation rates, for example, taxonomic papers, of which 11 from this study were published in the journal *Zootaxa*, generally have low citations because once a species is described, it is rarely necessary subsequently to cite the original journal article. This is also reflected in the relatively low impact factor of that journal: 0.945. The use of citation metrics for the field of taxonomy has been considered as an impediment to biodiversity research (e.g. Valdecasas, Castroviejo and Marcus 2000). Thus, citation metrics should be used cautiously as a measure of scientific impact for R4D organizations.

Another noteworthy point is the apparent value of collaborative research. The literature suggests that international collaboration and interdisciplinarity are positively related to increased publication and citation (Bornmann and Daniel 2005; Vogel, Hattke and Petersen 2017). This generality seems to hold true in the current study, which showed bi-lateral and multi-lateral projects on average tended to produce a larger number of journal articles and attracted more citations per funded dollar compared to projects conducted in a single partner country (Figure 5). The scope of this article did not allow us to analyze why this is the case. Nonetheless, we can say that multi-lateral projects are likely to have access to a larger pool of resources, and so the resulting publications could be more consequential. Further, publications arising from larger data sets across multiple countries tend to offer more generalizable lessons, which attract greater citations. As Katz and Martin (1997) suggested, R4D organizations should assess the costs of research partnerships and collaboration with partner countries and associated benefits to ensure that the desired outcomes of research are achieved and use of resources optimized.

Analyzes of the correlations between the funding quantum to projects and overall publication output yielded mixed results. We found that well-funded Fisheries projects tended to produce more refereed publications per funded dollar than modestly funded projects, but funding had little effect on this measure of productivity for Aquaculture or Shared projects. This disparity could be due, in part, to the fact that fisheries are often geographically broad and projects require greater funding to gather publishable data than aquaculture projects that can require fewer resources to complete a publishable study. Projects with smaller amount of funds actually produced refereed articles with greater total number of citations per funded dollar in the Aquaculture and Shared sectors, but not so for Fisheries projects. In this case, smaller (i.e. less funded) aquaculture projects might already have physical resources (hatchery, tanks, and ponds) and a greater proportion of funding goes to research and data collection, whereas projects requiring greater infrastructure expenditure have larger budgets and so produce fewer publications per funded dollar. Importantly, the amount of funding to projects did not significantly affect the scientific 'impact' of project publications per funded dollar.

Many World Bank project leaders believe that longer and greater-funded projects are more successful (Ika, Diallo and Thuillier 2012). However, empirical analysis contradicts this claim, although caution in interpretation is needed since, for example, 'difficult' projects might take longer to complete and are more likely to have poor impact (Denizer, Kaufmann and Kraay 2013). In our study, project duration did not significantly affect any of the three bibliometric measures used. While longer-term projects might build greater capital in data and technological advances compared to shorter projects, the publication quality and scientific impacts per funded dollar are clearly not enhanced per funded dollar. Thus, while there might be administrative benefits to funding agencies in commissioning longer projects, there is little evidence of improved scientific-output 'bang for the buck' over shorter projects.

As we found, bibliometrics results explain only one part of the scientific impacts of R4D organizations (see Moed 2005; Confraria, Mira and Wang 2017). Scientific impact also encompasses the creation of new knowledge, technologies, products, and services; and capacity building (Davis et al. 2008; SFC, HEFCE and HEFCW 2012). Therefore, measuring the scientific impact of R4D organizations should include an understanding of (1) the socio-cultural, political, and economic contexts under which the various research projects were carried out; (2) activities, including capacity building, carried out by the Program in the partner countries; and (3) the different actors involved in the research projects.

6.2 The context in partner countries, capacity building, project leaders' motivation and scientific impact

As discussed earlier, through the Program, ACIAR commissioned various Australian research institutions and IARCs to conduct research in Southeast Asia, Pacific Islands countries, and PNG. Hence, the Program operated in an environment that involved diverse people and projects within complex social, institutional, and political milieus. These complexities are likely to have influenced the overall scientific output and outcomes of projects in different countries (Confraria, Mira and Wang 2017).

For example, our analyzes showed a marked variation in the publication output from projects based within different regions/countries. We found relatively large differences between the number of journal articles and citations per funded dollar among countries. Notably, projects based in PNG and the Philippines had a much lower journal article output and citations per funded dollar. This might be due to a complex set of factors including socio-cultural, political, and economic contexts and the institutional and personal capabilities in those countries. Nonetheless, the findings suggest that if publication metrics are considered an important yardstick for scientific impact, assessments of this nature could help to shape policy on target countries to maximize scientific impact and identify countries in which scientific publication training could be most influential.

The evaluation revealed that a significant amount of financial and human resources were invested by the Program to build the capacity of researchers and their institutions in partner countries through educational tours and visits, mentoring, workshops and seminars, training, Masters, and PhD programs. The high frequency of joint-authorship with developing-country scientists and students accounts for further impact in the long term. In other R4D research fields, research is often led by developed-country researchers, and publication often follows a 'mentor-apprentice model', whereby developing-country researchers tend to be co-authors (Boshoff 2009). In our study, the prevalence of peer-reviewed publications led by project leaders and scientists from developed countries (especially highly cited papers) reflects the fact that research organizations commissioned by ACIAR to lead projects were from Australia or other international agricultural research centres. Developing country researchers often lacked skills and/or found it difficult to allocate time to write and prepare manuscripts to a level required by high-tier journals. Nonetheless, co-authorship and mentoring constitute important capacity development for research institutions in partner countries. These capacity building activities are expected to have had a broader scientific impact by way of facilitating the creation of new knowledge, changes in the communities and their livelihoods, and changes in policy and practices, which are unlikely to be reflected in publication metrics of the Program's projects (Confraria, Mira and Wang 2017).

Qualitative data indicated that the primary drivers of publication outputs were the motivation of the project leaders and their institutional culture. We anticipate that project leaders' interest in reinforcing their global recognition in the scientific community, combined with their desire to gain significant standing and recognition within their institutions, together influence their publishing behaviour. This suggests that a funding institution's policies and procedures might not exert much influence over the publication outputs of its projects. Improving the scientific impact of funded projects might need transparent and enforceable policies in this regard, including the development of policies that incentivize collaborative work and co-authorship. If the capacity development benefits for partner countries are to be sustained, there is a need for R4D agencies to encourage the development of policies and programs that facilitate co-authorship and collaboration between developed and developing country researchers. The argument posed in Section 5.1 that the socio-cultural, political and economic contexts of research need to be considered in measuring the scientific impact of R4D organizations is of relevance here as well. Therefore, we suggest that the scientific impact of R4D agencies, in contrast to purely academic institutions, must be measured against a broad suite of metrics.

6.3 Bibliometrics and the scientific impact of R4D organizations: some reflections

Earlier, we discussed the utility of formal publication metrics as a departure point in benchmarking R4D organizations. We also found that other factors (e.g. new knowledge, capacity development, and context) are important considerations, which traditional metrics may not necessarily capture. One argument posed in the literature (e.g. Garfield 1999), and by the Expert Panel, was that publication outputs, citations and impact factors are discipline-biased and context-dependent. Our earlier discussion about the field of taxonomy is a case in point. Bibliometrics, therefore, is not an all-encompassing indication of scientific impact. Unless projects are explicitly designed to focus on scientific research, and not policy or technology development outcomes, publication metrics can be inadequate.

Publishing in high impact factor journals and attracting citations does not always guarantee high readership by the public, practitioners, and policy-makers, and thus, does not necessarily translate to high-impact policy and practice. Indeed, Flickinger et al. (2014) argue that there is no significant relation between practical relevance and citation rates. Chavarro, Tang and Ràfols (2017) further concluded that, in some cases, non-mainstream journals with low impact factors might offer better opportunities for scientists in developing countries to disseminate knowledge to local communities. Others have also argued that the ranking of a journal might not necessarily be a good indicator of the practical relevance of scholarship (Adler and Harzing 2009). Papers published in highly ranked journals might be popular among, and frequently cited by, academics and researchers who speak with each other circuitously. However, such publications might be inaccessible to the general public, practitioners, and policy-makers—especially in developing countries where English is not the first language.

With these arguments in mind, we contend that bibliometrics should be considered as 'a means', rather than 'an end', to the assessment of the scientific impact of R4D programs. Bibliometrics can indicate where scientific impacts fit and are progressing against defined benchmarks of relevance to the Program. However, such metrics are only one part of 'the aggregate effect of the changes in practice, products and policies by the final user groups' (Davis et al. 2008: 25).

Bibliometrics offers a normative and objective indication of the scientific impact of R4D organizations; however, caution must be exercised when comparing across disciplines. For example, a FWCI of 0.75 might reflect a high impact R4D-funded publication from a small-island state but could be considered low for research within a developed country. Differentiation between research conducted in and for developing countries and research conducted in and for the benefit of a developed country is therefore essential.

7. Conclusion

This study had two principal objectives: to provide an assessment of the scientific impact of ACIAR's Fisheries Program between 2000 and 2012 as a departure point for benchmarking similar R4D organizations and to stimulate discussion among the broader R4D community regarding appropriate measures of scientific impact.

With regard to the first objective, our methodology focussed on bibliometrics, including the number of peer-reviewed journal articles produced through Program projects and the number of citations these articles received. To critically analyze these data, we sought the opinions of an Expert Panel and gathered qualitative data on key drivers behind publication outputs. Our analysis found that bibliometrics offers useful initial measures to understand the scientific impact of an R4D Program and are often the metrics valued by funding agencies and employers of the researchers involved.

Our data also suggested that, without specific policies, R4D Programs might have a little direct bearing on the publication output of funded projects. Key publication drivers were the capacity and motivation of the project leaders and the institutional culture and requirements of their employers. We also found that, in some cases, there were significant relationships between publication/citation metrics and the sector funded, the country in which the project was based and the total project funding. Projects involving multiple partner countries tended to have a higher scientific impact, but the outcomes were variable. Generally speaking, projects with greater funding or longer time frames did not produce greater 'bang for the buck'—inferring that small and modestly-funded projects were (roughly) as productive in scientific impacts per funded dollar. We can, therefore, conclude that R4D institutions require clear and enforceable policies and procedures about publication outputs if their funding programs are to be assessed and benchmarked upon bibliometric measures.

Our analysis also suggests that the metrics used to evaluate the scientific impacts of R4D organizations need to be broader and more nuanced than those of the researchers and their respective commissioned academic institutions. Such measures should include aspects of capability development and recognition of the benefits of research to the R4D agency's home country. R4D organizations are also advised to engage with their key stakeholders to agree on the measures of importance, and how relevant data are to be collected and analyzed.

Notes

1. As at census date of 17 March 2016, publication data had not been provided for 17 of the 90 projects.
2. With the exception of a small number of project leaders who had retired, were non-contactable or had passed away.
3. The project leaders interviewed were also project leaders on a further 15 ACIAR Fisheries projects that were not included in our publications analysis because they either (1) ended after the end of 2012 or (2) were ongoing at the time of the interview.

4. The John Allwright Fellowship is an ACIAR Fellowships Scheme that provides the opportunity for partner-country scientists involved in ACIAR-supported collaborative research projects to obtain postgraduate qualifications at Australian tertiary institutions. The Scheme aimed at enhancing research capacity in ACIAR's partner country institutions.

Supplementary data

Supplementary data is available at *Research Evaluation* Journal online.

Acknowledgements

The authors wish to thank the ACIAR Fisheries Program and RDS Partners Pty Ltd for providing the funding and time required to conduct the evaluation and write this paper. The authors would also like to thank the Project Leaders and members of the Expert Panel who generously gave their time and knowledge during the review process. The authors would also like to thank the editor and the reviewers for providing such a valuable and constructive comments that have improved the quality of the paper.

Funding

This work was supported by the Australian Centre for International Agricultural Research [Project No. C2016/108 and FIS/2010/096 to S.W.P.].

Conflict of interest statement. None declared.

References

Adler, N. J., and Harzing, A.-W. (2009) 'When Knowledge Wins: Transcending the Sense and Nonsense of Academic Rankings', *Academy of Management Learning & Education*, 8/1: 72–95.

Bartlett, A. G. (2016) 'Evaluating Relative Success of Donor-Funded Collaborative Research Projects', *Research Evaluation*, 25/4: 405–15.

Bertocchi, G. et al. (2015) 'Bibliometric Evaluation vs Informed Peer Review: Evidence from Italy', *Research Policy*, 44/2: 451–66.

Bollen, J., Van de Sompel, H., Hagberg, A., and Chute, R. (2009) 'A Principal Component Analysis of 39 Scientific Impact Measures', *PloS one*, 4/6: e6022. doi: 10.1371/journal.pone.0006022.

Bornmann, L., and Daniel, H.-D. (2005) 'Does the h-Index for Ranking of Scientists Really Work?', *Scientometrics*, 65/3: 391–2.

Boshoff, N. (2009) 'Neo-Colonialism and Research Collaboration in Central Africa', *Scientometrics*, 81/2: 413–34.

Bozeman, B., and Youtie, J. (2017) 'Socio-Economic Impacts and Public Value of Government-Funded Research: Lessons from Four US National Science Foundation Initiatives', *Research Policy*, 46/8: 1387–98.

Braun, T., Glänzel, W., and Schubert, A. (2006) 'A Hirsch-Type Index for Journals', *Scientometrics*, 69/1: 169–73.

Brody, T., Harnad, S., and Carr, L. (2006) 'Earlier Web Usage Statistics as Predictors of Later Citation Impact', *Journal of the Association for Information Science and Technology*, 57/8: 1060–72.

Bryman, A. (2008) *Social Research Methods*, 3rd edn. New York: Oxford University Press.

Campbell, D. et al. (2010) 'Bibliometrics as a Performance Measurement Tool for Research Evaluation: The Case of Research Funded by the National Cancer Institute of Canada', *American Journal of Evaluation*, 31/1: 66–83.

Chavarro, D., Tang, P., and Ràfols, I. (2017) 'Why Researchers Publish in Non-Mainstream Journals: Training, Knowledge Bridging, and Gap Filling', *Research Policy*, 46/9: 1666–80.

Colledge, L. (2017) *Snowball Metrics Recipe Book*, 3rd edn. Amsterdam, the Netherlands: Snowball Metrics program partners < https://www.snowballmetrics.com/reports/> accessed 4 April 2019.

Confraria, H., Mira, G. M., and Wang, L. (2017) 'Determinants of Citation Impact: A Comparative Analysis of the Global South versus the Global North', *Research Policy*, 46/1: 265–79.

Costas, R., Zahedi, Z., and Wouters, P. (2015) 'Do "Altmetrics" Correlate with Citations? Extensive Comparison of Altmetric Indicators with Citations from a Multidisciplinary Perspective', *Journal of the Association for Information Science and Technology*, 66/10: 2003–19.

Davis, J., Gordon, J., Pearce, D., and Templeton, D. (2008) 'Guidelines for Assessing the Impacts of ACIAR's Research Activities', *ACIAR Impact Assessment Series No. 58*, 120pp.

Denizer, C., Kaufmann, D., and Kraay, A. (2013) 'Good Countries or Good Projects? Macro and Micro Correlates of World Bank Project Performance', *Journal of Development Economics*, 105: 288–302.

DFID (2014) *What is the Evidence on the Impact of Research on International Development? DFID Literature Review*. UK: Department for International Development.

Farmer, B., Duncan, R., Enright, T, and Jarvie, W. (2013) *Independent Review of the Australian Centre for International Agricultural Research (ACIAR)*, Canberra: Australian Centre for International Agricultural Research.

Flickinger, M., Tuschke, A., Gruber-Muecke, T., and Fiedler, M. (2014) 'In Search of Rigor, Relevance, and Legitimacy: What Drives the Impact of Publications?', *Journal of Business Economics*, 84/1: 99–128.

Garfield, E. (1999) 'Journal Impact Factor: A Brief Review', *Canadian Medical Association Journal*, 161/8: 979–80.

Godin, B., and Doré, C. (2004) *Measuring the Impacts of Science: Beyond the Economic Dimension*, CSIIC Working Paper <http://sta.uwi.edu/rdifund/documents/GodinDoreImpacts.pdf> accessed 3 July 2018.

Harzing, A.-W. K., and Van der Wal, R. (2008) 'Google Scholar as a New Source for Citation Analysis', *Ethics in Science and Environmental Politics*, 8/1: 61–73.

Herdt, R. W. (2012) 'People, Institutions, and Technology: A Personal View of the Role of Foundations in International Agricultural Research and Development 1960–2010', *Food Policy*, 37/2: 179–90.

Hicks, D., Tomizawa, H., Saitoh, Y., and Kobayashi, S. (2004) 'Bibliometric Techniques in the Evaluation of Federally Funded Research in the United States', *Research Evaluation*, 13/2: 76–86.

Ika, L. A., Diallo, A., and Thuillier, D. (2012) 'Critical Success Factors for World Bank Projects: An Empirical Investigation', *International Journal of Project Management*, 30: 105–16.

Katz, J. S., and Martin, B. R. (1997) 'What Is Research Collaboration?', *Research Policy*, 26/1: 1–18.

Mazzoleni, R., and Nelson, R. R. (2005) *The Roles of Research at Universities and Public Labs in Economic Catch-Up*, LEM Working Paper Series, No. 2006/01, Scuola Superiore Sant'Anna, Laboratory of Economics and Management (LEM), Pisa.

McGillivray, M., Feeny, S., Hermes, N., and Lensink, R. (2006) 'Controversies over the Impact of Development Aid: It Works; It Doesn't; It Can, But That Depends . . .', *Journal of International Development*, 18/7: 1031–50.

Meho, L. I., and Yang, K. (2007) 'Impact of Data Sources on Citation Counts and Rankings of LIS Faculty: Web of Science versus Scopus and Google Scholar', *Journal of the Association for Information Science and Technology*, 58/13: 2105–25.

Moed, H. F. (2005) *Citation Analysis in Research Evaluation. Information Science and Knowledge Management*. Dordrecht, The Netherlands: Springer.

Mullen, J., Gray, D., and Meyer, J. (2015) 'Evaluating the impact of capacity building by ACIAR', *The 59th AARES Annual Conference, February 10–13, 2015*, Rotorua, New Zealand.

Patton, M. Q. (2015) *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 4th edn. Thousand Oaks, CA: SAGE Publications.

Pearson, S., Chudleigh, P., Simpson, S., and Schofield, N. (2012) 'Learning to Invest Better: Using Ex-Post Investment Analysis on Agri-Environmental Research and Development', *Research Evaluation*, 21/2: 136–51.

Sarewitz, D., and Pielke, R. A. (2007) 'The Neglected Heart of Science Policy: Reconciling Supply of and Demand for Science', *Environmental Science & Policy*, 10/1: 5–16.

Schwarz, A.-M. et al. (2011) 'Vulnerability and Resilience of Remote Rural Communities to Shocks and Global Changes: Empirical Analysis from Solomon Islands', *Global Environmental Change*, 21/3: 1128–40.

Seglen, P. O. (1997) 'Why the Impact Factor of Journals Should Not Be Used for Evaluating Research', *British Medical Journal*, 314/7079: 498–513.

SFC, HEFCE, and HEFCW (2012) *Assessment Framework and Guidance on Submissions (REF 02.2011)* <http://www.ref.ac.uk/2014/pubs/2011-02/> accessed 20 July 2016.

Thelwall, M., Haustein, S., Larivière, V., and Sugimoto, C. R. (2013) 'Do Altmetrics Work? Twitter and Ten Other Social Web Services', *PloS One*, 8/5: e64841. doi: 10.1371/journal.pone.0064841.

Valdecasas, A. G., Castroviejo, S., and Marcus, L. F. (2000) 'Reliance on the Citation Index Undermines the Study of Biodiversity', *Nature*, 403/6771: 698.

Vogel, R., Hattke, F., and Petersen, J. (2017) 'Journal Rankings in Management and Business Studies: What Rules Do we Play by?', *Research Policy*, 46/10: 1707–22.

Waltman, L. (2016) 'A Review of the Literature on Citation Impact Indicators', *Journal of Informetrics*, 10/2: 365–91.

White, H. (1992) 'The Macroeconomic Impact of Development Aid: A Critical Survey', *The Journal of Development Studies*, 28/2: 163–240.

Wilsdon, J. et al. (2015) *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. doi: 10.13140/RG.2.1.4929.1363.

Wouters, P., and Costas, R. (2012) *Users, Narcissism and Control: Tracking the Impact of Scholarly Publications in the 21st Century*. Utrecht: SURFfoundation.

Wouters, P. et al. (2015) *The Metric Tide: Literature Review (Supplementary Report I to the Independent Review of the Role of Metrics in Research Assessment and Management*. HEFCE. doi: 10.13140/RG.2.1.5066.3520.