

RESEARCH METHODS AND TECHNOLOGY RESEARCH ARTICLE

Feasibility of common bibliometrics in evaluating translational science

M. Schneider^{1*}, C. M. Kane², J. Rainwater³, L. Guerrero^{4,5}, G. Tong⁶, S. R. Desai⁷
and W. Trochim²

¹ School of Social Ecology, Institute for Clinical and Translational Science, University of California, Irvine, CA, USA

² Cornell Office for Research on Evaluation, Weill Cornell Medical College, Ithaca, New York, USA

³ Clinical and Translational Science Center, University of California Davis, Davis, California, USA

⁴ General Internal Medicine and Health Services Research, David Geffen School of Medicine at UCLA, Los Angeles, California, USA

⁵ UCLA Clinical and Translational Science Institute, University of California, Los Angeles, CA, USA

⁶ Clinical and Translational Research Institute, University of California, San Francisco, CA, USA

⁷ Division of Cardiology, UC San Diego Department of Medicine, San Diego, California, USA

Journal of Clinical and Translational Science (2017), 1, pp. 45–52 doi:10.1017/cts.2016.8

Introduction. A pilot study by 6 Clinical and Translational Science Awards (CTSAs) explored how bibliometrics can be used to assess research influence.

Methods. Evaluators from 6 institutions shared data on publications (4202 total) they supported, and conducted a combined analysis with state-of-the-art tools. This paper presents selected results based on the tools from 2 widely used vendors for bibliometrics: Thomson Reuters and Elsevier.

Results. Both vendors located a high percentage of publications within their proprietary databases (>90%) and provided similar but not equivalent bibliometrics for estimating productivity (number of publications) and influence (citation rates, percentage of papers in the top 10% of citations, observed citations relative to expected citations). A recently available bibliometric from the National Institutes of Health Office of Portfolio Analysis, examined after the initial analysis, showed tremendous potential for use in the CTSA context.

Conclusion. Despite challenges in making cross-CTSA comparisons, bibliometrics can enhance our understanding of the value of CTSA-supported clinical and translational research.

Received 15 July 2016; Revised 4 August 2016; Accepted 11 October 2017; First published online 31 January 2017

Key words: Metrics, CTSA, Publications.

Introduction

Research publications are effectively the currency of scientific work. They represent the public cumulative record of science, documenting empirical results and providing a forum for theorizing, debate, and the gradual advance of scientific knowledge. Although publications are not the end point of scientific research, it is impossible to imagine translating research into knowledge or practice without the manifestation of research in publications. Because of this, no serious evaluation of the influence of translational research can ignore the importance of publications as intermediate outcomes.

The National Institutes of Health (NIH)-funded Clinical and Translational Science Award (CTSA) initiative [1] is perhaps the largest single cross-

institutional scientific center grant initiative currently funded by the NIH. Data from the Department of Health and Human Services appropriations justification (<https://ncats.nih.gov/files/FY15-justification.pdf>) indicate that NIH allocated \$471,719,000 in the 2015–2016 fiscal year to support CTSA institutions that have the broad mission to encourage the translation of biomedical research from discovery to use and impact in the population. Evaluating success of the CTSA Consortium is an ambitious and complex endeavor [2–6], and metrics related to research publications supported by the Consortium have been frequently discussed as an essential element of that evaluation.

Because assessing outcomes such as scientific productivity, efficiency, influence, and collaboration is of central importance for the CTSAs, we conducted a pilot study on the feasibility and utility of some available bibliometric approaches. Each of the 62 CTSAs has a core function that provides internal evaluation designed to monitor and improve operations. While each internal evaluation team develops its own approaches to providing internal evaluative feedback, increasingly the emphasis of the CTSA evaluation effort has turned to evaluating cross-center outcomes and to assessing the degree to which the CTSA initiative is making progress in advancing biomedical science [3, 4, 7].

* Address for correspondence: M. Schneider, Ph.D., 258 Social Ecology I, University of California, Irvine, Irvine, CA 92697, USA.

(Email: mls@uci.edu)

What are Bibliometrics?

The quantitative analysis of publications is known as bibliometrics. Bibliometrics can be defined as “the application of quantitative analysis and statistics to publications such as journal articles and their accompanying citation counts” [8]. Although the term “bibliometrics” was apparently not used until 1969, researchers had been counting written research products as far back as the ancient library in Alexandria [9]. Whereas the number of publications and citation counts offer an indication of the productivity of a researcher or institute, they do not indicate how influential a publication is, or how much it has been incorporated into subsequent scientific thought and work. The introduction of the measurement of citations of scientific publications (ie, the number of times an article has been cited by subsequent publications) through the development of the Science Citation Index made it possible to assess the degree to which a publication influences others [10, 11]. Subsequent indexing of publications and citations in large databases such as PubMed, Psychological Abstracts, the Web of Science, and Scopus, have enabled the development of a broad array of bibliometric indices and metrics that assess research productivity, quality, influence, collaboration, and multidisciplinary.

Some researchers have used bibliometric analyses beyond citation counts to better understand and evaluate “big science.” Within the HIV/AIDS clinical research networks, it was found that researchers were producing highly recognized work, engaging in extensive interdisciplinary collaborations, and having an impact across several areas of HIV-related science [12]. Bibliometric analyses including publication quality and transdisciplinarity were included as part of a pilot evaluation of the Transdisciplinary Tobacco Use Research Centers [13], and they were used in the evaluation of the impact of the National Breast Cancer Foundation [14]. Yet, the emerging inter-related specialties of bibliometrics, informetrics, and scientometrics [15] are complex and not without their fallibilities [16].

Academic disciplines vary in their citation culture, and low citation rates in some may not be the result of a lower average number of citations per paper. Instead, the lower rates may be due to the low fraction of citations that reference papers indexed in the specific archive used to generate the bibliometric [17, 25]. Similarly, a focus on a journal’s impact factor does not necessarily signal the significance of a particular article [18, 19], and paper-level impact measures may not capture whether the papers actually pushed science forward with novel ideas and groundbreaking research [20]. Some have gone as far as to argue that increased attention on bibliometric indices distorts researcher incentives and contributes to the increase in the number of papers that are retracted [21].

Cross-CTSA Pilot Study

The primary aim of this pilot study was to assess the feasibility of collecting, aggregating, analyzing, and utilizing bibliometric data across multiple CTSA institutions. The guiding research questions were as follows:

- How feasible is it for CSA’s to compile comparable lists of research publications in a standardized format?
- What proportion of the supplied publications can be matched to the bibliometric databases, and what evidence is there regarding potential bias due to incomplete matching?
- How comparable are the results of bibliometric analyses from the 2 vendors?
- What does the bibliometric analysis show with respect to (a) productivity and (b) influence?

A secondary aim was to compare the user experience and bibliometrics associated with the services offered by Elsevier to those associated

with the services offered by Thomson Reuters. The substantive results provide a cross-CTSA baseline that may be useful going forward.

Bibliometrics have emerged as an area of interest among the CSA’s for several reasons. First, the reporting required by the NIH has featured a publication count each year. Many CSA sites, independently but in parallel, have consequently devoted considerable resources to tracking publications, and have begun to explore bibliometrics that are more sophisticated than simple counts. Second, a cross-CTSA working group to develop common metrics further fueled interest in identifying methods for using information derived from publications to evaluate the annual progress of individual CSA’s.

On a macro level, bibliometrics offer an opportunity to provide evidence for the impact of the national consortium of CSA’s. However, the absence of a common methodology across individual CSA’s has limited progress toward this goal. Multiple CSA’s simultaneously pursue their own analytic strategies, resulting in data sets that are not comparable owing to differences in methods of data collection and computation. As one example, CSA’s have generally aligned with either Elsevier or Thomson Reuters to compute bibliometrics, resulting in nonequivalent sets of metrics. Each publishing company employs its own proprietary publication database, metrics, and algorithms.

Methods

This pilot study was a collaboration of 6 CSA’s and was designed to explore the feasibility of conducting bibliometric analysis of research publications that would both serve the internal monitoring needs of each center and be capable of cross-center aggregation and summary. The 6 participating institutions included the 5 CSA’s funded within the University of California system (Davis, Irvine, Los Angeles, San Diego, and San Francisco) as well as Weill Cornell Medical College in New York. The pilot project was coordinated by the Weill Cornell Medical College team, which also took the lead in data analyses. All 6 participating institutions contributed to all aspects of the project, including data preparation, analysis, interpretation, and dissemination. The pilot project took place from March 2014 to May 2016.

This CSA bibliometric collaboration worked with 2 of the largest and most comprehensive vendors of bibliometric information: Thomson Reuters, the owners of the Web of Science database, and Elsevier, the owners of the Scopus database. All 6 CSA’s provided a list of their CSA-related publications in a standardized format. These lists were aggregated and sent to both vendors. The vendors provided comprehensive bibliometric results in a format that enabled subsequent within and across-center analysis. For purposes of this pilot project both vendors agreed to provide results at no cost. As costing models for their services change considerably over time and are often a factor in determining feasibility, cost issues are not considered in this paper and the interested reader is encouraged to contact the vendors for current pricing.

Thomson Reuters provided the results in the form of a summary spreadsheet and a Microsoft Access database file (.mdb), along with the software tools needed for manipulating the database and generating reports (Xite). The Excel spreadsheet included 1 line per publication, with the carried descriptive variables, the institutional ID, and a number of publication-level bibliometric indicators. The Access database included tables for each publication, author, citing and cited publication, and publication abstract. Elsevier provided the results through the SciVal web interface, which offers user-friendly query features but does not provide direct access to the underlying database, although it is relatively easy to export the underlying data for any query should more extensive analyses be desired.

During the write-up phase of this project, the team became aware of a new bibliometric service being offered through the NIH Office

of Portfolio Analysis (OPA) called iCite. Because of the relevance of such a system for NIH-related research in particular, we felt it was important to include this new arrival in this project. We submitted the same database given to Elsevier and Thomson Reuters to the iCite system for a basic bibliometric analysis. However, because their analysis was done so much later than the main analysis for this project and would be affected by the longer passage of time (eg, higher citation counts), we did not integrate their metrics into the primary analysis. Instead, because of its potential importance, especially for NIH-related work, we include in this paper only the results from their key metric as described in Bibliometrics section.

Aggregation of Publication Data

Each of the 6 CTSA institutions provided a list of publications for their own center using a standardized spreadsheet format. To the extent possible, each institution provided a comprehensive list of publications from the inception of their center through the most recent Annual Progress Report (APR, operation reports submitted to the NIH annually) to the NIH before an arbitrary deadline of April 1, 2014. The lists included all publications that met the requirements for inclusion in the CTSA APR at the time of submission to the NIH. There were 4201 publications in total, with individual CTSAAs responsible for from 100 to over one thousand publications each.¹

The spreadsheet from each institution included 4 columns corresponding to the following publication characteristics, as available: PubMed Identification Number, PubMed Central Identification Number (PMCID), Digital Object Identifier, Institutional ID (a unique identifying number for each of the 6 institutions). In addition, 3 columns were dedicated to CTSA programs that provided specific types of support for the underlying research: education and training, pilot funding, or clinical research services. One of the participating CTSAAs was unable to characterize the publications by source of CTSA support, so these data were available for 5 out of the 6 CTSAAs.

Match Rates

Match rate refers to the extent to which each vendor's database of publications included publications from the list provided by the 6 CTSAAs. Deriving match rates for the 2 vendors required slightly different processes in working with each of them. Thomson Reuters provided a supplementary Excel file with additional metrics requested by our collaborative research group that contained a list of those articles which they were able to match to records in Web of Science. Each article was labeled according to institution and subgroup, permitting us to compute match rates in comparison to the original number of articles submitted to the vendors. Later correspondence yielded a separate Excel file containing citations of all those publications for which Thomson Reuters was unable to find a match in the Web of Science. In contrast, Elsevier provided match rates by institution and subgroup and also supplied a list of the records matched in Scopus and those not matched and the reason why the publication was not matched. Thomson Reuters and Elsevier matched the publications with citations in their respective proprietary databases using 1 or more of the identifiers provided with each publication (eg, PubMed Identification Number, PMCID, or Digital Object Identifier). The extent of match rates and their effect on this study are discussed in the Results section.

Bibliometrics

A wide variety of bibliometrics was provided by the vendors in the final results. The final set was negotiated by the collaborative team and the

vendors. In all, 29 metrics were available from the Thomson Reuters materials [8], and 24 metrics were provided by Elsevier [22]. After reviewing the available metrics for redundancy, interpretability, and utility, we arrived at a list of 4 unique measures that were available from both vendors. Table 1 shows the 4 bibliometrics selected for reporting, along with the specific terminology and definitions relevant to each vendor. In addition to a simple count of publications, a measure of productivity, and a simple count of first-generation citations, a measure of influence, we selected 2 metrics that reflect publication influence relative to the field: percentage of publications situated within the top 10% of the field based on number of citations, and what we have labeled the comparative citation ratio (CCR).

The CCR compares the citation rates of a target set of publications with the citation rates of a "comparable" set of similar publications. Virtually every major vendor of bibliometrics has 1 or more CCR metrics, and all have the same fundamental structure:

$$\text{CCR} = \frac{\text{Observed citation rate}}{\text{Expected citation rate}}$$

where the expected citation rate is computed from a "comparable" set of publications to the target publication. Although the observed citation rates are identical across vendors, the different CCR metrics differ in terms of how the expected citation rate is calculated. A CCR metric for a single publication equals 1 if the publication has the same number of citations that would be expected from the comparison group. It is >1 if it has received more citations to date than the comparison, and it is <1 if it has received fewer citations than expected. Because it is a ratio it is meaningful to say about a publication with a CCR value of 2.3 that the publication received 2.3 times as many citations as would be expected, regardless of how many citations the publication actually received. Moreover, because it is a ratio, one can average the CCR values for a set of publications to get an estimate of their relative influence as a group. One advantage of this type of metric is that each publication is compared with a comparison group that is constructed from publications similar to that publication, but the result can meaningfully be aggregated across publications from different fields or disciplines (because each publication is standardized to or adjusted for its local field-based norm).

In this study, we examined 1 CCR from each of the 2 vendors. For Elsevier, we used the Field-Weighted Citation Impact (FWCI) [22]. To compute the expected citation rate for a target publication, those publications in the Scopus database that have the same publication year, publication type, and discipline, as represented by the Scopus journal classification system, are used as the comparison set. The time frame used for computing the average citation rate is the publication year plus the following 3 years. If the publication's journal were assigned to more than 1 discipline, the harmonic mean would be used to compute the expected citation rate. Thomson Reuters did not have a discipline-level CCR available, although they did have a similar metric at the journal level: the category expected citation rate. To develop a metric that was at a comparable level to the FWCI, we calculated what we term a Category-C Index or (Cat-C Index for short). The Cat-C Index for a set of publications is the ratio of the sum of the citations for that set of publications divided by the sum of the category expected citation rates.

Results

Feasibility of Compiling Publication Lists

Table 2 shows the number of publications that were reported by each participating CTSA overall and in the 3 program categories: pilot studies, education, and clinical research resources. The total number of publications reported by each CTSA ranged from 101 to 1319. Out of the 5 institutions that classified publications by program, 4 reported the most publications generated by investigators who used the clinical

¹As a result of changes in the APR requirements over time, we observed some variability in publication reporting volume, especially in the initial years of the CTSA.

Table 1. Bibliometrics included in the cross Clinical and Translational Science Award study

Metric	Number of publications	Average number of citations per publication	Percentage of publications in the top 10% per citations	Comparative citation ratio
Used for Terminology	Productivity	Influence	Relative influence	Relative influence
<i>TR</i>	Counts of papers	Average citations per paper	Publications in top journal percentiles	Cat-C Index
<i>Elsevier</i>	Scholarly output	Citation per publication	Outputs in top percentiles	Field-Weighted Citation Impact
Description*				
<i>TR</i>	Paper counts, which measure productivity, are the most basic bibliometric measure and provide raw data for all citation analysis	Citations per paper is computed by dividing the sum of citations to some set of papers for a defined time period by the number of papers (paper count)	A paper percentile is determined by taking the year and journal category of the paper, creating a citation frequency distribution for all the papers in that year and category, and determining the percentage of papers at each level of citation. The percentile then indicates how a paper has performed relative to others in its field	The Cat-C Index for a set of publications is the ratio of the sum of the citations for that set of publications divided by the sum of the average expected citations for the set of publications where that expected average is computed across journal categories
<i>Elsevier</i>	Scholarly output in SciVal indicates the productivity of an entity: How many publications does this entity have indexed in Scopus?	Citations per publication in SciVal indicates the average citation impact of each entity's publications: How many citations have this entity's publication received on average?	Outputs in top percentiles in SciVal indicates the extent to which an entity's publications are present in the most-cited percentiles of a data universe: How many publications are in the top 1%, 5%, 10%, or 25% of the most-cited publications?	Field-Weighted Citation Impact in SciVal indicates how the number of citations received by an entity's publications compares with the average number of citations received by all other similar publications in the data universe

TR, Thomson Reuters; Cat-C Index, Category-C Index.

iCite's Relative Citation Ratio not included in this table.

*Except for the Cat-C Index, which was computed specifically for this project, all descriptions are taken from white papers released by Elsevier [22] and Thomson Reuters [8].

research unit. The proportion of publications generated by recipients of pilot awards varied considerably, from 4% (institution 2) to 39% (institution 4). Institutions that had been receiving CTSA funds for longer tended to report higher number of publications, but there was still considerable variability among the 3 sites that had all been receiving CTSA support for 4 years (institutions 2, 4, and 5). The 3 older institutions, defined as having received the grant award for more than 5 years, reported 146 publications on average each year, compared with 62 among the younger institutions.

Match Rate Analysis

Overall match rates were very high for both vendors (see Table 2), although Elsevier consistently located a higher proportion of publications

as compared with Thomson Reuters. When examined according to publication program subsets (education support, pilot study funding, and clinical research services), the overall higher match rate for Elsevier persisted (data not shown), and the subset of publications attributed to the education and training program had the lowest match rate according to Thomson Reuters (90%).

Examining the lists of records not matched offers some interesting information about the types of articles that might be excluded depending on vendor. A substantial proportion of the articles that were unmatched by Thomson Reuters ($n = 252$) were published in Open Access journals or were published in journals that were highly specialized and might be expected to have minimal reach, for example, *Journal for Social Action in Counseling and Psychology* and *Journal of Pathology Informatics*. In addition to providing a list of unmatched

Table 2. Publications submitted for analysis and percent matched [n (% matched Thomson Reuters/% matched Elsevier)]

	Institution						
	ALL	1	2	3	4	5	6
Age (y)		7	4	8	4	4	8
Size		Large	Small	Small	Large	Medium	Large
No. of publications	4201 (94/99)	842 (94/99)	245 (93/99)	1291 (91/99)	403 (96/99)	101 (94/99)	1319 (96/100)
Education	537	179	25	296	27	10	—
Pilot	407	91	11	118	159	28	—
CRS	682	194	90	107	261	30	—

CRS, clinical research services; CTSA, Clinical and Translational Science Awards.

Institution 6 was unable to provide the breakdown of publications by CTSA subset. Age is the number of years the CTSA had been receiving CTSA funding. Size category corresponds to the definitions provided in the National Center for Advancing Translational Sciences Funding Opportunity Announcement 2014: small (total anticipated CTSA <\$6 million), medium (total anticipated CTSA \$6–8 million), and large (total anticipated CTSA >\$8–10 million).

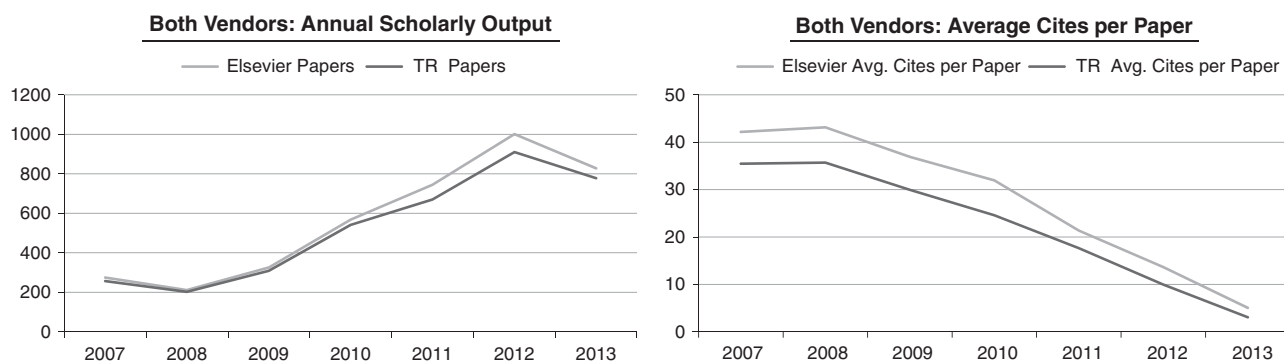


Fig. 1. Scholarly output and average cites per paper for Thomson Reuters (TR) and Elsevier: all publications (2007–2013).

publications ($n = 34$), Elsevier provided reasons such as not yet in Scopus ($n = 3$), journal year not indexed in Scopus ($n = 7$), journal not indexed in Scopus ($n = 22$), and book title ($n = 1$). Overall, then, it would appear that unmatched publications are likely to have relatively low citation counts either because they have been published in small-circulation journals or because they have not been in publication long enough to generate citations.

Research Productivity and Influence of CTSA

Fig. 1 shows the cumulative publication counts, or “scholarly output” and the average cites per paper for all institutions and both vendors over time. The graphs in Fig. 1 show several aspects worth noting. First, the results from the 2 vendors track each other well, with Elsevier consistently showing higher number of publications, and a higher number of average cites per paper. In Fig. 1, we also see an intuitive pattern regarding scholarly output: little or no output at first, followed by a rapid increase in productivity as each CTSA matured into a fully operational institution, and then a small dip in the most recent year, most likely an artifact of the arbitrary cut-off date for this project, which caught some CTSA institutions in the middle of their reporting cycle. On the right, we can see the normal downward trend of average cites per paper over time. This is due in part to the simultaneous increase in the overall number of publications and to the increasing number of citations in the first few years after those papers were published. Subsequently, the number of citations begins to decline as expected because the papers are increasingly recent in time and have not built up their expected citation base. The average cites per paper over time adjusts for the increasing number of publications in the first few years and shows the more traditional decline in citations for publications over time.

Both vendors traditionally use calendar year for time series data on bibliometric results. However, as we began to review our productivity data by institution, it became apparent that it would be far more useful and accurate to look at “project” year versus calendar year. As the 6 participating CTSA began operating in different calendar years, we adjusted our analysis to compare each vendor’s scholarly output by project year, that is, year 1 of operation, year 2, and so on.

Fig. 2 shows scholarly output for both vendors and all institutions by project year. As with Fig. 1, the results from the 2 vendors track each other closely with Elsevier (dashed lines) yielding a greater number of publications for all institutions than Thomson Reuters (solid lines). With the exception of institution 3, all institutions showed little or no publications in project year 1, with consistent increases in scholarly output over time. Institution 3 stands in contrast to these predominant trends, and illustrates one limitation of conducting this analysis by project year for this data set. As noted earlier in Table 2, institutions 3 and 6 are the 2 “oldest” CTSA in the group. In the foremost years of the CTSA initiative, the criteria for including publications in the annual

report to the NIH were far less restrictive. These earlier, more flexible reporting parameters were more open to interpretation and allowed for the variability in publication counts seen in the older institutions’ first project years. However, as annual CTSA reporting requirements have become more focused and stable over time,² we would expect to see far less variability if this type of bibliometric analysis were to be expanded to the current 60+ CTSA, or even repeated with the same set of 6 institutions in subsequent years. For instance, if we were to look solely at project year 4 in Fig. 2, again with the exception of institution 3, the volume of scholarly output correlates directly with institution size, with the 3 largest institutions reporting in the top half of the group, and the 3 smaller institutions reporting in the bottom half of the group for publication counts in that project year.

The need to adjust our analysis for institution project year versus calendar year is strictly limited to raw productivity and influence measures such as scholarly output and average number of citations. When measuring relative influence, bibliometrics such as Cat-C Index, FWCI, and percentile rankings are all structured as comparative measures, weighted by and compared with a greater set within each database (by field, journal, etc.). Accordingly, these comparative metrics are presented by calendar year.

Comparative Influence of CTSA Research

Fig. 1 showed the average number of citations for all papers. A better indicator of the influence of a set of papers than publication counts or average citation counts is to compute the proportion of papers that are in the top 10% of those cited in their respective field. If a set of papers has average influence, we would expect that about 10% of them are in the top 10% of cited articles.

The papers in all of the CTSA institutions tended to be over-represented in the top 10% of cited papers, an indicator that they exert a disproportionately high influence. The set of all papers in this collaborative analysis scored 32.40 using Elsevier’s “Outputs in Top 10 percent—Percentages” and 26.68 using Thomson Reuters’ “Percent of Papers (10%).” In other words, CTSA papers tend to be cited well above the expected percentage rate. This comparative metric adjusts for the tendency to have different citation rates in different fields and specialties. It is also noteworthy that the absolute percentage rates differ between the 2 vendors. Elsevier’s higher rates for each institution is likely because it indexes more journals. That said, the 2 sets of vendor metrics tracked well by institution.

²The current (2015) criteria for including a publication in the APR are that the publication be registered in PubMed Central (ie, have a PMCID number) and that the researcher directly benefited from the resources of the CTSA. In the first few years of CTSA funding, the criteria for inclusion in the APR were more relaxed, so earlier publications were not necessarily registered in PubMed Central [Congress mandated the submission of federally funded research to PubMed Central with the Consolidated Appropriations Act of 2008 (H.R. 2764), and the CTSA program began to require all reported publications to have PMCIDs in 2011].

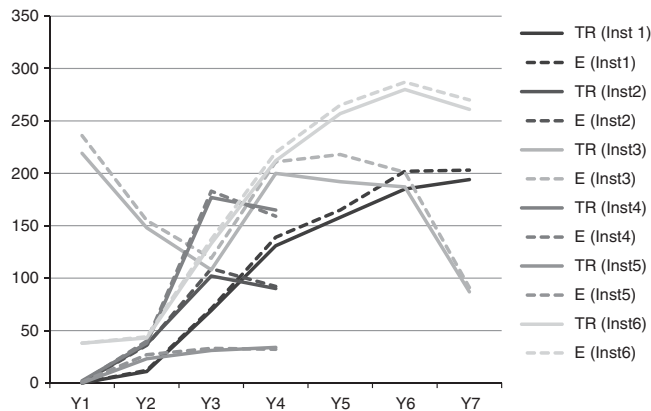


Fig. 2. Thomson Reuters (TR) and Elsevier (E) scholarly output by institution and project year.

Recently, the OPA at NIH has developed its own tool to estimate research influence, iCite (<https://icite.od.nih.gov/>), which calculates a comparative influence measure, the Relative Citation Ratio (RCR), with NIH-supported publications as the benchmark [23]. The FWCI, Cat-C Index, and RCR are very similar in that they are all CCRs, a ratio of observed/expected citations at the publication level that can then be aggregated across a set of publications, such as all CTSA-supported research. In computing the RCR, the article (publication) for which the RCR is computed is referred to as the reference article (RA). As in all CCRs, the numerator is the number of citations to date for an RA. The denominator, however, is unique and is computed from construction of a “co-citation network” defined for each RA that consists of all articles that are also cited by articles that cite the RA. The co-citation network represents an article’s area of influence and is unique to each RA. In addition, and of special importance for NIH-related work, the expected citation rate is also adjusted using regression analysis for comparable NIH R01-related publications [23]. Consequently, the expected citation rate is the average citation rates of the articles’ journals in the co-citation network when adjusted relative to NIH R01 publications.

In Fig. 3, publication influence is measured by CCR over time using the FWCI from Elsevier, the Cat-C Index from Thomson Reuters, and RCR from OPA at NIH. By all 3 measures we can see that CTSA institutions typically received between 2 and 3 times as many citations as would be expected for a comparable set of publications in their respective fields. As with the percentile metrics discussed above, the Elsevier results tended to be a little higher on average, most likely because Scopus indexes more journals. As CCR are calculated by indexing primary mainline journals, this means that Scopus probably includes a set of journals with lower average overall citations, thereby lowering the expected citations denominator and slightly raising the FWCI values over what the Cat-C Index or RCR would find for the same set of publications.

Fig. 3 also shows that the RCR tracks very closely to the Cat-C Index. It is important to note that the RCR analysis was conducted separately from all others and over a year later, and therefore encompasses more citations. Although this would tend to inflate absolute counts of number of citations, it is gratifying to see that, as expected, it does not appear to distort the relative influence indicator.

Another noteworthy detail in Fig. 3 is the steep spike in the FWCI for the year 2010. This spike was traced back to a single publication that had an extremely high number of citations. Because of the way the comparison group expected average citation value is computed and the fact that Scopus indexes a broader number of journals, this value influenced the graph more for the FWCI than for the RCR or Cat-C Index, although its influence is still faintly visible in the latter.

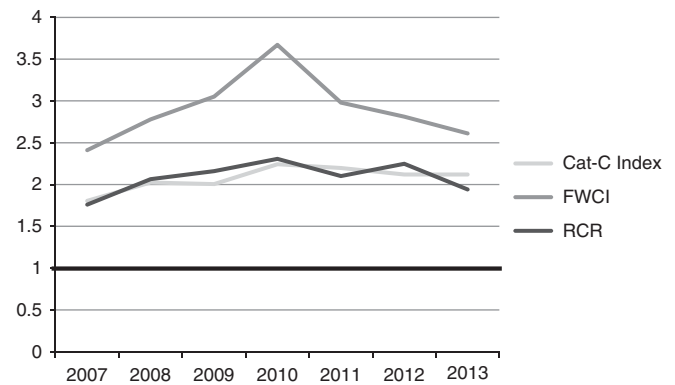


Fig. 3. Thomson Reuters Category-C Index (Cat-C Index), Elsevier Field-Weighted Citation Impact (FWCI), and iCite Relative Citation Ratio (RCR) by Year for all publications. The black line at 1.0 shows the expected rate for the respective comparative citation ratios.

Discussion

Our pilot project is the first to compile and analyze a set of CTSA-supported publications for a group of CTSA institutions using 2 widely used vendors for bibliometrics. Our study indicates that sharing and analyzing publication data are feasible and potentially useful for the development of common metrics on the productivity and influence of the CTSA initiative. Elsevier’s Scopus and Thomson Reuters’ Web of Science contain the vast majority of publications reported by CTSA and the vendors’ key metrics are comparable. A number of considerations, however, should be examined before proceeding with developing common metrics for CTSA publications.

First, we achieved somewhat limited success in identifying the subsets of publications attributable to the different CTSA programs in education and training, pilot grants, and clinical research support. Each of the 6 institutions in the pilot project approached this task somewhat differently, with one able to code each of its publications to 1 of the 3 subsets and one able to code none. Clear protocols for determining which publications result from, say, CTSA pilot grants, do not yet exist but are a necessary step in the progress toward common metrics.

A second caveat is that the publications analyzed here are those reported to the NIH as part of the annual reporting requirement. The NIH requirements for reporting publications as part of the annual reporting process for CTSA have changed annually, becoming more restrictive in recent years. This dynamic reporting environment results in data artifacts, such as the dramatic drop in publications reported by institution 3 after the first year of funding (see Fig. 2). In the early days of the CTSA initiative, there were no restrictions on which publications institutions chose to report as having “benefitted from the CTSA” (the only guideline provided by NIH). Many of the CTSA were awarded to institutions that previously had housed General Clinical Research Centers (GCRCs). Initially, the CTSA were framed as a replacement for the GCRCs, and for those institutions who had been home to GCRCs the new CTSA was viewed as, in part, an extension of this previous initiative. Accordingly, publications that emerged from research that had been supported by the GCRCs were frequently included in the initial publication reports of early funded CTSA. As the program matured, however, NIH provided increasingly narrow criteria for reporting publications, and the requirement that publications must have received support from the CTSA per se was imposed. The apparent drop in publication count in institution 3 in Fig. 2 is related to this artifact of the evolution in reporting criteria.

Even if the requirements become more stable in the future, it is important to review the NIH reporting requirements to know whether they have substantially changed when comparing across years. We also found that differences in CTSA age and size affect the number

of publications and their citations, although not necessarily in a linear fashion. A study of CTSA publications between the years of 2006 and 2010 found that “there is considerable variability across CTSA in terms of the number of publications that have been reported by individual CTSA [24] (p. 5)” and “considerable variability across institutions in the effect of elapsed time on publications [24] (p. 5).” In particular, the first cohort of CTSA was an exception to the pattern of reporting small numbers in the first year of funding. Members of the first cohort reported large number of publications on average in their first year compared with other cohorts, with especially large variability in the number of reported publications across institutions in the first 2 years [24].

We also found that the sophistication of data systems for tracking and thus reporting publications varied across CTSA institutions. A larger CTSA will not necessarily report more publications if their tracking system is not widely implemented or utilized. Another valid concern is that bibliometrics that measure research influence by counting citations will necessarily favor older CTSA and understate the impact of more recent research. Establishing subsets of appropriate comparison CTSA of similar age and size for an analysis is one way to take this variability into account when comparing bibliometrics across institutions.

Despite their limitations, bibliometric assessments of scientific productivity and influence are gaining favor at the NIH and elsewhere. At the level of the individual investigator, the Hirsch index (H-index)³ is widely used as a metric that conveys the approximate influence generated by a set of publications. Because the H-index is, by definition, strongly influenced by the number of publications within a data set, and by both time period and discipline, it does not lend itself easily to being used to compare across different CTSA hubs varying in size, age, and disciplinary focus. Recent advances in adjusting for citation norms across disciplinary fields have made bibliometrics more rigorous as indicators of scientific influence. Indices that adjust for the different rates of citation in different fields or journals have clear advantages over more simple measures. Comparative influence or “field normalized” metrics, such as the FWCI (Elsevier), the Cat-C Index (Thomson Reuters), and the NIH iCite RCR control better for differences between fields of research. Because they improve the utility of bibliometrics for evaluating interdisciplinary research, we recommend more reliance on these comparative influence measures.

However one computes it, and whatever vendor’s index is used, it is apparent that the publications from the participating CTSA have considerably more influence in terms of the 3 CCRs—on the order of 2 to 3 times more—than their benchmark comparison groups do. Although citations are not a direct measure of publication impact (it is difficult to conceive of how one might achieve such a measure), it is reasonable to argue that the data support the notion that these publications, as a set, are disproportionately more frequently noted in their fields than a similar set of comparison publications.

These comparative influence metrics, because they take into account variation in citation practices, may also partially ameliorate the argument that using bibliometric indicators to evaluate research output discourages researchers from working in interdisciplinary teams. Given the emphasis of the CTSA program on multidisciplinary research and publications, and the scope of CTSA-funded research (both bench science and clinical trials), evaluating publications and citations in the context of disciplinary differences is critical [25]. It should also be noted that, in our study, we did not consider the type

³The H-index is a distribution-based indicator that corresponds to the number of papers at or above a given citation level equal to the value of the citation threshold. This bibliometric indicator is used to show both productivity (number of papers) and impact (number of citations) in a single number. The H-index is useful because it discounts the disproportionate weight of highly cited papers, or papers that have not yet been cited [8].

or quality of publications, although some, such as major systematic reviews or guidelines for clinical practice, obviously make particularly important contributions. Presumably this influence is reflected in the CCRs. We also did not evaluate the timeliness or cost of CTSA publications, which NIH and others may consider important in evaluating the return on investment for CTSA.

In general, bibliometrics alone will be insufficient for overall judgments of the value of the CTSA initiative. Only some of what science discovers, and what CTSA accomplish, gets published. However, it is unlikely that any successful translation in the last 50 years that went from scientific discovery to impact did *not* pass through the published literature in some way and at some point. And, most such translations leave a long trail of publications along the way. Publications are the closest thing to a “currency” in science, and bibliometrics that are valid and vetted can play a critical role in judging the value of CTSA-supported work [26].

Conclusion

The substantial investment in research initiatives such as the NIH CTSA has driven the enhancement and use of evaluation methodologies to evaluate research impact, such as bibliometrics. This pilot study demonstrates that bibliometrics can be used as an evaluation method for assessing the amount and influence of CTSA-supported research. Results of analyses by 2 widely used vendors for bibliometrics were quite comparable for ascertaining both the scientific productivity of CTSA and the influence of CTSA publications. We recommend that bibliometrics, especially those that take into account relative citation patterns within disciplines (eg, comparative citation rates), be included as one reflection of the influence of CTSA and other large research initiatives. Publications, which are annually reported by CTSA institutions to NIH, and the resulting bibliometrics, would be relatively easy to include in a set of common metrics across institutions with little additional burden. Bibliometrics is a complex research field of its own with many different metrics that can be appropriately used; we have presented here a few select bibliometrics that offer a practical and feasible way to evaluate the CTSA initiative.

Acknowledgments

The authors thank Thomson Reuters and Elsevier for their contributions to the data analysis for this paper. This project was supported by funds from the following National Center for Advancing Translational Sciences grants: University of California Los Angeles CTSA grant no. UL1TR000124; University of California Davis CTSA grant no. UL1TR000002; University of California Irvine CTSA grant no. UL1TR000153; University of California San Diego CTSA grant no. UL1TR000100; University of California San Francisco CTSA grant no. UL1TR000004; Weill Cornell CTSA grant no. UL1TR000457.

Declaration of Interest

The authors declare that they have no conflicts of interest.

References

1. **National Institutes of Health.** Institutional Clinical and Translational Science Award [Internet]. 2008 [cited Mar 18, 2016]. (<http://grants.nih.gov/grants/guide/rfa-files/RFA-RM-08-002.html>)
2. **Dilts DM, Rosenblum D, Trochim WM.** A virtual national laboratory for reengineering clinical translational science. *Science Translational Medicine* 2012; **4**: 118cm2.
3. **Rubio DM, et al.** Developing common metrics for the Clinical and Translational Science Awards (CTSA): lessons learned. *Clinical and Translational Science* 2015; **8**: 451–459.

4. **Trochim WM, et al.** Evaluation guidelines for the Clinical and Translational Science Awards (CTSAs). *Clinical and Translational Science* 2013; **6**: 303–309.
5. **Scott CS, et al.** Expanding assessments of translational research programs: supplementing metrics with value judgments. *Evaluation & the Health Professions* 2014; **37**: 83–97.
6. **Pincus HA, et al.** Evaluation and the NIH Clinical and Translational Science Awards: a “top ten” list. *Evaluation & the Health Professions* 2013; **36**: 411–431.
7. **Madlock-Brown C, Eichmann D.** Assessing the translational capacity of five CTSA institutions, 2015. *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* 2015; 768–769.
8. **Thomson Reuters.** Using Bibliometrics: A Guide to Evaluating Research Performance with Citation Data. Philadelphia, PA: Thomson Reuters, 2008.
9. **Broadus RN.** Early approaches to bibliometrics. *Journal of the American Society for Information Science* 1987; **38**: 127–129.
10. **Garfield E.** Citation Indexing: Its Theory and Application in Science, Technology, and Humanities. New York: Wiley, 1979.
11. **Garfield E.** Commentary—fifty years of citation indexing. *International Journal of Epidemiology* 2006; **35**: 1127–1128.
12. **Rosas SR, et al.** Evaluating research and impact: a bibliometric analysis of research by the NIH/NIAID HIV/AIDS clinical trials networks. *PLoS One* 2011; **6**: e17428.
13. **Trochim WM, et al.** The evaluation of large research initiatives—a participatory integrative mixed-methods approach. *American Journal of Evaluation* 2008; **29**: 8–28.
14. **Donovan C, et al.** Evaluation of the impact of National Breast Cancer Foundation-funded research. *Med J Aust.* 2014; **3**;200(4): 214–218.
15. **Sengupta IN.** Bibliometrics, informetrics, scientometrics and librmetrics—an overview. *Libri* 1992; **42**: 75–98.
16. **Adams J.** Bibliometrics: the citation game. *Nature* 2014; **510**: 470–471.
17. **Marx W, Bornmann L.** On the causes of subject-specific citation rates in Web of Science. *Scientometrics* 2014; **102**: 1823–1827.
18. **Campbell P.** Escape from the impact factor. *Ethics in Science and Environmental Politics* 2008; **8**: 5–6.
19. **Schekman R.** How journals like Nature, Cell and Science are damaging science. *The Guardian* 2013; **9**: 12.
20. **Werner R.** The focus on bibliometrics makes papers less useful. *Nature* 2015; **517**: 245.
21. **Schekman R.** Reforming research assessment. *Elife* 2013; **2**: e00855.
22. **Colledge L, Verlinde R.** *SciVal Metrics Guidebook* [Internet]. Amsterdam: Elsevier, 2014. (<https://www.elsevier.com/research-intelligence/resource-library/scival-metrics-guidebook>).
23. **Hutchins BI, et al.** Relative Citation Ratio (RCR): a new metric that uses citation rates to measure influence at the article level. *PLoS Biology* 2015; **14**: e1002541. doi:10.1371/journal.pbio.1002541.
24. **Slaughter A, Frechtling J.** Initial Report on CTSA-Supported Publications: 2007–2010. *Westat Report to the National Center for Research Resources*. Rockville, MD: Westat, 2011.
25. **Rafols I, et al.** How journal rankings can suppress interdisciplinary research: a comparison between innovation studies and business & management. *Research Policy* 2012; **41**: 1262–1282.
26. **Lauer MS.** Thought exercises on accountability and performance measures at the National Heart, Lung, and Blood Institute (NHLBI): an invited commentary for circulation research. *Circulation Research* 2011; **108**: 405–409.