REFERÊNCIA

MARICATO, João Melo; SANTOS, Priscila Rodrigues dos. Highly Cited Publications index (HCP) for individual evaluation: applications in Brazilian information metric studies. **Library Philosophy and Practice**, art. 6394, 2021. Disponível em: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=12160&context=libphilprac. Acesso em: 18 nov. 2022.

Highly Cited Publications Index (HCP) for individual evaluation: applications in Brazilian Information Metric Studies

João Melo Maricato
University of Brasília, jmmaricato@gmail.com

Priscila Rodrigues dos Santos
University of Brasília, priscilarssantos@outlook.com

Highly Cited Publications Index (HCP) for individual evaluation: applications in Brazilian Information Metric Studies

João de Melo Maricato
University of Brasilia, Faculty of Information Science, Brasilia, Brazil
E-mail:jmmaricato@gmail.com
Orcid: 0000-0001-9162-6866

Priscila Rodrigues dos Santos
University of Brasilia, Faculty of Information Science, Brasilia, Brazil
E-mail: priscilarssantos@outlook.com
Orcid: 0000-0002-1897-8838

Abstract
The use of different bibliometric indicators to analyze the individual scientific production of researchers constantly raises questions among several authors. Based on this problem, Waltman, Van Eck and Wouters (2013) propose a methodology that values highly cited papers (high impact) over low cited papers (low impact). To this end, the authors developed a stylized Highly Cited Publications (HCP) index aiming to minimize some limitations inherent to the metrics used to evaluate scientists. The present work proposes to carry out the empirical application of the HCP index stylized by the authors to researchers in the field of Information Metrics Studies (IMS) in Brazil in order to identify its potentialities and limitations. The research data was limited to 101 researchers in the area of IMS in Brazil, who had their production and citation data extracted from Google Citations in the month of July 2020, being evaluated different data sets in two distinct time periods (1941-2020 versus 2010-2020), for the purpose of comparative analysis, being possible to observe that according to Person's correlation, the use of different data sets, as long as they respect the same variables, do not show significant changes in the ranking of the researchers' HCP indexes. The use of the stylized HCP index is interesting for the identification of highly cited publications and can be used to complement the evaluation indicators, but it presents limitations that may inhibit its use, for example, the negative indexes of specific researchers.

Keywords: Bibliometrics, Information metrics studies, Individual evaluation, Highly cited publications (HCP)

DECLARATIONS

Authors' contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by João de Melo Maricato and Priscila Rodrigues dos Santos. The first draft of the manuscript was written by Priscila Rodrigues dos Santos and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgement

We are grateful for the contributions of Dr. Rodrigo Costas Comesana, a CWTS researcher at the University of Leiden, during the course that included his participation at the University of

Brasília in 2020.

1. INTRODUCTION

In modern science, Information Metrics Studies (IMS) are applied to different academic contexts in order to evaluate scientific impacts. In the evaluation of scientific production, the data sources are usually books, articles, book chapters, journals, papers presented at conferences, among others (Mueller, 2008). However, due to the various facets of science, the indicators currently available are not able to map and evaluate the various aspects of scientific production (Dainesi & Pietrobon, 2007), with limitations to studies that use bibliometric, scientometric, and altmetric methods. IMS, in this paper, is used to group researchers who work in the areas of bibliometrics, scientometrics, informetrics, patentometrics, webometrics, altmetrics and related fields, strongly linked to library and information science.

Indicators to evaluate researchers' productivity are widely discussed by authors with different approaches (Hicks et al., 2015; Hirsch, 2005; Thomaz, Assad & Moreira, 2011; Mattedi & Spiess, 2017). Among the most commonly used indicators are the h-index and the g-index, which are metrics commonly used to evaluate individual researchers, by relating indicators of published scientific production and citations in a given period of time. The use of the HCP is commonly justified in the literature due to the limitations of the h-index, however, in this research we understand that these indices have different objectives and purposes and that, therefore, they should be understood as complementary.

Bibliometric and scientometric indicators can be used in different contexts and bring several benefits to scientific communication, but there are also discussions and criticisms about the criteria used, applications, and their uses. Any metric used to measure scholarly and scientific activities is susceptible to criticism, whether its applications are at the level of journals, countries, institutions, documents, or individual researchers.

One way to measure the individual scientific production of a researcher, the central object of analysis in this research, discussed by authors Waltman, Van Eck and Wouters (2013), is the Highly Cited Publications (HCP) index. The authors address the hypothesis that the more a researcher produces, the higher the quality their publications present, i.e. if the greater quantity of publications results in a greater quantity of citations.

In this context, HCP arises as a proposal to evaluate researchers and research individually, seeking to correct inconsistencies of the h-index, by valuing the most cited publications in a knowledge area. The HCP index, proposed by the authors, uses a formula that relates production and citations, starting from the pre-established theoretical premises that

10% of publications in an area have high impact (highly cited) and 90% have low impact (low cited) (Waltman, Van Eck & Wouters, 2013).

The present study proposes to empirically test the HCP formula adapted by Waltman, Van Eck and Wouters (2013), using data from researchers linked to Brazilian institutions working in the area of Information Metrics Studies (IMS). Thus, this study aims to test the formula proposed by the authors in order to analyze the possibilities of use and limits of application of HCP in the real Brazilian context, as well as to situate these researchers in the area of IMS, according to this methodology.

Considering the formula application with data collected from Google Citation, one may also discuss this data source for the application of the HCP formula. Researchers test this formula on different sets and subsets of production and citation data from works in the field of IMS in Brazil, investigating the possibilities and limits of HCP application from different perspectives. We performed the tests in two different periods (1941-2020 and 2010-2020) enabling the comparison of the application of HCP and the visualization of researchers in the IMS field in two different historical moments.

The tests are performed considering analyses of four subsets of data: 1) total papers published, related or not to IMS, by researchers active in the field; 2) total papers published, related or not to IMS, published by researchers active in the field, excluding those not cited; 3) only IMS-related papers published by researchers active in the field; 4) only IMS-related papers published by researchers active in the field, excluding those not cited.

Given the above, from the two temporal datasets and the respective four subsets of data, comparisons are made and Pearson correlation indices are calculated in order to answer: what are the main difficulties and problems associated with the calculation of HCP using Google Citation production and citation data of Brazilian researchers? Which production and citation datasets are more appropriate for the calculation of HCP? Is it possible to use data without prior thematic classification of research to identify which are and which are not from the subject area of analysis of a group of researchers in this case, the area of IMS? Is it more appropriate to calculate the HCP considering the papers with zero citations or the totality of papers, regardless of whether they were cited or not? Has the Brazilian area of IMS suffered important changes in the positions of researchers in the rankings, from the point of view of the HCP index?

2. EVALUATING RESEARCHERS: INDICES AND THE HCP PROPOSAL

The results of scientific research are, in general, communicated through scientific publications. The acceptance by the scientific community of these works is strongly related to the citations received, and from these it is possible to make approximations with the quality of the research developed. The quantitative evaluation of researchers, universities, countries, graduate programs, among others, with the use of the IMS, frequently occurs by studying the quantitative production, especially of scientific articles and their citations in subsequent publications.

Various methodologies, indexes, and quantitative indicators have been developed over the years to evaluate journals, articles, and researchers. Many of these initiatives are criticized for evaluating publications by taking into consideration the number of citations of the set of articles published by a given journal in a given period. These indicators are normally used to score the vehicles, whose indexes or scores are attributed, by inference, to the productions and researchers individually. Some examples of these indexes are the Impact factor (IF) (Clarivate Analytics), CiteScore (Elsevier), Eigenfactor (University of Washington) and the Normalized Impact per Paper (SNIP) of Leiden University. Following a similar logic, but using an alphanumeric classification, one can mention the Qualis reference, produced in Brazil by the Coordination for the Improvement of Higher Education Personnel (Capes).

Faced with the criticism of using journal indicators, based on the logic illustrated, new proposals of indexes and indicators have been developed to evaluate researchers and papers, such as the h-index, the h10-index, g-index, RG Score (ResearchGate) and the Altmetric Score (Digital Science), among others. These indicators make it possible to evaluate individuals and their productions individually, but they are also subject to criticism (Chart 1).

Chart 1 Most commonly used bibliometric indicators to assess individual researchers and scientific production.

Index	Proposed	Characteristic
h-index	Jorge Hirsch (2005)	It is a single number indicator that evaluates the quantity and impact of a researcher's scientific production. Despite being an indicator of easy understanding, the h-index presents inconsistency in the results, since one of its limitations is not taking into account highly cited articles.
g-index	Leo Egghe (2006)	Developed to correct the h-index in relation to the most cited articles, it gives greater weight to them. Even the g-index corrects this limitation, the authors Huang and Chi (2010) consider not to be the most suitable for comparing the scientific productions of researchers of different ages, due to the restriction of publication time.
i10-index	Google (2011)	This measure is used by Google Scholar, whose value corresponds to the number of publications, in the last 5 years, with at least 10 citations (Nascimento, 2016). Although it is useful to identify researchers with a certain level of impact in a given field, the i10-index is limiting because it does not account for publications with fewer than 10 citations (Aithal, 2017), as well as disregards highly cited publications in its calculation.
RG Score	ResearchGate (2009)	A bibliometric/altmetric indicator from ResearchGate, a proprietary academic social platform. The basis for calculating the RG Score is any contribution that the researcher

		makes in his/her profile (published articles, projects, questions and answers, among other interactions). One of the main problems with its use relates to the difficulty of application in other contexts and the impossibility of auditing the data.
Altmetric score	Digital Science (2010)	It is a weighted count of the mentions that an academic product receives in different sources, carried out by the Altmetric platform. The dissemination of the product, like articles, receives different weights depending on the media. More weight is given to dissemination in news journals compared to blog posts, Twitter and other media. However, the altmetric score is proprietary and its methodology is not clearly disclosed (Trueger, 2015).

Source: Self elaborated (2020).

The vast number of indexes and indicators have a strong interest in correcting or mitigating limitations of their predecessors, including the possibility of measuring results in different channels and sources of information. In turn, the HCP, a calculation methodology elaborated to cope with limitations of the h-index, was developed to measure high impact publications, given that "HCPs are considered a symbol of scientific excellence and superior performance" (Ma & Li and Shang, 2020).

For Bornmann (2013), highly cited publications are those that are part of the top 10% of the most referred publications within their subject area. Garfield (1977) and Moral-Munoz et al. (2016) already talked about highly cited papers using the nomenclature "citation classics." According to the authors, these classics are important since, from them, it is possible to understand the current, past, and future scientific structure of a given field of knowledge.

Highly cited publications have unique aspects, since citation rates vary between fields of knowledge, hence Glanzel and Shubert (1992) acknowledge there is no standard criterion to define highly cited publications. Clarivate Analytics (2020) indicates some characteristics that influence high citation rates, such as the means used for the divulgation of results, the disciplinary area and the nature of the discovery. For this reason, it is believed that thematic classification of scientific production may be necessary, decreasing the possibility of biases when calculating HCP.

The HCP is used in databases that evaluate journals, such as the InCites platform, which hosts a service called Essential Science Indicators (ESI), which makes use of the index to rank its publications. According to InCites, publications ranked as highly cited through the HCP are considered indicators of scientific excellence and high performance and comparisons between scientific fields can be used internationally (Incites, 2019).

One of the advantages of the HCP indicator pointed out by Waltman and Van Eck (2012) is that, despite having similarities with the h-index, it does not produce inconsistent rankings in its results, as do those that occur with the h-index. Bornmann, De Moya-Anegón, and Leydesdorff (2011) consider as the most obvious advantages offered by HCP the possibility of comparing the reference value (10% expected), with the actual value measured. Nevertheless, the HCP, in its original formulation, may present incongruence in its results, such as "[...] overestimating the scientific impact of researchers who focus on producing many publications without paying much attention to the impact of their work. " (Waltman, Van Eck & Wouters 2013, p. 4)

Given the inconsistencies in the traditional HCP calculation, especially since the calculation of the index in the traditional way is highly influenced by the high number of publications, Waltman, Van Eck and Wouters (2013) make adaptations in the HCP index in order to eliminate observed restrictions. In their study, a theoretical article dwelling on the use of a stylized model of the HCP index, based on a relation of production and citations. The stylized formula proposed by the researchers to measure HCP is constructed in two steps. The first part aims to establish the relationship between scientific impact and citations, and to this end, the authors create the following parameter: α ($0 \leq \alpha \leq 0.09$), where 0 is assigned to low cited publications and 0.09 to highly cited publications. Thus, the weighting of highly cited publications is $wHC = \frac{0.1\alpha - 0.09}{\alpha - 0.09}$ and, for the low quoted publications $wLC = \frac{0.1\alpha}{\alpha - 0.09}$. When applying these formulas, the resulting values are for wHC = 4.15 and for wLC = - 0.35. However, these values, according to the authors, can be conventionally adjusted due to the characteristics of the different areas of knowledge (Waltman, Van Eck & Wouters, 2013). In the present research, we chose to adopt them in the way they were calculated by the authors, without any kind of change.

The HCP value is obtained through the following formula HCP = nLCwLC +nHCwHC. While wHC = 4.15 and wLC = - 0.35, the values of nLC and nHC correspond to the quantities of the authors' publications that were low cited (nLC) and highly cited (nHC). A highly cited publication corresponds to those that, of the total number of publications in a knowledge area, are part of the 10% most cited publications. The remaining publications (90%) are considered as low cited publications. Counting the highly cited and less cited publications of a researcher in a certain knowledge area and applying this to the formula (HCP = nLCwLC + nHCwHC) produces the HCP index. In the authors' modified HCP index the wLC is negative. Therefore, low cited publications are given a negative weight, so the less cited publications someone has, the lower their HCP value will be. (Waltman, Van Eck & Wouters, 2013).

3. METHODS

To apply and perform experiments with Waltman, Van Eck and Wouters' (2013) stylized HCP formula, production and citation data of researchers working in the field of IMS in Brazil were selected and analyzed. The identification of authors, their production and respective citations took place in their profiles in the Google Citations platform, and if they did not have profiles in Google Citations, the data were extracted through Harzing's Publish or Perish tool.

The selection of researchers active in the field of IMS in Brazil occurred by consulting previous research that studied the area (Urbizagastegui, 2016; Gabriel Junior, Moraes & Oliveira, 2018; Lima, Soares & Oliveira, 2011; Reis, Nogueira & Oliveira, 2019; Oliveira, 2013; Gabriel Junior, Freitas & Bufrem, 2011). In these, 108 individuals were identified, where 70 (65%) had their profiles identified in Google Citation, 31 (29%) had no profile in Google Citation (they were identified with Harzing's Publish or Perish tool) and 6 (6%) were not identified, being discarded. Thus, 101 researchers and their productions and citations (including self-citations) were considered in this study.

It is believed that, with this methodology, we reached a representative part of the scientific community linked to Brazilian institutions in the field of IMS and, thus, of the productions and citations in the area. However, it is necessary to recognize that it is not possible to guarantee that all researchers active in the field and their respective productions were identified. The scientific production of the researchers was not segmented by document type, i.e. all types of production (dissertations, theses, papers presented at events, books and chapters, etc.) were considered for the HCP calculation. The researchers' production was not previously delimited by year of publication, having been collected in the month of July 2020.

To identify the highly cited publications in the area of IMS in Brazil, the production of the 101 researchers was ranked decreasingly by citations. Thus, it was possible to identify the high impact publications in the area of IMS in Brazil (10% of the most cited publications by the area) and the low impact ones (90% less cited) and to perform the subsequent application of the HCP formula proposed by Waltman, Van Eck, and Wouters (2013).

The HCP calculation followed the analytical structure considering: 1) two temporal cut-offs (being the first period with productions and citations from 1941 to 2020 and the second between 2010 and 2020); 2) two data sets; 3) two data subsets; and, 4) four data variables. A data set was structured referring to all published papers (100% of production and citation) for the period 1941-2020, which was named "100% Data Sets (1941-2020)". From this data set a subset of data was structured referring to the 10% most cited papers in the same period, which was named "Subset of data of approximately 10% (1941-2020)". Following the same principle, a data set was structured with all published papers (100% of production and citation) in the period between 2010 and 2020, called "Data sets of 100% (2010-2020)" from which the subset called "Data subset of approximately 10% (2010-2020)" was extracted. The temporal clippings allow the temporal comparison of the application of the HCP and visualization of researchers working in the area of IMS in two distinct historical moments.

Each of the data sets and subsets were analyzed considering different data variables, enabling analysis of the HCP from various perspectives. Each of the data sets and subsets were analyzed, in the two time periods, considering the following variables: Variable 1 - all published papers, related or not to IMS, by researchers active in the field; Variable 2 - all published papers, related or not to IMS, published by researchers active in the field, excluding the papers that did not receive citations; Variable 3 - only papers related to IMS published by researchers active in the field; Variable 4 - only papers related to IMS published by researchers active in the field, excluding the papers that did not receive citations.

The data variables were structured with the totality of the papers published by the researchers and, also, with publications only from the IMS area, in order to compare the HCP index in these two contexts. Since many researchers who publish in the field of IMS also publish in other fields of knowledge, it was necessary to analyze the publications individually in order to classify them thematically. IMS publications were considered to be those related to topics not limited to, but including: scientific communication, databases, university-company relations, sociology of science, information retrieval, information organization, science policy, technology foresight, and scientific divulgation.

Another necessary methodological procedure concerns the selection of the 10% of highly cited productions. It was necessary to approximate upwards, since it is not possible to select precisely the 10% papers, since the number of minimum citations that a paper receives may coincide among different researchers. This difficulty of working with precise percentiles, for HCP calculation purposes, is also a concern of other researchers (Schreiber, 2013).

Finally, starting from these of the two subsets of data and the respective variables, comparisons are made and Pearson correlation indices are calculated in order to answer the questions guiding this research.

4. RESULTS AND DISCUSSIONS

First, analyses are performed of the publications and citations referring to the "100% datasets (1941-2020)" and then the "100% datasets (2010-2020)". From the sets, the identification and analysis of highly cited (10%) and low cited (90%) publications is performed for the respective periods: "Approximately 10% subset of data (1941-2020)" and "Approximately 10% subset of data (2010-2020)". Once these two preliminary analysis steps are completed, the authors' HCP calculations are analyzed considering the respective temporal cut-offs.

4.1. Production and citation data (1941-2020)

Table 1 shows the total production and citation data of researchers working in the field of IMS in Brazil and their respective data variables for the period 1941 to 2020 ("100% full version"). During this period, researchers published a total of 14,146 papers, which received 103,748 citations (average of 7.3 citations per paper), with the most cited paper receiving 1,484 citations. As the data variables are limited, the number of papers decreases. When the publications that did not receive citations ("100% full version without 0 citations") are disregarded, the number of papers decreases to 7,312, but the total number of citations and the most cited paper does not change, because in this data variable only the papers that did not receive citations were excluded. However, the average number of citations per paper increases considerably to 14.2.

In Table 1, the data variable "Full version of 100% IMS only", groups papers published by researchers who are active in the field and the respective papers also have a relationship to IMS (in this data set, papers that had no relationship to IMS were discarded). Thus, the total number of papers published by researchers in the area related to IMS was 5,928 and the respective citations totaled 56,627. The average number of citations for this data variable is 9.6, showing that papers related to IMS receive more citations than papers published by the same researchers on other subjects. When papers that did not receive any citations are excluded ("100% full version of IMS only without 0 citations"), we arrive at a total of 3,470 papers and 56,627 citations (16.3 citations per paper on average).

Table 1 100% dataset of production and citation of researchers working in the field of IMS in Brazil (1941-2020)

Data sets of 100% (1941-2020)	Papers (T)	Citations (C)	Mean (C/T)	C Máx	C Min
100% complete version (100% of all papers, cited or not, in the whole period, including those not in the IMS area, including undated ones)	14145	103748	7,3	1484	0
100% full version without 0 citation (100% of all papers, including those not in the IMS area, excluding those not cited, in the entire period, including those without date)	7312	103748	14,2	1484	1
100% complete version of IMS only (100% of all cited and uncited IMS papers from the entire period, including undated ones)	5928	56627	9,6	1484	0
Full version of 100% IMS only without 0 citation (100% of all papers in the IMS area, excluding those not cited, for the entire period, including undated ones)	3470	56627	16,3	1484	1

Source: Research data (2020).

From Table 1, Table 2 is elaborated, which contemplates the data about the highly cited papers published by researchers in the field of IMS in the period between 1941 and 2020. As indicated in Table 2, the 10% of all highly cited papers ("Full version of approximately 10%") totals 1,432 and the respective citations were 73,155; the most cited paper obtained 1,484 citations and the least cited of this set was 18. The average of highly cited publications by researchers working in the field of IMS in Brazil is 51.1 citations per paper, significantly higher than the general average presented in Table 1 (average of 7.3 citations per publication). The highest number of citations presented in Table 2 coincides with those presented in Table 1 in all variables, because the most cited article is related to IMS.

Among the possibilities of analysis of Table 2, it is noteworthy that the average number of citations to papers related to IMS by researchers working in the field (regardless of the data variables) are substantially higher than those of papers published considering the total set, including those not from the field of IMS; e.g. the data variable "Full version of approximately 10% IMS only" the average number of citations per paper is 93.5. This demonstrates the numerical importance of citations of papers from the IMS area, to the detriment of other themes that researchers also work on.

The number of highly cited publications considered for the HCP calculation was between 10.12% and 10.20% (column "% considered" in Table 2). Despite not being represented in Table 2, the low cited production in each of the data sets and respective variables can be calculated by subtracting the data presented in the "Papers (T)" column of Table 2 and Table 1. Thus, for example, the amount of low cited papers referring to the data variable "Full version of approximately 10%" in Table 2 is 12,713 publications, or 89.9%. The impossibility of precisely defining the top 10% most cited is considered to be a limitation of the HCP, also observed by Schreiber (2013) who noted that there is an ambiguity in the calculation of the percentiles because generally papers with the same citation count are found on the border between the percentile classes.

Table 2 Subset of data from the highly cited production (10%) of researchers working in the field of IMS in Brazil (1941-2020)

Data subset of approximately 10% (1941-2020)	Papers (T)	Citations (C)	Mean (C/T)	C Max	C Min	% considered
Full version of approximately 10% (10% of all papers ranked decreasingly by citation, for the entire period, including non-IMS papers, cited or not, including undated ones)	1.432	73.155	51,1	1.484	18	10,12%
Full version of approximately 10% without 0 citation (10% of all papers ranked decreasingly by citation, including those not in IMS, excluding those not cited, for the entire period, including those without date)	745	57.317	76,9	1.484	33	10,18%
Complete version of approximately 10% IMS only (10% of all papers ranked descending by citation, from the area of IMS cited and not cited, in the entire period, including undated ones)	600	39.596	66,0	1.484	21	10,12%
Full version of approximately 10% IMS-only papers without 0 citation (10% of all papers ranked decreasingly by citation, from the IMS area, excluding those not cited, for the entire period, including undated ones)	354	33.100	93,5	1.484	35	10,20%

Source: Research data (2020).

4.2. Production and citation data (2010-2020)

Table 3 and Table 4 follow the same principle as Table 1 and Table 2. The difference between them refers to the time frame. Table 3 presents the 100% of papers produced by researchers working in the field of IMS in Brazil in the period from 2010 to 2020 ("100% datasets (2010-2020)"), and their respective data variables. In the last 11 years researchers produced 7,722 papers, more than half compared to the entire collected period (1941-2020).

These 7,722 papers were cited 42,131 times (average of 5.5 citations per paper), lower than the average number of citations to publications in the 1941-2020 dataset presented in Table 1 (average of 7.3 citations per publication). A similar behavior occurs across time cut-offs and data variables from production, citation, and averages perspectives. In general, the production data show the growth and strengthening of the area of IMS in Brazil in the last decades. The fact that papers, on average, were less cited in the second period (2010-2020) compared to the first (1941-2020) may be due to the publication period, with more possibility of accumulating citations over time when compared to more recent papers. For Hu et al. (2018), it does not make sense to compare the citation frequency between two papers published in different years or areas, since the citation count of a paper is sensitive to citation time windows, publication types, and research areas.

In order to apply the formula indicated by authors Waltman, Van Eck and Wouters (2013) to calculate the HCP, for the period between 2010-2020, publications that did not have a date were discarded. It is noted that the most cited paper in the period between 2010-2020 received 453 citations and, due to the fact that it is related to the area of IMS, it appears in all data variables (Table 3). Furthermore, it should be noted that the average number of citations in each of the data variables presented in Table 3 is quantitatively less discrepant when compared to the same averages presented in Table 1, i.e. although IMS-related publications in both periods have higher citation averages, the data indicate these differences are less important in the second period (2010-2020).

Table 3 100% dataset of production and citation of researchers working in the field of IMS in Brazil (2010-2020)

data set of 100% (2010-2020)	**Papers (T)**	**Citations (C)**	**Mean (C/T)**	**C Max**	**C Min**
2010 version of 100% total (100% of all papers, cited or not, published from 2010 onwards, including those not in the IMS field, excluding undated ones)	7.722	42.131	5,5	453	14
Version 2010 of 100% without 0 citation (100% of all papers, published from 2010 onwards, including non-IMS papers, excluding those not cited and those without date)	4.206	42.129	10,0	453	1
Version 2010 of 100% only IMS with 0 citation (100% of all papers, published from 2010 onwards, from the area of IMS cited and not cited, excluding and undated)	3.510	21.849	6,2	453	0
2010 version of 100% only IMS without 0 citation (100% of all papers in the area of IMS, published from 2010 onwards, excluding those not cited and those without date)	2.096	21.849	10,4	453	1

Source: Research data (2020).

Table 4 shows the approximate 10% of the data presented in Table 3 ("Approximate 10% data subset (2010-2020)"), i.e. the data of highly cited papers published by researchers working in the field of IMS in the different data variables between 2010 and 2020. The highly cited papers in the period were 792 ("2010 version of approximately 10% with 0 citations"), which received 28,043 citations (average citations per paper of 35.4). The minimum citations in the subset of data presented in Table 4 range between 14 and 22 citations per publication.

The percentages of highly cited publications, which will be used for the HCP calculation in this second time frame (2010-2020), presented in the dataset in Table 4 are somewhat higher when compared to those in Table 2, reaching 10.68% in the data variable "2010 version of approximately 10% only IMS with 0 citations". This limitation for calculating the HCP (difficulty in precisely selecting the 10% of most cited papers) tends to grow gradually when the number of papers becomes smaller, as occurred in this research when using the 11-year time frame. In this research, the difference was not so great (10.68%), but the use of HCP in shorter periods or in smaller scientific communities may be unfeasible.

Table 4 Subset of data from the highly cited production (10%) of researchers working in the field of IMS in Brazil (2010-2020)

Data subset of approximately 10% (2010-2020)	Papers (T)	Citations (C)	Mean (C/T)	C Max	C Min	% considered
2010 version of approximately 10% with 0 citations (10% of all papers ranked decreasingly by citation, as of 2010, including non-IMS papers, whether cited or not, excluding undated)	792	28043	35,4	453	14	10,26%
2010 version of approximately 10% without 0 citation (10% of all papers ranked decreasingly by citation, published from 2010 onwards, including those not in the IMS field, excluding those not cited and those without date)	424	21817	51,5	453	22	10,08%
2010 version of approximately 10% only IMS with 0 citation (10% of all papers ranked decreasingly by citation, published as of 2010, from the IMS area, cited and not cited, excluding undated)	375	14485	38,6	453	15	10,68%
2010 version of approximately 10% IMS-only papers without 0 citation (10% of all papers ranked decreasingly by citation from the IMS area, published from 2010 onwards, excluding those not cited and those undated)	214	11643	54,4	453	22	10,21%

Source: Research data (2020).

4.3. Calculation of HCP for data subsets

From the data presented in the previous section, in which the highly cited (approximately 10%) and low cited (approximately 90%) papers of researchers working in the field of IMS in Brazil were identified, it was possible to calculate the HCP indexes and investigate the Pearson correlations of each of the rankings according to the data variables. We first analyze the publications in Table 2 (for the period 1941-2020) and then those presented in Table 4 (for the period 2010-2020).

4.4. Calculation of HCP in subsets and data variables (1941-2010)

In the subset of data for the period between 1941 and 2010, there are different HCP rates of the researchers due to the data variables. The largest HCPs in each data variable were respectively: in the "Full version of approximately 10%" the largest HCP was 176.7; in the "Full version of approximately 10% without 0 citation" 133.55; in the "Full version of approximately 10% only IMS" 143.35; and in the "Full version of approximately 10% only IMS without 0 citation" 98.55. Despite the differences in the largest HCPs, they cannot be considered in themselves a problem, since eventual uses of the index should occur in an isonomic manner.

On the other hand, the researchers' smallest HCPs presented negative values. The lowest HCPs in each data variable were respectively: "Full version of approximately 10%" the lowest HCP was -191.25; in "Full version of approximately 10% without 0 citation" -34.8; in "Full version of approximately 10% only IMS" -52.9; and, in "Full version of approximately 10% only IMS without 0 citation" -22.2. The negative index of the researchers' HCPs (approximately 50% in all data sets) causes strangeness and may lead to questioning and criticism. The negative value can cause misinterpretations and resistance in using the HCP for evaluation purposes and researchers.

What authors Waltman, Van Eck and Wouters (2013) consider a positive point of their proposed formula for calculating HCP can also be seen as a limitation. However, the wLC values can be adjusted, according to the authors themselves. One can, in other analyses, increase the value of wLC (remembering that the one proposed by the authors was - 0.35), until one reaches a value considered adequate for the context of HCP use (area, countries, size of the scientific community, etc.). In the context analyzed in this research, it is considered that the value should be adjusted, but in other research, for example Praus (2019), despite also having negative HCPs, this was not reported a factor to be corrected.

Despite this limitation, it is understood to be more relevant to understand whether there are differences in the data variables. To investigate this, the Pearson correlation between the four data variables of the "Approximately 10% (1941-2020) subset of data" was analyzed. The correlations of the authors' rankings can be seen in Table 5. The correlations between all the data variables are either strong or very strong, the lowest being the correlation between the data variables "Complete version of approximately 10%" and "Complete version of approximately 10% only IMS without 0 citation"; even then the correlation is considered strong according to Pearson's scale (0.77).

Table 5 Correlation between the rankings of researchers working in the field of IMS in Brazil according to their HCPs index by data variables (1941-2020)

Data Variables	Full version of about 10%	Full version of about 10% without 0 quotes	Full version of about 10% IMS only	Full version of approximately 10% only IMS without 0 quotes
Full version of about 10%.	1	0,84	0,90	0,77
Full version of about 10% without 0 quotes		1	0,81	0,90
Full version of about 10% IMS only			1	0,86
Full version of approximately 10% only				1

IMS without 0 quotes					

Source: Research data (2020).

In turn, in the period between 1941-2020, the highest correlation indices of the authors' rankings (disregarding the correlations of the data variables with themselves), are between the "Full version of approximately 10%" and "Full version of approximately 10% only IMS "(correlation of 0.90) and that between the "Full version of approximately 10% without 0 citation" and the "Full version of approximately 10% only IMS without 0 citation", also with a correlation of 0.90 both considered very strong correlations.

In view of the analyses of the data of researchers working in the IMS field who published in the period 1941-2020, it is possible to state that, for the calculation of the authors' HCP, one can select the papers that are or are not from the IMS field, because the correlations were very strong (0.9). Similarly, the data show that the use of the data variables for HCP calculation purposes, with or without citation, are indifferent, because the correlation was also very strong between these two sets of data (0.9). Therefore, there is consistency in calculating the authors' HCP with the totality of publications, regardless of whether or not they are from the field of IMS, or if they are cited or not (zero citations), without major losses to the evaluation and ranking of researchers.

In view of the correlations presented, we highlight that the rankings do not present significant differences between the positions of the authors due to the data variables used. Variations in the positions among the authors may occur, but, in general, they vary little, the lowest of which is found in the case of the correlations identified as 0.90. Since the correlations are not perfect (1), naturally, the data set analyzed qualitatively to identify publications that are or are not related to the area of IMS, present more precision for the calculation of the area's HCP. Nevertheless, the finding is important because the classification of papers according to subject matter (in this case related and unrelated to IMS) is time-consuming and may make large-scale studies unfeasible.

We have chosen, as an example, to present the authors and their respective HCPs calculated from the data variable "Complete version of approximately 10% IMS only" referring to the "Subset of approximately 10% data (1941-2020)". Table 6 presents the first 60 researchers working in the field of IMS in Brazil who published papers related to the area, regardless of whether they were cited, in this thematic between 1941-2020 and their respective HCPs in descending order.

Table 6 Ranking of the HCP indexes of 60 researchers working in the field of IMS in Brazil and their respective publications, in the theme, highly cited and lightly cited (1941-2020).

Researchers	Highly cited (IMS)	Low cited (IMS)	Total Publications (IMS)	HCP
Velho, Lea	45	124	169	143,35
Mueller, Suzana Pinheiro Machado	22	52	74	73,1
Leta, Jacqueline	23	66	89	72,35
Rossoni, Luciano	18	26	44	65,6
Meneghini, Rogério	16	42	58	51,7
Marteleto, Regina	15	41	56	47,9
Guimarães, José Augusto	19	114	133	38,95
Kobashi, Nair Yumiko	13	48	61	37,15
Ensslin, Leonardo	17	102	119	34,85
Ensslin, Sandra Rolim	19	126	145	34,75
Riccio, Edson Luiz	11	38	49	32,35
Stumpf, Ida Regina	14	84	98	28,7
Witter, Geraldina Porto	11	50	61	28,15
Beuren, Ilse Maria	14	87	101	27,65
Alvarenga, Lídia	9	29	38	27,2
Mugnaini, Rogério	12	69	81	25,65
González de Gomez, Maria Nélida	7	18	25	22,75
Cesar Júnior, Roberto M.	6	9	15	21,75
Guarido Filho, Edson Ronaldo	6	12	18	20,7
Hocayen-da-Silva, Antônio João	6	14	20	20
Pinheiro, Lena Vania	10	62	72	19,8
Mena-Chalco, Jesus P.	11	76	87	19,05
Vanti, Nadia Aurora	6	25	31	16,15
Vanz, Samile Andréa de Souza	12	97	109	15,85
Parreiras, Fernando Silva	6	30	36	14,4
Cruz, Ana Paula Capuano da	5	19	24	14,1
Braga, Gilda Maria	5	20	25	13,75
Hayashi, Maria Cristina Piumbato Innocentini	13	115	128	13,7
Araújo, Carlos Alberto Ávila	10	80	90	13,5
Noronha, Daisy Pires	9	70	79	12,85
Matheus, Renato Fabiano	4	14	18	11,7
Santos, Raimundo Nonato Macedo dos	13	121	134	11,6
Targino, Maria das Graças	5	29	34	10,6
Espejo, Márcia Maria dos Santos Bortolocci	5	33	38	9,2
Silva, Márcia Regina da	4	24	28	8,2
Ferreira Júnior, Amarílio	3	14	17	7,55
Dias, Eduardo Jose Wense	3	15	18	7,2
Lima, Ricardo Arcanjo de	2	5	7	6,55
Población, Dinah Apparecida de Mello Aguiar	6	54	60	6
Lima, Maycke Young de	2	8	10	5,5
Digiampietri, Luciano Antonio	6	57	63	4,95
Oddone, Nanci	6	57	63	4,95
Brambilla, Sônia Domingues Santos	2	10	12	4,8
Maia, Maria de Fatima Santos	3	22	25	4,75
Walter, Silvana Anita	5	47	52	4,3
Pizzani, Luciana	2	13	15	3,75
Lima, Regina Célia Montenegro	1	3	4	3,1
Bello, Suzelei Faria	2	15	17	3,05
Miranda, Antonio Lisboa Carvalho de	1	7	8	1,7
Vilan Filho, Jayme Leiro	3	31	34	1,6

Pecegueiro, Cláudia Maria Pinho de Abreu	1	8	9	1,35
Robredo, Jaime	3	32	35	1,25
Fujino, Asa	4	44	48	1,2
Maricato, João de Melo	3	33	36	0,9
Oliveira, Robson Ramos	1	10	11	0,65
Oliveira, Silas Marques de	1	12	13	-0,05
Caldeira, Paulo da Terra	1	13	14	-0,4
Calabro, Luciana	2	26	28	-0,8
Castro, Júlio Vitor Rodrigues de	0	3	3	-1,05
Hyodo, Tatiana	0	6	6	-2,1

Source: Research data (2020).

It can be seen that high productivity does not necessarily result in a high HCP index. Thus, researchers who have, in average, more publications may obtain lower indexes than specific researchers who have published less, but who have more expressive amounts of papers among the highly cited ones. The purpose of the HCP index proposed by Waltman, Van Eck, and Wouters (2013) is to value higher impact publications over low impact publications. The index seeks to be an alternative to the h-index which, oppositely, does not value highly cited papers in its calculation. Researchers with higher h-indexes are generally more productive, since the more one produces, the greater the possibility of having works cited. Kelly and Jennions (2006) found that there is a high correlation between the h-index and a researcher's total number of publications. The use of the h-index, unlike the HCP, tends to favor the logic of productivism.

4.5. HCP calculation on data subsets and variables (2010-2020)

Several similarities are observed between the two periods (2010-2020 and 1941-2020) and respective data variables. The highest HCPs in each subset of data were: "Complete version of approximately 10%" the highest HCP was 209.75; "Complete version of approximately 10% without 0 citation" 123.5; "Complete version of approximately 10% only IMS" 85.65 and "Complete version of approximately 10% only IMS without 0 citation" 62.85. It is noteworthy that, in general, the HCP indexes and citation averages are higher in the first period and lower in the second period, reinforcing the hypothesis that a paper published longer ago tends to receive more citation than a more recent one.

As occurred in the time frame from 1941 to 2020, the HCP indexes in the second period were also negative in some cases. The lowest indices in each of the data variables analyzed were: "Complete version approximately 10%" -41.7; "Complete version approximately 10% without 0 citation" -20.8; "Complete version approximately 10% only IMS" -32.6; and, in the "Complete version approximately 10% only IMS without 0 citation" the lowest index was -15.75. As already stated, the possibility of a negative HCP index can be considered a weakness of the formula proposed by authors Waltman, Van Eck and Wouters (2013), but which can be easily adjusted.

Analyzing the correlations between the rankings of researchers working in the field of IMS in Brazil according to their HCP indexes and data variables between the period 2010-2020 (Table 7), they are between moderate and strong. The lowest correlations are between the subsets of data "2010 version of approximately 10% without 0 citation" and "2010 version of approximately 10% only IMS with 0 citation" (correlation of 0.65) and that between "2010 version of approximately 10% with 0 citations" and "2010 version of approximately 10% without 0 citation", (0.69). The highest correlations are found between the dataset "Approximately 10% 2010 version with 0 citations" and "Approximately 10% 2010 version only IMS with 0 citations" (correlation of 0.89), and next the correlations between "Approximately 10% 2010 version without 0 citations" and "Approximately 10% 2010 version only IMS without 0 citations" (0.86).

Table 7 Correlation between the rankings of researchers working in the field of IMS in Brazil according to their HCP indexes by data variables (2010-2020)

Data Variables	2010 version of about 10% with 0 citations	2010 version of approximately 10% without 0 citation	2010 version of approximately 10% only IMS with 0 citations	2010 version of approximately 10% IMS only without 0 citation
2010 version of about 10% with 0 citations	1	0,69	0,89	0,71
2010 version of approximately 10% without 0 citation		1	0,65	0,86
2010 version of approximately 10% only IMS with 0 citations			1	0,78
2010 version of approximately 10% IMS only without 0 citation				1

Source: Research data (2020).

When the correlations of the first period (1941-2020) of Table 5 are compared with the correlation of the 2010-2020 period (Table 7), they have great similarities despite being weaker in the second period. This indicates that the application of the HCP presents similar behaviors in both periods, strengthening the arguments in favor of its application with the variables in which the correlations are higher in both time intervals. Thus, as observed in the first period, it is understood to be possible to choose the data set to calculate the HCP (only IMS papers with citation or zero citation or the totality of papers published with citation or zero citation). This, as noted earlier, may favor the collection of large volumes of data, without the need for thematic classification of publications for the purposes of calculating the

HCP. One can choose, for HCP calculation, the variables with the total production ("Full version of approximately 10% with 0 citation" for the 1941-2020 time frame or the "2010 version of approximately 10% with 0 citations" for the 2010 to 2020 time frame) or the version containing only selected papers from the IMS area ("Full version of approximately 10% only IMS with 0 citation" for the 1941-2020 time frame or the "2010 version of approximately 10% only IMS with 0 citation" for the 2010 to 2020 time frame). In view of the higher correlation, these data variables should preferably be used for the HCP calculation.

Table 8 presents the top 60 researchers working in the area of IMS and their respective papers published on the subject, regardless of whether they were cited or not in the period 2010-2020 and their respective HCPs in descending order.

Table 8 Ranking of the HCP index of 60 researchers working in the field of IMS in Brazil and their respective publications, on the theme, highly cited and less cited (2010-2020)

Researchers	Highly cited (IMS)	Low cited (IMS)	Total Publications (IMS)	HCP
Velho, Lea	23	28	51	85,65
Ensslin, Leonardo	26	77	103	80,95
Ensslin, Sandra Rolim	26	93	119	75,35
Mena-Chalco, Jesus P.	14	63	77	36,05
Leta, Jacqueline	11	36	47	33,05
Ribeiro, Henrique César Melo	15	87	102	31,8
Rossoni, Luciano	7	13	20	24,5
Quoniam, Luc	16	133	149	19,85
Mugnaini, Rogério	8	48	56	16,4
Milanez, Douglas Henrique	7	37	44	16,1
Stumpf, Ida Regina	7	37	44	16,1
Beuren, Ilse Maria	9	61	70	16
Meneghini, Rogério	5	14	19	15,85
Faria, Leandro Innocentini Lopes de	9	62	71	15,65
Guarido Filho, Edson Ronaldo	4	4	8	15,2
Gregolin, José Angelo Rodrigues	5	17	22	14,8
Araújo, Carlos Alberto Ávila	8	53	61	14,65
Cesar Júnior, Roberto M.	4	6	10	14,5
Espejo, Márcia Maria dos Santos Bortolocci	5	23	28	12,7
Oliveira, Ely Francina Tannuri de	11	95	106	12,4
Cruz, Ana Paula Capuano da	4	12	16	12,4
Digiampietri, Luciano Antonio	7	50	57	11,55
Kobashi, Nair Yumiko	4	17	21	10,65
Pizzani, Luciana	3	7	10	10
Parreiras, Fernando Silva	4	19	23	9,95
Bello, Suzelei Faria	3	8	11	9,65
Duarte, Emeide Nóbrega	4	21	25	9,25
Vanz, Samile Andréa de Souza	9	81	90	9
Matheus, Renato Fabiano	2	0	2	8,3
Mueller, Suzana Pinheiro Machado	4	24	28	8,2
Brambilla, Sônia Domingues Santos	2	6	8	6,2
González de Gomez, Maria Nélida	2	8	10	5,5
Robredo, Jaime	2	9	11	5,15
Walter, Silvana Anita	4	33	37	5,05
Noronha, Daisy Pires	3	23	26	4,4

Lima, Maycke Young de	1	1	2	3,8
Grácio, Maria Cláudia Cabrini	10	108	118	3,7
Lima, Ricardo Arcanjo de	1	2	3	3,45
Oliveira, Silas Marques de	1	2	3	3,45
Santos, Jane Lucia Silva	3	26	29	3,35
Riccio, Edson Luiz	2	15	17	3,05
Vanti, Nadia Aurora	2	15	17	3,05
Maricato, João de Melo	3	27	30	3
Alvarenga, Lídia	2	16	18	2,7
Targino, Maria das Graças	2	17	19	2,35
Guimarães, José Augusto	6	67	73	1,45
Calabro, Luciana	2	20	22	1,3
Dias, Eduardo Jose Wense	1	9	10	1
Marteleto, Regina	2	21	23	0,95
Población, Dinah Apparecida de Mello Aguiar	1	10	11	0,65
Fujino, Asa	2	24	26	-0,1
Braga, Gilda Maria	0	1	1	-0,35
Miranda, Antonio Lisboa Carvalho de	0	1	1	-0,35
Oliveira, Eloísa da Conceição Príncipe	1	13	14	-0,4
Caldeira, Paulo da Terra	0	2	2	-0,7
Castro, Júlio Vitor Rodrigues de	0	2	2	-0,7
Hyodo, Tatiana	0	2	2	-0,7
Pecegueiro, Cláudia Maria Pinho de Abreu	0	3	3	-1,05
Hoffmann, Wanda Aparecida Machado	1	15	16	-1,1

Source: Research data (2020).

With the comparative analysis between the rankings of the HCP index of researchers in each time frame (1941-2020 versus 2010-2020), we observe in some cases the alternation and in others the maintenance of the researchers' position in the ranking. Some researchers occupy the top positions in any data set and period (the researcher Lea Velho, for example, occupies the first position in both periods). Other researchers, who had an insignificant position in the first period, began to stand out in the second period because of their highly cited publications (for example, Jesus P. Mena-Chalco, who was in the 22nd position in the first period and rose to the fourth position in the second). In the same way, researchers who were present in prominent positions in the first periods are no longer protagonists in the second period (for example, the researcher Suzana Pinheiro Machado Mueller, who was in the second position in the first period and who in the second period is in the 30th position).

This dynamic can be better understood numerically by analyzing the Pearson correlation of the positions in the two rankings. It can be seen that the correlation varied between the data sets between 0.53 and 0.58 (moderate correlation). In other words, the area of IMS in Brazil, from the HCP point of view, has evolved slowly, with no radical changes, in general, between the positions of researchers in the area from the point of view of highly cited publications.

5 CONCLUSIONS

One of the main difficulties and problems associated with the use of the HCP refers to the impossibility of defining precisely which are the 10% most cited papers in a field of knowledge, being necessary to use approximate data, since generally papers with the same citation count are found on the border between the percentiles, and there are several papers with the same number of citations. In this research, the highest approximation was 10.68%, which is not an important difference. However, as the number of papers is reduced (for example, due to the size of the scientific community and the time frame), this difference may increase, and may make the application of the index unfeasible.

Regarding the use of Google citation as a source of information to calculate the HCP, the search engine has the ability to cover the production and citation information better when compared to international databases, such as the Web of Science. On the other hand, the low quality of the data, absence of dates, and the lack of a more precise classification of areas may influence HCP calculations. Not all researchers working in IMS in Brazil have profiles on the platform, and it is not possible to guarantee the identification of all publications. Therefore, the use of this tool remains a dilemma.

Another limitation of the use of the stylized HCP by the authors refers to the possibility that researchers may have negative indexes (approximately 50% of the researchers had negative HCPs). This fact can cause misinterpretations and resistance in using HCP for evaluation purposes and researchers. What authors Waltman, Van Eck and Wouters (2013) consider a positive point of the proposed formula, can also be seen as a limitation. However, the authors believe that the values can be adjusted and are therefore not an insurmountable problem.

Regarding the most suitable production and citation datasets for the HCP calculation, one can select papers from both within and outside the IMS area, as the correlations were very strong (0.9). Similarly, the datasets of papers with zero citations or all papers are indifferent, as the correlation was also very strong between them (0.9). Therefore, there is consistency in calculating the authors' HCP with all publications, regardless of whether they are from the field of IMS or not, or whether they are cited or not (zero citations), without major losses to the evaluation and ranking of researchers. Although the correlation is not perfect (1), these findings are important, because the classification of papers according to theme (in this case related and unrelated to IMS) is time-consuming and may make large-scale studies unfeasible.

When the evolution of the Brazilian IMS area was analyzed from the HCP index point of view, the area grew and strengthened. The growth was sustainable, without major disruptions between the two periods analyzed. This dynamic could be verified through the Pearson correlation analysis of the positions of the two rankings (1941-2020 and 2010-2020). The correlation varied between the data sets from 0.53 to 0.58 (moderate correlation); in other words, the area of IMS in Brazil, from the HCP point of view, has been changing slowly, with no radical changes, in general, between the positions of researchers in the area from the perspective of highly cited publications.

In this research, we chose to present the data obtained from the formula proposed by the authors, instead of changing or proposing another "optimal" parameter. We did some tests, and it is evident that the number of researchers with negative HCP gradually decreases as the weight of the wLC is lower than originally proposed (-0.35). In order not to have HCP negative researchers, obviously, wLC should be removed from the formula. Any new wLC parameter must be analyzed on a case-by-case basis, according to the objectives of applying the formula, and it is not possible, in our view, to say which parameter is the best. If the objective is to compare a group of researchers with a certain homogeneity (as was the case in our research), to be published in a ranking on the Web, e.g., we believe that the best thing would be to remove the wLC from the formula. However, if the objective were to find more efficient researchers with a specific objective, whose results would be used internally by a committee, the wLC= -0.35 could be an option to be considered.

The use of indicators based on highly cited publications, like any indicator, has positive and negative points. The HCP calculation, with or without the stylization proposed by Waltman, Van Eck and Wouters (2013), has strengths and weaknesses. Depending on the objectives one has in mind, they can be useful. One can, for example, use such indicators to evaluate graduate programs, select researchers, direct research funding, among other uses. However, one should be cautious, because this indicator can create entry barriers for young researchers in the academic community. The HCP is not capable of replacing commonly used indexes and indicators.

REFERENCES

Aithal, P. S. (2017). Comparative study of various research indices used to measure quality of research publications. *International Journal of Applied and Advanced Scientific Research*, *2*(1), 81-89.

Bicalho, L. M. (2009). *As relações interdisciplinares refletidas na literatura brasileira da informação*. 269 f. Tese (Doutorado em Ciência da Informação) - Programa Pós-Graduação em Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte. https://repositorio.ufmg.br/bitstream/1843/ECID-7UUQ69/1/teselucineia_versaodefinitiva.pdf

Bornmann, L. (2013). A better alternative to the h index. *J. Informetrics*, *7*(1), 100. https://doi.org/10.1016/j.joi.2012.09.004

Bornmann, L., Moya-Anegón, F. de and Leydesdorff, L. (2020). *The new excellence indicator in the world report of the scimago institutions rankings.* 2011. https://doi.org/10.1016/j.joi.2011.11.006

Clarivate Analytics. (2020). *Web of science core collection help*. https://images.webofknowledge.com/images/help/WOS/hs_citation_applications.html.

Dainesi, S. M. and Pietrobon, R. (2007). Scientific indicators of productivity: Time for action. *Revista Brasileira de Psiquiatria*, *29*(2). https://doi.org/10.1590/s1516-44462007000200002

Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, *69*(1), 131-152. https://doi.org/10.1007/s11192-006-0144-7

Gabriel Junior, R. F., Freitas, J. L. and Bufrem, L. S. (2011). Relações temáticas e de autorias em estudos métricos da informação representados na brapci. *XII Enancib: Políticas de informação para a sociedade*. Brasília: Universidade de Brasília. http://200.20.0.78/repositorios/handle/123456789/2161?show=full

Gabriel Junior, R. F., Moraes, M. A. D., and Oliveira, E. F. T. D. (2018). A elite e a frente de pesquisa dos EBBCs (2012 até 2016). Encontro Brasileiro de Bibliometria e Cientometria (6.: 2018 jul. 17-20: Rio de Janeiro, RJ). Anais [recurso eletrônico]. Rio de Janeiro, RJ: UFRJ, 2018.

Garfield, E. (1977). *Introducing citation classics-human side of scientific reports*. http://www.garfield.library.upenn.edu/essays/v3p001y1977-78.pdf.

Harzing, A. W. (2020). *Publish or perish 4 user's manual***.** Available: https://harzing.com/pophelp/metrics.htm#gindex

Hicks, D. et al. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, *520*(7548), 429-431. https://doi.org/10.1038/520429a

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences*, *102*(46), 16569-16572. https://arxiv.org/abs/physics/0508025

Hu, Z., Tian, W., Xu, S., Zhang, C. and Wang, X. (2018). Four pitfalls in normalizing citation indicators: An investigation of ESI's selection of highly cited papers. *Journal of Informetrics*, *12*(4), 1133-1145. Available: https://www.sciencedirect.com/science/article/pii/S1751157718301147?casa_token=EK1Nn3

Huang, M. and Chi, P. (2020). A comparative analysis of the application of h-index, g-index, and a-index in institutional-level research evaluation. *Journal of Library and Information Studies*, *8*(2), 1-10. https://core.ac.uk/download/pdf/26994985.pdf

Incites. **(**2019). *Highly Cited Papers.* http://help.prod-incites.com/inCites2Live/indicatorsGroup/aboutHandbook/usingCitationIndicatorsWisely/highlyCitedPapers.html

Jordan, K. (2015). *Exploring the ResearchGate score as an academic metric: Reflections and implications for practice*. https://doi.org/10.1016/j.tree.2006.01.005

Kelly, C. D. and Jennions, M. D. (2006). The h index and career assessment by numbers. *Trends in Ecology & Evolution*, *21*(4), 167-170. https://doi.org/10.1016/j.tree.2006.01.005

Lima, L. S., Soares, C. F. and Oliveira, E. F. T. (2011). Investigação da Produção Científica no tema "Estudos Métricos" na base de dados Brapci: Uma análise bibliométrica. *EDICIC***,** v. *1*(4), 299-310. http://hdl.handle.net/11449/115232

Mattedi, M. A. and Spiess, M. R. (2017). A avaliação da produtividade científica. *História, Ciências, Saúde-Manguinhos*, *2*(43), 623-643. https://www.scielo.br/pdf/hcsm/v24n3/0104-5970-hcsm-24-03-0623.pdf

Moral-Munoz, E., Cobo, M. J., Chiclana, F., Collop, A. and Herrera-Viedma, E. (2016). Analyzing highly cited papers in intelligent transportation systems. *IEEE Trans Intell Transport Syst***,** *17*(4), 993-1001.

Mueller, S. P. M. (2008). Métricas para a ciência e tecnologia e o financiamento da pesquisa: algumas reflexões. *Encontros Bibli: Revista eletrônica de biblioteconomia e ciência da informação*, *13*(1), 24-35. https://brapci.inf.br/_repositorio/2010/09/pdf_04810f633d_0011925.pdf

Nascimento, A. G. (2016) *Altmetria para bibliotecários***:** *Guia prático de métricas alternativas para avaliação da produção científica.* Porto Alegre: Simplíssimo.

Oliveira, E. F. T. (2013). *Análise de domínio em "estudos métricos" no Brasil***:** *Produção, impacto e visibilidade em âmbito nacional e internacional*. São Paulo: Cultura Acadêmica. https://www.marilia.unesp.br/Home/Publicacoes/estudos-metricos-da-informacao-no-brasil---e-book.pdf

PRAUS, Petr. (2019) High-ranked citations percentage as an indicator of publications quality. *Scientometrics*, *120*(1), 319-329. https://link.springer.com/article/10.1007/s11192-019-03128-6.

Reis, E. V., Nogueira, E. C. T. and Oliveira, E. F. T. (2019). Tendências de pesquisa do GT-7: Comparando citações em dois períodos. *XX Encontro nacional de pesquisa em ciência da informação*. Florianópolis: Universidade Federal de Santa Catarina. https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/938/935

Researchgate. (2020). *RG Score.* https://explore.researchgate.net/display/support/RG+Score

Schreiber, Michael. (2013). How much do different ways of calculating percentiles influence the derived performance indicators? A case study. *Scientometrics*, *97*(3), 821-829. https://doi.org/10.1007/s11192-013-0984-x

Thomaz, P. G., Assad, R. S. and Moreira, L. F. P. (2011). Uso do fator de impacto e do índice h para avaliar pesquisadores e publicações. *Arquivos Brasileiros de Cardiologia, 9*(62), 90-93. https://repositorio.usp.br/item/002164190

Trueger, N., Thoma, B., Hsu, C. H., Sullivan D., Peters, L. and Lin, M. (2015). The altmetric score: A new measure for article-level dissemination and impact. *Annals of emergency medicine, 6*(65), 549-553.

Urbizagastegui, R. (2016) Bibliometria, infometria, cienciometria e outras" métricas" no Brasil. *Encontros Bibli*: Revista eletrônica de biblioteconomia e ciência da informação, *21*(47), 51-66. https://doi.org/10.5007/1518-2924.2016v21n47p51

Waltman, L. and Van Eck, N. J. (2012). The inconsistency of the h-index. *Journal of the American Society for Information Science and Technology, 63*(2), 406-415. https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.21678

Waltman, L., Van Eck, N. J. and Wouters, P. (2013). Counting publications and citations: Is more always better?. *Journal of Informetrics, 7*(3), 635-641. https://openaccess.leidenuniv.nl/handle/1887/20468