Bibliometrics of Cambodian Science: subjects, researchers and impact in the Science Citation Index Expanded

Julián Monge-Nájera; https://orcid.org/0000-0001-7764-2966
Siyan Yi; https://orcid.org/0000-0002-3045-5386
Yuh-Shan Ho; https://orcid.org/0000-0001-5090-3732

1. Laboratorio de Ecología Urbana, Vicerrectoría de Investigación, Universidad Estatal a Distancia, 2050 San José, Costa Rica; julianmonge@gmail.com
2. Saw Swee Hock School of Public Health, National University of Singapore, Singapore; KHANA Center for Population Health Research, Phnom Penh, Cambodia; siyan@doctor.com
3. Trend Research Centre, Asia University, No. 500 Lioufeng Road, Wufeng, Taichung 41354, Taiwan; ysho@asia.edu.tw (Correspondence*)
4. Center for Global Health Research, Touro University, Vallejo, CA, USA.

Received 08-II-2021. Corrected 06-V-2021. Accepted 19-V-2021.

Abstract. Introduction: Cambodia is a small agricultural tropical country for which only two small scientometric studies, published five years ago, were available until now. Objective: To identify, for Cambodian research, subjects, outlets, authors, institutions, citations and recommendations. Methods: The data were retrieved from the Science Citation Index Expanded (January 04, 2021) using the word “Cambodia” for the period 1972 to 2019. Results: We retrieved 3 689 documents: for the half century covered, the yearly presence of Cambodia in the index has grown strongly, to 325 articles and 6 555 authors in 2019 alone. Most documents are articles in English about health, particularly infectious and tropical diseases. Most international collaboration is done, by country, with the USA, France, and Thailand; and by institution, with Mahidol University, the Pasteur Institute and Oxford University. The most productive institutions are the Cambodian Ministry of Health and the Cambodian National Center for Parasitology. The main outlets are PLoS One, Malaria and PLoS Neglected Tropical Diseases. Cambodian articles are cited for up to 33 years, with a peak of 4.5 citations within the first two years. Those in English, or from well-funded foreign projects, have more citations in this particular index; especially if they are about malaria, hepatitis or influenza. Conclusions: The nature and impact of Cambodian science outside the SCI-EXPANDED remain unknown, but publications in that index have increased, concentrate on solving local problems, and depend heavily on international collaboration, following a well-known pattern of science in tropical countries. We suggest a funding system based on international peers who assign funds to the most productive researchers with minimal bureaucracy, so that local research is done on a greater variety of topics and with less participation of Cambodian researchers as low-level members in foreign projects.

Key words: scientific productivity of tropical countries; top research areas in the tropics; tropical health and diseases; scientific output; science and economic development.

Additionally to the problems of corruption and bureaucracy, science and technology in tropical countries are affected by the mismatch between extraordinary natural resources and the scarcity of funds to study, conserve and use them (Reboredo, Romano, & Armsworth, 2020). Similarities among tropical countries include the absence, in the Web of Science, of a large number of articles that are not in English but that are locally important, and the numerical dominance of articles produced by international projects led by well-financed...
institutions in the USA and Europe (Bah, Fu, & Ho, 2019; Calahorrano, Monge-Nájera, Wang, & Ho, 2020). The differences include the number of local scientific journals whose publications and citations are not recorded by the SCI-EXPANDED (from 18 in Guatemala to 130 in El Salvador, Monge-Nájera & Ho, 2017a; Monge-Nájera & Ho, 2018) and citation lifespans that range from 15 years in Panama to 70 years in Nicaragua (Monge-Nájera & Ho, 2015; Monge-Nájera & Ho, 2017b).

Here we continue the exploration of Asian scientific presence in the SCI-EXPANDED with Cambodia, an Asian tropical country with a population of 15.6 million in a small territory of approximately 181,000 km². The economy is based on agriculture and the country has a low per capita yearly income of US $1,643 (World Bank, 2020). Apparently, there are only two previous studies about Cambodian science; one, by Turpin, Zhang, Burgos, and Amaradsa (2015) considered Cambodia as part of a general survey of science in Southeast Asia and Oceania, and found that Cambodian science was growing rapidly but had a low level of development that made it highly depending on foreign researchers. Another study from the same year, covering the period 2000 to 2012, reported that 628 articles had been published, mainly on health topics, and that most were descriptive studies with a significant increase over time; this period was important because, in 2010, the Cambodian government established several policies to strength science (Goyet et al., 2015).

In this article, our goal is to present a larger study, covering half a century of Cambodian scientific output, and centering on subjects, authors, institutions and outlets, as well as citation, but limited to the Web of Science, because there is no equivalent to analyze Cambodian publications published in the official language of the country or in journals not covered by the Web of Science.

MATERIALS AND METHODS

The data was acquired from the Science Citation Index Expanded, Clarivate Analytics, Web of Science Core Collection (updated January 04, 2021). We carried out an advanced search using the word “Cambodia” in the field country (CU), limited to the period 1972 to 2019, and retrieved 3,689 publications. We retrieved document type (article, letter, book chapter, etc.), subject, language, institutions and authors. Data were analyzed with Excel. The journal impact factors (IF 2019) were extracted from the 2019 Journal Citation Reports (JCR).

In this database, the corresponding author is nominated as the ‘reprint author’. As a result, we will use the term ‘corresponding author’ in this study. In a single-author article where authorship is undefined, the author is categorized as both the first and corresponding author. Similarly, in a single-institution article, the institution is also listed as the first-author and the corresponding-author institution. In multiple corresponding-author articles, only the last corresponding-author, institution, and country are considered. In single-author articles only the Cambodia institution or the first Cambodia institution are considered to be corresponding institution (corresponding author is most likely to appear first or last in the byline, Mattsson, Sundberg & Laget, 2011).

Institutions in Cambodia were checked and merged, for example Inst Pasteur Cambodia, Inst Pasteur Cambodia, Inst Pasteur, Pasteur Inst Cambodia, Reseau Int Inst Pasteur, and Inst Pasteur Int Network were reclassified as Inst Pasteur Cambodia (Pasteur Institute in Cambodia). Affiliations in England, Scotland, Northern Ireland, and Wales were reclassified as in the United Kingdom (UK). Affiliations in Hong Kong prior to 1997 were included under the heading of China. Affiliations in French Guiana were reclassified as in France. Affiliations in Greenland were reclassified as in
Denmark. Affiliations in Czechoslovakia were checked and reclassified as in Czech Republic. Similarly, Czechoslovak Acad Sci (Czechoslovak Academy of Sciences) was reclassified as Czech Acad Sci (Czech Academy of Sciences). Of all document types, only articles were analyzed in detail, because they represented the majority of document types, as well as whole research ideas and results.

We applied three citation indicators to each article: the number of citations from the Web of Science Core Collection in the most recent year, $C_{2019}$ or the number of citations in 2019; the total citations (TC) from the Web of Science Core Collection from publication year to the end of 2019 were recorded as $TC_{2019}$; and we also calculated citations per publication ($CPP_{2019} = TC_{2019}/TP$ where TP is Total of Publications). We also measured Article lifespan as the number of years for which articles continue to be cited after publication.

**RESULTS**

**Historical output, coauthorship and length:** The Web of Science expanded index included 3,689 publications from Cambodia for the period 1972-2019, with a clear tendency to grow. Less than 23 articles were included before 2003, and a sharp increase took place in 2003, to reach 325 articles in 2019 alone.

It also increased from less than 70 total authors per year before 2002, to 6,555 authors having their work indexed in 2019 alone; the number of coauthors per article also increased significantly, from 2 or 3 in the 1970s, to 20 coauthors per article in 2019. The length of articles varied greatly but has had a mean of ten pages per document in recent years (Digital appendix 1). The mean number of authors per publication was 14 in articles, 9 in meeting abstracts and 12 authors in reviews (Digital appendix 2).

**Article types:** There were 13 document types, most of them articles (77%), followed by meeting abstracts, reviews and letters (Digital appendix 2). Note: The Web of Science can classify a document in more than one type, for example, 59 proceedings articles were also classified as articles, and seven reviews were also classified as book chapters, and thus the sum of percentages can be higher than 100%.

**Language:** Language of publication is one of the basic factors in bibliometric studies of big data. Cambodian research that reached the index was published in four languages, but 99% were in English (2,830 articles), followed distantly by French (0.67%), German (0.18%), and Spanish (0.035%); articles in English had far more coauthors than those in other languages (Digital appendix 3).

**Subjects:** On the one hand, most Cambodian articles deal with infectious diseases; tropical medicine; and health, while agronomy and zoology occupy the last places by frequency; on the other hand, publications tend to concentrate on relatively few journals (Fig. 1, further details in Digital appendix 4), the top exceptions being pharmacology, environmental sciences and health which appear in a larger number of journals (Fig. 1, Digital appendix 5).

All the top subject categories are increasing in output over the years, particularly the study of infectious diseases (Fig. 2).

**International collaboration:** The majority of collaboration was done with the USA (29%), France (22%), and Thailand (21%) (Digital appendix 6).

Most international collaboration articles with published with Mahidol University in Thailand (25%); the Pasteur Institute in France (17%), and the University of Oxford in the United Kingdom (17%) (Digital appendix 7). International collaboration is increasing rapidly, led by collaboration with the USA (Fig. 3).

**Institutions:** The most productive institutions are the Pasteur Institute (20%); the Cambodian Ministry of Health (10%); and the Cambodian National Center for Parasitology, Entomology and Malaria Control (8%) (Digital appendix 8).
Article lifespan and citation trends:
Cambodian articles can cite for three decades after their publication, and citation increased more rapidly in the first two years after publication, with a mean of one citation within the publication year, and a peak 4.5 in the two full years (Fig. 4).

Articles published in 2009 had the highest accumulated citations, \( CPP_{2009} \) with 50 citations. The maximal cumulative citations per publication \( CPP_{2009} \) was in reviews, with 54 citations per review, which was 2.6 times the citation rate for articles (Digital appendix 2).

Articles not in English had fewer citations, with \( CPP_{2009} \) of 2.9, versus a \( CPP_{2009} \) of 21 for English (Digital appendix 3). Overall, the citation trend for Cambodian articles is positive and shows an increase over time (Digital appendix 9), with the normal delay in citations for recent articles which have not had the time to accumulate more citations (Fig. 5). There is no visible effect of the 2010 government policies designed to boost scientific output (see Introduction, and Fig. 5).

Effect of country, institution and journal on citation:
Articles with researchers from Switzerland, the United Kingdom and Vietnam had the most citations (Digital appendix 6), particularly when done in collaboration with the World Health Organization, Mahidol University, and the University of Oxford (Digital appendix 7). Locally, the institutions with most citations were the National Center for Parasitology, Entomology and Malaria Control; the Pasteur Institute, and the WHO Cambodia.

Fig. 1. Distribution of Cambodian scientific production in the top 20 productive Web of Science categories in the Science Citation Index Expanded. Left vertical articles: percent articles on each category; right vertical axis: number of journals where articles on that category were published.
Fig. 2. Development of the top categories over time.

Fig. 3. Collaboration trends of the top four countries.
Fig. 4. Citation life span for Cambodia articles.

Fig. 5. Number of articles and citations per publication by year. The Cambodia National Science and Technology Master Plan was implemented in 2014 but there is no change in the growth curve as a result.
Representative Office (Digital appendix 8). When the research team leaders are not from Cambodia, the number of citations of the articles is higher (Fig. 6). For the top journals, citations increased over time (Digital appendix 10). The same applies to the most cited articles (Digital appendix 11).

DISCUSSION

The increasing number of Cambodian publications in the Science Citation Index Expanded is not different from the pattern found for every other tropical country previously studied (Crespo-Gascón, Tortosa, & Guerrero-Casado, 2019), and it possibly reflects both a real increase in scientific output (Turpin et al., 2015) and the particularly poor or even absent coverage of literature by SCI-EXPANDED database for some regions and years, a problem that was detected over a decade ago (Falagas, Pitsouni, Malietzis, & Pappas, 2008; Larsen, & Ins, 2010) and continues to be unsolved by the Web of Science (Mongeon & Paul-Hus, 2016). In the last two decades, the number of articles from that reach the Web of Science has steadily grown for all tropical studied (Trang, Monge-Nájera, & Ho, 2020), and we found the same trend in Cambodia, but expected an additional increase as a result of the Cambodian National Science and Technology Master Plan 2014-2020. However, our results do not show any change in the general trend; this can mean that the plan takes longer than five years to have any visible effect; that it affected only publications in the Khmer language, which are important but not counted by the Web of Science; or that the plan had no effect on the number of publications reaching the Web of Science.

The scientific output and citations of Cambodian science in the official language, Khmer, is locally important but missing in the
SCI-EXPANDED (Biddulph, 2010; Goyet et al., 2015; Turpin et al., 2015); this explains why publications in the database are almost all in English. The predominance of articles is also normal and has been found in many other tropical countries (Monge-Nájera & Ho, 2012; Monge-Nájera & Ho, 2018; Bah et al., 2019; Trang et al., 2020) and reflects the fact that articles are the unit of scientific literature (Ho, Satoh, & Lin, 2010).

The predominance of research on health, particularly infectious and tropical diseases under state health and parasitology bureaus and done with cooperation with the Pasteur Institute, is typical of the less industrialized countries all over the tropics, and results from a strong local need to improve their low health indices, still affected by tropical parasites and other diseases that are not important in the industrialized world (Monge-Nájera & Ho, 2017c; Tchuifen, Fu, & Ho, 2017; Monge-Nájera & Ho, 2018; Trang et al., 2020).

The presence of the USA, France, and Thailand as the countries which collaborate the most with Cambodian science, also follows a known tropical pattern called by Monge-Nájera and Ho (2017c) as the Affinity Model; that is, that tropical countries tend to participate in joint research with countries that are geographically and culturally close, including those that were their colonial powers in the past. This pattern has also been identified outside the tropics regarding the distance that knowledge flows between territories (Abramo, D’Angelo, & Di Costa, 2020). The USA, as world leader in science and technology, is also among the top three collaborative countries all over the tropics (Monge-Nájera & Ho, 2015; Tchuifen et al., 2017; Trang et al., 2020).

The 30 year citation lifespan for Cambodian articles places it closer to tropical countries that have a fluid scientific development, like Panama, with its 20 year lifespan (Monge-Nájera & Ho, 2015), than to less developed countries like Nicaragua, where a slow scientific advance is marked by a 70 year period in which the same articles continue to be cited (Monge-Nájera & Ho, 2017a).

Like everywhere in the tropics, the highest citation of Cambodian articles on health and led by foreign researchers is also explained by the fact that international health megaprojects are more likely to be done in areas of general interest where much research is done, published and cited (Tahamtan, Afshar, & Ahmazadeh, 2016); it is a matter of numbers rather than of quality and shows a large imbalance in detriment of small countries (Bah et al., 2019; Chinchilla-Rodriguez, Sugimoto, & Larivière, 2019). Along this line, the finding that projects led by Cambodians receive less citations in the Web of Science is not surprising, because Cambodian-led research is based on smaller projects and can be published in languages other than English and in journals not covered by the Web of Science (Goyet et al., 2015; Turpin et al., 2015).

In conclusion, the nature and impact of Cambodian science outside the SCI-EXPANDED remain unknown, but publications recorded by the SCI-EXPANDED have increased in recent years, concentrate on solving local problems, and depend heavily on international collaboration, following a well-known pattern of science in tropical countries. According to Allik, Lauk and Realo (2020), the best way to produce good science is to establish a strong and lasting democracy in a small and well governed country. Our recommendation would be that the Cambodian government establishes a funding system based on international peers who assign funds to the most productive researchers with minimal bureaucracy, so that local research is done on a greater variety of topics and with less participation of Cambodian researchers as low-level members in foreign projects.

**Ethical statement:** authors declare that they all agree with this publication and made significant contributions; that there is no conflict of interest of any kind; and that we followed all pertinent ethical and legal procedures and requirements. All financial sources are fully and clearly stated in the acknowledgements.
section. A signed document has been filed in the journal archives.

ACKNOWLEDGMENTS

We thank three anonymous reviewers for very useful suggestions to improve an earlier draft and Carolina Seas for her assistance with manuscript preparation.

RESUMEN

Bibliometría de la ciencia camboyana: temas, investigadores e impacto en el Science Citation Index Expanded

Introducción: Camboya es un país tropical asiático pequeño y “subdesarrollado”, con una economía basada en la agricultura, para el que hasta ahora solo estaban disponibles dos pequeños estudios cienciométricos, publicados hace cinco años. Objetivo: Identificar, para la investigación camboyana, qué se estuda; quién hizo la investigación; dónde y cuándo se publicó; y los factores que afectan su citación. Métodos: Usamos el Science Citation Index Expanded (4 de enero de 2021) utilizando la palabra “Camboya” y se limitaron al período 1972 a 2019. Resultados: Hallamos 3 689 documentos; durante el medio siglo cubierto, la presencia anual de Camboya en el índice ha aumentado considerablemente, con 325 artículos y 6 555 autores incorporados tan solo en 2019. La mayoría son artículos en inglés sobre salud, particularmente enfermedades infecciosas y tropicales. La mayor parte de la colaboración internacional se realiza, por país, con EE. UU., Francia y Tailandia; y por institución, con la Universidad Mahidol, el Instituto Pasteur y la Universidad de Oxford. Las instituciones más productivas son el Ministerio de Salud de Camboya y el Centro Nacional de Parasitología de Camboya. Las principales revistas son PLoS One, Malaria y PLoS Neglected Tropical Diseases. Los artículos camboyanos se citan hasta por 33 años, con un máximo de 4.5 citas en los dos primeros años. Los que son en inglés, o de proyectos dirigidos por el extranjero, tienen más citas en esta base de datos; particularmente si se trata de malaria, hepatitis o influenza. Conclusiones: La naturaleza y el impacto de la ciencia camboyana fuera del SCI-EXPANDED siguen sin conocerse, pero las publicaciones en ese índice han aumentado, se concentran en resolver problemas locales y dependen en gran medida de la colaboración internacional, siguiendo un patrón bien conocido en los países tropicales. Sugerimos un sistema de financiación basado en pares internacionales que asignen, con burocracia mínima, fondos a los investigadores más productivos, de modo que la investigación local se realice en una mayor variedad de temas y con una menor participación de investigadores camboyanos como miembros de bajo nivel en proyectos extranjeros.

Palabras clave: productividad científica de los países tropicales; principales áreas de investigación en los trópicos; salud y enfermedades tropicales; producción científica; ciencia y desarrollo económico.

REFERENCES


and public health priorities in low income countries: evidence from a systematic literature review focused on Cambodia. *Implementation Science, 10*(1), 1-12.


Larsen, P.O., & Ins, M.V. (2010). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics, 84*(3), 575-603. DOI: 10.1007/s11192-010-0202-z


See Digital Appendix at: / Ver Apéndice digital en: revistas.ucr.ac.cr