Journal of Informetrics 11 (2017) 823–834

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Review

Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the Literature

Gali Halevi[a,*], Henk Moed[b], Judit Bar-Ilan[c]

[a] Icahn School of Medicine at Mount Sinai, Department of Medicine, New York, NY, USA
[b] Universita degli Studi di Roma La Sapienza, Rome, Italy
[c] Bar-Ilan University, Department of Information Science, Ramat Gan, Israel

ARTICLE INFO

Article history:
Received 24 February 2017
Received in revised form 24 May 2017
Accepted 18 June 2017

ABSTRACT

As Google Scholar (GS) gains more ground as free scholarly literature retrieval source it's becoming important to understand its quality and reliability in terms of scope and content. Studies comparing GS to controlled databases such as Scopus, Web of Science (WOS) and others have been published almost since GS inception. These studies focus on its coverage, quality and ability to replace controlled databases as a source of reliable scientific literature. In addition, GS introduction of citations tracking and journal metrics have spurred a body of literature focusing on its ability to produce reliable metrics. In this article we aimed to review some studies in these areas in an effort to provide insights into GS ability to replace controlled databases in various subject areas. We reviewed 91 comparative articles from 2005 until 2016 which compared GS to various databases and especially Web of Science (WOS) and Scopus in an effort to determine whether GS can be used as a suitable source of scientific information and as a source of data for scientific evaluation. Our results show that GS has significantly expanded its coverage through the years which makes it a powerful database of scholarly literature. However, the quality of resources indexed and overall policy still remains known. Caution should be exercised when relying on GS for citations and metrics mainly because it can be easily manipulated and its indexing quality still remains a challenge.

Contents

* Corresponding author.
E-mail addresses: gali.halevi@mssm.edu (G. Halevi), hf.moed@gmail.com (H. Moed), judit.bar-ilan@biu.ac.il (J. Bar-Ilan).

http://dx.doi.org/10.1016/j.joi.2017.06.005
1751-1577/

1. Introduction

Since its inception in 2004, Google Scholar (GS) drew the attention of the scientific community as scholarly literature retrieval and citations tracking tool. Emerging as a free tool for scholarly literature retrieval, it was the focus of dozens of studies examining its accuracy and coverage by comparing it to established databases, especially Scopus and Web of Science (WOS). These studies examined GS on a variety of levels including coverage, citations, author profiles and metrics to name a few. As more and more studies examining GS are published and with each focusing on a specific aspect of the database, it's becoming important to synthesize the results and draw a picture of GS's ability to be used as a reliable source of scholarly content.

Additionally, since these studies compare GS to different databases and platforms with findings and conclusions that vary depending on the benchmarked source and discipline, there is also a need to map the advantages and disadvantages of GS when compared to specific disciplinary databases.

The purpose of this paper is to provide a review of selected literature and synthesize the results in order to offer an overarching view of their methodologies, findings and conclusions per topic. In addition, considering that many of the studies of GS focus on comparisons to various platforms and databases, we aimed to provide a set of advantages and disadvantages of each when compared to GS. We believe that this review enables an informed use of GS for different scholarly purposes.

In order to ensure the representation of studies on the topic we conducted multiple searches on different platforms and databases including WOS, Scopus, PubMed and Google Scholar. Overall we reviewed 91 articles from 2005 through 2016. Our literature searches focused on studies that examined GS in terms of (1) Source-level coverage (2) Citations tracking and (3) Author level comparisons

(1) Source-level coverage is defined as the extent to which GS covers journals, books, conference proceedings etc. Studies in this area focus on examining the completeness of such coverage and whether there are gaps or delays in indexing them.
(2) Citation tracking. This area of research examines GS capability to reliably track citations to scholarly artifacts it indexes. Citation tracking is an important metric used for productivity and impact examination. Therefore, studies in this area focus on comparing GS to other databases and examine its ability to reliably track citations and be used as a research evaluation tool.
(3) Author level comparisons relate to the vast literature related to the calculation and analysis of publication and citation counts for individual authors. GS author profiles, h- indices and comparisons between GS and Web of Science and/or Scopus play a crucial role.

The review is organized according to the topics above which will be discussed in 3 separate sections. The discussion and conclusions section offers a summary of the methodologies and conclusions reached by the reviewed studies.

2. Source coverage

In order to be able to compare GS scope and content to other databases, several studies focused on finding a methodology that could assist in estimating the number of records it indexes. Unlike other databases, the actual size of GS was never disclosed by Google and there is no official documentation to account for the number of records and sources it includes or its updates frequency and guidelines. Several studies attempted to develop a methodology that will enable an estimation of the size of GS database.

The approaches to such estimation ranged from benchmarking GS against controlled databases such as Microsoft Academic Search (Khabsa & Giles, 2014) or WOS (Orduña-Malea, Ayllón, Martín-Martín, & López-Cózar, 2014). These studies used randomized samples of queries and examined the level of overlap between GS and the other databases. Their results estimated the number of English scientific publications on the web at approximately 144 million out of which GS covered an estimate of 100 million. Compared to WOS, GS was found to have tripled the amount of documents which include non-English publications. These estimations must be considered as a major factor when comparing GS to other databases. In order to place GS coverage in context, the following databases mostly used to compare GS index, a much lower number of papers:

- Web of Science 65,171,195 (based on all dates search in WOS: [YEAR PUBLISHED: (1700–2017)]
- Scopus: 43,186,550 million records (based on all dates search in Scopus: [REFPUBYEAR > 1700]
- Microsoft Academic Search: 168 million records (Hug & Brändle, 2017)
- PubMed 27,226,186 (based on an all dates search in PubMed: ["1700"[Date – Publication]: "3000"[Date – Publication]

As can be seen, the sheer amount of documents covered by GS far surpasses those of the other scientific literature database. This must be taken into consideration when comparative studies are conducted. Most controlled databases are also selective in their coverage. Working under quality and often topical guidelines, these databases do not aim to cover all the scholarly universe of works but rather offer a focused and reviewed access to a body of knowledge. Unlike these databases, GS is essentially an enormous web crawler which collects and indexes scientific output either through publishers' websites or from scientific repositories of universities, library catalogs and other websites that contain academic and scientific publications. Bearing in mind the main differences between controlled databases and GS, coverage comparisons have been the focus of many studies in this area. Within this body of work, several studies compared GS coverage with that of Scopus, Web of Science, PubMed and others. In its early years GS was the subject of several studies that compared it to information retrieval databases in different disciplines. These earlier studies mostly found significant gaps between its perceived and actual coverage (Jacso, 2005; Mayr & Walter, 2008; Neuhaus, Neuhaus, Asher, & Wrede, 2006).

Despite of its crawling capabilities and the fact that many publishers allowed GS to crawl their websites in these early years, GS was found to omit relevant publications that were found in publishers' websites and had noticeable delay in updating published articles from paid-for platforms. This was probably due to early inception roadblocks experienced by GS when negotiating with publishers. One example is American Chemical Society (ACS) which sued GS for the use of the word "Scholar" in their name. ACS saw that as a direct competition to their SciFinder Scholar and filed a lawsuit early in GS launching ("ACS sues Google over Scholar | The Scientist Magazine®," n.d.). Open access sources however, were found to be covered adequately in GS even in the first few years after its inception.

However, GS was found to better cover subject specific sources when compared to established databases such as WOS and Scopus (Levine-Clark & Gil, 2008; Levine-Clark & Kraus, 2007; Mayr & Walter, 2008; Walters, 2007). These studies found that multidisciplinary subject areas and the social sciences were better covered by GS than by WOS or Scopus. Yet, despite of its coverage advantage of social sciences, all the early studies above found that GS updates remained problematic with the database showing irregularity in its updating capabilities. It should be noted that these studies were conducted in GS early years and within 2–4 years after its inception. Therefore it should not be surprising that there were areas such as coverage of publishers' content behind paid walls and regular updates were somewhat problematic.

In addition to mere coverage examination, bibliometric studies aimed to also measure the citations tracking capabilities of GS compared to WOS and Scopus. Citation tracking is discussed in detail in a following section. In the current section some findings are presented that bear direct relevance for the assessment of GS' source coverage.

The main reason for the growing interest in GS citation tracking capabilities stems from the fact that unlike Scopus or WOS it does not have a defined coverage policy. As mentioned above, most databases are limited to a specific list of sources, subjects, document types and other restrictions. WOS and Scopus index peer reviewed scholarly literature only. GS, however, indexes all output regardless of whether or not it is peer-reviewed. This means that even sources such as research reports, presentations, white papers, blogs, theses (PhD or Masters), magazine mentions are captured as citations in GS. This, of course, has an enormous impact on authors, departments and institutions. When measured for "impact", GS will display many more citations than any other database that restricts the types of output it covers. Several studies recommended, therefore, that GS will be used for citations tracking in conjunction with controlled databases (Levine-Clark & Gil, 2008; Levine-Clark & Kraus, 2007; Meier & Conkling, 2008). While GS showed delays in indexing peer-reviewed, paid-for publisher journals, it did manage to track more citations in several subject areas that were not well covered by WOS or Scopus.

One important field in which this was found to be true is engineering. A comparison of GS and Compendex was conducted by (Cusker, 2013). This study used a previously used methodology (Meier & Conkling, 2008) to examine whether GS advanced in ways that could make it the primary indexing tool for engineering literature. Repeating the strings of queries used in the 2008 study and selected the top 25 articles by using the Relevance ranking function within Compendex. The articles were then searched for in GS. Unlike the previous study, the authors found that despite of the fact that GS does not have the advanced features of some of other host software offering such as ranking, sorting and filtering options it "performed better overall against Compendex in this study than it did in that earlier paper" (p.242). That said; the authors believe that GS can be used as a primary literature search engine for Engineering. The main reason for the favorable conclusion is the cost of Compendex when compared to GS being a free tool. The shortcomings of GS were not found to be significant enough compared to the fact that it's free and covers the literature. Therefore, using citations counts from Scopus and comparing them to GS seems to better capture the citations universe of scholarly output and provides a dimension of "impact" that goes beyond peer-reviewed citations tracking.

These recommendations are seen to be concluded in several subject areas ranging from medicine (Falagas, Pitsouni, Malietzis, & Pappas, 2008; Mastrangelo et al., 2010), psychology (García-Pérez, 2010), earth sciences (Mikki, 2010), life sciences (Beckmann & Wehrden, 2012; Kirkwood & Kirkwood, 2011), education (van Aalst, 2010), and library science (Lewandowski, 2010) with the main reasons for using GS in addition to subject-specific or controlled databases being the lack of clarity around quality control and scope of indexing which make GS incapable of being used as a single source of citations tracking. The position of social sciences and humanities is further discussed in the next section.

Regardless of the subject areas studied in these papers, the vast majority of studies point to several inherent challenges for using GS as a sole source of literature among which are: lack of quality control in coverage and general lack of clarity around which sources are indexed. GS retrieval of non-peer-reviewed sources, gaps in coverage of publishers' content and its inability to rank results in a meaningful way were among the drivers for the general recommendation to use GS in conjunction with topic-specific, controlled databases (Martín-Martín, Orduna-Malea, Ayllón, & López-Cózar, 2016; Shariff et al., 2013). Lewandowski (2010) also reemphasized that "while lots of LIS articles can be found using Google Scholar, it is not a substitute for abstracting and indexing databases that give complete coverage of all articles of the journals selected for indexing" (p. 9).

In 2010, several questions were raised regarding GS ability to distinguish spam or 'junk publications' from real scientific literature (Beel & Gipp, 2010; Jacsó, 2011). In these studies, fake documents were uploaded to the web and picked up by GS as valid, demonstrating that despite of its size, GS doesn't hold any quality control policies. A well-known study (Lopez-Cozar, Robinson-Garcia, & Torres-Salinas, 2012) describing the creation of six fake documents written by a faked author was reported in 2012. In this study the authors uploaded the fake documents to a website created by them and attributed to themselves. The authors have seen an increase of 774 citations which were mentioned in 129 papers, which, in turn, increased the authors and journals H index. Despite their attempts to alert GS and even after removing these documents from the website, the citations are still tracked and maintained by GS. This study demonstrated how easily GS results can be manipulated by anyone wishing to do so because of lack of quality control or selectiveness in its coverage.

GS results containing errors in meta-data and duplicated records were also tracked by comparative studies which found that GS not only contained duplicate items but also copies of articles with errors in author names, dates and issues (Adriaanse & Rensleigh, 2011; Boeker, Vach, & Motschall, 2013; Nourbakhsh, Nugent, Wang, Cevik, & Nugent, 2012; Adriaanse & Rensleigh, 2013; Tober, 2011). Using recall, precision and rankings measures, the studies also noted that compared to WOS and Scopus, GS did not have adequate search tools that could assist with more specific searches thus resulting in recommending that GS will be used for general searches rather than in-depth literature searches. More recent studies show that despite of the fact that GS improved its coverage over the years, especially in the Social Sciences, its search capabilities are still lacking (Martín-Martín, Orduña-Malea, Ayllón, & López-Cózar, 2014).

The usefulness of Google Scholar (GS) as a bibliographic database for biomedical systematic review was the topic of study as well (Bramer, Giustini, Kramer, & Anderson, 2013). The authors utilized 21 biomedical systematic reviews and used their reported search strategy on GS and PubMed. The results were compared to the articles cited in the reviews. This study found that GS had higher coverage and recall compared to PubMed. The re-run searches performed well on GS with 80% of references returned compared to 68% in PubMed. However, a similar study conducted by (Giustini & Boulos, 2013) disclosed different results. Using a sample systematic review on web 2.0 and population health, Giustini & Boulos (2013), conducted a search of 506 publications covered in the review. They found that GS indexed 95% of the papers. The 30 papers that were not found via GS were available using PubMed or a regular Google search. Although the number of retrieved articles was high, the authors did not recommend using GS as the only source of literature retrieval for systematic reviews.

GS coverage of repositories was examined by Arlitsch and O'Brien (2012) who studied several institutional repositories and their content compared to GS. At the time of the study, these repositories were not well covered by GS due to metadata issues and GS incapability to retrieve them. More recent studies show significant improvement of GS coverage compared to its early years. DeGraff, DeGraff, and Romesburg (2013) demonstrated that the growth in the number of open-access journals and institutional repositories increases the number of articles readily available via Google Scholar in the area of geosciences. However, when examined for coverage of Open Access (OA) journals indexed in Latin American repositories, GS and Google have shown low rates of indexing and therefore are not representative of that body of literature (Orduña-Malea & López-Cózar, 2015)

Moed, Bar-Ilan, and Halevi (2016) found substantial differences in source coverage between subject fields. For articles in a selected set of journal articles they found that in chemistry and virology, citation rates in GS and Scopus are statistically similar, whereas in political science the citation rates are in GS almost 4 times those in Scopus. An important aspect of source coverage and citation tracking is indexing speed. In their study comparing GS with Scopus, Moed, Bar-Ilan and Halevi (2016) found a median Scopus indexing delay of two months compared to GS. They added that this delay is partly due to not reporting cited references in articles in press in Scopus.

Source type level comparisons focus mainly on journals. Most of the studies in this area examine the extent to which GS covers journal's content and the rate by which new content is indexed compared to other databases. The importance of these studies lies in the inherent limitation of large bibliographic databases. Scopus, WOS and others use careful selection criteria of sources to index. Journals, conference proceedings and books are evaluated before indexing and thus create an inherent limitation to the breadth and depth of the database. In addition, each bibliographic database has certain disciplinary biases whether intended or unintended. Some databases are designed to cover certain subject areas (i.e. Compendex) and as such are limited in scope. Yet, even larger databases such as Scopus and WOS have shown to have limited coverage of certain disciplines (Norris & Oppenheim, 2007).

In addition to coverage, studies also looked at the journals ranking scores as they are calculated using different databases. Since journals' rankings are closely linked to the overall evaluation of researchers' work, calculating citations rates is crucial (Delgado López-Cózar & Robinson-García, 2012; Delgado-López-Cózar & Cabezas-Clavijo, 2013). This becomes even more critical in disciplines such as Arts & Humanities and Social Sciences (Hodge & Lacasse, 2011; Jacobs, 2011) which might be less covered in general bibliographic databases such as Scopus and WOS.

An example can be found in a study on the citations coverage of publications by 360 soil researchers (Minasny, Hartemink, McBratney, & Jang, 2013) compared GS to WOS and Scopus. The authors concluded that the number of papers per researcher was 2.3 times higher in GS than the other two databases. Their main assertion is that the GS h-index is a good measure for journals in this area of study and captures the journal's citation performance better than the Impact Factor because it includes more citations than those covered by the other databases. This study concurs with Harzing and Van Der Wal, (2009) who recommended using GS h-index to measure journals in Management and International Business journals.

GS book citations were studied by (Kousha, Thelwall, & Rezaie, 2011). This study compared citations to 1000 books using GS and Scopus. The study found that book citations in GS were almost 3 times larger than those in Scopus. The authors conclude that in book-oriented disciplines in the social sciences, arts, and humanities, online book citations may be sufficiently numerous to support peer review for research evaluation using GS. However, considering that Scopus did not cover books comprehensively at the time, this conclusion is somewhat problematic.

Latest studies focusing on comparisons of GS to WOS and Scopus are seen to be still examining the issues of coverage (Bramer, Giustini, & Kramer, 2016; Harzing & Alakangas, 2016; Trapp, 2016; Chan, Chan, Tong, & Zhang, 2016; Jacobs, 2016) and citations (Levay, Ainsworth, Kettle, & Morgan, 2016) in different disciplines. Interestingly, studies focusing on coverage issues found that GS is measuring up to disciplinary controlled databases such as MEDLINE and Embase. This is an indication that GS has been able to grow the amount of publishers who are now allowing the engine to index their sources. Studies on journal rankings advise to use GS as a source of citations tracking in the social sciences arena such as sociology (Chan, Chang, & Chang, 2013; Delgado-López-Cózar & Cabezas-Clavijo, 2013; Jacobs, 2011).

That said, precision and lack of advanced search options remain a challenge and hinder GS from becoming a sole tool for scholarly activities such as systematic reviews. Recent studies on citations tracking conclude that GS does retrieve more citations per article but caution against its lack of quality control, duplications and citations counts of different versions, which inflate the scores and are unreliable as a sole source of research evaluation. Moed et al. (2016) found a huge dispersion in GS source titles. In their study comparing GS source coverage with that of Scopus, unique GS sources not covered in Scopus come from Google Books and/or from large book publishers, and from large disciplinary and institutional repositories. They present a list of the most important GS sources not covered in Scopus.

3. Citations tracking comparisons

By far the most studies comparing Google Scholar to other databases focus on citations tracking. These studies examine the levels of citations coverage of GS and other databases. The main purpose of these studies was to find whether GS can replace databases such as WOS, Scopus and others as a source of citations tracking. Citations' tracking is used for various purposes including the calculation of author, article and journal impacts. Therefore, the main challenge with reviewing and synthesizing this body of literature is the fact that these studies focus on different purposes and disciplines. Citations behavior varies from one discipline to another and while this is true, the overall results of the comparisons show some similarities that can be generalized.

Early studies in this area were focused on citation counts comparisons which tried to determine whether GS can be used as a reliable source of citations tracking (Bakkalbasi, Bauer, Glover, & Wang, 2006; Bauer & Bakkalbasi, 2005; Gardner & Eng, 2005; Neuhaus et al., 2006). These relatively early studies found that in areas of social sciences GS retrieved more citations than WOS or Scopus. However, there was a delay in citations retrieval of new articles. The delay in citations tracking in the first few years after GS was launched can also be explained by the fact that not all publishers were allowing GS to crawl their content. In these disciplines the studies concluded that although GS retrieved more citations; researchers should consult GS in addition to WOS or Scopus while acknowledging that "Google Scholar is still in beta testing, so it has the potential to improve significantly before it becomes fully operational" (Bar-Ilan, 2008, p260). SocStudies conducted early in GS development which examined citations tracking across multiple disciplines (biology, chemistry, physics, computing, sociology, economics, psychology, and education) found that the increased citations counts in GS are field – dependent showing strength in covering social sciences and conference articles when they are published online (Kousha & Thelwall, 2007, 2008; Meho & Yang, 2007). The conclusions of these studies concur with ones of the same time frame in cautioning users in using Google Scholar as a sole source for citations counts. Later studies were able to better define GS strengths and weaknesses as a citations tracking tool.

Researchers were able to track GS growth, improve their ability to harvest GS data and use the improved search capabilities were some of the reasons that these studies were able to paint a more accurate picture. The subject areas studied span from social sciences to medicine (Bar-Ilan, 2010; Kousha & Thelwall, 2008; Kulkarni, Aziz, Shams, & Busse, 2009; Levine-Clark & Gil, 2008; Mingers & Lipitakis, 2010; Amara & Landry, 2012).

These studies also demonstrated that GS strength lays in its citations coverage of social sciences and humanities compared to WOS and Scopus. Mingers & Lipitakis (2010) studied citations tracking in the areas of business and management and concluded that whilst somewhat unreliable, has a much better coverage of these disciplines. Meho & Yang (2007), for example showed that in the area of Library & Information Science (LIS), GS has higher coverage of conference proceedings as well as international, non-English language journals compared to other databases. In addition, Bar-Ilan (2010), using a well-known book "*Introduction to Informetrics*" as an example, found that there are significant differences in coverage between WOS, Scopus and GS. While GS missed about 30% of the citations covered by Scopus and WOS, it did discover additional citations that were not covered either by Scopus or by WOS and thus demonstrated the coverage differences between them. The differences in coverage was also found by Kulkarni et al. (2009) who discovered that Google Scholar had significantly fewer citations to group-authored articles and articles with declared industry funding while both Scopus and GS tracked more citations to non-English journals and proceedings than what was found in WOS.

However, it should also be noted that some of the earlier studies indicated that 70% of Google Scholar citations came from non-journal documents that are available on the web (Kousha & Thelwall, 2008; Bornmann et al., 2009; Kulkarni, 2009; Levine-Clark & Gil, 2009). These studies were conducted a few years after GS was launched. Not all publishers let GS crawl their content at that time which could explain the predominantly free content GS was indexing. In addition, these studies also showed that GS citations data differed greatly from those tracked by citations from the fee-based databases. They called into question GS validity in terms of accuracy and completeness. The vast majority of the authors recommended using a combination of citations counts sources such as GS, Scopus and WOS when evaluating faculty because GS data was incomplete and required lengthy collection time and data cleansing.

Later studies also warned about the use of GS as a sole source for citations tracking due to the fact that it can be easily manipulated while not offering quality control or regulation. To prove this, Lopez-Cozar et al. (2012) conducted an experiment to analyze GS capacity to detect citation counting manipulation. Six documents were uploaded to an institutional web domain authored by a false researcher and were detected by GS citations profiles of the authors therefore proving that the lack of standardization and control of GS data can easily be manipulated. Similarly Bohannon (2014) warned of the ease by which Google Scholar can be gamed because it includes sources from across the Internet— not only vetted journals—and has no human curators.

A study by Ebrahim et al., 2014, compared 101 highly cited papers from Malaysian universities in the field of engineering and compared their citations counts in GS and WOS. In addition, this study sought to discover whether the number of versions in GS influences the number of citations that articles received. Overall the researchers were able to track more citations per article in GS and recommend using GS for tenure and promotion purposes in conjunction mainly because of GS being a free source while Scopus and especially WOS are costly.

The lack of transparency with regards to its covered sources and the designers' unwillingness to allow data exports for analysis presents difficulty in assessing its accuracy and usefulness as a source for evaluation metrics which involve citations counts (Ortega and Aguillo, 2014). More recent studies were able to address GS coverage practices in an effort to discover what types of resources are more likely to be covered by it. Aguillo (2012) performed a webometric analysis of academic web domains and found that 63.8% of the records are hosted in generic domains like .com or .org, confirming that most of the GS data is from large commercial or non-profit sources. The authors concluded that the lack of transparency with regards to the sources covered by GS is hindering it from becoming a replacement tool for databases such as Scopus and WOS. In addition the authors found that the larger coverage GS provides consists in some cases of items not comparable with those provided by similar databases especially from large international or national databases, or repositories. Studies as the ones referred to above have been critical of GS lack of transparency especially as far as having a clear indexing policy that could be used by the bibliometric community to conduct citation analyses. The large coverage range offered by GS is viewed as beneficiary as long as there are consistent indexing practices that can guide the data analysis. Thus far, this issue remains a challenge and GS still does not provide indexing policy or allows for large downloads of citations data. These issues still prevent it from being used as a main source of citation analyses especially for research evaluation purposes.

Overall it appears that GS improved its coverage over the years. This could be the result of publishers allowing GS to access their data and present their records in the GS search results. However, GS search capabilities remain an issue especially in areas that require specific fields such as clinical research. On the one hand GS has a simple and what looks like a straightforward search interface. This could be misleading when specific resources are needed. Orduna-Malea, Martín-Martín, and López-Cózar (2014) demonstrated that GS can be used to find gray literature in a clinical context and can be a powerful tool as long as the user is aware of its more advanced capabilities. Searching for specific content types is not as straightforward as GS seems to be overall. While its search page looks as simple as the Google search engine, GS requires more advanced knowledge of searching and refining results if specific results are needed (Orduña-Malea et al., 2014). Other issues that did not seem to change over the years include lack of transparency with regards to its covered sources and unwillingness to allow data exports for analysis that presents difficulty in assessing its accuracy and usefulness as a source for evaluation metrics that involve citation counts.

One of the features of Google Scholar is the listing of various versions per source when available. Many times the versions may contain a reference to the final published article on a publisher website which can be behind a pay wall. Many times, however, these versions may contain pre-print versions in full text format. Studies into the citations to full text versions of articles on GS found that not only do full text versions found via ResearchGate or other educational repositories receive more citations but that there is a correlation between the number of full text versions found and the number of citations the article receives (Jamali & Nabavi, 2015). Moed et al. (2016) reported that the linear correlation between GS and Scopus citation counts at the article level in a set of 12 journals in 6 subject fields is high: Pearson's r is in the range of 0.8-0.9. They also found that, although in GS different versions of an article are indexed, citation counts are hardly affected by double counts. However, the authors found that GS citations can be prone to errors especially because of insufficient bibliographic information in GS indexed sources, errors in linking citations to articles. For example the study found that one of the three most frequently cited targets articles in Journal of Virology is cited in GS 270 times. However, a secondary analysis revealed that 180 of these were linked erroneously to this target article and were all extracted from a particular (Brazilian) journal available via a Cuban website. The authors conclude that GS indexing strategy seems to favor speed over quality.

4. Author level comparisons

In 2012, GS launched its researcher profiles and offered several h-index calculations. Creating accurate author profiles using populated data has been a challenge to many information providers. Name disambiguation is the main barrier to aggregating an author's body of work in a complete manner. GS rankings are significantly different mainly because of wider coverage of resources not indexed in the other two databases. While these resources generate more citations, it is difficult to predict rankings as GS does not have a clear indexing policy; an issue that has been pointed to in several other studies (Bar-Ilan, 2008; Gray et al., 2012; Jacsó, 2012; Wildgaard, 2015).

The h-indices of highly cited researchers based on GS were considerably different from the values obtained using WOS or Scopus. Bar-Ilan (2008) examined the variation of authors' h- index when using different citations indexes. She compared the h-indices of a list of highly-cited Israeli researchers based on citations counts retrieved from the WOS, Scopus and GS and found that they vary quite significantly from one database to another. Thor and Bornmann (2011) examined the issue using a single author example showing that automated h-index calculations could not capture the true citations counts attributed to his work regardless of the database used. A comparison of bibliometric indicators for computer science scholars (Franceschet, 2010) concluded that GS computes significantly higher indicators' scores than WOS. This is not surprising. The fact that GS indexes a larger number of resources types, languages and its coverage of conference proceedings contribute to the higher h-index score calculated by GS. Nevertheless, citation-based rankings of both scholars and journals do not significantly change when compiled on the two data sources, while rankings based on the h-index show a moderate degree of variation (Cabezas-Clavijo & Delgado-López-Cózar, 2013; Delgado & Repiso, 2013; Doğan, Şencan, & Tonta, 2016; Li, Sanderson, Willett, Norris, & Oppenheim, 2010).

h-index comparisons in the area of nursing was conducted by De Groote & Raszewski (2012) showing that Scopus, WOS and GS provided different h-index ratings for authors. The study found that the 3 databases provided some unique citations. Therefore the authors recommended to use a combination of databases to conduct author-level citations analysis mainly because one tool alone cannot be relied on to provide a thorough assessment of a researcher's impact. In the area of business and management Amara and Landry (2012) examined the research outputs of scholars in business schools and show how their performance assessment is significantly affected when using data extracted either from WOS or from GS. The study found that accounting professors tend to publish a smaller proportion of their work in GS covered resources than their colleagues in information management, finance and economics. GS was found to have a significant advantage over WOS and Scopus for calculating h- index scores for Spanish social sciences researchers (Etxebarria & Gomez-Uranga, 2010). While WOS and Scopus demonstrated limited coverage of social sciences overall and especially those produced by Spanish speaking researchers, GS demonstrated its encompassing ability to index different document types in various language which resulted in the ability to calculate higher h- indices for these researchers.

In 2012 (Jacsó, 2012) set out to evaluate the GS author profiles, their advantages and disadvantages. Despite the fact that GS offered at the time enhanced features to track authors' output, Jacsó deemed them as "too little, too late". The main reasons for GS author profile being inadequate according to this study are the lack of data quality control which is manifested in duplicate records, multiple versions and incorrect affiliations listings.

However, studies in specific areas or countries were not as conclusive as the above statement. A more focused study on the h-index results of Iranian nuclear medicine scientists was conducted by (Zarifmahmoudi & Sadeghi, 2012). The h-index per researcher was compared using GS, WOS and Scopus. The study found that despite of the fact that Scopus, WOS, and GS provided different h-indices for each researcher; the overall rankings were similar. The authors stated that based on these results the three databases could be used interchangeably. The main reasons for this recommendation is the fact that each of the databases retrieved unique papers per each researcher while providing a complete list that balanced the h-index calculations once each set was considered. This was recently confirmed by Wildgaard (2015), who also found that certain areas of science (Astronomy, Environmental Science, Philosophy and Public Health), are better covered by GS and produce higher author rankings than others fields. The main recommendation in the study was for authors to be aware of the indexing coverage of each tool and not rely on one to compute their author-level impact indicators.

Comparing Google Scholar Citations (GSC) and Microsoft Academic Search (MAS) author profiles, Ortega and Aguillo (2014) sought to discover their accuracy and completeness. Comparing over 700 profiles, the authors found that GSC retrieves more citations and documents compared to MAS but shows bias towards computer and information sciences. While MAS retrieved fewer documents it was described as better balanced as far as disciplinary coverage was concerned. However, the authors noted that MAS suffers from duplicate profiles. This study also recommends using both databases for evaluation purposes. Conducting a study on the precision and recall levels of GS using specific names of authors found that GS provides better recall and provides more access to full-text than PubMed, but poor precision. When testing actual search strings used by these clinicians GS failed to retrieve them (Wakimoto, 2014).

Harzing (2014) examined Nobel Prize winners in chemistry, economics, medicine and physics and their citations impact in GS, Scopus and WOS. She found that GS displays considerable stability over time and that coverage for disciplines that have traditionally been poorly represented in GS (chemistry and physics) is increasing rapidly. Lastly, Harzing and Alakangas (2016) published the latest report of their longitudinal comparison between WOS, Scopus and GS. Examining 146 senior academics in five disciplines they found that the three databases display stable growth as far as the number of publications. However, the authors did find that GS still presents challenges especially in its inclusion of non-peer reviewed sources as citations; retrieval of duplicate, and thus redundant, documents in different versions which cause "stray citations" and it is possible to manipulate.

5. Other comparisons

An evaluative comparison on institutional level was conducted by Bornmann, Thor, Marx, and Schier (2016). Focusing on the humanities and social sciences, this study conducted an analysis on GS data on an institution in 2009 and compared them to data from WOS and Scopus. 122 publications were included in the study. Each of the papers was searched for in GS resulting in more than 1000 citations per each paper. These citations were normalized by the world average of citations to the same document type in the same year. The study found that for publications in the humanities and social sciences, GS is an appropriate source of citations tracking because of its wider coverage of these disciplines and its ability to retrieve more citations across the dataset. WOS and Scopus, in this case did not cover these subjects as well.

6. Discussion and conclusions

Reviewing the literature on GS when compared to controlled databases reveals not only its evolution as a scientific literature retrieval platform but also its strengths and weaknesses. Despite its development through the years and the inclusion of scientific evaluation tools such as citations tracking, author profiles and the like, some of its weaknesses remain constant after more than a decade. In order to track the main challenges, advantages and evolution of GS, the synthesized results are summarized in Table 1.

As can be seen from Table 1, early studies in the area of GS and comparisons with controlled databases began with investigation of social sciences. This is probably the result of their lesser representation in large databases such as WOS and Scopus. Most of the studies focused on coverage issues. The results mainly showed that GS lacks coverage of major publishers' content. Yet, it should be noted that in its early years, publishers did not open their content for GS crawling. This situation has changed in the following years when GS is shown to increase its coverage of major publishers' content as well as retrieve more scholarly artifacts from sources not covered by controlled databases and peer-reviewed journals.

While GS coverage is seen to be growing, studies are recommending that GS will be used in combination with controlled databases and even as a replacement to them in certain areas of research such as arts & humanities and engineering. However, results are showing that lack of indexing policy and transparency, duplications, lack of advanced search options and inability to apply quality control to its indexing capabilities prevent GS from becoming a sole source of scholarly retrieval. Data manipulations have been shown to be easy to achieve by simply uploading fake documents to the web which are then crawled by GS and counted for citations tracking. There is also the issue of data availability. GS limits the amount of data downloads. Therefore it becomes difficult to download data easily and be able to calculate accurate h-index scores also hinders the use of GS as a sole source of citations tracking. Finally, lack of advanced search capabilities makes systematic reviews especially in clinical and medical sciences, which rely on accuracy for comprehensive reviews are difficult to perform on GS. Finally, the unwillingness of GS to allow large-scale downloads of citation and publication data makes using its data for bibliometric analysis very difficult.

GS came along way since its launch in 2004. From a small venture led by two researchers it developed into a scholarly indexing giant which includes over 160 million artifacts. According to an interview with its founders Alex Verstak and Anurag Acharya, GS was initiated as a solution to enable academics to find scholarly items when they search the web and not as a separate platform. It's popularity and ability to crawl numerous sources and apply relevancy rankings to them made GS into the platform it is today. When asked "How do you know what literature to index?" Dr. Van Noorden replies as follows:

> "*'Scholarly' is what everybody else in the scholarly field considers scholarly. It sounds like a recursive definition but it does settle down. We crawl the whole web, and for a new blog, for example, you see what the connections are to the rest of scholarship that you already know about. If many people cite it, or if it cites many people, it is probably scholarly. There is no one magic formula: you bring evidence to bear from many features*" (Van Noorden, 2014, p.1)

This statement embodies the major challenges faced by scholars using GS as a source of not only literature retrieval but also citations tracking.

The main conclusions of studies conducted on the topic of comparison of GS controlled databases are therefore:

1. GS has both coverage and citations advantage over controlled databases in some areas of social sciences, in arts & humanities and engineering & computer science. This is mainly the result of WOS and Scopus not exhaustively covering these areas. GS also has an advantage in these areas because specialized databases in these disciplines are highly selective and have an inherent language bias towards English, and/or to journals as document sources.
2. GS also shows strength in citations tracking across disciplines because of its inclusion of non-English, free and open access resources as well as conference proceedings and books which are carefully selected and included in controlled databases.
3. GS cannot serve as a sole source of literature retrieval or scholarly benchmarking because of its lack of quality assurance and lack of transparency about the resources it covers.

In conclusion; GS is becoming a legitimate source for scholarly information retrieval as well as citations tracking tool. Its coverage of non-peer-reviewed and non-English artifacts has been realized to be an advantage especially for social sciences, arts & humanities as well as life sciences. However, the vast majority of studies caution users not to rely on GS alone for their scientific discovery. In the area of citations calculations, GS is shown to be easily manipulated, and includes duplicate artifacts and citations counts. The fact that the data cannot be easily downloaded and analyzed makes for serious reliability issues to be raised.

In addition, GS does not supply any information on whether its coverage of data sources is systematic, or whether there are gaps to be watched for. Several studies found that GS still lags in coverage of some publishers' content despite of these

Table 1
Summary of results.

Year of Study	Disciplines	Focus area	Main findings	Compared databases	Main recommendations
2005–2006	Social Sciences (LIS, Education) Earth Sciences	Coverage of journals articles	GS fails to consistently retrieve content from publishers' websites but is able to retrieve open access and non- peer reviewed content as well as more proceedings.	Scopus, WOS, PsycINFO, Social Science Citation Index, and ERIC as well as publishers' websites	Use GS in combination with controlled databases and publishers' websites to ensure complete retrieval of articles.
2007–2008	Biology, Chemistry, Physics, Computing, Sociology, Biomedicine, Economics, Psychology, Education Business and Economy, Social Sciences, LISlater-life migration Authors h-index	Coverage, citations tracking, h-index	GS is more usable for social sciences and for tracking conference proceedings. GS still displays gaps in coverage of publishers' content. GS is found to provide new areas of "impact" that go beyond peer-reviewed citations.	Scopus, WOS SCI, SSCI, AH and German social sciences literature database (SOLIS) PubMed, Compendex	Use GS with combination with controlled databases. More endorsement of GS ability to retrieve more artifacts especially in social sciences and arts & humanities.
2009–2010	Chemistry Management and International Business Medicine Computer Science Authors Social sciences (Psychology, LIS, Education)	Coverage, citations tracking	GS is found to track more citations to journal articles especially to non-English journals and proceedings as well as book chapters. For researchers, GS ability to track citations to various document types in various language results in higher h-index	Science Citation Index, Scopus, Chemical Abstracts, Scopus, WOS, PsycINFO PubMed	GS is gaining notice as a complimentary scholarly tool. Most studies recommend using GS h-index to measure journals.
2011–2012	Environmental Sciences Social sciences (social work information management finance and economics) Biology Life sciences Nursing Clinical Sciences Nuclear medicine	Coverage h-index Citations Search capabilities Authors	GS is catching up with coverage of both publishers' content as well as OA content Gaps in coverage are still found but in lower rates GS researcher h-index is found to be able and provides faculty with an additional tool to document the quality of the venues in which they publish.	Scopus, WOS BIOSIS, PubMed ScienceDirect.	Use GS in combination with controlled databases for author h-index calculations In social sciences GS can be used interchangeably with controlled databases Journal rankings can be calculated using GS citations
2013–2014	Environmental Sciences Systematic biomedical systematic review Engineering Communication Studies Population Health Chemistry, Economics, medicine and physics Engineering Nephrology	Coverage Search Capabilities Citations tracking Author impact Precision and recall	GS is shown to retrieve 80%-90% of publishers' content There are inconsistencies in the retrieved results especially for systematic reviews GS performs better in areas of Engineering and social sciences GS reduces the coverage bias toward English language Areas of Chemistry and Physics were shown to be increasingly growing in GS coverage Higher h-index is found to be calculated for authors. Google Scholar provides better recall and provides more access to full-text than PubMed	Scopus, WOS Medline PubMed Compendex Microsoft Academic Search	Clinical systematic reviews cannot rely on GS alone because of gaps in coverage that still exist. For author h-index GS provides higher scores but they need verification for duplications. GS is recommended to be used as supplementary source because of its ability to retrieve non-English artifacts that has less language bias
2015–2016	Astronomy, Environmental Science, Philosophy Public Health Humanities Social sciences Chemistry, Economics, Medicine Physics	Coverage Citations Author h-index	GS is measuring up to disciplinary controlled databases Precision and lack of advanced search options remain a challenge. Lack of indexing policy and transparency are still a concern especially for citations counts.	WOS, Scopus	GS results still need verification against controlled databases especially in medicine. Journal rankings can use GS as a source of citations tracking in the social sciences.

websites allowing it to crawl their content, although this is not as common now. There is lack of clarity around document types covered and updates frequencies. This lack of transparency, diminishes GS reliability not only when citations calculations are concerned but also as a source of consistent reliable literature retrieval tool.

Finally, to this day GS does not offer comprehensive, advanced search capabilities. The advanced search is too basic to accommodate complex questions and results refinements that are crucial for systematic reviews of any type.

Therefore, we conclude that GS must be used in conjunction with controlled databases. Results from each source should be carefully examined for duplications. Document type should also be examined closely when using GS to ensure that they comply with scholarly standards.

Reference managers could assist with the detection of duplicated records and examination of the source titles collected. For citations counts, the situation is more complex. In order to standardize citations collection, one will need to closely compare citations counts from controlled databases to those found in GS one record at a time. For sources not covered by controlled databases, one will need to examine each citing source for accuracy, quality and legitimacy. This issue remains a challenge for bibliometric research especially when citations analysis is concerned.

References

Adriaanse, L. S., & Rensleigh, C. (2011). Comparing web of science, Scopus and Google Scholar from an environmental sciences perspective. *South African Journal of Libraries & Information Science*, *77*(2).

Adriaanse, L. S., & Rensleigh, C. (2013). Web of science, Scopus and Google Scholar: A content comprehensiveness comparison. *The Electronic Library*, *31*(6), 727–744.

Amara, N., & Landry, R. (2012). Counting citations in the field of business and management: Why use Google Scholar rather than the Web of Science. *Scientometrics*, *93*(3), 553–581.

Arlitsch, K., & O'Brien, P. S. (2012). Invisible institutional repositories: Addressing the low indexing ratios of IRs in Google Scholar. *Library Hi Tech*, *30*(1), 60–81.

Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and web of science. *Biomedical Digital Libraries*, *3*(1), 7.

Bar-Ilan, J. (2008). Which h-index?—A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, *74*(2), 257–271.

Bar-Ilan, J. (2010). Citations to the Introduction to informetrics indexed by WOS, Scopus and Google Scholar. *Scientometrics*, *82*(3), 495–506.

Bauer, K., & Bakkalbasi, N. (2005). *An examination of citation counts in a new scholarly communication environment*. D-Lib Magazine. Retrieved from. http://www.citeulike.org/group/400/article/322773

Beckmann, M., & Wehrden, H. (2012). Where you search is what you get: Literature mining–Google Scholar versus Web of Science using a data set from a literature search in vegetation science. *Journal of Vegetation Science*, *23*(6), 1197–1199.

Beel, J., & Gipp, B. (2010). On the robustness of Google Scholar against spam. In *Proceedings of the 21 st ACM conference on hypertext and hypermedia* (pp. 297–298). Retrieved from. http://dl.acm.org/citation.cfm?id=1810683

Boeker, M., Vach, W., & Motschall, E. (2013). Google Scholar as replacement for systematic literature searches: Good relative recall and precision are not enough. *BMC Medical Research Methodology*, *13*(1), 1.

Bohannon, J. (2014). Google Scholar wins raves-but can it be trusted. *Science*, *343*(6166), 14.

Bornmann, L., Marx, W., Schier, H., Rahm, E., Thor, A., & Daniel, H.-D. (2009). Convergent validity of bibliometric Google Scholar data in the field of chemistry—Citation counts for papers that were accepted by Angewandte Chemie International Edition or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts. *Journal of Informetrics*, *3*(1), 27–35.

Bornmann, L., Thor, A., Marx, W., & Schier, H. (2016). The application of bibliometrics to research evaluation in the humanities and social sciences: An exploratory study using normalized Google Scholar data for the publications of a research institute. *Journal of the Association for Information Science and Technology*, http://dx.doi.org/10.1002/asi.23627, n/a–n/a

Bramer, W. M., Giustini, D., Kramer, B. M., & Anderson, P. F. (2013). The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: A review of searches used in systematic reviews. *Systematic Reviews*, *2*(1), 1.

Bramer, W. M., Giustini, D., & Kramer, B. M. R. (2016). Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: A prospective study. *Systematic Reviews*, *5*(1) http://dx.doi.org/10.1186/s13643-016-0215-7

Chan, J. Y., Chan, K. C., Tong, J. Y., & Zhang, F. (2016). Using Google Scholar citations to rank accounting programs: A global perspective. *Review of Quantitative Finance and Accounting*, *47*(1), 29–55. http://dx.doi.org/10.1007/s11156-014-0493-x

Cusker, J. (2013). Elsevier Compendex and Google Scholar: a quantitative comparison of two resources for engineering research and an update to prior comparisons. *The Journal of Academic Librarianship*, *39*(3), 241–243.

De Groote, S. L., & Raszewski, R. (2012). Coverage of Google Scholar, Scopus, and Web of Science: A case study of the h-index in nursing. *Nursing Outlook*, *60*(6), 391–400.

DeGraff, J. V., DeGraff, N., & Romesburg, H. C. (2013). Literature searches with Google Scholar: Knowing what you are and are not getting. *GSA Today*, *23*(10), 44–45.

Delgado López-Cózar, E., & Robinson-García, N. (2012). *Repositories in Google Scholar Metrics or what is this document type doing in a place as such?* Retrieved from. http://digibug.ugr.es/handle/10481/22019

Delgado-López-Cózar, E., & Cabezas-Clavijo. (2013). Ranking journals: Could Google scholar metrics be an alternative to journal citation reports and Scimago journal rank? *Learned Publishing*, *26*(2), 101–113.

Ebrahim, N. A., Salehi, H., Embi, M. A., Danaee, M., Mohammadjafari, M., Zavvari, A., . . . & Shahbazi-Moghadam, M. (2014). Equality of Google Scholar with Web of Science citations: Case of Malaysian engineering highly cited papers. *Modern Applied Science*, *8*(5), 63.

Etxebarria, G., & Gomez-Uranga, M. (2010). Use of Scopus and Google Scholar to measure social sciences production in four major Spanish universities. *Scientometrics*, *82*(2), 333–349.

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, web of science, and Google scholar: Strengths and weaknesses. *The FASEB Journal*, *22*(2), 338–342.

Franceschet, M. (2010). A comparison of bibliometric indicators for computer science scholars and journals on Web of Science and Google Scholar. *Scientometrics*, *83*(1), 243–258.

García-Pérez, M. A. (2010). Accuracy and completeness of publication and citation records in the Web of Science, PsycINFO, and Google Scholar: A case study for the computation of h indices in Psychology. *Journal of the American Society for Information Science and Technology*, *61*(10), 2070–2085.

Gardner, S., & Eng, S. (2005). Gaga over Google? Scholar in the social sciences. *Library Hi Tech News*, *22*(8), 42–45.

Giustini, D., & Boulos, M. N. K. (2013). Google Scholar is not enough to be used alone for systematic reviews. *Online Journal of Public Health Informatics*,. Retrieved from. http://journals.uic.edu/ojs/index.php/ojphi/article/view/4623

Harzing, A.-W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics*, *106*(2), 787–804.
Harzing, A.-W., & Van Der Wal, R. (2009). A Google Scholar h-index for journals: An alternative metric to measure journal impact in economics and business. *Journal of the American Society for Information Science and Technology*, *60*(1), 41–46.
Harzing, A.-W. (2014). A longitudinal study of Google Scholar coverage between 2012 and 2013. *Scientometrics*, *98*(1), 565–575.
Hodge, D. R., & Lacasse, J. R. (2011). Ranking disciplinary journals with the Google Scholar h- index: A new tool for constructing cases for tenure, promotion, and other professional decisions. *Journal of Social Work Education*, *47*(3), 579–596.
Jacobs, J. A. (2011). Journal rankings in sociology: Using the H index with google scholar. *The American Sociologist*, 1–33.
Jacobs, J. A. (2016). Journal rankings in sociology: Using the H index with google scholar. *American Sociologist*, *47*(2–3), 192–224. http://dx.doi.org/10.1007/s12108-015-9292-7
Jacsó, P. (2011). Google Scholar duped and deduped-the aura of robometrics. *Online Information Review*, *35*(1), 154–160.
Jacsó, P. (2012). Google Scholar Author Citation Tracker: Is it too little, too late? *Online Information Review*, *36*(1), 126–141.
Jacso, P. (2005). As we may search-Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. *Current Science-Bangalore*, *89*(9), 1537.
Jamali, H. R., & Nabavi, M. (2015). Open access and sources of full-text articles in Google Scholar in different subject fields. *Scientometrics*, *105*(3), 1635–1651.
Khabsa, M., & Giles, C. L. (2014). The number of scholarly documents on the public web. *PLoS One*, *9*(5), e93949. http://dx.doi.org/10.1371/journal.pone.0093949
Kirkwood, H. P., & Kirkwood, M. C. (2011). Researching the life sciences: BIOSIS previews and google scholar. *Online*, *35*(3), 24–28.
Kousha, K., & Thelwall, M. (2007). Google Scholar citations and Google Web/URL citations: A multi-discipline exploratory analysis. *Journal of the American Society for Information Science and Technology*, *58*(7), 1055–1065.
Kousha, K., & Thelwall, M. (2008). Sources of Google Scholar citations outside the Science Citation Index: A comparison between four science disciplines. *Scientometrics*, *74*(2), 273–294.
Kousha, K., Thelwall, M., & Rezaie, S. (2011). Assessing the citation impact of books: The role of Google Books, Google Scholar, and Scopus. *Journal of the American Society for Information Science and Technology*, *62*(11), 2147–2164.
Kulkarni, A. V., Aziz, B., Shams, I., & Busse, J. W. (2009). Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *JAMA*, *302*(10), 1092–1096.
Levay, P., Ainsworth, N., Kettle, R., & Morgan, A. (2016). Identifying evidence for public health guidance: A comparison of citation searching with Web of Science and Google Scholar. *Research Synthesis Methods*, *7*(1), 34–45. http://dx.doi.org/10.1002/jrsm.1158
Levine-Clark, M., & Gil, E. L. (2008). A comparative citation analysis of web of science, scopus, and google scholar. *Journal of Business & Finance Librarianship*, *14*(1), 32–46.
Levine-Clark, M., & Kraus, J. (2007). Finding chemistry information using google scholar: A comparison with chemical abstracts service. *Science & Technology Libraries*, *27*(4), 3–17.
Lewandowski, D. (2010). Google Scholar as a tool for discovering journal articles in library and information science. *Online Information Review*, *34*(2), 250–262.
Lopez-Cozar, E. D., Robinson-Garcia, N., & Torres-Salinas, D. (2012). *Manipulating google scholar citations and google scholar metrics: Simple, easy and tempting*. arXiv Preprint arXiv:1212.0638. Retrieved from. http://arxiv.org/abs/1212.0638
Martín-Martín, A., Orduña-Malea, E., Ayllón, J. M., & López-Cózar, E. D. (2014). *Does Google Scholar contain all highly cited documents (1950–2013)?* arXiv:1410.8464 [Cs] [Retrieved from]. http://arxiv.org/abs/1410.8464
Martín-Martín, A., Orduna-Malea, E., Ayllón, J. M., & López-Cózar, E. D. (2016). *The counting house: Measuring those who count. Presence of bibliometrics, scientometrics, informetrics, webometrics and altmetrics in the google scholar citations, ResearcherID, ResearchGate, mendeley & twitter*. arXiv Preprint arXiv:1602.02412. Retrieved from. http://arxiv.org/abs/1602.02412
Mastrangelo, G., Fadda, E., Rossi, C. R., Zamprogno, E., Buja, A., & Cegolon, L. (2010). Literature search on risk factors for sarcoma: PubMed and Google Scholar may be complementary sources. *BMC Research Notes*, *3*(1), 131.
Mayr, P., & Walter, A.-K. (2008). Studying journal coverage in google scholar. *Journal of Library Administration*, *47*(1–2), 81–99.
Meho, L. I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, *58*(13), 2105–2125.
Meier, J. J., & Conkling, T. W. (2008). Google Scholar's coverage of the engineering literature: An empirical study. *The Journal of Academic Librarianship*, *34*(3), 196–201.
Mikki, S. (2010). Comparing google scholar and ISI web of science for earth sciences. *Scientometrics*, *82*(2), 321–331.
Minasny, B., Hartemink, A. E., McBratney, A., & Jang, H.-J. (2013). Citations and the h index of soil researchers and journals in the Web of Science, Scopus, and Google Scholar. *PeerJ*, *1*, e183.
Mingers, J., & Lipitakis, E. A. (2010). Counting the citations: A comparison of Web of Science and Google Scholar in the field of business and management. *Scientometrics*, *85*(2), 613–625.
Moed, H. F., Bar-Ilan, J., & Halevi, G. (2016). A new methodology for comparing Google Scholar and Scopus. *Journal of Informetrics*, *10*(2), 533–551. http://dx.doi.org/10.1016/j.joi.2016.04.017
Neuhaus, C., Neuhaus, E., Asher, A., & Wrede, C. (2006). The depth and breadth of Google Scholar: An empirical study. *Portal: Libraries and the Academy*, *6*(2), 127–141.
Norris, M., & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. *Journal of Informetrics*, *1*(2), 161–169.
Nourbakhsh, E., Nugent, R., Wang, H., Cevik, C., & Nugent, K. (2012). Medical literature searches: A comparison of PubMed and Google Scholar. *Health Information & Libraries Journal*, *29*(3), 214–222.
Orduña-Malea, E., Ayllón, J. M., Martín-Martín, A., & López-Cózar, E. D. (2014). *About the size of Google Scholar: Playing the numbers*. arXiv Preprint arXiv:1407.6239. Retrieved from. http://arxiv.org/abs/1407.6239
Ortega, J. L., & Aguillo, I. F. (2014). Microsoft academic search and Google scholar citations: Comparative analysis of author profiles. *Journal of the Association for Information Science and Technology*, *65*(6), 1149–1156.
Shariff, S. Z., Bejaimal, S. A., Sontrop, J. M., Iansavichus, A. V., Haynes, R. B., Weir, M. A., & Garg, A. X. (2013). Retrieving clinical evidence: A comparison of PubMed and Google Scholar for quick clinical searches. *Journal of Medical Internet Research*, *15*(8), e164.
Thor, A., & Bornmann, L. (2011). The calculation of the single publication h index and related performance measures: A web application based on Google Scholar data. *Online Information Review*, *35*(2), 291–300.
Tober, M. (2011). PubMed, ScienceDirect, Scopus or Google Scholar–Which is the best search engine for an effective literature research in laser medicine? *Medical Laser Application*, *26*(3), 139–144.
Trapp, J. (2016). Web of Science, Scopus, and Google Scholar citation rates: a case study of medical physics and biomedical engineering: what gets cited and what doesn't? *Australasian Physical & Engineering Sciences in Medicine*, 1–7.
van Aalst, J. (2010). Using Google Scholar to estimate the impact of journal articles in education. *Educational Researcher*, *39*(5), 387–400.
Van Noorden, R. (2014). Google Scholar pioneer on search engine's future. *Nature*,. Retrieved from. http://www.citeulike.org/group/15400/article/13421985

Wakimoto, D. K. (2014). Google scholar retrieves twice as many relevant citations as PubMed and provides greater full-text access for quick, clinical nephrology searches. *Evidence Based Library and Information Practice*, *9*(1), 36–38.

Walters, W. H. (2007). Google Scholar coverage of a multidisciplinary field. *Information Processing & Management*, *43*(4), 1121–1132.

Wildgaard, L. (2015). A comparison of 17 author-level bibliometric indicators for researchers in astronomy, environmental science, philosophy and public health in web of science and Google Scholar. *Scientometrics*, *104*(3), 873–906.

Zarifmahmoudi, L., & Sadeghi, R. (2012). Comparison of ISI web of knowledge, SCOPUS, and Google Scholar h-indices of Iranian nuclear medicine scientists. *Iranian Journal of Nuclear Medicine*, *20*(1), 1–4.