AUTHORS:
Kylie de Jager[1]
Chipo Chimhundu[1]
Yolande X.R. Harley[2]
Tania S. Douglas[1]

AFFILIATIONS:
[1]Division of Biomedical Engineering, University of Cape Town, Cape Town, South Africa
[2]Faculty Research Office, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Tania Douglas

EMAIL:
tania.douglas@uct.ac.za

DATES:
Received: 12 Mar. 2019
Revised: 30 Nov. 2019
Accepted: 07 Jan. 2020
Published: 26 Mar. 2020

HOW TO CITE:
De Jager K, Chimhundu C, Harley YXR, Douglas TS. Collaboration and citation impact: Trends in health sciences research at the University of Cape Town. S Afr J Sci. 2020;116(3/4), Art. #6132, 8 pages. https://doi.org/10.17159/sajs.2020/6132

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Jane Carruthers

KEYWORDS:
bibliometrics, relative citation rate, impact factor, co-authorship

FUNDING:
None

Collaboration and citation impact: Trends in health sciences research at the University of Cape Town

Against a background of substantial growth in publication output in health sciences at the University of Cape Town (UCT) over the past two decades, we examined the relationship between collaboration with domestic and foreign institutions and resulting citations of co-published work. We report on trends in authorship and citation impact for health sciences research at UCT across three 3-year periods: 1999–2001, 2006–2008 and 2013–2015. We examined numbers of collaborative publications with domestic and foreign co-authors; the status of collaboration with other African countries; the location of the 'drivers' of the research (with 'drivers' indicated by first or last authorship); and expected and observed citation counts – used as an indicator of impact – over time. We found that the relative citation rate of the set of UCT health sciences publications has increased; the set of 1999–2001 publications was less frequently cited than expected for the journals in which the publications appear, while the 2006–2008 and 2013–2015 sets were cited more frequently than expected. Relative citation rates were greater for papers for which UCT shared international co-authorship than for papers with UCT-only or domestic co-authorship. Our findings confirm reports in the literature of higher citation of internationally co-authored publications. We additionally found that the publications with the highest relative citation rates were driven by authors from foreign institutions.

Significance:

- Methods are presented for extracting, measuring, analysing and representing the citation impact of collaborative research.
- The relative citation rate of health sciences publications produced by UCT has increased and co-publication with international authors has increased.
- The findings confirm reports in the literature of higher citation of publications co-authored with international collaborators.
- An apparent influence of foreign drivers on citation impact, holds risk for South African science.

Introduction

International collaboration for health-related research is encouraged by governments, funding agencies and university executives, and is sought by researchers. Motives for collaboration include access to equipment, infrastructure, knowledge, expertise and funding, as well as raising research and researcher profiles. In addition, bibliometric research showing associations between international co-authorship and research quality indicators[1,2] creates the expectation that international partnerships increase research impact.

Since the first democratic elections in 1994, the South African government has emphasised the development of science and innovation policy, in an effort to use science and technology as levers for socio-economic development.[3] South African funding agencies extol the benefits of partnership. The South African Medical Research Council, the National Research Foundation, and the Department of Science and Technology, in their recent strategic plans, encourage collaboration and international partnership.[4-6] These agencies also co-fund international research partnerships with foreign agencies such as the US National Institutes of Health and the UK Medical Research Council. A study comparing South Africa's publication volumes for the periods 1990–1994 and 2004–2008 has suggested that the substantial increases in the latter period could mainly be attributed to the strengthening of collaboration with foreign partners.[3] More recently, in a report on the state of research in South Africa, Mouton et al.[7] have shown that, among papers published by South African authors, the proportion of papers co-published with collaborators in Africa and in the rest of the world has increased steadily over the period 2000–2016, while single-author papers and those with South African collaborators only, have declined. Mouton et al.[7] also report that South Africa's publication output since 2000 has shown an average growth rate of 2.9% annually, with the country's contribution to global research output increasing from 0.4% in 2000 to 0.91% in 2016. The Mouton et al.[7] report further shows that the citation impact of South African publications has increased over the period covered by their analysis. Citations are a measure of the acknowledgement by researchers of the work published by their peers.

At South Africa's University of Cape Town (UCT), the annual number of publications in health sciences (journal articles and reviews) listed in Scopus[8] has increased from 408 to 1729 between 1999 and 2015. Little is documented about the nature of the international collaborations that have contributed to these publications, and whether and how they have changed over time. The aim of our study was to examine the relationship between these collaboration patterns and the citation impact of the health sciences research outputs of UCT. We report on trends in collaboration, authorship and citation impact for health sciences research at UCT across three 3-year periods: 1999–2001, 2006–2008 and 2013–2015, using co-publication as an indicator of collaboration.[1] We examined numbers of collaborative publications with domestic and foreign co-authors over the periods studied, the status of collaboration with other African countries, the location of the 'drivers' of the research, and expected and observed citation counts – used as an indicator of impact – over time, from the perspective of a research-active university with considerable access to international collaboration.

Methods

We were interested in publications generated by UCT in the broad field of health sciences. Relevant publications were sourced using Scopus. Titles in Scopus are classified under four broad subject clusters: health sciences; physical sciences; life sciences; and social sciences including humanities. These clusters are further divided into 27 major, and more than 300 minor, subject areas.[9] Of the four broad subject clusters, the health and life sciences clusters, defined by Scopus to consist of 11 major subject areas, contain the publications of interest.

Articles in the health sciences and life sciences clusters with UCT-affiliated authors were extracted by searching Scopus for versions of the name "University of Cape Town", limiting the search to articles and reviews as document types and only considering publications from three time periods: 1999–2001, 2006–2008 and 2013–2015. The search string for 1999 was: AFFILORG("U* of Cape Town" OR "U* of capetown" OR "U* Cape Town" OR "U* capetown" OR uct) AND DOCTYPE (ar OR re) AND PUBYEAR = 1999 AND SUBJAREA (mult OR medi OR nurs OR vete OR dent OR heal OR agri OR bioc OR immu OR neur OR phar). A Scopus search was carried out for each year of the three time periods.

The publication lists identified by Scopus were exported along with citation data (C – all citations; CX – excluding self-citations) and PubMed ID (if present) for each publication. The search was carried out on 18 August 2017 for all publication years of interest. The data for 2014 showed <1% difference between C and CX data (expected to be ~20%). A new Scopus search was conducted on 15 May 2018 for 2014 only and corresponding publication and citation data exported.

Despite limiting our article search to two of the Scopus clusters, the extraction yielded a topic coverage wider than our area of interest. Therefore, to focus the data set, we retained from our Scopus search only those publications that had a PubMed ID. PubMed's subject coverage includes our area of interest only – biomedicine and health – as recorded in the MEDLINE database. Unlike Scopus, PubMed does not provide citation data, hence the use of both databases to extract all relevant data (including citation data), while limiting the data to our area of interest using the PubMed ID.

Our analysis also required SNIP (source normalised impact per publication) and IPP (impact per publication) journal indicators for the three time periods; this information was obtained from the CWTS Journal Indicators website.[10] These indicators have been calculated based on the Scopus bibliographic database.

Data processing

The publication and citation data were exported by Scopus in separate files, as were the CWTS Journal Indicators data. For each publication, Scopus provides a list of author names only, a list of author names with their respective affiliation data, affiliation data only reported as a unique list, a 'corresponding author' name and affiliation list, article title, PubMed ID (PMID), Scopus electronic ID (EID), journal title, International Standard Serial Number (ISSN), volume, issue and page numbers. The CWTS data provide journal title, ISSN, volume and issue number in addition to the SNIP and IPP values.

MATLAB (Mathworks, Natick, MA, USA) was used to perform string comparisons (publication year, ISSN, journal title, volume, issue, article title) on the exported bibliographic and citation information. Matched publications were merged into a single spreadsheet. Discrepancies were flagged, manually checked and corrected. It was not always possible to match each publication with the CWTS Journal Indicators; such publications were excluded from further analysis.

MATLAB was also used to check selected information for each publication, and any inconsistencies were flagged for manual inspection and cleaned:

1. Duplicate article titles, PMIDs and EIDs were identified.
   - Multiple publications with identical PMIDs were found to have identical titles, digital object identifiers (DOIs) and journals (source title, volume, issue and page numbers). Citation counts, however, could be different. Such publications were sourced in Scopus using their title. The citation list for each publication was extracted and compared. Typically, the publications did not have overlapping citations. Consequently, the two duplicate entries were merged into a single entry (title, journal, DOI, PMID) and the total citation count retained (i.e. citation counts for the duplicate publications were added).
   - Duplicate titles were kept as separate entries if they had different DOIs, journals (i.e. differences in source, volume, issue or page number) and PMIDs. Citation counts were retained as separate counts for such duplicates.
   - No duplicate EIDs were found.
2. The list of affiliations associated with each publication retrieved from Scopus was compared with the list of affiliations provided for each co-author. Both lists were then independently examined for occurrences of known country names, and the identified countries from both lists compared. Any discrepancies found through the comparison, as well as any publications that did not contain 'South Africa' in the affiliation lists, were flagged for manual inspection.
   - Online databases (PubMed, Scopus, Google) were used to check affiliation data for the relevant publications as well as typographical errors or missing country information, for manual correction.
   - Some publications were found to not have a South African affiliation (typically due to an incorrectly identified UCT affiliation, see Point 3 below). Such publications were removed from the data set.
3. Variations of the name University of Cape Town (as used in the Scopus search terms) were used to flag publications that contained affiliation names which did not definitively represent the University of Cape Town. For instance:
   - UCT – could be an abbreviation for something other than the University of Cape Town.
   - U of Cape Town – the U could stand for something other than University
   - Unit Cape Town – does not match "University of Cape Town"

   Such publications were manually checked. Non-"University of Cape Town" publications were removed from the data set.
4. The three Scopus lists – author names only, authors with affiliation data and affiliation data only – were compared and differences were flagged for manual inspection to correct for inconsistencies:
   - Publications were found which did not include a separate affiliation for each author. In such cases, the original publication was consulted, and the affiliation list corrected accordingly.
   - In instances in which Scopus affiliation data were missing for a middle author, the original publication was also consulted and the affiliation list corrected accordingly. If the original publication was still found to not show affiliation information for the author, any one of the affiliations of other authors would be used instead for completeness, as it would not affect the driver analysis or country representation information as provided by the publication.
   - If the only affiliation provided was that of the corresponding author for both Scopus data and the original publication, it was assumed that the affiliation was the same for all authors.

Co-authorship trends

Co-author countries were extracted from the affiliation data from each publication. The locations of authors for UCT's health sciences publications were indicated on a geographic heat map using Tableau Public, a freely available data visualisation software tool.[11]

Research drivers

Authorship order usually indicates the level of contribution for each author listed on a publication. In the health sciences and public health fields, the first author is typically the one making the largest contribution, with the last author usually having a more senior or supervisory role, often contributing to the inception of the research project.[12] Lead authorship is determined by the level of responsibility for the manuscript, research contribution and in some cases the responsibility of correspondence after publication. Although there is no universal rule for author listing of publications, often the first and the last authors have a leading role in the direction of the work. As such, this study considered an author to be a driver or leader of the research if they appeared as the first or last author.

In some cases, authors may simply be listed alphabetically without considering level of contribution. To determine the proportion of such cases, publications with alphabetical author lists, where the publication had four or more co-authors, were identified and counted. Due to the small percentage of such publications, the impact of these publications on investigating research driver patterns was considered to be negligible, and the publications were included in the analysis.

The countries driving the research were considered those with which the first and last authors are affiliated; in the case of authors with multiple affiliations from different countries, all such countries were considered to be drivers. The association between driver location and citations was examined.

Publication sets

For each of the three time periods considered (1999–2001, 2006–2008, 2013–2015), the data were separated into mutually exclusive affiliation and driver sets as outlined in Table 1. The affiliation sets considered the affiliations of all co-authors present in a publication, while the driver sets (subsets of the affiliation sets) only considered the affiliations of the first and last authors. The affiliation sets described in Table 1 were defined to determine whether co-authorship included UCT authors only, or domestic (within South Africa) or international partnerships as defined by the U, D and I affiliation sets, respectively. All combinations of driver sets were created as a means of identifying which of the co-authorship combinations within each affiliation set had more influence on the research conducted. Dominance of the IUd publications in the international set, for example, would communicate that, although there is international collaboration, UCT tends to lead the collaborative activity, while a dominance of IFd would suggest that international publications with UCT are largely driven by foreign entities.

Citation impact

Questions concerning whether self-citations should be excluded from citation analyses have been raised[13,14], as self-citations are known to have an impact on certain types of analyses. Definitions of self-citation vary slightly depending on the nature of the data, but, in most cases, self-citation is considered an instance in which both the citing and the cited paper have at least one author in common.[13] This situation is more likely to occur with highly collaborative publications due to the higher number of co-authors. However, the more authors present on a publication, the more likely it is to be cited in general, at two additional citations on average per additional author, with the increasing number of self-citations making a small contribution to increased citation rates.[13]

Some scholars agree that self-citations are problematic for low-level analysis such as at the level of the institution[13,14], and that self-citations have less impact for analyses involving larger groups such as the comparison of citation impact across countries. The share of self-citations in a data set influences whether the exclusion of self-citations should be considered prior to further analysis. Self-citation rates vary across fields due to differences in citation norms and tend to be low in fields such as clinical medicine.[13] Country-level analyses have been conducted using data sets with self-citation shares between 24% and 28%.[2] However, proportions of up to 20% are generally considered permissible for the inclusion of self-citations in citation impact assessments.[13]

The share of self-citations across our studied time periods was between 19% and 22%, which is comparable to previous work that has used data with self-citations ranging from 20% to 28% for country-level analyses.[13,14] Our analysis is at the international level, for a broad scientific field, so is unlikely to be affected significantly by the inclusion of self-citations. However, we chose to exclude self-citations to enable additional comparison of the smaller UCT and domestic publication sets to the much larger international publication sets.

For a publication set, the mean observed citation rate (MOCR) is the total number of citations accumulated in Year Y for publications of Years Y-1, Y-2 and Y-3, divided by the total number of publications, *n*, in that same 3-year publication window. Mean expected citation rate (MECR) is the average number of citations per paper accumulated in Year Y for the journals represented in the publication set, for the same 3-year publication window.[2,15]

The MECR and MOCR, both excluding self-citations, were calculated for all publication sets considered, and were used to compare citation impact across the affiliation and driver publication sets. MOCR and MECR (Equations 1 and 2) were modified for our data set from the

Table 1: Definition of affiliation and driver publication sets

Affiliation set		Description	Author affiliation		Driver set	
			First	Last		
U	UCT	All co-authors have only a UCT affiliation	U	U	U	UCT driven
D	Domestic	Co-authors are affiliated with UCT and at least one South African (non-UCT) affiliation. A single co-author can have multiple affiliations to both UCT and the other South African entity.	U	U	DUd	Domestic-UCT driven
			D	D	DNd	Domestic-non-UCT driven
			U and D in combination		DCd	Domestic-UCT co-driven
I	International	Co-authors are affiliated with UCT and at least one foreign (non-South African) entity. A single co-author can have multiple affiliations including an affiliation with another South African (non-UCT) entity.	U	U	IUd	International-UCT driven
			D	D	IDd	International-domestic driven
			F	F	IFd	International-foreign driven
			U and D in combination		IUDd	International-UCT-domestic co-driven
			U and F in combination		IUFd	International-UCT-foreign co-driven
			D and F in combination		IDFd	International-domestic-foreign co-driven
			U, D and F in combination		IMd	International multi-driven

U, UCT affiliation; D, non-UCT South African (domestic) affiliation; F, non-South African (foreign) affiliation

definitions in previous studies[2,15] through the use of a 3-year publication window instead of a 2-year window, and the use of the Scopus IPP instead of the Clarivate Analytics Journal Citation Reports (JCR) Garfield factor (also known as the journal impact factor).

$$MOCR = \frac{\sum_{i=1}^{n} CX_i}{n} \quad \text{Equation 1}$$

$$MECR = \frac{\sum_{i=1}^{n} IPP_i}{n} \quad \text{Equation 2}$$

IPP differs from the Garfield factor mainly in that IPP is calculated on the basis of papers published in the previous 3 years (Y-1, Y-2 and Y-3) instead of the previous 2 years. The IPP of Year Y, based on a 3-year publication window, would therefore mean collecting citations in Year Y for papers published in Years Y-1, Y-2 and Y-3. As an example, citations for papers published in 2006, 2007 and 2008 would be collected to evaluate the IPP for the year 2009. The longer publication window used by Scopus compared to the window used by Journal Citation Reports is believed to help reduce the impact of differences in citation trends between fields and/or journals.[16]

The MOCR/MECR ratio, known as the relative citation rate (RCR), is a more precise measure of impact than journal-based metrics, as it accounts for citation trends of various fields. It is field-independent and indicates if a publication has been cited as expected.[17] However, the ratio retains an element of journal-level influence; this influence could be eliminated with normalisation by all publications in the field, for which data were not readily available, rather than expected journal citation rate. While the RCR is not explicitly used in this study, citation impact was investigated by plotting MOCR against MECR in a relational chart, as suggested by Schubert and Braun[18]. The MOCR=MECR (or RCR =1) line is the line at which the observed citation rate for the publication set meets the expected citation rate for the journals in which the articles of the publication set appear. The line can be regarded as the boundary between lower and higher than expected impact, but 'not without an element of arbitrariness'[18].

Results

Table 2 shows the attributes of the UCT (U), domestic (D) and international (I) affiliation publication sets. An upward trend in the number of publications in each set is evident, with an approximate doubling of outputs from one 3-year window to the next for the retained data set. For the U, D and I data sets, the number of publications is seen to increase by factors of 1.6, 5.8 and 8.0, respectively, between 1999 and 2015. While an increase in the number of UCT authors is consistent with the overall increase in number of authors per publication, the overall proportion of UCT authors per publication as well as the average number of UCT authors per publication have generally decreased with time. This finding is also in agreement with the downward trend in the UCT affiliation set, which constitutes 42% of the retained publications for 1999–2001, 20% for 2006–2008, and only 13% for 2013–2015.

While very few papers were published in collaboration with other African countries during the 1999–2001 period (eight countries), a general upward trend in African collaboration is observed, with a marked increase between the first two periods such that the 2006–2008 period had four times the number of African collaborators than did the 1999–2001 period.

Co-authorship trends

Figure 1 shows the locations of countries that have co-authored health sciences publications with UCT in the 1999–2001, 2006–2008 and 2013–2015 periods. Each country is colour coded, with cool colours representing few co-publications and hot colours representing a high number of co-publications with UCT. Figure 2 is similar but focuses on the African continent, specifically to show African co-authorship trends. Figures 1 and 2 therefore show two heat maps, representing global collaborations (green to red scale) and African collaboration (light blue to dark blue scale).

In both Figure 1 and Figure 2, an increase in co-publication is observed, consistent with the values in Table 2. The global maps (Figure 1) suggest that co-publication has occurred most frequently with North America, Europe and Australia, and has also increased over time, while the African maps (Figure 2) suggest that intensity of co-publication has increased, particularly with countries in southern and East Africa, with Cameroon, Nigeria and Ghana also showing increased co-publication with UCT over time. As time passed, the African countries present in earlier periods persisted in collaborating with UCT; the increasing co-publication rates may also be the result of new collaborations having been fostered.

Table 3 shows the top 10 countries worldwide (excluding South Africa) that have co-published with UCT, ranked by number of health sciences co-publications. The top two positions have been maintained by the USA and the UK. Most of the countries in the 3rd to 9th places have persisted but with a shuffling of positions.

Table 2: Overview of UCT health sciences publications for the 1999–2001, 2006–2008 and 2013–2015 periods

	Publication set attributes	1999–2001	2006–2008	2013–2015
Publications	Total number of publications	1184	2535	4844
	Proportion of self-citations in retained publications*	19%	22%	22%
	Total number of retained publications (with IPP, SNIP and PMID values) (% of total number of publications)	656 (55%)	1684 (66%)	3298 (68%)
	Total citations of retained publication set, excluding self-citations	1940	8392	17 413
	UCT publications (% of retained publications)	273 (42%)	345 (20%)	442 (13%)
	Domestic publications (% of retained publications)	96 (15%)	320 (19%)	558 (17%)
	International publications (% of retained publications)	287 (44%)	1019 (61%)	2298 (70%)
Authors	Number of authors	3166	9829	30 750
	Average number of authors per publication	5	6	9
	Number of UCT authors	1838	4149	8007
	Average number of UCT authors per publication	3	2	2
	Proportion of UCT authors	58%	42%	26%
	Proportion of publications with four or more authors in alphabetical order	1.1%	1.0%	0.5%
Countries	Number of countries	51	99	148
	Number of African countries (excluding South Africa)	8	32	43

IPP, impact per publication; SNIP, source normalised impact per publication; PMID, PubMed reference number
**Retained publications had complete affiliation data, PubMed IDs, CWTS Journal Indicators data and were also verified to have at least one UCT author amongst the listed authors.*

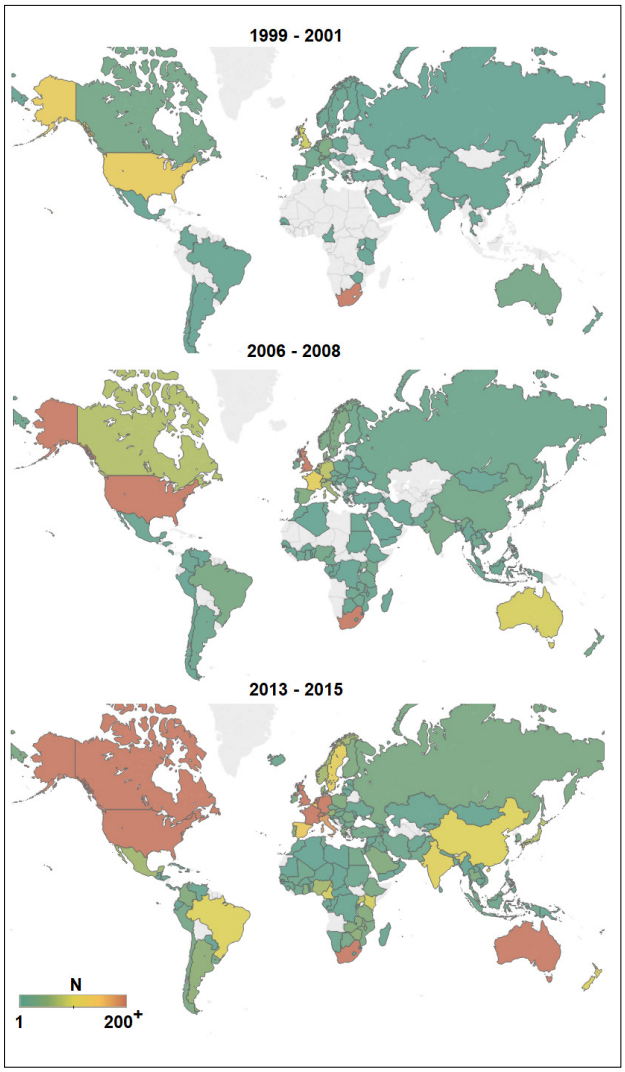


Figure 1: Global footprint of UCT health sciences co-authors. Number of co-publications (N) increase in the direction green to red, with co-publications of 200 or more shown in red.

Table 3 also shows the number of co-publications with countries on the African continent (excluding South Africa) – as these were very few in the first time period, the ranking display is limited to four ranks. For 1999–2001, the highest number of co-publications observed in the data set was two. As African collaboration increased with time, ranks became more clearly defined with the top four ranking countries being Nigeria, Uganda, Kenya and Cameroon in the latter two time periods, although positions varied.

Figure 3 shows the numbers and proportions of publications in each of the publication sets defined in Table 1, with the U, D and I sub-totals matching those listed in Table 2.

In accordance with Table 2, Figure 3 shows that the proportion of UCT-only authored papers has decreased over time. This decrease is accompanied by a substantial increase in internationally co-authored papers. With respect to the drivers of the research (first/last authorship), in the domestic sub-category, the domestic co-driven (DCd) publications are more prevalent than the UCT-only driven (DUd) and the domestic non-UCT driven (DNd) ones. At the international level, the foreign-only driven papers (IFd) and those co-driven by UCT and a foreign entity (IUFd), constitute the largest contribution to the international co-publications, together accounting for (in roughly equal measures) approximately 72% of the international publication set over the studied period.

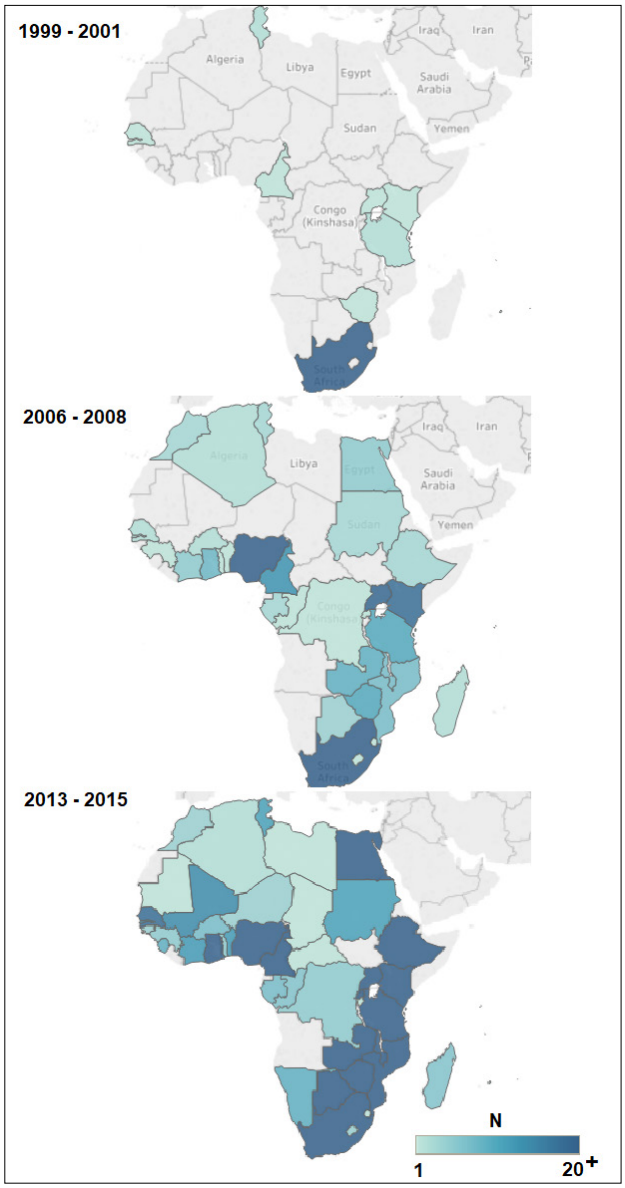


Figure 2: African footprint of UCT health sciences co-authors. Number of co-publications (N) increase in the direction light blue to dark blue, with co-publications of 20 or more shown in dark blue.

Citation impact

The documents with IPP were used for MOCR and MECR calculation. Figure 4 plots MOCR against MECR, with each marker representing a publication set. The publication sets are shown in relation to the MOCR=MECR line. As time progresses, most of the markers move closer to the line, except for the IFd subset (differentiated from all other subsets with a diamond marker). The IFd subset is seen to move upward, rightward and further from the MECR=MOCR line.

Discussion

Co-publication is a collaborative activity[2,19,20], and the qualifier for authorship is all authors contributing, in various degrees, to the design, investigation, manuscript writing and approval of the final product[12]. Co-publication can therefore be seen as a proxy for collaboration, although this relationship can be distorted.[21] Our data set shows a notable increase over time in the number of health sciences publications (Table 2), consistent with the observation that, in general, South Africa's scientific publication output has climbed steeply since around 2004[7] after fairly constant levels had been maintained from the mid-1980s[3]. The rate at which the international publication set increases is higher than that for the UCT and domestic sets.

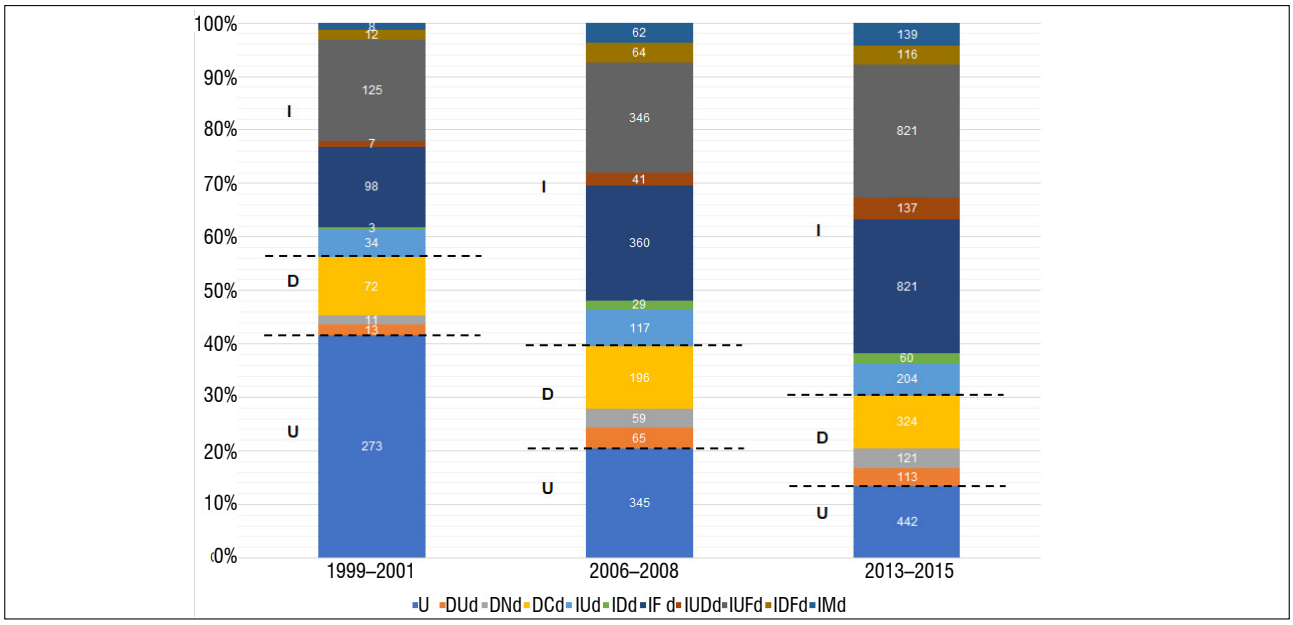
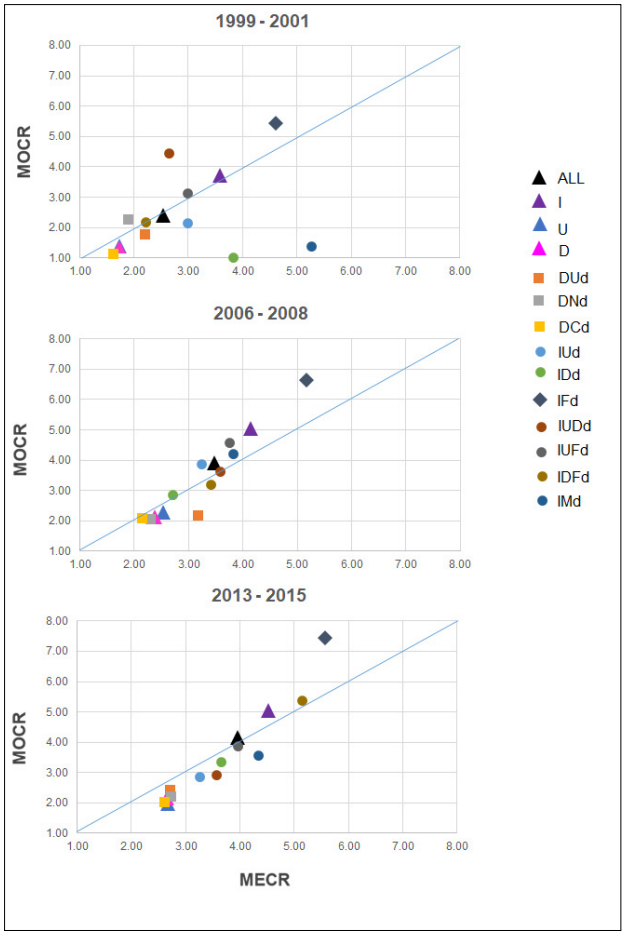


Figure 3: Breakdown of UCT health sciences publication driver sets over time. The number of publications is shown on the bars.



MOCR, mean observed citation rate; MECR, mean expected citation rate
Key as per Table 1

Figure 4: Relational charts over time for UCT health sciences publications.

Coupled with a growth in volume of output, UCT's health sciences research has therefore shifted to increased domestic and especially international co-publication. UCT's collaboration with African countries, specifically, has also grown. Mouton et al.[7] have shown an increase in international co-publication for South African research outputs in health sciences, from 59% in 2000 to 64% in 2016, whereas our data set shows an increase in international co-publication from 44% for 1999–2001, to 70% for 2013–2015 (retained publications in Table 2).

The Mouton report attributes high international collaboration rates for health sciences to internationally supported clinical trials and other large projects. Mouton et al.[7] also describe the existence of 'mega-research' institutes, which are substantially funded by international organisations, as a distinctive feature of the South African health research landscape. These facilities have the potential to expand the country's health research competencies and leadership and the ability of its researchers to drive collaborative research agendas. The assembly of effective and productive local research teams is not, however, limited to such large facilities. Further analysis of the composition and contributions of co-authors of publications with high citation impact, might reveal a range of strategies for building successful collaborative teams.

Table 2 shows that the number of authors of any affiliation in the full set of publications has increased at a faster rate (approximately 1:3:10) than the number of authors with UCT affiliations specifically (approximately 1:2:4). This reflects the increase in external collaborators, with the observed proportion of UCT authors across the publication sets decreasing from 57% to 36% across the studied periods. In addition, while the total number of UCT authors has increased with time, the average number of UCT authors per publication has remained around 2. As also evidenced in Table 2, the number of citations (excluding self-citations) has increased at a faster rate (approximately 1:4:9) than has the number of retained publications (approximately 1:3:5) across the three studied periods (1999–2001:2006–2008:2013–2015). As the number of citations generally increases at a rate of two citations per additional author[13], the faster growth in citations is expected.

UCT's co-authors show a wide geographical spread across continents. There is a general increase in collaboration worldwide, including with other African countries, as shown in Figures 1 and 2. Throughout the studied period, the top two countries collaborating with UCT have been the USA and the UK (Table 3). Sooryamoorthy[22] similarly found the USA and the UK to be the leading co-publishing countries with South Africa for papers in the Web of Science for the period 1945–2010. UCT's top nine co-publishing countries have remained consistent over the 16-year period, comprising mostly European countries.

Where UCT's co-publication is purely domestic, our findings reveal that it is more common for UCT and a local entity to jointly drive the research. However, with international collaboration, the more likely nature of co-publication is for either a foreign entity to completely drive the research

Table 3: Top countries (excluding South Africa) by number of health sciences co-publications with the University of Cape Town

	1999–2001		2006–2008		2013–2015	
Rank	**Country**	**Number of publications**	**Country**	**Number of documents**	**Country**	**Number of publications**
1	USA	124	USA	465	USA	1125
2	United Kingdom	92	United Kingdom	381	United Kingdom	836
3	Germany	34	France	113	Netherlands	296
4	Switzerland	24	Australia	96	Australia	295
5	Australia	23	Netherlands	82	France	261
6	Canada	22	Germany	81	Germany	260
7	Netherlands	22	Canada	76	Switzerland	256
8	France	16	Switzerland	64	Canada	252
9	Italy	15	Italy	63	Italy	179
10	Spain	13	Denmark	41	Belgium	154
Top African countries						
1	Cameroon, Tanzania, Tunisia	2	Nigeria	20	Cameroon, Kenya	95
2			Uganda	19		
3			Kenya	18	Uganda	80
4	Kenya, Mauritius, Senegal, Uganda, Zimbabwe	1	Cameroon	13	Nigeria	68

or for UCT to jointly drive the research with the foreign entity (Figure 3). Foreign research drivers may be supported by funding from non-South African agencies, and these publications would then reflect projects on which UCT researchers are co-investigators rather than principal investigators. The growth in UCT-foreign co-driven publications may reflect an increase in access to international funding by UCT researchers on projects that they co-lead with foreign researchers. Examination of funding attributions on publications would clarify funding sources for different types of co-publication.

The full set of UCT publications ('All' in Figure 4) lies below the MOCR=MECR line in the first time period, but above the line in the second and third periods. This pattern indicates an increase in the relative citation rate of UCT publications from lower than to higher than expected (for the journals in which the publications appear) over time. Identification of the publications with the highest relative citation rates would enable analysis of the topics receiving high citation attention and consideration of the health sciences impact of such publications, as well as deeper examination of the nature of the collaboration involved. Such analysis was beyond our scope but presents an area for future study.

Foreign partnerships strengthen UCT's citation impact: in all three time periods the internationally collaborative publication set (I) is above the MOCR=MECR line, whereas U and D are below the line. Thus, the relative citation rate is greater for papers on which UCT has international co-authorship than it is for papers with UCT-only or domestic co-authorship. The mean observed citation rate is also higher for I than it is for U or D in all time periods, indicating that UCT's internationally collaborative publications typically have a higher citation rate than internal or domestic publications. This finding supports what has been described in the literature. For example, Abramo et al.[1] found a correlation between journal impact factor and co-publication with foreign co-authors in the Italian research system. Mouton et al.[7] found an increase in the mean normalised citation score of South Africa's health sciences papers over the period 2000–2016, from 0.9 to 1.3; this increase means that South Africa's health sciences publications have on average shifted over time from being cited slightly less frequently to more frequently than the world average in the relevant fields. The growth in citation impact for the UCT health sciences example reflects the situation in South Africa more broadly, suggesting that increased citation impact at the national level may at least partly be due to increased international collaboration, consistent with findings in other countries.[1,2]

Higher citation of internationally co-authored publications may be attributed to an audience effect, in which more authors from a greater range of countries provide access to a larger community of citing researchers, and the growth in international collaboration may reflect preferential attachment to international co-authors based on their status and reputation.[23] Kahn[3] expressed concern that South African science might be vulnerable should there be weakening of collaborations with foreign partners, given that foreign collaboration had likely driven increases in publication volumes in recent years. However, in a study using data from Scopus and Web of Science, Wagner et al.[23] found evidence to suggest that international collaborations suppress novelty and produce conventional outputs. The authors cite obstacles generated by collaboration, such as communication barriers and costs that limit creativity, as possible reasons for this finding. If science and technology are to be drivers of innovation and socio-economic development in South Africa, a lack of novelty in actively encouraged and rapidly proliferating international collaborations is a concern.

The motivations for international collaboration proposed by Wagner et al.[23] do not specifically address the driving authors of collaborative research. As indicated by the highest MOCR in all three time periods, UCT publications that are internationally collaborative and driven by the foreign entity (IFd) tend to be cited more frequently than publications with other driver combinations, and are also cited at higher rates than expected for the journals in which these papers are published (Figure 4, diamond marker). Despite the growth in publication productivity, the apparent influence of foreign drivers on, and the likely associated dependence on foreign principal investigators for high citation impact, holds risk for South African science.

Acknowledgements

We thank Dr Robyn May and Ms Esmari Huysamen for early assistance with data collection and analysis.

Authors' contributions

T.S.D. led and managed the project. All authors contributed to conceptualisation. K.d.J. collected the data. K.d.J. and C.C. developed the methodology, curated the data and analysed the data. C.C. validated the data. K.d.J. and C.C. wrote the initial draft. All authors contributed to the writing of the final manuscript.

References

1. Abramo G, D'Angelo CA, Di Costa F. Research collaboration and productivity: Is there correlation? High Educ. 2009;57(2):155–171. https://doi.org/10.1007/s10734-008-9139-z
2. Wang L, Thijs B, Glänzel W. Characteristics of international collaboration in sport sciences publications and its influence on citation impact. Scientometrics. 2015;105(2):843–862. https://doi.org/10.1007/s11192-015-1735-y
3. Kahn MJ. A bibliometric analysis of South Africa's scientific outputs – some trends and implications. S Afr J Sci. 2011;107(1/2), Art. #406, 6 pages. https://doi.org/10.4102/sajs.v107i1/2.406

4. South African National Research Foundation. National Research Foundation strategy 2020 [document on the Internet]. c2019 [cited 2019 Nov 04]. Available from: https://www.nrf.ac.za/sites/default/files/documents/NRF%20Strategy%20Implementation.pdf
5. South African Medical Research Council. SAMRC strategic plan, for the fiscal years 2015/16–2019/20 [document on the Internet]. c2015 [cited 2019 Nov 04]. Available from: http://www.mrc.ac.za/flipbooks/StratPlan/mobile/index.html#p=1
6. South African Department of Science and Technology. Department of Science and Technology strategic plan for the fiscal years 2015–2020 [document on the Internet]. c2015 [cited 2019 Nov 04]. Available from: https://www.dst.gov.za/images/Attachments/DST_Strategic_Plan_2015-2020_web_pdf.pdf
7. Mouton J, Basson, I, Blackenberg J, Boshoff N, Prozesky H, Redelinghuys H, et al. The state of the South African research enterprise. Stellenbosch: DST-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy; 2019. Available from: http://www0.sun.ac.za/scistip/wp-content/uploads/2019/08/state-of-the-South-African-research-enterprise.pdf
8. Elsevier. Scopus [database on the Internet]. c2004 [cited 2019 Mar 06]. Available from: https://www.scopus.com/
9. Elsevier B.V. Scopus: Content coverage guide [document on the Internet]. c2017 [cited 2019 Nov 04]. Available from: https://www.elsevier.com/__data/assets/pdf_file/0007/69451/0597-Scopus-Content-Coverage-Guide-US-LETTER-v4-HI-singles-no-ticks.pdf
10. Leiden University's Centre for Science and Technology Studies. CWTS Journal Indicators [homepage on the Internet]. c2018 [cited 2019 Mar 05]. Available from: http://www.journalindicators.com/
11. Tableau. Maps [software on the Internet]. c2003 [cited 2019 Nov 04]. Available from: https://www.tableau.com/solutions/maps
12. Tarkang EE, Kweku M, Zotor FB. Publication practices and responsible authorship: A review article. J Public Health Africa. 2017;8(723):36–42. https://doi.org/10.4081/jphia.2017.723
13. Aksnes DW. A macro study of self-citation. Scientometrics. 2003;56(2):235–246. https://doi.org/10.1023/A:1021919228368
14. Glänzel W, Thijs B. The influence of author self-citations on bibliometric macro indicators. Scientometrics. 2006;59(3):281–310. https://doi.org/10.1023/B:SCIE.0000018535.99885.e9
15. Vinkler P. Relations of relative scientometric indicators. Scientometrics. 2003;58(3):687–694. https://doi.org/10.1023/B:SCIE.0000006888.69146.24
16. Moed HF. Measuring contextual citation impact of scientific journals. J Informetr. 2010;4(3):265–277. https://doi.org/10.1016/j.joi.2010.01.002
17. Hutchins BI, Yuan X, Anderson JM, Santangelo GM. Relative citation ratio (RCR): A new metric that uses citation rates to measure influence at the article level. PLoS Biol. 2016;14(9), e1002541, 25 pages. https://doi.org/10.1371/journal.pbio.1002541
18. Schubert A, Braun T. Relative indicators and relational charts for comparative assessment of publication output and citation impact. Scientometrics. 1986;9(5–6):281–291. https://doi.org/10.1007/BF02017249
19. Meyer M, Bhattacharya C. Commonalities and differences between scholarly and technical collaboration. Scientometrics. 2004;61(3):443–456. https://doi.org/10.1023/B:SCIE.0000045120.04489.80
20. Pouris A, Ho Y-S. Research emphasis and collaboration in Africa. Scientometrics. 2014;98(3):2169–2184. https://doi.org/10.1007/s11192-013-1156-8
21. Kahn M. Co-authorship as a proxy for collaboration: A cautionary tale. Sci Public Policy. 2018;45(1):117–123. https://doi.org/10.1093/scipol/scx052
22. Sooryamoorthy R. Transforming science in South Africa: Development, collaboration and productivity. Basingstoke: Palgrave Macmillan; 2015: p. 101–134. https://doi.org/10.1057/9781137493071
23. Wagner CS, Whetsell TA, Mukherjee S. International research collaboration: Novelty, conventionality, and atypicality in knowledge recombination. Res Policy. 2019;48:1260–1270. https://doi.org/10.1016/j.respol.2019.01.002