Bibliometrics as a Performance Measurement Tool for Research Evaluation: The Case of Research Funded by the National Cancer Institute of Canada

David Campbell[1], Michelle Picard-Aitken[1], Grégoire Côté[1], Benoît Macaluso[2], Jean-Pierre Robitaille[2], Nicolas Bastien[2], Marie-Claude Laframboise[2], Louis-Michel Lebeau[2], Philippe Mirabel[2], Vincent Larivière[2] and Éric Archambault[1]

[1] *david.campbell@science-metrix.com, m.picard-aitken@science-metrix.com, gregoire.cote@science-metrix.com, eric.archambault@science-metrix.com*
Science-Metrix, 1335A Avenue du Mont-Royal Est, Montréal, Québec H2J1Y6 (Canada)

[2] *macaluso.benoit@uqam.ca, robitaille.jean-pierre@uqam.ca, bastien.nicolas@courrier.uqam.ca, mclaframboise@gmail.com, lebeau.louis-michel@uqam.ca, mirabel.philippe@courrier.uqam.ca, lariviere.vincent@uqam.ca*
Observatoire des sciences et des technologies (OST), Centre interuniversitaire de recherche sur la science et la technologie (CIRST), Université du Québec à Montréal, 1290 rue Saint-Denis, Montréal, Québec H2X 3J6 (Canada)

Abstract

As bibliometric indicators are objective, reliable and cost-effective measures of peer-reviewed research outputs, they are expected to play an increasingly important role in research assessment as well as in the management of research, including supporting the selection of projects eligible to receive financial support. This paper presents the results of a project where bibliometrics was used as a performance measurement tool to supplement the evaluation of the National Cancer Institute of Canada (NCIC) funding program. The following questions were addressed: (1) Has the NCIC peer-review process been successful in selecting outstanding Canadian researchers in cancer research? (2) Has NCIC funding contributed to increasing the scientific performance of the papers authored by the researchers it supported? (3) How do NCIC-supported researchers compare to their US counterparts that received support from the US National Cancer Institute (NCI)? Citation analysis data provide evidence that the NCIC supported scientists who stand out among Canadian cancer researchers and that NCIC funding had a positive effect on the scientific impact of papers authored by the researchers it supported. Results also show that, although NCIC researchers have less scientific impact than NCI researchers, they are more cited than the US and Canadian average.

Introduction

The results of scientific research in the natural sciences and engineering (NSE, including the health sciences) are mainly disseminated through the publication of peer-reviewed papers in scientific journals (Larivière et al., 2006; Moed, 2005). Consequently, the scientific performance of researchers pursuing research in the health sciences is best assessed through the use of bibliometrics, a set of methods and procedures used in the quantification of bibliographic records (i.e., the basic units of measurement are bibliographic records of peer-reviewed publications).

Initially developed by information scientists, bibliometrics is now being applied more frequently, in combination with peer-review, to the field of research evaluation. The increased role of bibliometrics in research evaluation stems from a confluence of factors. For instance, governments have increasingly recognized the importance of accountability for public spending in research; meanwhile, the more traditional assessments based on peer review only can no longer be seen as impartial, as the recipients of research funding are typically those very 'peers' who participate in the review process (Hicks et al., 2004; King, 1987; Martin, 1996; Trochim et al., 2008). Because bibliometric indicators are based on a set of internationally recognized standards, they are considered to be the most objective and reliable

Proceedings of ISSI 2009, pages 301–312.
Edited by B. Larsen and J. Leta.

measures of academic research output, in addition to being cost effective in their usage (Archambault & Côté, 2008).

Research evaluation usually makes use of a variety of indicators to draw the most complete picture possible of the complex aspects that account for the performance of research organizations. This usually entails a combination of peer ratings and bibliometric indicators (King, 1987). This study's objective is not to provide an exhaustive overview of the multi-dimensional facets involved in assessing research performance but rather to discuss the use of bibliometrics as a performance measurement tool for the evaluation of a funding program, taking the National Cancer Institute of Canada's (NCIC) Operating and Program Project Grants as a case study.

The NCIC was established in 1947 as a joint initiative of the Department of National Health and Welfare (now Health Canada) and the Canadian Cancer Society. On February 1 2009, the Canadian Cancer Society integrated the operations of the NCIC, creating the Canadian Cancer Society Research Institute. Since 1988, the main goal set forth by the NCIC's Board of Directors has been "to undertake and support cancer research and related programs in Canada that will lead to the reduction of the incidence, morbidity, and mortality from cancer" (National Cancer Institute of Canada, 2009).

Among the various actions undertaken by the NCIC to achieve this mission, the support of cancer research through grants plays a pivotal role. As the funds awarded by the NCIC are raised exclusively through charitable donations to the Canadian Cancer Society and the Terry Fox Foundation, the Institute recognizes that it is accountable to the public for the value it creates through research. Thus, the NCIC launched an evaluation of its Operating and Program Project Grants in 2005 as part of its 2015 Strategic Plan (National Cancer Institute of Canada, 2005). So far, the NCIC has conceptualized and implemented an assessment of its awarding process, scrutinizing both the application and review of proposals. This has involved surveying both applicants (successful and unsuccessful) and members of the peer review panels. To supplement this evaluation, a performance assessment of the scientific research supported by the NCIC was performed.

The first part of the present study assesses the success of the peer-review process in selecting outstanding Canadian researchers for the attribution of NCIC grants by comparing the scientific impact, as measured using citation analysis, of NCIC-supported researchers with that of Canadian researchers (excluding papers by NCIC-supported researchers). The scientific impact of papers authored by researchers while they were receiving funds from NCIC (i.e., *supported papers*) is then compared to that of their papers authored while they were not funded by NCIC (i.e., *non-supported papers*) to establish whether the support provided by the NCIC might have had a positive effect on the scientific performance of the supported researchers. The analysis extends over 13 years worth of NCIC Operating and Program Project Grants (1994–2006).

Within an evaluative context, bibliometric analysis must also include comparables of the assessed entity to allow for conclusions to be drawn about its relative strengths and weaknesses. In this study, the US National Institute of Health's National Cancer Institute (NIH-NCI) was chosen as the comparable for the NCIC. Both the size and impact of NCIC-funded researchers' scientific production are compared to those of NCI-funded researchers' production. As data are only available on researchers receiving NCI grants after 2000, the analysis is limited to the 2000–2006 period.

Methods

This section describes the construction of the datasets used in producing reliable indicators of scientific production for NCIC-funded applicants (core dataset) and NCI funded applicants

(benchmark dataset). It also presents the taxonomy used in determining the domains of activity of NCIC and NCI researchers, the bibliometric indicators used to quantify scientific outputs, and the statistical analysis performed on bibliometric indicators.

Databases

In this study, Thomson Reuters' Web of Science (WoS), which includes three databases (the Science Citation Index Expanded™ [SCI Expanded], the Social Sciences Citation Index™, and the Arts & Humanities Citation Index™) covering nearly all fields of science was used to produce statistics on the scientific production of NCIC researchers and their comparables. Although the vast majority of papers relevant to their area of practice (i.e., cancer research) are indexed in SCI Expanded, using all of WoS allowed for the retrieval of additional papers in other areas of interest, such as the socioeconomics of cancer.

The WoS data were received from its provider (Thomson Reuters) as flat text files and were then transformed and imported into a relational SQL Server database. Although the WoS lists several types of documents, only articles, research notes, and review articles were retained in producing the bibliometric indicators.

Constitution of the Core Dataset: NCIC-Funded Applicants

A bibliometric dataset for an institution is usually built by retrieving papers in which the name of the institution is found in the authors' addresses. Because the NCIC supports research as opposed to being an institute which performs intramural research, its name is not expected to be found in the authors' addresses on published papers. As a result, to build a dataset of NCIC supported papers, a publication portfolio had to be reconstituted for each researcher who received funding from the NCIC.

The construction of the dataset followed a two-part process. First, the NCIC provided a list of principal investigators (PIs) who were awarded operating and program project grants over the last decade (685 PIs). The names of these researchers were then used to retrieve their scientific output indexed in the WoS. To avoid overestimates created by homograph problems, each researcher's paper portfolio was manually cleaned to remove false positives (i.e., papers belonging to another researcher with the same surname and initials).

Bulk researchers' name query: Before executing the automatic retrieval of papers by NCIC researchers, the names as they appear in the NCIC's list were transformed to match the format of author names in the WoS. Author names in the WoS do not include the first name of the authors, only their initials. For example, "John W. Smith" is transformed into "Smith-JW" and also into "Smith-J". The latter form ensures that publications wherein the middle name (or its initial) is omitted are retrieved. Subsequently, the formatted names are queried against the database to retrieve, for each researcher, all of the papers bearing his/her name as an author between 1994 and 2006. The search is limited to papers bearing a Canadian address to minimize the occurrence of false positives resulting from homographs in researchers' names.

Due to the prevalence of homograph problems, the automatic query overestimated the number of publications in many paper portfolios, especially for researchers with a common surname (e.g., Smith). Since there is no a priori regarding which researchers will be overestimated and which will not, the papers retrieved automatically were validated manually for each researcher.

Portfolio cleaning: Cleaning the publication portfolios consisted of manually removing papers authored by homographs that were erroneously assigned to a researcher by the automatic query. In so doing, careful attention was paid to the disciplines and specific topics of papers belonging to a publication portfolio. Several questions arise when analysing

whether or not a set of papers belong to a given researcher (e.g. Are those papers consistent with respect to the discipline of the researcher as revealed by his/her departmental affiliation? Is the scope of those papers broader than the products of only one individual researcher?). For example, the attribution of an engineering paper to a biologist, or a physics paper to an historian would be seriously questioned. However, given the commonness of multidisciplinarity in science, it is not sufficient to rely mechanically on departmental affiliations of researchers to validate the publications of their portfolio. For example, a philosopher may publish articles dealing with medical ethics in clinical medicine journals, and an engineer may collaborate on papers dealing with environmental problems published in biology or earth sciences journals. The institutional addresses may provide additional clues, since they often include the authors' departments. However, the WoS does not contain any information on the relation between the authors' names and their institutional addresses.

In cases where the previous actions failed to discriminate whether a paper should or should not be considered a part of a researcher's portfolio, the publication was downloaded when it was electronically available through libraries or open access. The article's signatures on the paper itself often provide a link between each author's name and her/his institutional address (including departmental affiliation), which normally allows one to unambiguously identify false positives.

Besides false positives, another issue relates to false negatives; papers authored by a researcher that were not retrieved by the automatic query. These "absent papers" reflect the fact that the WoS only covers a fraction of all the work published worldwide by researchers. For example, journals of local interest, books, and various publications that are generally referred to as "grey literature" (such as in-house research reports) in addition to most conference and symposium proceedings are not indexed in Thomson Reuters' scientific databases. Therefore, the publications in the WoS do not encompass the entire CV of researchers funded by the NCIC. Of the scientific output (mostly papers) compiled in the NCIC's annual reports, 22% were not published in journals indexed in the WoS (data not shown). Nevertheless, the WoS indexes the portion of their publications that is the most visible and the most frequently cited by the scientific community. Another factor that can contribute to false negatives is that the name of a researcher as it appears in the list of PIs provided by the NCIC may not be identical to the name under which he or she published articles. This can results from inaccuracies in the names in the input file and from changes in the name of female researchers due to marriage/divorce, to mention just a few.

As the cleaning of publication portfolios involves judgement on the part of individuals performing the task, errors inevitably occur. In this respect, we previously performed a validation of this procedure, which demonstrated that when working with aggregated portfolios (i.e. a number of researchers associated with a given organization), the error rate is negligible (<1%) enabling the production of reliable indicators of scientific production (unpublished data). Altogether, manual cleaning of publication portfolios is a time- and resource-consuming process requiring careful attention. Yet it is the only way to guarantee that results are sufficiently robust to evaluate important questions such as the impact of funding on specific groups of researchers.

In building the dataset for NCIC-funded applicants, the cleaning process removed half of the initial number of papers automatically assigned to researchers, leaving 24,208 papers authored by NCIC-supported researchers at any given time between 1994 and 2006. Papers were considered to be *NCIC-supported* if they were published between the year after the start of the grant and the year after the end of the grant. For example, if researchers were supported by the NCIC from 1997 to 2000, their papers published between 1998 and 2001 were counted as *NCIC-supported papers*. As the study period begins in 1994, papers are considered as of

1995 to allow for the one-year lag for 1994 funding. Between 1995 and 2006, a total of 22,793 papers were authored by NCIC-supported researchers, of which 54% (12,244) were *NCIC-supported papers*. The balance (22,793 – 12,244 = 10,549) are referred to as *non-supported papers* (these are all other papers by NCIC researchers, i.e., papers written before receiving the funding or more than one year after the end of the funding period). It should be highlighted that researchers receiving NCIC funding could also be receiving grants from other funders; papers considered as *NCIC-supported* may thus have been also partly supported by other sources of funding. Therefore, the authors of this article do not assume that the performance of these papers, in terms of scientific impact, can be entirely attributed to NCIC. However, because NCIC-funded researchers also receive funding from other sources when they do not receive support from NCIC, everything being equal, the protocol suggested here takes into account the variation in financing that should accrue to NCIC. The approach used to tag papers as supported or non-supported is discussed at length in the *Discussion*.

Constitution of the Benchmark Dataset: NCI-Funded Applicants

The benchmark dataset consists of a set of papers from an organization comparable to the NCIC, namely the US NCI. As was the case for the NCIC, inclusion of the NCI in the address field of papers published by the researchers it funds is not expected (i.e., for the extramural researchers considered here). Therefore, a publication portfolio had to be reconstituted for each researcher in a sample of PIs with support from the NCI using the method described above for NCIC-supported researchers. The list of funded researchers was made available to the NCIC by the NCI.

To compare the scientific output of NCIC- and NCI-supported researchers, a random sample of NCI supported researchers was paired to the population of NCIC-supported researchers to obtain an equal number of researchers on both sides of the comparison. Data on researchers who received NCI grants were only made available from 2000 onward. Because the list of NCI-funded researchers did not provide information on the type and amount of grants awarded to them, researchers having received funding from 2000 to 2006 from different types of NCI grants were selected by randomly sampling NCI supported researchers. As such, to obtain the comparable population of NCIC-supported researchers, Canadian researchers were selected so long as they obtained NCIC funding from 2000 to 2006 regardless the type of grants they received (NCIC Program Project grant or NCIC Operating grant). A population of 523 NCIC-supported researchers was thus obtained. Consequently, 523 NCI-supported researchers were randomly selected from those funded between 2000 and 2006 to match the population size of NCIC supported researchers over this seven-year period. All papers published by both groups of researchers from 2000 to 2006 were counted regardless of whether these researchers were funded over the whole period or not.

Disciplinary Classification

The categories and methods used to delineate the various domains of activity of NCIC and NCI researchers are, by and large, those used by the US National Science Foundation (NSF) in the Science and Engineering Indicators series (see for example, National Science Foundation, 2006); the taxonomy is a journal-based classification and has been in use since the 1970s. Because the NSF classification does not entirely satisfy the needs in the SSH, we modified this taxonomy with our own classification of journals for the social sciences.

The resulting taxonomy has one important advantage over other classifications (such as that used by Thomson Reuters); it is mutually exclusive, which means that each paper is attributed to a single field or subfield based on the journal in which it is published. One limitation of this classification is that papers published on a subject, such as, for example, the environment, but in a journal specialized in chemical engineering, would be classified as

belonging to the field of chemistry and the subfield of chemical engineering, even though its subject is the environment. The anomalies have little effect when large numbers are considered; however, their impact is greater when the number of papers considered is small (e.g., below 30). Some of the subfields are categorized as general (e.g., general biomedical research), and this reflects the fact that in many fields there are some journals that address a broader readership.

Bibliometric Indicators

Using researcher portfolios built using the aforementioned methods as well as papers computed at the country level (for Canada and the US) the following indicators were calculated:

Number of papers: Whole counting of scientific papers written by authors associated with a funding organization (i.e., NCIC or NCI) based on author names or with a country based on author addresses.

Average of relative citations (ARC): This is an indicator of the *observed* scientific impact of papers produced by an entity (i.e., a country, a group of researchers) which takes into account inter-field variations in citedness and removes self-citations. Citations for each paper were counted over a 3-year window (i.e., for papers published in 1995, citations were counted during the 1995–1997 period). The exceptions are papers published in 2005, which have a citation window of two years (2005 and 2006), and in 2006, which have a citation window of one year.

Statistical Analyses

To establish whether there were significant differences between the scientific impact of various entities, a series of statistical tests were performed in SPSS. For each statistical test, the difference in scientific impact was considered to be significant at $p < 0.05$, very significant at $p < 0.01$ and highly significant at $p < 0.001$. Because data on scientific impact are not normally distributed, non-parametric tests were used. The Mann-Whitney U test was used to compare the ARC of two independent samples whereas the Z-test for two proportions was used to compare the proportion of papers of two independent samples that are in the 5% of papers with the highest citation counts.

Results

The results examine the research output of researchers supported by the NCIC over the 1994–2006 period (n=685, of which 679 published at least one paper during the period) to address three questions raised by the evaluation of the NCIC operating and program project grants: Has NCIC peer review process been successful in selecting outstanding Canadian researchers performing cancer research?; Has NCIC funding contributed to increasing the scientific performance, in terms of scientific impact, of the papers authored by the researchers it supported?; How do NCIC-supported researchers compare, in terms of scientific output, to their US counterparts which received support from the NCI? Data regarding the first two questions are presented in the following sections: *Total Scientific Production of NCIC-Supported Researchers* and *Scientific Production of NCIC-Supported Researchers in Cancer-Related Journals*. Data regarding the last question are presented in the following section: *NCIC-Supported Researchers Compared to NCI-Supported Researchers*.

Total Scientific Production of NCIC-Supported Researchers

Overall, just over half of scientific papers by researchers supported by the NCIC at any time over the 1994–2006 period were published with financial support from the NCIC: 12,244 out

of 22,793 papers by NCIC-funded researchers were published after their first year of NCIC funding up until the year after the end of the grant. As shown in Figure 1, the proportion of NCIC-supported papers by NCIC-researchers has been shrinking since 1995, as the number of non-supported papers (i.e., all other papers by NCIC researchers) has increased steadily, while the number of NCIC-supported papers published per year has been more stable.

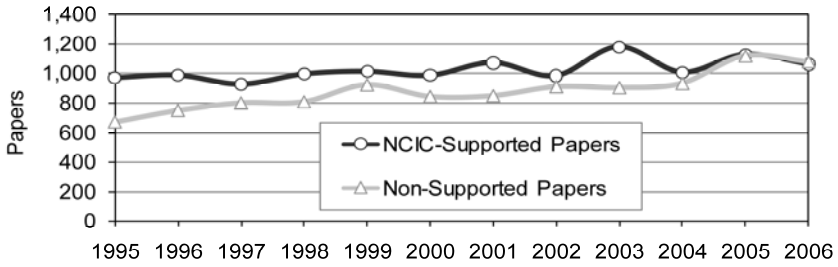

Figure 1. Yearly number of papers written by NCIC-supported researchers.

Besides the number of published papers, an important indicator of scientific performance is the actual impact that the published papers have had on the scientific community. The average of relative citations (ARC) is an indicator of the number of times an entity's papers (e.g., a group of researchers, a country) are actually cited relative to the world average and can therefore be used as a measure of *observed* scientific impact. All entities considered in this section score above the world average (i.e., ARC above 1).

Researchers who were supported by the NCIC at any one time between 1994 and 2006 have a substantially higher scientific impact than Canadian researchers (excluding papers by NCIC-supported researchers) in all fields (highly significant for the three sets of papers [all papers by NCIC researchers, NCIC-supported papers, and non-supported papers], $p < 0.001$; Figure 2). The impact of their papers when they were supported by the NCIC (i.e., those published one year after the start of the grant period until one year after the end of the grant period) was even greater than the impact of their papers authored without the support of the NCIC (highly significant, $p < 0.001$). However, the difference between the scientific impact of NCIC-supported papers and those without NCIC support was slightly smaller in the five most recent years; the latter set of papers had even greater impact than the former in 2006.

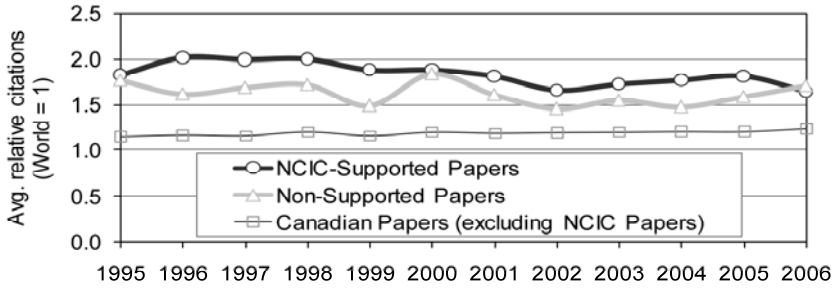

Figure 2. Average of relative citations of papers by NCIC and Canadian researchers.

Scientific Production of NCIC-Supported Researchers in Cancer-Related Journals

Researchers who received support from the NCIC at any given time from 1994 to 2006 contributed to nearly half of the Canadian papers published in the subfield of cancer research (delineated using specialist oncology journals based on the NSF classification, see Methods) from 1995 to 2006 and 26% of Canadian papers in cancer research were produced by these authors while they were receiving funds from NCIC (data not shown). Within this subfield, the scientific impact of papers by researchers who were supported by the NCIC (NCIC-

supported papers together with non-supported papers) was higher than the scientific impact of Canadian papers and the difference was highly significant ($p < 0.001$) for the three following sets of papers: all papers by NCIC researchers, NCIC-supported papers, and non-supported papers (Figure 3). As was the case for the total output of NCIC researchers (all subfields included), the impact of their papers in cancer research when they were supported by the NCIC was generally (for 10 out of 13 years) higher than when they were not supported, and the difference was significant ($p < 0.05$; Figure 3).

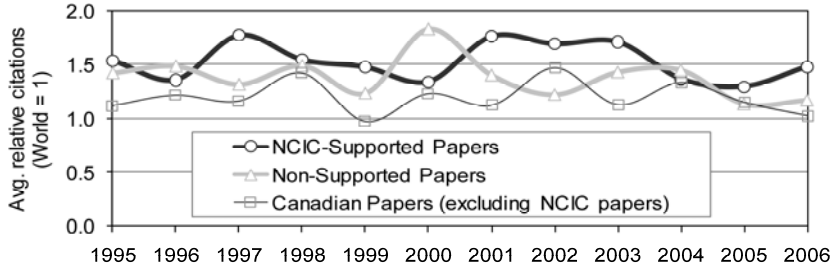

Figure 3. Average of relative citations of papers by NCIC and Canadian researchers in cancer research.

NCIC-Supported Researchers Compared to NCI-Supported Researchers

This section presents a comparative analysis of the output of researchers funded by the NCIC with that of researchers supported by the NCI. First, funding trends from both institutes are examined. Then, the output of NCIC-supported researchers is compared to a paired sample of NCI-supported researchers (random selection, see Methods).

For the 2001–2007 period, the average level of funding awarded per researcher by the NCI was 4.5 times that awarded by the NCIC (data not shown). On average, NCI-supported researchers received about US$600,000 at Purchasing Power Parity (PPP) in 2006 compared to about US$150,000 PPP for NCIC-supported researchers. It should be noted, however, that NCI grants also include PI salaries whereas NCIC grants do not provide for personal salary support of the PIs and/or co-applicants. Furthermore, based on financial data provided by both institutes, grants awarded to US researchers by the NCI covered, on average, 3.85 years of research out of 7, compared to 3.64 years out of 7 for Canadian researchers who received NCIC grants. Thus, although the amount of funding is disproportionately different, the length of the funding is highly comparable.

The scientific production of the two groups was similar throughout the 2000–2006 period, although the output of NCI-supported researchers experienced a slightly stronger increase (net increase of 25%) than the output of NCIC-supported researchers (net increase of 21%) (Figure 4). In total, NCIC-supported researchers published nearly as many papers as did NCI-supported researchers (11,019 versus 11,794), despite a level of funding that is nearly five times smaller (considering only financial support from the NCI and the NCIC).

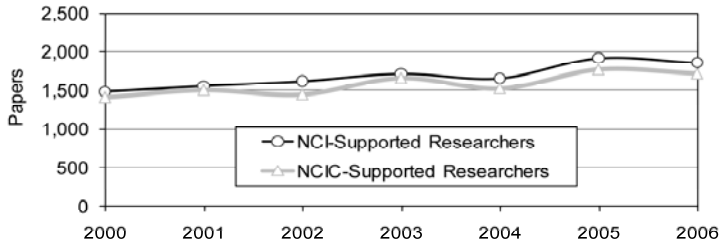

Figure 4. Yearly scientific output, NCIC population versus NCI-paired sample.

On the other hand, the papers produced by NCI-supported researchers have had stronger scientific impact than the papers published by NCIC-supported scientists (highly significant, $p < 0.001$), with the exception of the year 2000, during which both groups had similar ARC values (Figure 5). For the 2000–2006 period, papers by NCI-supported researchers received, on average, about 120% more citations than the average world paper, while those of NCIC researchers received about 70% more citations. Both groups produced papers that achieved greater scientific impact than their respective nation's average paper (highly significant, $p < 0.001$, Figure 5).

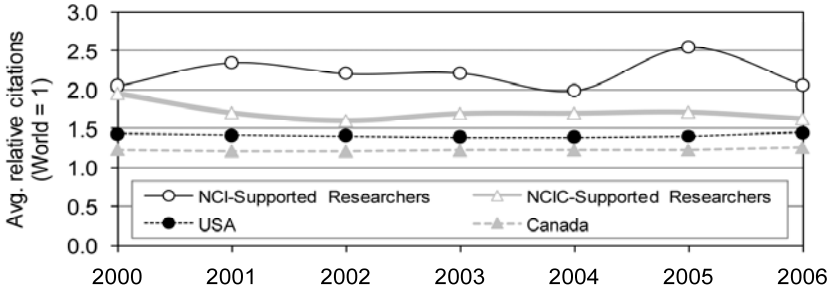

Figure 5. Average of relative citations of papers, NCIC population versus NCI-paired sample.

Discussion

A bibliometric approach was used to assess the performance of a funding program, the NCIC Operating and Program Project Grants, as it relates to (1) the successful attribution of NCIC grants to outstanding Canadian researchers performing cancer research, (2) the contribution of NCIC funding to increasing the scientific performance of the papers authored by the researchers it supported, and (3) the comparison of the scientific output of NCIC-supported researchers with that of their US counterparts that received support from the NCI.

To address these questions, the papers authored by Canadian researchers who received funding from the NCIC at any given time from 1994 to 2006 had to be classified as *NCIC-supported papers* or *non-supported papers*. This was achieved by tagging papers published between the year after the start of an NCIC grant and the year after the end of an NCIC grant as *NCIC-supported papers* and the balance as *non-supported papers*. As the study period begins in 1994, papers were considered as of 1995 to allow for the one-year lag.

The choice of the time window for tagging papers as being supported or non-supported is arbitrary and may lead to the inclusion of a number of *non-supported papers* into *NCIC-supported papers* and vice versa. However, it is our belief that this approach, when used with a large number of researchers (e.g., in this project n=685 PIs) over a long time period (12 years), provides a representative samples of papers published with the financial support of the NCIC. Other studies have used funding acknowledgement for comparing sets of supported and non-supported papers (Cronin and Shaw, 1999; Rangnekar, 2005). Although this approach can identify more precisely supported papers, its recall rate can be as low as 50% since authors of scientific papers do not systematically acknowledge sources of funding (Lewison & Carding, 2003). This can, in turn, lead to the inclusion of supported papers into non-supported papers. Because the papers' acknowledgements were not indexed in the database used to perform this study and because it was not feasible, given the large data sample used (> 20,000 papers), to manually download each paper and check the acknowledgements section, the former approach was chosen. As neither of these two approaches is perfect, care must be taken when interpreting differences observed between supported and non-supported papers. To strengthen the analysis, future work could examine how supported and non-supported papers classified based on grant periods compare to supported and non-supported papers classified based on funding acknowledgment. To

perform this comparison, the dataset used in this study could be reduced to a manageable size when capturing data on funding acknowledgement (e.g., using the three most recent years).

Researchers receiving NCIC funding also receive grants from other funders such that papers considered as *NCIC-supported* may have been also partly supported by other sources of funding. In fact, the NCIC awarded about 25% (over $60M in 2005) of all funds for cancer research in Canada distributed by members of the Canadian Cancer Research Alliance (CCRA), which includes federal and provincial governments as well as voluntary funding organizations like NCIC (Canadian Cancer Research Alliance, 2007). As such, the authors of the present paper do not assume that the performance of *NCIC-supported papers*, in terms of scientific impact, can be entirely attributed to the NCIC.

During the 1995–2006 period, just over half of the total scientific production of researchers who received funding from the NCIC at any time between 1994 and 2006 was produced with financial support from the Institute. These researchers were authors on nearly half of the Canadian papers published in cancer-related journals from 1995 to 2006; moreover, the results of this study suggest that a quarter of Canadian papers in these journals were published with NCIC support. There is therefore no doubt that the NCIC's financial contribution (25% of Canada's funding for cancer research; CCRA, 2007) resulted in a proportionately sized scientific output (25% of Canada's production in cancer journals) by the researchers it supported.

The NCIC strives to advance cancer research by supporting outstanding researchers selected through peer review. A number of studies found a positive correlation between peer ratings and citation scores of researchers in grant competitions (Lawani and Bayer, 1983; McAllister, Anderson and Narin, 1980). When these measures of scientific excellence point in the same direction, the selection of awardees is generally regarded as being more reliable than when only peer ratings are considered (King, 1987). Because peers are making increased use of bibliometric indicators in rating researchers, as appears to be the case among members of the NCIC's review panel (according to NCIC representatives), positive correlations between peer ratings and citation scores will appear even more frequently in the future, especially if reviewers make formal use of bibliometric indicators (using measures provided by statistical offices, companies, or specialized university departments) rather than using their tacit knowledge about how well a researcher performs with respect to these indicators. If the NCIC's peer-review process has been efficient in selecting leading researchers, papers authored by successful applicants while they were not financially supported by the NCIC are expected to have, on average, greater scientific impact than other Canadian papers. In the present case, the results, using the average of relative citations (ARC), support this hypothesis when considering papers in all fields as well as when limiting the analysis to papers in cancer journals. Considering that the scientific impact of NCIC researchers is significantly higher than that of other Canadian researchers (i.e., excluding papers by NCIC-funded researchers), who themselves score above the world average, it is quite clear that researchers funded by the NCIC are of high calibre.

The analysis of the *observed* scientific impact also supports the hypothesis that the funds provided by the NCIC contributed to increasing the scientific performance of successful applicants: the ARC of NCIC-supported papers is significantly higher than that of non-supported papers. While some citations in scientific papers may be non-essential, they generally substantiate a finding or assumption, and in cases where the cited work is incorrect or outdated, the act of citing still reflects a stimulatory effect of the cited work (King, 1987). Thus, the fact that the ARC of NCIC-supported papers is higher than that of non-supported papers suggests that the NCIC contributed to increasing the usefulness, for the scientific community, of the knowledge produced by NCIC awardees.

The results of the current study provide preliminary evidence of the successful selection of outstanding Canadian researchers by the NCIC peer-review panel and of the NCIC's positive effect on the scientific impact of the researchers it supported. However, a more detailed analysis of the research outputs and of funding at the level of individual NCIC researchers is required to fully understand the meaning of these results. The current study compared the scientific impact of NCIC-supported researchers with that of other Canadian researchers using scientific articles published in cancer journals. It was recently shown that only 50% of cancer-related research is being published in specialist oncology journals and that oncologists publish their most influential work in more general journals or in journals covering other specialities (López-Illescas, Moya-Anegón & Moed, 2008). Thus, it is fairly safe to assume that about 50% of cancer research was sampled in this study, such that the scientific impact of NCIC-supported researchers would likely remain above that of other Canadian researchers if the field was expanded to include cancer-related research not published in specialist oncology journals. Nevertheless, it could be interesting to compare the results of the current study with those obtained with an expanded dataset.

Another issue relates to the type of research supported by the NCIC or the type of research generally performed by the researchers it supported. For instance, if NCIC supports a specific type of cancer research, such as clinical investigations, then efforts should be made to examine the citation habits of researchers performing basic research as opposed to clinical research. Indeed, it might be found that clinical research is generally more cited than basic research, in which case the difference in the scientific impact of NCIC-supported researchers and other Canadian researchers would relate to citation patterns rather than to the scientific excellence of the performed research.

Compared to a random sample of researchers funded by the NCI of the same size as the population of NCIC-funded researchers, NCIC-supported researchers produced, on average, nearly as many papers per year from 2000 to 2006, despite receiving about five times less funding per researcher (taking into account only financial support from the NCI and the NCIC). Considering that the NCI provides approximately 70% of the total public funding for cancer research in the US (Jönsson & Wilking, 2007), compared to the NCIC's contribution of about 25% in Canada (Canadian Cancer Research Alliance, 2007), NCIC researchers likely have access to only about one sixth of the total public funds available to NCI researchers. Since, on average, NCIC researchers likely have less financial resources, the present result could indicate that they are more productive (more papers produced per dollar investment) than NCI researchers.

However, as investigations into the causes of cancer and potential cures are becoming more and more high-tech, the cost of cancer research has increased significantly, requiring ever greater investment on the part of funding bodies. As such, access to greater financial resources may not directly translate into more scientific publications, but might instead result in projects that could not have been accomplished otherwise (e.g., large-scale, long-term or innovative projects, such as the Human Genome Project). The higher scientific impact of the papers produced by NCI researchers (compared to those published by NCIC researchers) might then be explained, at least in part, by NCI researchers performing more research projects of this type, which likely attract more attention from the scientific community. Here again, more data is required on the type of research performed and on the distribution of available resources (e.g., infrastructure, financial and human resources) among researchers to provide a detailed interpretation of the results. In spite of the limitations discussed above, this study demonstrates the usefulness of bibliometrics as a performance measurement tool for the evaluation of research and of funding programs.

Acknowledgments

The authors which to thank the National Cancer Institute of Canada (NCIC) and the US National Institute of Health's National Cancer Institute (NIH-NCI) for having provided listings of the principal investigators they have funded and Dr. Henk F. Moed of CWTS for his constructive comments on this study. This work was supported by the NCIC.

References

Archambault, É. & Côté, G. (2008). Better bibliometrics needed for greater accountability of research expenditures. *Research Money*, 12, 8.

Canadian Cancer Research Alliance. (2007). *Cancer Research Investment in Canada, 2005*. Retrieved January 16, 2009 from: http://www.ccra-acrc.ca/PDF%20Files/CCRA_07_EN_Full%20Report.pdf.

Cronin, B. & Shaw, D. (1999). Citation, funding acknowledgment and author nationality relationships in four information science journals. *Journal of Documentation*, 55, 402-408.

Hicks, D., Tomizawa, H., Saitoh, Y. & Kobayashi, S. (2004). Bibliometric techniques in the evaluation of federally funded research in the United States. *Research Evaluation*, 13, 78-86.

Jönsson, B. & Wilking, N. (2007). Cancer research and development and the drug development process. *Annals of Oncology*, 18: iii49-iii54.

King, J. (1987). A review of bibliometric and other science indicators and their role in research evaluation. *Journal of Information Science*, 13, 261-271.

Larivière, V., Archambault, É., Gingras, Y. & Vignola-Gagné, É. (2006). The place of serials in referencing practices: comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, 57, 997-1004.

Lawini, S.M. & Bayer, A.E. (1983). Validity of citation criteria for assessing the influence of scientific publications: new evidence with peer assessment. Journal *of the American Society for Information Science*, 34: 59-74.

Lewison, G. & Carding, P. (2003). Evaluating UK research in speech and language therapy. *International Journal of Language & Communication Disorders*, 38, 65-84.

López-Illescas, C., de Moya-Anegón, F. & Moed H.F. (2008). The actual citation impact of European onclogical research. *European Journal of Cancer*, 44, 228-236.

Martin, B.R. (1996). The use of multiple indicators in the assessment of basic research. *Scientometrics*, 36, 343-362.

McAllister P., Anderson, R.C. & Narin, F. (1980). Comparison of peer and citation assessment of the influence of scientific journals. *Journal of the American Society for Information Science*, 31, 147-152.

Moed H.F. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer.

National Cancer Institute of Canada. (2005). NCIC 2015: *Driving excellent cancer research to improve cancer control*. Retrieved January 16, 2009 from: http://www.ncic.cancer.ca/~/media/NCIC/Files%20list/English%20files/ PDF/Strategic%20plan%20-%20EN%20-%20PDF%20-%202005_530236530.ashx.

National Cancer Institute of Canada. (2009). *NCIC Website*. Retrieved January 16, 2009 from: http://www.ncic.cancer.ca/About%20us%20and%20news/History%20and%20mission.aspx.

National Science Foundation. (2006). *Science and Engineering Indicators*. Retrieved January 16, 2006 from: http://www.nsf.gov/statistics/seind06/pdfstart.htm.

Rangnekar, D. (2005). Acknowledged: analysing the bibliometric presence of the Multiple Sclerosis Society. *Aslib Proceedings*, 57, 247-260.

Trochim, W.M., Marcus, S.E., Mâsse, L.C., Moser, R.P. & Weld, P.C. (2008). The evaluation of large research initiatives: a participatory integrative mixed-methods approach. *American Journal of Evaluation*, 29, 8-28.