Journal of Scientometric Res. 2019; 8(2):86-93
http://www.jscires.org

Research Article

# The Prevalence of Software Deployment in Persian Scientometric Studies: A Meta-analysis Approach

Mohammad Alaee Arani, Faramarz Soheili, Afshin Mousavi Chelak, Ali Akbar Khasseh*

*Department of Library and Information Science, Payame Noor University, Tehran, IRAN.*

**ABSTRACT**
This study aims at conducting a meta-analysis on the use of software packages in scientometric-related articles published in Iranian journals. The study consisted of 170 research papers selected based on their credibility within the survey. The articles published from 2013 to 2016 in the field of scientometrics in Iranian journals. Firstly, the assumptions of homogeneity and publication bias were checked; then the fixed effect size was interpreted according to Cohen's model. Results showed that two general statistical software, Microsoft Excel and IBM SPSS, were mostly used in the articles. Finding revealed the homogeneity of the effect size and the lack of bias in the studies investigated. The mean effect size of the prevalence of professional software use in scientometric research was 0.557 ($P = 0.000$). The results also showed from 2013 to 2016, the use of professional software increased. The meta-analysis results indicated that the most frequently used tools are not professional software. This is because of the inability to import the overall output of the retrieved results and their mismatch with the researcher's specific and diverse goals. It seems that the employment of the software depends on what the researchers want to do and what sort of analysis they want to perform. Furthermore, in many cases, researchers seem to have little knowledge about the special tools in this regard. It seems that the employment of the software depends on what the researchers want to do and what sort of analysis they want to perform. Furthermore, in many cases, researchers seem to have little knowledge about the special tools in this regard.

**Keywords:** Meta-analysis, Scientometric Software Packages, Research evaluation.

**Correspondence**
**Ali Akbar Khasseh**
Department of Library and Information Science, Payame Noor University, Tehran, IRAN.
Email: khasseh@gmail.com

Received: 09-07-2018
Revised: 31-10-2018
Accepted: 16-07-2019
**DOI:** 10.5530/jscires.8.2.14

## INTRODUCTION

Scientometric studies have been developed as a subsidiary branch of Library and Information Science over time[1] and have attracted many interested researchers to this field. Bibliometrics, Scientometrics, Informetrics, Webometrics and Technometrics are concepts that can be found in many specialized Information Science journals. Nalimov defined scientometrics as "the quantitative methods of research on the development of science as an informational process"[2]

Scientometric studies are considered as a fully-fledged area of study and there has been an increase in research publications in this field in the last few years. The number of published articles in the specialized journals of this area in 2010 is four times more than the number of articles published ten years ago. Although in the 1980s and 1990s, scientometric studies were beginning to find their place among other disciplines, focusing more on information science, recently, the area of scientometric studies, as an independent field, has partially evolved its social identity In other words, researchers and professional Figures of this area have established an independent community and it seems that the circle has reached an acceptable stability. The academic publication of this field is also independent of the Information Science and has achieved an acceptable cognitive distinction.[3]



Most of Iranian researchers in Library and Information Science publish their works in Persian journals, as well as English journals. An increase in the number of scientometric studies within Persian journals has been evident in recent years. Due to the emergence and gradual development of scientometric studies in Iran, there is an urgent need for an extensive and comprehensive description of the state of the research in this field. Therefore, in this study, the use of software packages in scientometric studies is to be investigated.

Initially, scientometric studies merely supported the graphical representation based on software tools such as SPSS and Pajek.[3] However, it is necessary that the new software applications, in addition to data analysis in larger scales, take into account data visualization.

Scientists believe that software plays a critical role in their research.[4] The use of scientometric software in various processes of simulation, implementation, extraction, homogenization, rapid information processing, drawing networks, among others, is part of the scientometric studies and like other areas of science, it has benefited from them with the development of the technology. On the one hand, the software is based on the features and the nature of the hardware of their own era and on the other, they depend on the type and attitude of software manufacturing technology, as well as the user's expectation and demand. Therefore, through the evaluation of the state of software utilization in scientometric studies, this study aims to extract the tools and software of measurement and visualization, considering the frequency of use in scientometric-related articles published in Iranian journals (both Persian and English languages).

The priority of use and application of existing software by researchers allow users of these software systems to make better decisions in choosing the program that suits their needs. Any decision made regarding the training of the software, designing and implementation of appropriate software for the Iranian researchers and scientific outcomes in the Persian language requires the knowledge of the current state of software use among the target community.

In order to know and understand the actual needs of the researchers for the software and their features to develop appropriate software tools for scientometric research, it is necessary to review the recent studies within the field. However, it should be reminded that the choice of software is partially related to the database and the type of data.

The main questions of the research are:

Which software tools have researchers used in scientometric-related articles published in Iranian journals?

What percentage of scientometric-related articles published in Iranian journals have employed specialized scientometric software?

What are the most widely used software tools in scientometric-related articles published in Iranian journals?

What is the frequency usage of any software in scientometric-related articles published in Iranian journals?

Who are the most frequent researchers in using specialized software?

What is the effect size of the specialized software tools in scientometrics?

## Literature Review

Scientists put a lot of effort into developing software and their software benefits scientific community.[5] In recent years, several researchers have paid attention to the software use and their impact on scientific literature. For instance, by examining how software is mentioned and cited among 9548 articles published in PloS One, Pan, Yan and Hua investigated the use and impact of software in 12 defined disciplines. They showed that the practice of software citations varies noticeably at the discipline level and software that is free for academic use is more likely to receive citations than commercial software.[5] Duck *et al.* indicated that 97.7 % of BMC Bioinformatics papers used software/database.[6] Huang *et al*. revealed that researchers tend to choose software that is widely used by others in their community and prefer software that is free for academic use.[7] Zhao and Wei showed the lack of standard reference for users for software.[8]

In addition, there have been several studies on the motivations of software development and sharing. Results showed that scientists participate in developing and sharing scientific software for extrinsic benefits such as earning citations and advancing careers. Also, academic reputation and monetary rewards motivate scientists to make their software free for academic use.[9–12]

On the other hand, there are many studies within the scientometric domain, which have conducted by meta-analysis technique. In a meta-analysis study conducted by Bornmann *et al*. (2011), the correlation between the h-index and the other related indexes (37 indices) were evaluated in 32 independent studies. A total of 135 correlations were reported in this study. The findings revealed a high correlation between the h-index and the 37 relevant indices.[13]

In a research entitled Science mapping software tools: review, analysis and cooperative study among tools, Cobo *et al*. explained the overall process of science mapping analysis and introduced several techniques and tools for mapping and visualizing science and their usage.[14] Sernoko examined 108 scientometric studies in the area of knowledge management using the meta-analysis method. Based on the results of this study, most of the articles were published in journals irrelevant to the area of knowledge management. In the second part of the study, the contribution of different countries to the production of these articles and interdisciplinary collaboration was described.[15] Bornmann used a meta-analysis method to study research in three categories of altemetrics, including microblogs, online reference managers and blogging. The correlation between the number of altmetrics and the number of citations of studies was the main criterion of the meta-analysis. This study considered the benefit of altmetrics in evaluating the research along with the traditional citation-based metrics. This correlation was reported for microblogs to be negligible, for blogs little and for online resource management bookmarks medium to high.[16]

Generally, software developers are interested to know the use and impact of their software.[16] As the review of the literature indicates, meta-analysis is an important method in the field. The present study, however, is distinct from previous studies in terms of population, the measured factor and the meta-analysis method. It also is distinct in terms of identifying the effect size and their combination with the previous studies. Finally, the sample size is more comprehensive.

## MATERIALS AND METHODS

This study has employed the meta-analysis method. In such a method, the design and findings of a number of studies are recorded in the form of quantitative data and then fed into statistical programs. After employing statistical methods, new and coherent results are extracted. In this study, similar research questions in all studies conducted in scientometrics are analyzed in a specific time frame, from the perspective of software utilization. The results can illustrate the general picture in the population of studies and also show the gap among them.[17]

The research population includes all articles in the area of scientometric published in Iranian journals (both Persian and English languages). This list was reviewed in three academic years from the fall of 2013 to the summer of 2016. Since the titles of the articles in the field were diverse, the search engine method did not seem to be an appropriate way of achieving all of them, leading to an incomprehensive body of articles. However, the journals were reviewed based on their titles. In addition, the scientometric terms were searched in several combinations to get the complete sample size. Accordingly, out of the 200 articles retrieved, 170 scientometric-related articles, which are totally published in Iranian journals, were extracted and meta-analyzed (30 articles have been excluded due to the relevancy issues). The data source for this study is the Islamic World Science Citation Center (ISC), which is a leading source of scholarly research data in Iran (https://isc.gov.ir/en).

Next, all articles were investigated to answer the main questions of this meta-analysis. More precisely, the abstract section and methodology section of all articles were studied in detail. The findings of each were extracted for further examination. For analyzing the data, a Comprehensive Meta-analysis Software version 2 (CMA2) was used.

The main method of meta-analysis is based on the combination of results, which is usually used after converting the statistics to the index (r) and estimating the effect size. To analyze the inferential data, the assumptions of the meta-analysis were to be assumed; first, that is, using a funnel plot and Egger's linear regression method, the Begg and Mazumdar Rank Correlation, the publication bias was to be analyzed. The Q test was also employed to assume the heterogeneity of the studies. Due to the homogeneity among the studies in question, the fixed effects model was used to combine the results to achieve the effect size.

## RESULTS

Table 1 summarizes the overall statistics obtained from the preliminary analysis of the content of the articles in question.

The content analysis of scientometric studies and especially their research methodology section showed that, in total, 35 software were used 254 times. In 36 articles, no software was used. In 113 articles, the general software was used and in other studies specialized software were employed. Accordingly, the average software utilization in each study is approximately two cases.

Table 2 shows the amount and variety of the use of specialized software in scientometric studies. Among the specialized scientometric software, the Ucinet, Vosviewer and HistCite were most frequently used for analyzing the data. In total, these three software were used for 60 times.

The results of the meta-analysis indicate that a total of 7 general (non-specialized) software tools were used. As shown in Table 3, among the general software, the Microsoft Excel and IBM SPSS were used more than others. The frequency of the use and the percentage frequency of usage to the total software utilization are presented in Table 3.

To answer the 5th research question, the usage frequency of the specialized software used by researchers was also investigated based on the names of the researchers. The results indicate that the use of specialized software in studies conducted by Osareh, Soheili, Erfanmanesh, Tavakkolizadeh and Khasseh was higher than the others.

To answer the 6th research question, in analyzing the inferential data, the assumption of the homogeneity of the meta-

**Table 1: Summary of the status of the articles examined in the meta-analysis.**

| Time Frame | Number of articles | Number of times software mentioned in the text | Number of unique software titles used | Frequency of the general software utilization | Frequency of specialized software |
|---|---|---|---|---|---|
| First period (Fall 2013 to Summer 2014)) | 54 | 83 | 21 | 36 | 47 |
| Second period (Fall 2014 to Summer 2015) | 54 | 72 | 15 | 37 | 35 |
| Third period (Fall 2015 to Summer of 2016) | 62 | 100 | 22 | 39 | 61 |

**Table 2: The amount and variety of using specialized software in scientometric studies.**

| Software Name | Usage Frequency | % |
|---|---|---|
| UCINET | 25 | 9.8 |
| HISTCITE | 19 | 7.5 |
| VOSVIEWER | 16 | 6.3 |
| NETDRAW | 12 | 4.7 |
| NODEXL | 9 | 3.5 |
| COAUTH | 7 | 2.8 |
| RAVAR Matrix (Premap) | 7 | 2.8 |
| PAJEK | 7 | 2.8 |
| BIBEXCEL | 6 | 2.4 |
| ISI.EXE | 5 | 2.0 |
| BIBINDEX | 4 | 1.6 |
| CITESPACE | 3 | 1.2 |
| TEXT COLLECTOR | 3 | 1.2 |
| QSB | 3 | 1.2 |
| Network Workbench (NWB) | 2 | 0.8 |
| TH.EXE | 2 | 0.8 |
| TH4.EXE | 2 | 0.8 |
| GEPHI | 1 | 0.4 |
| LUTKA SOFT | 1 | 0.4 |
| RPYS | 1 | 0.4 |
| PATREF3.EXE | 1 | 0.4 |
| USPTOL.EXE | 1 | 0.4 |
| USPTO2-4.EXE | 1 | 0.4 |
| TEXT TO PAJEK | 1 | 0.4 |
| FIND STRING (Dr Beglou) | 1 | 0.4 |
| SCI2 | 1 | 0.4 |
| WORD STAT | 1 | 0.4 |
| SPRING EMBEDER | 1 | 0.4 |
| Total | 143 | 55.9 |

**Table 3: The variety and amount of using general software in scientometric studies.**

| No. | Software name | Usage frequency | % |
|---|---|---|---|
| | EXCEL | 68 | 26.8 |
| | SPSS | 39 | 15.4 |
| | AMOS | 1 | 0.4 |
| | I-GRAPH | 1 | 0.4 |
| | LISREL | 1 | 0.4 |
| | R | 1 | 0.4 |
| | MINITAB | 1 | 0.4 |
| | Total | 112 | 44.1 |

analysis was examined first. Then, employing the funnel plot and Egger's linear regression method, the Begg-Mazumdar Rank Correlation, the publication bias was investigated. The Q test was used to examine the heterogeneity of the studies. After that, due to the homogeneity among the studies under investigation, the fixed effect model was used to combine the results to achieve the effect size.

## Examining the assumption of homogeneity among studies

One of the main assumptions in the meta-analysis is the homogeneity test of the studies. To test this assumption, the Q test was employed. The results of this test are presented in Table 4.

According to the results of the test (Q = 2.233, P 0.01), the null hypothesis for the homogeneity of the studies is confirmed and the hypothesis of the heterogeneity is rejected. In other words, the non-significance of the Q index indicates that there is a homogeneity in the effect size of the studies under investigation. But this index is sensitive to the increase in the effect size, i.e., when effect size increases, the probability of the rejection of the homogeneity increases too. Squared *I* is another index which is used for the same purpose. The squared coefficient *I* has a value of zero to 100% and in fact, represents the inhomogeneity value as a percentage. The closer the value of this coefficient to 100%, the more the heterogeneity of the effect size of the early studies would be. The results of the squared coefficient *I* in this study indicate that about 10% of the total variation of the studies is related to their heterogeneity. Therefore, their integration with the random effects model is not justified and the fixed effects model should be used to combine the results. In fact, this test indicates that the prevalence of using specialized software is not different in terms of the features and characteristics of the studies.

## Examining the publication bias assumption

Another major assumption of the meta-analysis is the publication bias assumption that happens due to the publication of published research papers, the non-publication of the unpublished research papers and different types of collection errors. One of the problems that would distort the validity of the outcomes is the lack of access to all studies conducted within a specific time interval around the subject in question. To examine this assumption, a funnel plot and Egger's linear regression method, the Begg-Mazumdar Rank Correlation, was used.

## Funnel plot

The funnel plot is one of the commonly used methods to investigate the publication bias. The funnel plot of the studies is Figure 1.

The results of the inverted funnel plot indicate the relative symmetry of the studies, but there cannot be a decisive evaluation for such a result and the relevant statistical tests (Egger's linear regression method, the Begg and Mazumdar Rank Correlation) should be employed for this purpose. In this

**Table 4: The results of the homogeneity test (Q)**

| | Effect size and 95% interval | | | Test of null (2-tail) | | Heterogeneity | | | |
|---|---|---|---|---|---|---|---|---|---|
| Number Studies | Point estimate | Lower limit | Upper limit | Z-value | P-value | Q-value | df (Q) | P-value | I-squared |
| 3 | 0.557 | 0.495 | 0.617 | 1.794 | 0.073 | 2.233 | 2 | 0.328 | 10.415 |
| 3 | 0.556 | 0.491 | 0.619 | 1.679 | 0.093 | | | | |

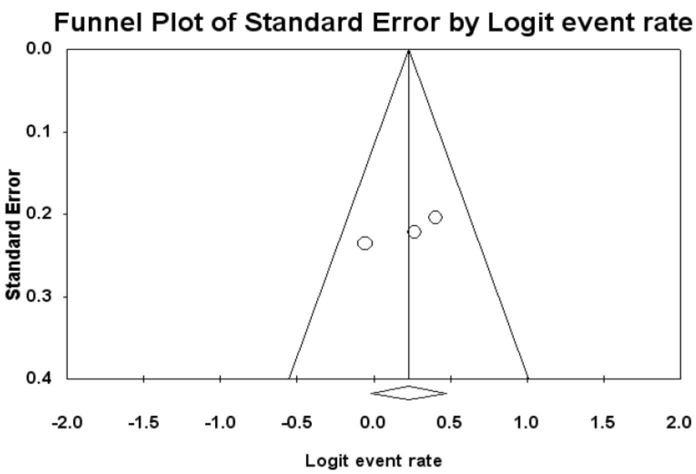



**Figure 1:** The publication bias status in scientometric studies.

**Table 5: The results of Begg and Mazumdar Rank Correlation analysis.**

| *Statistical index* | *Kendall correlation value (tau)* | *Z-Value* | *Significance level (P-Value)* |
|---|---|---|---|
| *Results* | *0.666* | *1.04* | *0.14* |

**Table 6: The results of Egger's linear regression method.**

| Statistical index | Cut off (B) | Standard Error (SE) | *t*-Value | Significance level (*P*-Value) | |
|---|---|---|---|---|---|
| | | | | One-tailed | Two-tailed |
| Results | 13.99 | 4.07 | 3.433 | 0.90 | 0.180 |

method, the null hypothesis (H0) indicates the symmetry of the graph and the non-bias publication and the alternative hypothesis (H1) indicates the asymmetry of the funnel plot and the publication bias.

## Results of Begg-Mazumdar Rank Correlation

Begg-Mazumdar Rank Correlation and Kendall's Rank Correlation tests determine the correlation between the standard effect size and the variance of these effects. The zero value in this correlation is due to the lack of a relationship between the effect size and accuracy. The deviation from zero indicates that there is a relationship. If the asymmetry is due to the publication bias, then it is expected to have more bias in relation to the larger effect size. The results of the Begg-Mazumdar Rank Correlation analysis to examine the publication bias are presented in Table 5.

According to Table 5, the value of Kendall's tau is 0.666. Considering the significance value ($P$ = 0.14), it can be admitted that although there is a relationship between the effect size and the accuracy of the relationship, this relationship is not significant and the null hypothesis indicating that the funnel is symmetrical and does not bias the publication is confirmed.

## Results of Egger's linear regression method

In the absence of the publication bias, small-scale studies result in small standard effect and large-scale studies in large standard effect. This creates a regression line that is a cut-off point for the original regression line. If the cut-off point for linear regression differs from the expected level, the cause might be the bias in publication. The results of Egger's linear regression method analysis to examine the bias of the publication is presented in Table 6.

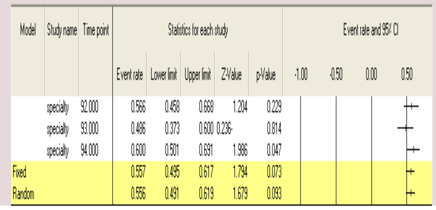
Based on the results of Egger's linear regression method, the cut-off point equals to 13.99 and the confidence interval of 95% is 3.433. Since the one-tailed $P$ value is 90.0 and the two-tailed $P$ value is 0.18, the null hypothesis base on the symmetry of the funnel and the non-bias of the publication is confirmed.

On the left side of Table 7, there is a summary of the data as well as the evidence regarding the specialized software utilization that were used during the three years, based on the share of the specialized software used in the studies. In the middle of the Table (Statistics for each study), basic statistics for each year are presented. Based on the results in all studies, since the standard value (Z-value) in 2015 studies is beyond the range of 1. 96 to 1.96, the significance level ($P$-Value) of studies in this year has been at an error level less than 1%, with a confidence value of 99%.

Combining the effect size in the meta-analysis can be completed by applying one of the two models of fixed effects or random effects. The difference between these two models is that in the fixed effects, it is assumed that the studies under investigation share a common true effect size (meaning that there is an effect size) and the true effect size difference is only due to the sampling error. However, in the random effects model, unlike the fixed effects model, it is assumed that there is a distribution of the effects size and the difference in the

**Table 7: A summary of the meta-analysis on scientometric studies**

| Model | Study name | Time point | Statistics for each study | | | | | Event rate and 95% CI |
|---|---|---|---|---|---|---|---|---|
| | | | Event rate | Lower limit | Upper limit | Z-Value | *p*-Value | |
| | specialty | 92.000 | 0.566 | 0.458 | 0.668 | 1.204 | 0.229 | |
| | specialty | 93.000 | 0.486 | 0.373 | 0.600 | 0.236 | 0.814 | |
| | specialty | 94.000 | 0.600 | 0.501 | 0.691 | 4.986 | 0.047 | |
| Fixed | | | 0.557 | 0.495 | 0.617 | 1.794 | 0.073 | |
| Random | | | 0.556 | 0.491 | 0.619 | 1.679 | 0.093 | |



effects size among the studies is not only due to the sampling error alone, but also due to other factors such as the measurement error and the intrinsic difference among the studies. In other words, in the case of homogeneity of the studies, the fixed effects model would be the base model and the results of the model with random effects, in heterogeneous conditions, will have more generalizability than the model with fixed effects.

As shown in Table 8, the effect size for the models with fixed effects and random effects are 0.557 and 0.556 respectively. Reporting either of those two effects depends on the homogeneity or heterogeneity of the studies under investigation, as discussed below.

It should be noted that the obtained point estimation in both models of fixed effects and random effects is based on Cohen's scale, indicating that the effect is high. Subsequently, this means that the prevalence of using specialized software is high.

In the first step, the meta-analyses should examine the most important assumptions of this kind of study, namely, the homogeneity of the carried-out studies and the examination of the publication bias. Based on this model, the effect size of the prevalence of the specialized software utilization (0.557) is interpreted to be in the third interval, i.e., high.

**Table 8: System of Interpreting Effect Size Due to Cohen's Mode.**

| Effect size | r | d |
|---|---|---|
| Low | 0.1 | 0.2 |
| Medium | 0.3 | 0.5 |
| High | 0.5 | 0.8 |

**Table 9: The effect of the prevalence of using specialized software in scientometric studies.**

| | Number of usage | Effect size | Low level | High level | Z-Value | *P*-Value |
|---|---|---|---|---|---|---|
| Specialized software | 143 | 0.557 | 0.495 | 0.617 | 1.179 | 0.073 |

## Effect size correlation

After examining the assumptions of the software meta-analysis, it was concluded that due to the homogeneity of studies, the fixed effects model should be used to combine the results in order to report the effect size. Thus, the effect size of the carried-out studies with the fixed model is presented in Table 9.

The statistical calculations indicate that the average size of the fixed effect in relation to the prevalence of the specialized software utilization in the sample of this study is 0.555. As the estimated size is in the confidence range, it can be concluded that it is statistically significant.

# DISCUSSION

A large number of specialized software have not been used in any of the studies in question. In addition, the greatest effect size in scientometric studies belongs to the two-general statistical software (non-specialized): Excel and SPSS. These two software are mentioned 107 times (42 percent of the total usage). The annual review of the usage of these two software also suggests the increase in their use each year's indicating no decrease in the acceptance of these two software by the researchers in the field. In other words, the maximum effect size in this category belongs to these two non-specialized software. This means that the software used by the researchers is limited in number and there is a tendency to use general statistical software. As the results of this study show, it should be noted that although Excel and SSPS software meet many of the computational requirements of scientometric studies, there is still a need for specialized software that can perform preprocessing, data extraction, normalization, mapping and visualization. Proper cluster analysis, for example, is only conducted by SSPS, while it requires the use of several general or specialized software and this kind of analysis can be poorly done with Ucinet.

A number of software have been produced abroad that meet the needs of their manufacturing organizations and do not fully answer the needs of Iranian researchers. However, in many cases, the researcher at best can perform the statistical analysis with the SPSS package. Of course, in such a condition, one should have citation data as well as a set of well-predicted variables and parameters as a precondition. What is evident in many scientometric articles is the lack of a well-designed and appropriate methodology. In this study, the articles in the field of scientometrics, conducted in Iran, were selected and analyzed according to the criteria presented in the methodology section. Considering the obtained averages, the prevalence of software utilization is outstanding in the studies, but the effect size of the specialized software is not significantly different from that of the non-specialized software. The number of the specialized software utilization for carrying out scientometric studies is through provoking, if compared with the utilization of the other statistical analysis software. As the results of the meta-analysis show, during the time frame considered in this study, 35 software was used 254 times for data processing, of which the number of specialized software utilization frequency was only 143 cases. The proximity of the effect size of the specialized software to the other statistical software is actually an evidence to the claim. About 44% of the studies used non-specialized software. This shows that many scientometric studies need the use of non-specialized software when it comes to analysis. In some cases, the specialized software is available in the market but the researcher did not have any preference in using them and in some other cases no new software was replaced.

It is imperative that the first requirements for analyzing the data be considered. We need to know what kind of information or data is needed and what the main issues regarding publication research, scientific growth, impact factor, among others, are. Any special bibliographic or scientometric tools do not enable one to work in a complete and precise manner (For example, the ALSCAL section of the SPSS software is useful for visualizing and mapping of data).

The use of non-specialized software is most commonly found in those studies whose statistical population consists of a journal, a specific domestic organization, or scientific productions in Persian. There are several reasons why students and researchers are eager to use such tools: a) the lack of uniformity among collected information from the Persian-language databases, b) the lack of prediction of the means of transferring information retrieved from databases to specialized software in standard formats and d) the inability to enter Persian data in many scientometric software. There are few researchers with expert knowledge of specialized software and they are the only ones who mostly utilize the software in the field.

Due to the lack of the scientometric tools for analyzing Persian data, the only software developed by Iranian researchers, which can receive and analyze Persian data, is Ravar Matrix or Pre-Map (the new version). This software is only used in seven studies with a contribution of 8.2%. Of course, it should be noted that in addition to this software, Pajek VOSViewer, NetDraw and SPSS accept Persian data too.

In a number of studies, the non-specialized software has been used as a homogeneity making tool, as well as for the early and alternative analyses. There are several reasons for this, namely, the lack of uniformity of data, the impossibility of entering collected data directly into the specialized software, or the lack of familiarity with the possibilities of the specialized software tools.

Currently, many scientometric studies have gone beyond data mining tools and techniques for visualizing and mapping as well as clustering articles, writers and institutions. Some examples of these software are R-programming and Python Scripting language, which are easier to use than other programs and have been used in some of the studies investigated here.

The design and development of the scientometric software are particularly important to organize the statistical system and management of the scientific information of the country. Analyzing scientific data for reducing human error and diversifying the input and output of information will have a significant role in the area of science and technology policy and scientometric studies in Iran.

The analysis and comparison of the methodology section of the studies show that most of the data are processed in several discrete stages, each of which is constructed based on using various tools. After that, non-manual instructions have often been used to connect the processing stages and set up a continuous stream of data. An alternative solution to do this can be a defining interface for the transfer between layers and separate processing stages. It seems that the Pre-Map software has this latter purpose. However, these interfaces should be designed between each pair of the sequential layers of analyses in a particular way, depending on the software tools that researchers use, so that they are reusable and generalizable. Whether the data should be import automatically or manually by the researcher has yet to be investigated from different aspects and carefully examined at the practical level.

More research should be conducted on introducing and evaluating the software for visualization and evaluation of science in Iranian's research projects. Since scientometric studies are fresh compared with other areas of Information Science, there are still no comprehensive guidelines for the design and qualitative/quantitative evaluation of the scientometric tools. For this purpose, bearing in mind the principles of software evaluation, the basics and design requirements, there needs to be an attempt to provide an appropriate utilization pattern for the designers and producers of the specialized software.

## CONCLUSION

Although many statistics based on the search strategies can be extracted from a variety of citation database and then analyzed, a systematic research should use certain tools for analyzing the research findings for deeper evaluation with regard to the aims of each research. Nowadays, the software is of vital importance to scientific research and scientists put a lot of effort into developing software and their software benefits the scientific community. Therefore, the value of software has yet to be explored and recognized 5 the current study conducted a meta-analysis on the use of software packages in Persian scientometric research.

From the analytical perspective, the findings showed that the average of the effect size is 0.555. Based on Cohen's scale, this value indicates a high level of effect size. In other words, the specialized software for visualization and evaluation of science has played an important role as a tool for analyzing and visualizing data in scientometrics. The largest effect size was respectively attributed to the Ucinet (with a contribution of 9.8%), Vosviewer (with a contribution of 7.5%), HistCite (with a contribution of 6.3%) and Netdraw (with a contribution of 7.4%). On the other hand, ten specialized software are only used once in the studies.

A large number of specialized software have not been used in any of the studies in question. Besides, the greatest effect size in scientometric studies belongs to the two-general statistical software (non-specialized): Excel and SPSS. Furthermore, results indicated that 42 percent of the total usage was allocated to these two software in 107 cases. The annual review of the usage of these two software also suggests the increase in their use each year's indicating no decrease in the acceptance of these two software by the researchers in the field. In other words, the maximum effect size in this category belongs to these two non-specialized software. This means that the software used by the researchers is limited in number and there is a tendency to use the general statistical software. As the results of this study show, it should be noted that although Excel and SSPS software meet many of the computational requirements of scientometric studies, there is still a need for specialized software that can perform preprocessing, data extraction, normalization, mapping and visualization. Proper cluster analysis, for example, is only conducted by SSPS, while it requires the use of several general or specialized software and this kind of analysis can be poorly done with Ucinet.

For further studies, it is recommended that in order to upgrade scientometrics software services, present various output and analyze refined data, different kinds of case studies should be carried out.

To the software developers, it is suggested that instead of a static map, dynamic analysis via mapping should be developed. The focus of attention should be on chronological bases. Finally, it is necessary to mention that the normalization and simplification have great significance to researchers in non-English and complex scripts data analysis.

## ACKNOWLEDGEMENT

The authors would like to acknowledge the anonymous reviewers for their valuable suggestions and comments in enhancing the quality of the article.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## REFERENCES

1. Khasseh AA, Soheili F, Moghaddam HS, Chelak AM. Intellectual structure of knowledge in iMetrics: A co-word analysis. Information Processing and Management. 2017;53(3):705-20.
2. Nalimov V, Mulcjenko B. Measurement of Science: Study of the Development of Science as an Information Process. Washington DC: Foreign Technology Division. 1971;210.
3. Milejeciv S, Leydesdorff L. Information Metrics (iMetrics): A research specialty with a socio-cognitive identity?. Scientometrics. 2013;95(1):141-57.
4. Howison J, Bullard J. Software in the scientific literature: Problems with seeing, finding and using software mentioned in the biology literature. Journal of the Association for Information Science and Technology. 2016;67(9):2137-55.
5. Pan X, Yan E, Hua W. Disciplinary differences of software use and impact in scientific literature. Scientometrics. 2016;109(3):1593-610.
6. Duck G, Nenadic G, Brass A, Stevens RRDL. Bio NDS: Exploring bioinformatics' database and software use through literature mining. BMC Bioinformatics. 2013;14(1):1-13.
7. Huang X, Ding X, Lee CP, Lu T, Gu N, Hall S. Meanings and boundaries of scientific software sharing. In Proceedings of Conference on Computer Supported Cooperative Work. 2013;423-34.
8. Zhao R, Wei M. Impact evaluation of open source software: An Altmetrics perspective. Scientometrics. 2017;110(2):1017-33.
9. Crowston K, Wei K, Howison J, Wiggins A. Free/Libre open-source software development: What we know and what we do not know. ACM Computing Surveys. 2012;44(2):1-7.
10. Roberts JA, Hann IH, Slaughter SA. Understanding the motivations, participation and performance of open source software developers: A longitudinal study of the Apache projects. Management Science. 2006;52(7):984-99.
11. Howison J, Herbsleb JD. Scientific software production. In Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work- CSCW'11. 2011;513-22.
12. Poisot T. Best publishing practices to improve user confidence in scientific software. Ideas in Ecology and Evolution. 2015;8(1):50-4.
13. Bornmann L, Mutz R, Hug SE, Daniel HD. A multilevel meta-analysis of studies reporting correlations between the h-index and 37 different h index variants. Journal of Informetrics. 2011;5(3):346-59.
14. Cobo MJ, Lopez-Herrera AG, Herrera-Viedma E, Herrera F. Science mapping software tools: review, analysis and cooperation study among tools. Journal of the American Society for Information Science and Technology. 2011;62(7):1382-402.
15. Serenko A. Meta-analysis of scientometric research of knowledge management: discovering the identity of the discipline. Journal of Knowledge Management. 2013;17(5):773-812.
16. Bornmann L. Alternative Metrics in Scientometrics: a meta-analysis of research into three altmetrics. Scientometrics. 2015;103(3):1123-4.
17. Hartung J, Knapp G, Sinha BK. Statistical meta-analysis with applications. Hoboken, NJ: Wiley. 2008.