

[Home](#)[Table of Contents](#)[Titles & Subject Index](#)[Authors Index](#)

Citations to highly-cited researchers by their co-authors and their self-citations: How these factors affect highly-cited researchers' *h*-index in Scopus

[Leila Dehghani](#)

Faculty Member, Department of Medical Library and Information Science, Bushehr University of Medical Sciences, Bushehr, Iran. E-mail: leiladehghani (at) yahoo.com

[Reza Basirian Jahromi](#)

Faculty Member, Department of Medical Library and Information Science, Bushehr University of Medical Sciences, Bushehr, Iran. E-mail: rezabj (at) gmail.com

[Mazyar Ganjoo](#)

Department of Computer Science, Islamic Azad University- Bushehr Branch, Bushehr, Iran. E-mail: ganjoo (at) gmail.com

Received October 12, 2011; Accepted December 20, 2011

Abstract

The h-index is one of the recent indicators in the field of scientometric which introduced by Jorge Hirsch and measures both the productivity and impact of scholars' work. It has been widely used to show the relevance of the research work of prominent scholars. But, does h-index really show this relevance? The purpose of this study is to investigate the frequency of citations to highly-cited researchers by their co-authors, as well as to quantify the weight of self-citations of highly-cited researchers' in three fields of science: clinical medicine, computer science, and economics & business. Finally, we seek to evaluate the impact of these two elements on their h-index in the Scopus database. In this research, highly-cited researchers were selected from the three fields via [HighlyCited.com](#). The study population includes 999 highly-cited researchers in these three fields (280 authors in clinical medicine, 373 authors in computer science, and 346 authors in economics & business). The results show that the average rate of self-citations and co-author citations in clinical medicine is more than computer science, and in computer science is more than economics & business. The results show that there is a positive correlation between self-citation and total citations ($P<0.01$) in these three fields. The results also show that there is a positive correlation between "co-author citations" and total citations ($P<0.01$). Moreover, there is a positive correlation between self-citations and h-index of highly-cited researchers as well as co-author citations and h-index of highly-cited researchers in the above-mentioned three fields. It is seen that there is a significant difference in self-citations behavior and co-author citation behavior in different fields which relates to subject context. Co-author citation is an indicator which, on the one hand, not under control of own researcher and, on the other hand, represents the structure of scientific communications. Although the number of citations is not necessarily increased by raising the number of authors, but co-author citations increased by raising the number of co-authors.

Keywords

Highly-cited researchers; Self-citation; Co-author; Citation analysis; H-index; Scopus

Introduction

Scientometric indicators – like the number of published articles by a researcher during a period of time, the number of total citations to the articles, the average of citations to each article, the impact factor, and the rank of the journals where the articles have been published – are dependable criteria in evaluating the scientific output of researchers. [Hirsch](#) (2005) has proposed the *h*-index as a single number criterion to evaluate the scientific output of a researcher. According to his definition, "A scientist has index *h* if *h* of his/her *N_p* papers have at least *h* citations each, and the other (*N_p - h*) papers have fewer than *h* citations each".

This indicator has defined the number of scientific output of researchers as termed as scientific effectiveness of these researchers. Another characteristic and advantage of *h*-index includes that it is mathematically simple, it may be applied to any level of aggregation, and it is a robust indicator ([Rousseau](#), 2007). Although this indicator is considered for calculating the period of researcher's work but it calculates total works of a researcher and neglects his/her highly-cited and lowly-cited articles. [Kelly and Jennions](#) (2006) have detected that this indicator has been designed in order to improve the other measuring indicators of science like the number of total articles, and the number of total citations; therefore, it can differentiate between affective researchers and those who published many articles.

Although *h*-index has some advantages over similar indicators, it has some weaknesses. "Some criticisms of the *h*-index have include that it is bounded by the total number of publications, it does not recognize the context of the citations received, it is limited by the completeness of the database used to do the calculation, and it does not account for the number of authors on paper" (Brown, 2009). In addition, as *h*-index increases by increasing the number of citations, so two effective elements – self-citation and citing co-authors – are very important. Interestingly, Hirsch (2005) concluded that "self-citation would not significantly affect the *h*-index of an individual author. Although self-citations can obviously increase a scientist's *h*-index, he contended that their effect on the *h*-index is much smaller than on the total citation count" (Brown, 2009). This is an assumption that has been challenged by many studies (Kelly & Jennions, 2006; Schreiber, 2007a) that have shown average decreases in the *h*-index of 12 percent with self-citations excluded. Also some other researchers in their articles have considered the relationship between self-citation and *h*-index (Couto et al., 2009; Engqvist & Frommen, 2010; Gianoli & Molina-Montenegro, 2009; Schreiber, 2007a). Of course, it is reasonable that authors citing to the previous relevant works and using their previous experiments and results. Thijs & Glänzel (2006) stated that "self-citation has been shown to vary by field and by origin of publication but average self-citation rates between 10 and 40 percent are common". Cutto et al. (2009) stated that "average self-citation rates between 10 and 30 percent are common". According to Brown, self-citation is a powerful tool within the author's power to use and may be open to abuse to increase total citation count and, through much targeted citation, to specifically increase the *h*-index (Brown, 2009). So it seems reasonable that *h*-index be considered without self-citations included for a more unbiased comparison of authors in the same field.

Researches show that the effect of self-citation on *h*-index among lowly-cited authors is more (Fowler & Aksnes, 2007; Gianoli & Molina-Montenegro, 2009; Iglesias & Pecharroman, 2007; Schreiber, 2007b). Although in Gianoli and Molina-Montenegro, the *h*-index of highly-cited researchers has been considered, but they have just focused on one field or special field (Gianoli & Molina-Montenegro, 2009). However, the effect of citing to researchers by their co-authors and their self-citations on *h*-index has not been investigated.

The main purpose of this study is to indicate some deficiencies in previous researches with emphasis on *h*-index. These deficiencies include unknowing effect of co-authors citations on *h*-index of highly-cited researchers and investigating of self-citations among highly-cited researchers at international level. Therefore, the major objectives of this research are to investigate the frequency of citations to highly-cited researchers by their co-authors and their self-citations and to evaluate the effect of these elements on *h*-index of highly-cited researchers in Scopus.

Objectives

The objectives of this study are to determine:

- the rate of self-citations and co-author citations in three fields of science (i.e., *clinical medicine, computer science, and economics & business*);
- the difference between self-citations of highly-cited researchers and co-author citations in three fields of science;
- the relationship between self-citations of highly-cited researchers and co-author citations in three fields of science;
- the relationship between the number of articles, citations, and *h*-index;
- the effect of self-citations and co-author citations on *h*-index among highly-cited researchers; and
- the alternations of *h*-index by excluding self-citations and co-author citations among highly-cited researchers of three fields of science.

Materials and Methods

Databases which have been used for calculating *h*-index are ISI Web of Science, Scopus, and Google Scholar. Although the Scopus database is weaker than ISI Web of Science in terms of time coverage and is representing citation information of articles since 1996, but Science Citation Index (SCI) and Social Science Citation Index (SSCI) are representing citation information of articles since 1974. So, *h*-index which has been represented for each author in ISI Web of Science is more than the Scopus database (Jacso, 2008a, 2008b, 2008c, 2008d).

The Scopus database has the possibility for calculating *h*-index by excluding "self-citation" as well as calculating *h*-index by excluding "all of co-authors"; therefore, this database was selected. Note that this selection is one of the main limitations in our research.

The data of the present research were collected in February 14-30, 2011 via Scopus. To investigate the frequency of self-citations and the number of citations of co-authors among highly-cited researchers, the list of ISI highly-cited researchers has been used which can be accessed via HighlyCited.com. This database is produced by Thomson Reuters and has information about researchers who have the most citations by the other researchers. We chose three fields – *clinical medicine, computer science, and economics & business*– in three main disciplines – medicine, engineering, and humanities – and finally 999 highly-cited researchers in those three fields were selected (280 authors in clinical medicine, 373 authors in computer science, and 346 authors in economics & business) as research population.

To achieve recall and precision in searching authors' names which input variously (inconsistency in inputting author names), first we referred to HighlyCited.com and required data were extracted for 999 researchers. Then, we used "author search" box to accomplish our goals. The Scopus database can maximally select 15 statuses of authors' names; due to this limitation we decided to select records with maximum number of articles.

For data analyzing, first we analyzed collecting data by Excel; these data include: the number of articles in Scopus, *h*-index, the number of citations, *h*-index by excluding self-citation, the number of citations without self-citations, the number of self-citations, *h*-index by excluding "all of co-authors", *h*-index by excluding the citations of co-authors, the number of citations by excluding "all of co-authors", and the number of the citations of co-authors. As the required data in the Scopus database has been represented since 1996 and some of these researchers have had their more scientific activities before 1996, these researchers deleted from the same population because their *h*-index was low. Finally, out of those 999 scholars, 684 researchers were selected. After using Excel for primitive calculating and descriptive statistics, SPSS 16.0 was used for inferential statistics; these tests included: T-test, ANOVA test, Correlation test, and Regression test.

Results

The rate of self-citation among highly-cited researchers in clinical medicine, computer science, and economics & business

The results show that the average of self-citations in clinical medicine is more than computer science, and in computer science it is more than economics & business (Figure 1).

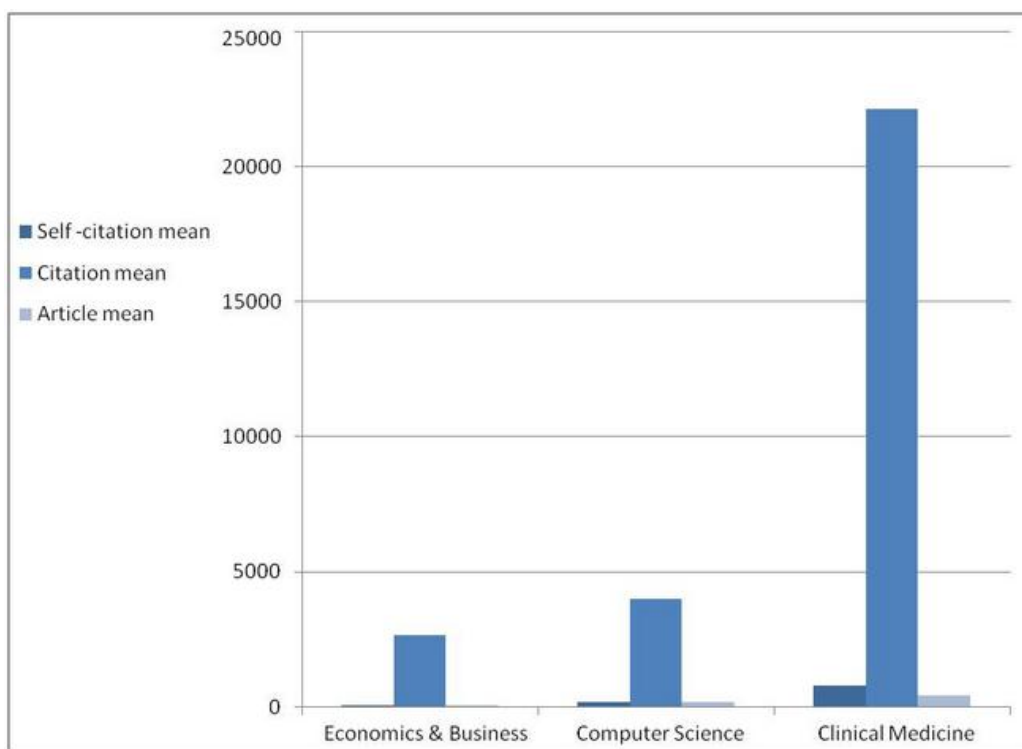


Figure 1. Average of the number of citations, articles, and self-citations in three fields

In these three fields, the results show that there is a positive correlation between self-citations and total citations ($P < 0.01$). Figure 2 indicates the correlation in three fields.

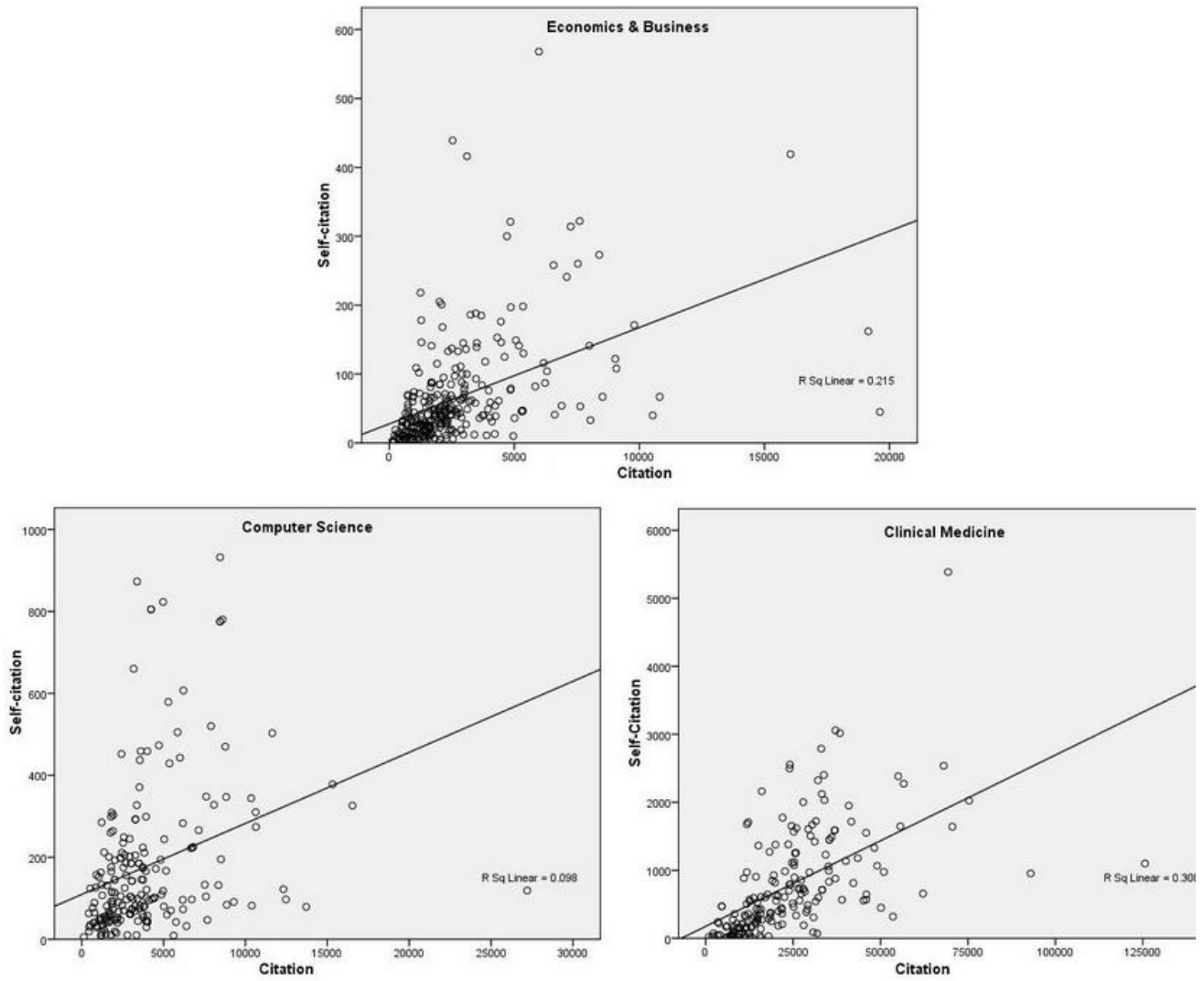


Figure 2. Correlation between self-citations and total citations in three fields

The percentage of "self-citations" to "total citations" is varying in these three fields; this ratio is between 1% and 17% in economics & business, between 1% and 26% in computer science, and between 1% and 14% in clinical medicine (Figure 3).

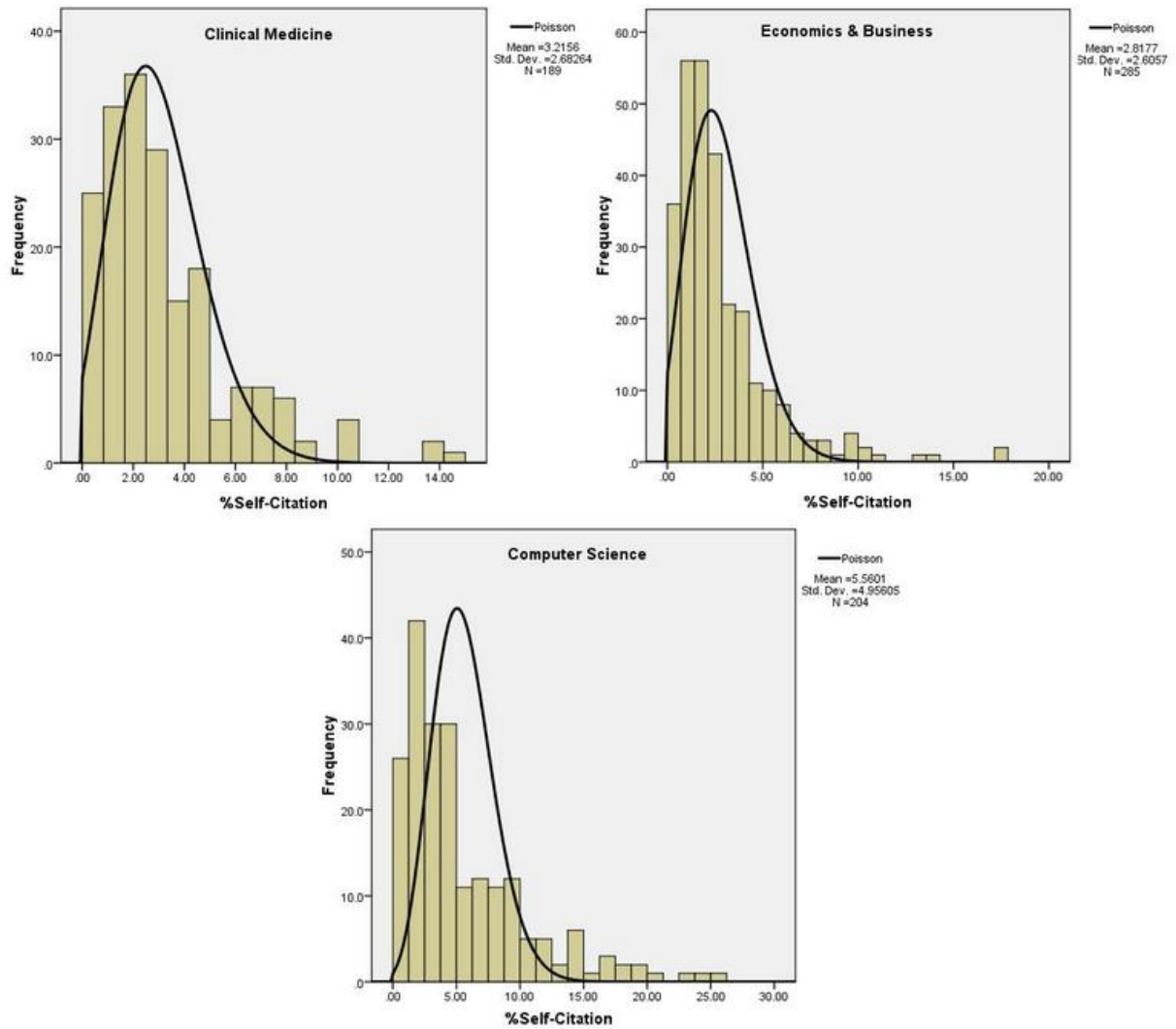


Figure 3. The average and percentage of self-citations to total citations in three fields

The rate of "co-author citation" among highly-cited researchers in clinical medicine, computer science, and economics & business

The results show that the average of "co-author citations" in clinical medicine is more than computer science, and in computer science it is more than economics & business (Figure 4).

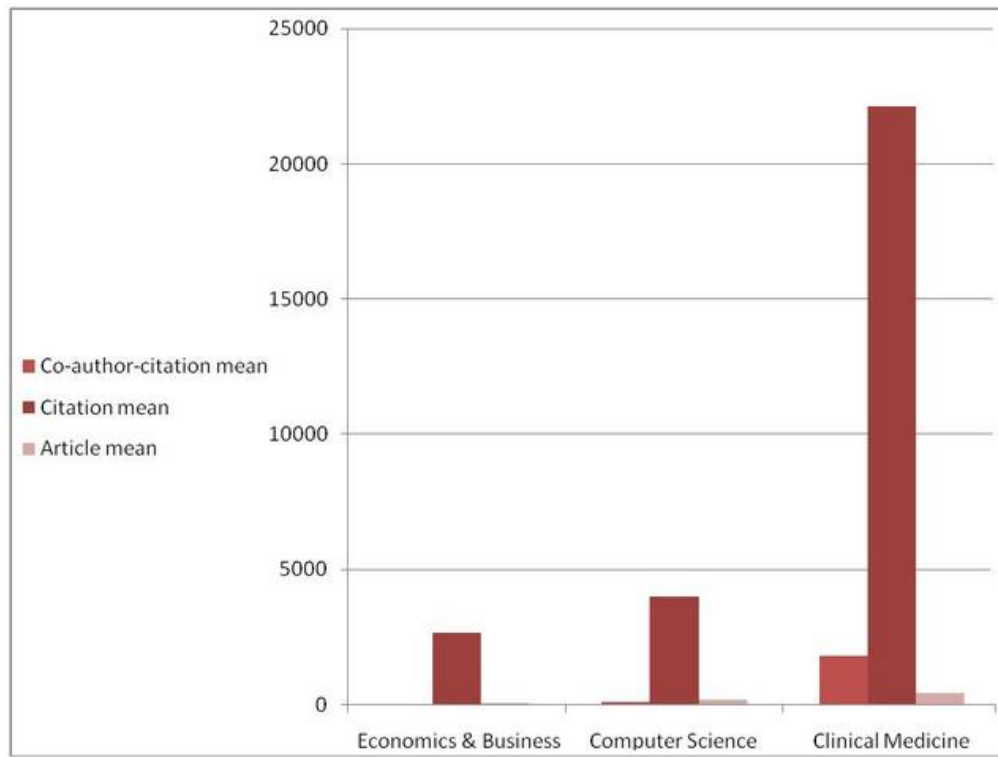


Figure 4. The average of number of citations, articles, and co-author citations in three fields

The results show that there is a positive correlation between "co-author citations" and total citations ($P < 0.01$) in these fields. Figure 5 indicates the correlation in three fields.

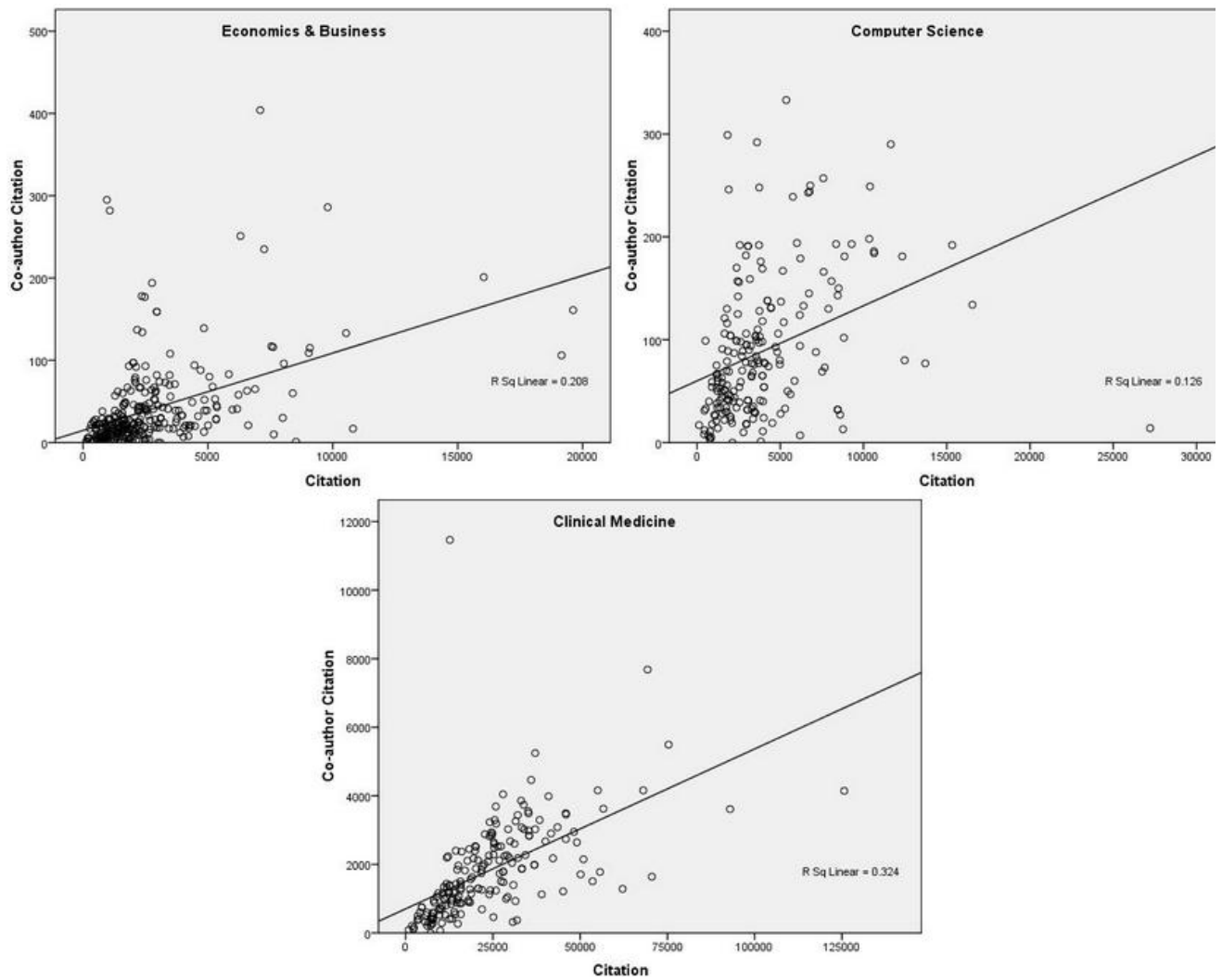


Figure 5. Correlation between co-author citations and total citations

The percentage of "co-author citations" to "total citations" is varying in these three fields; this ratio is between 0.1% and 8% in both economics & business and in computer science, and between 1% and 18% in clinical medicine (Figure 6).

The difference between self-citations and co-author citations among highly-cited researchers in three fields

Using ANOVA test, it is found that there is a significant difference between "self-citations" and "co-author citations" in these three fields ($P < 0.01$).

The relationship between self-citations and co-author citations

The results show that there is a positive correlation between self-citations and co-author citations ($P < 0.01$). Figure 7 indicates the correlation in these three fields.

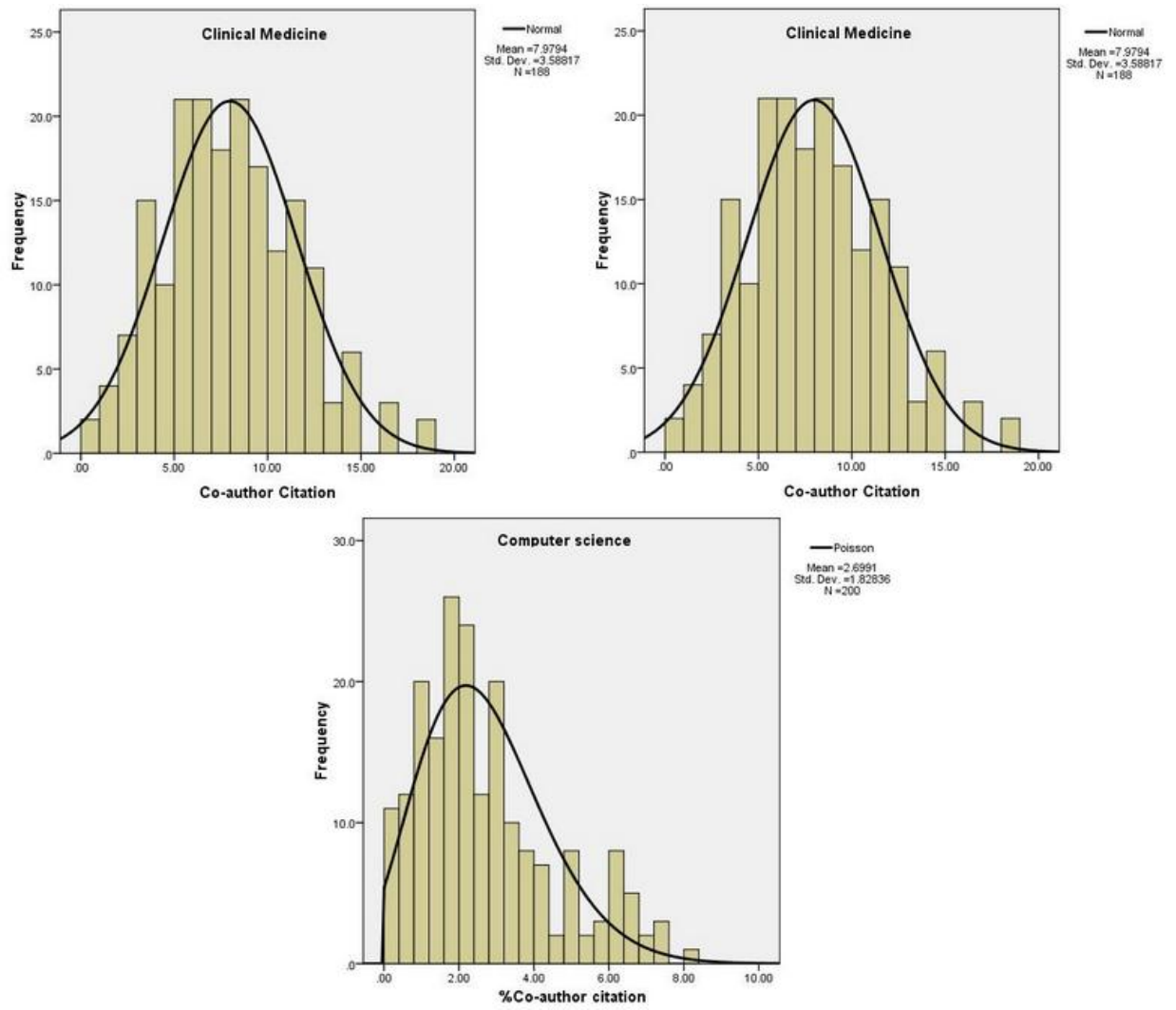


Figure 6. The average and percentage of co-author citations to total citations in three fields

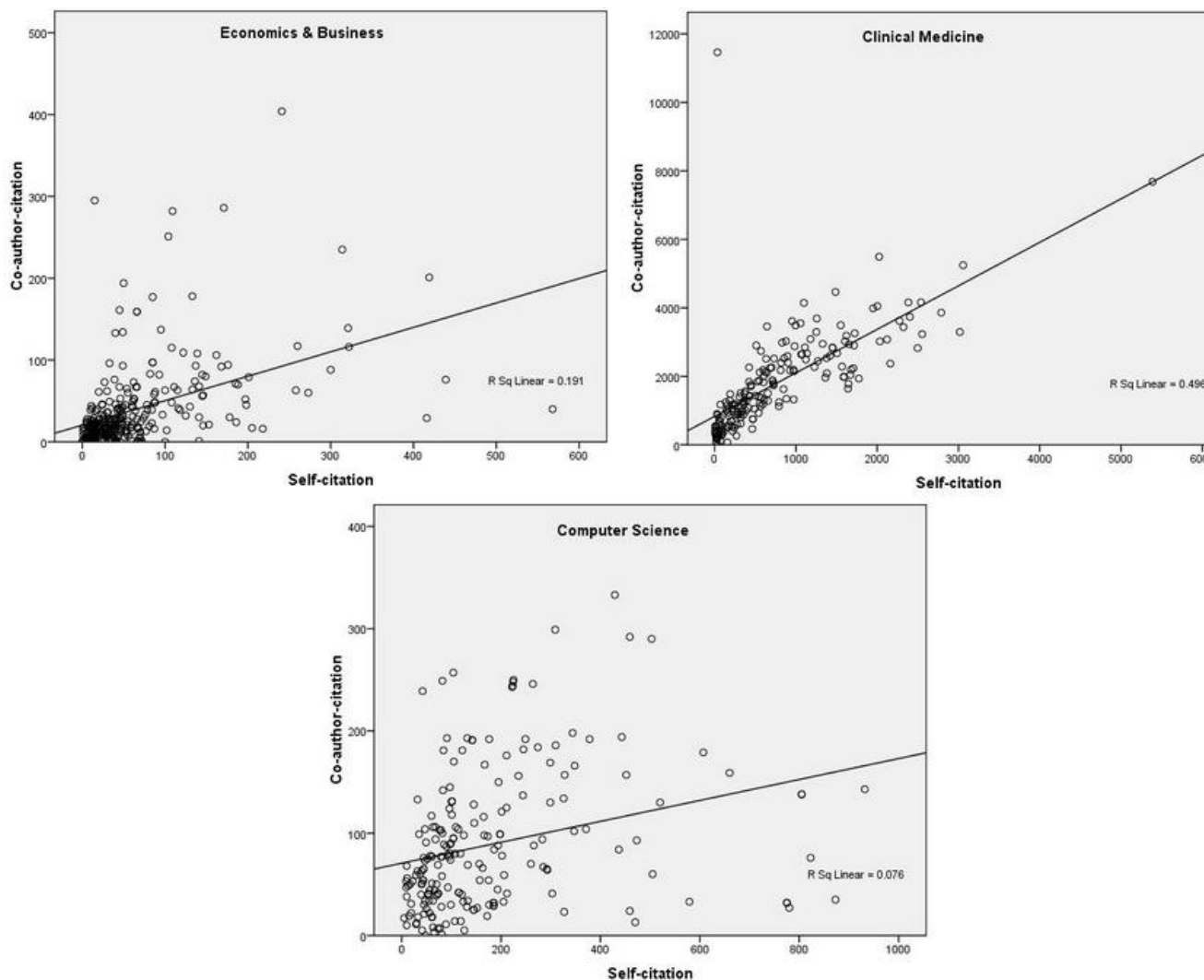


Figure 7. Correlation between the number of self-citations and co-author citations in three fields

The relationship between the effect of "self-citations" and "co-author citations" on h-index of highly-cited researchers in three fields

The results show that there is a positive correlation between self-citations and *h*-index of highly-cited researchers as well as co-author citations and *h*-index of highly-cited researchers in three fields (Table 1).

Table 1. The effect of self-citations and co-author citations on *h*-index of highly-cited researchers

	Economics & Business	Computer Science	Clinical Medicine
Effect of self-citations on <i>h</i> -index	0.277 P<0.001	0.609 P<0.001	0.740 P<0.001
Effect of co-author citations on <i>h</i> -index	0.195 P=0.195	0.263 P<0.001	0.303 P<0.001

The comparison of the effect of self-citations and co-author citations on h-index of highly-cited researchers in three fields

To answer this question, T-test was used and the results show that the average of *h*-index by excluding self-citations is more than the average of *h*-index by excluding co-author citations among highly-cited researchers (P<0.01). The effect of self-citations on increasing *h*-index among highly-cited researchers in computer science and economics & business is more than clinical medicine (Table 2).

Table 2. The comparison between the effect of self-citations and co-author citations on *h*-index

	Economics & Business	Computer Science	Clinical Medicine	Total
Average of <i>h</i> -index by excluding self-citations	0.34±0.77	0.8±0.96	1.21±1.42	0.72±1.10

Average of h-index by excluding co-author citations	0.15±0.47	0.51±0.62	1.67±1.50	0.68±1.11
P Value	P<0.001	P<0.001	P<0.001	P<0.001

Discussion

Scientometric indicators are used to measure scientific output, research performance, institutional performance, scientific collaboration, and knowledge growth. They are also used in university ranking, researcher ranking, and research policy-making.

Researchers require previous related researches for indicating continuity of communications between previous researches. Also, "co-author citation" indicates the scientific communications between scholar and his/her colleagues. "Self-citation" and "co-author citation" are parts of the process of scientific communications. The authors sum up the discussion with evidence from earlier research in the following ways:

- *The state of citations, self-citations, and co-author citations among highly-cited researchers in three fields:* The results show that the average of citations, self-citations, and co-author citations has decreased respectively in clinical medicine, computer science, and economics & business. There is a positive correlation between total citations and self-citations on the one hand, and there is a positive correlation between total citations and co-author citations in three fields on the other hand. Since for each researcher in clinical medicine, the average of the number of articles and number of citations is more than two other fields, so it is obvious that in this field the average of self-citation and co-author citation is more than two other fields. Previous research indicates that the differentiation in self-citation behavior and co-author citation behavior is common in different fields of science ([Costas, van Leeuwen, & Bordons, 2010](#); [Davarpanah & Amel, 2009](#)) as evident in this study.
- *Rates of self-citation and co-author citation:* The percentage of self-citation in clinical medicine is varying between 1% and 14%; this rate of self-citation is less than the rate of self-citation in computer science, and economics & business. Also, the percentage of co-author citation in clinical medicine is varying between 1% and 18% which is higher than two other fields. The reason of this matter is that the number of co-authors in each article in clinical medicine is high which is further evident from the study of [Costas, van Leeuwen, and Bordons \(2010\)](#) who have reported the relations between self-citations and the number of citations on the one hand, and co-author citation and the number of citations on the other hand. Previous researchers have demonstrated that average self-citation rates between 10 and 30 percent are common ([Aksnes, 2003](#)). Since the self-citation is represented by author himself/herself, it is better that researchers prevent using intentional self-citation; it means that the temptation to increase one's h-index oneself should be avoided. But co-author citation cannot be under controlled; it is done commonly and shows the scientific communications ([Costas et al., 2010](#)).
- *The relationship between self-citation and h-index; co-author citation and h-index:* The results show that there is a positive correlation between self-citation, co-author citation, and h-index in three fields; these findings corroborate previous research ([Davarpanah & Amel, 2009](#)).
- *The effect of self-citation and co-author citation on h-index:* The results show that the effect of co-author citation on h-index in clinical medicine is more than self-citation but the effect of self-citation on h-index is more than co-author citation in computer science and economics & business. As a result of the higher number of authors in clinical medicine, the average of co-author citation is more than self-citation. This shows the variety and expansion of scientific communications in clinical medicine.
- *Should we exclude self-citations to calculate h-index from total citations?* [Hirsch \(2005\)](#) proposed that self-citation does not have any impact on h-index obviously. Some researchers have been challenging this hypothesis in their articles. [Schreiber \(2007b\)](#) has reported that h-index has been decreased – between 6.6% and 25% – by excluding self-citation in different fields. In this research, change rates have been calculated by excluding self-citation in three fields: 2% in economics & business, 4.6% in computer science, and 2.3% in clinical medicine. This research shows that the effect of self-citation on h-index is low among highly-cited researchers. Findings of the previous studies indicated that the effect of self-citation on increasing h-index was little more among lowly-cited researchers ([Gianoli & Molina-Montenegro, 2009](#); [Schreiber, 2007b](#)). [Brown](#) proposed b-index – which means h-index by excluding self-citation ([Brown, 2009](#)). He proposed that h-index replaced by b-index. According to [Aksnes](#), self-citation is entirely reasonable because it avoids repeating experimental descriptions, theoretical patterns, and previous results; and it is reasonable because each researcher has more familiar with his/her previous researches, so he/she cites to his/her previous works to complete the researches ([Aksnes, 2003](#)). But sometimes, some researchers use self-citation to increase the number of their citations. However, the results of this study show that self-citations do not have any significant effect on h-index of highly-cited researchers. So, we propose that self-citations should not be deleted for calculating h-index of highly-cited researchers. However, self-citation does not have any significant effect on h-index of highly-cited researchers. Therefore, we strongly recommend that self-citations should not be excluded for calculating h-index of highly-cited researchers. We also propose that h-index remains in "status-quo" for highly-cited researchers but we should propose some moderated indicators for lowly-cited researchers.
- *Should we exclude co-author citations to calculate h-index from total citations?* In previous studies, researchers did not pay attention to co-author citations. Findings of this research show that the effect of co-author citation on h-index is 1% in economics & business, 3.1% in computer science, and 3.6% in clinical

medicine. Although the number of citations does not increase by raising the number of authors, but co-author citations increase by raising the number of co-authors. For example, in clinical medicine the percentage of co-author citation is more than self-citation because of the high number of co-authors and team working in this field. Of course, co-author behavior has no relation with author himself/herself, but continue to have relations with his/her co-authors. So, we propose that co-author citations should not be considered for calculating *h*-index of highly-cited researchers.

Conclusion

Scientometric indicators based on the number of publications, authors, references and citations are increasingly used to evaluate the pattern of research performed by researchers, institutions, and countries. This study aims to investigate the relationship between co-author citations and self-citations of highly-cited researchers and their *h*-indexes. Findings reveal that the number of citations is not necessarily increased by raising the number of authors, but co-author citations increased by increasing the number of co-authors. The results also show that there is a positive correlation between self-citations, co-author citations, and *h*-index of highly-cited researchers in clinical medicine, computer science, and economics & business.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments that helped to improve the paper.

References

- Aksnes, D.W. (2003). A macro study of self-citation. *Scientometrics*, 56(2), 235-246.
- Brown, R.J.C. (2009). A simple method for excluding self-citation from the h-index: The b-index. *Online Information Review*, 33(6), 1129-1136.
- Costas, R., Van Leeuwen, T.N., & Bordons, M. (2010). Self-citations at the meso and individual levels: Effects of different calculation methods. *Scientometrics*, 82(3), 517-537.
- Couto, F.M., Grego, T., Pesquita, C., & Verissimo, P. (2009). [Handling self-citations using Google Scholar](#). *Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics*, 13(2), paper 2. Retrieved September 15, 2011, from <http://cybermetrics.cindoc.csic.es/articles/v13i1p2.html>
- Davarpanah, M.R., & Amel, F. (2009). Author self-citation pattern in science. *Library Review*, 58(4), 301-309.
- Engqvist, L., & Frommen, J.G. (2010). New insights into the relationship between the h-index and self-citations? *Journal of the American Society for Information Science and Technology*, 61(7), 1514-1516.
- Fowler, J.H., & Aksnes, D.W. (2007). Does self-citation pay? *Scientometrics*, 72(3), 427-437.
- Gianoli, E., & Molina-Montenegro, M.A. (2009). Insights into the relationship between the h-index and self-citations. *Journal of the American Society for Information Science and Technology*, 60(6), 1283-1285.
- Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America*, 102(46), 16569-16572.
- Iglesias, J.E., & Pecharroman, C. (2007). Scaling the h-index for different scientific ISI fields. *Scientometrics*, 73(3), 303-320.
- Jacso, P. (2008a). The pros and cons of computing the h-index using Google Scholar. *Online information review*, 32(3), 437-452.
- Jacso, P. (2008b). The pros and cons of computing the h-index using Scopus. *Online Information Review*, 32(4), 524-535.
- Jacso, P. (2008c). The pros and cons of computing the h-index using Web of Science. *Online Information Review*, 32(5), 673-688.
- Jacso, P. (2008d). Testing the calculation of a realistic h-index in Google Scholar, Scopus, and Web of Science for FW Lancaster. *Library Trends*, 56(4), 784-815.
- Kelly, C.D., & Jennions, M.D. (2006). The h index and career assessment by numbers. *Trends in Ecology and Evolution*, 21(4), 167-170.
- Rousseau, R. (2007). The influence of missing publications on the Hirsch index. *Journal of Informetrics*, 1(1), 2-7.
- Schreiber, M. (2007a). A case study of the Hirsch index for 26 non-prominent physicists. *Annalen der Physik*, 16(9), 640-652.
- Schreiber, M. (2007b). Self-citation corrections for the Hirsch index. *Europhysics Letters*, 78(3), 1-6.
- Thijs, B., & Glänzel, W. (2006). The influence of author self-citations on bibliometric meso-indicators. The case of European universities. *Scientometrics*, 66(1), 71-80.

Bibliographic information of this paper for citing:

Dehghani, Leila, Basirian Jahromi, Reza, & Ganjoo, Mazyar (2011). "Citations to highly-cited researchers by their co-authors and their self-citations: How these factors affect highly-cited researchers' *h*-index in Scopus." *Webology*, 8(2), Article 91. Available at: <http://www.webology.org/2011/v8n2/a91.html>

