in period 2006-2010. The total number of citations of these articles ranged from 618-4,833 (median = 934.5) citations. These articles originated from 21 countries, with nearly half from the USA (n = 49), followed by Spain and Italy (n = 19). The University of Barcelona, Spain contributed the most articles (n = 13). Hepatology contributed the greatest number of articles (n = 19), with New England Journal of Medicine having the greatest number of citations (n = 21,276).

Seventy-five of the top 100 papers were original articles; most of these focused on genetics/pathophysiology (n = 27) or management (n = 24).

CONCLUSION: Our analysis serves as a reference for the most highly cited works in HCC and characterizes research trends in the field.

Key words: Bibliometrics; Citation classics; Liver cancer; Hepatocellular carcinoma; Scientometrics

© 2018 The Author(s). Published by ACT Publishing Group Ltd. All rights reserved.


INTRODUCTION

Hepatocellular carcinoma (HCC) is the most frequent primary liver cancer, and in 2015 it ranked as the sixth most common neoplasm and the fourth leading cause of cancer-related mortality[1]. Given this high global disease burden, it can be appreciated that significant resources are devoted to research in HCC. The association of several modifiable risk factors such as viral hepatitis infections and alcoholic liver disease with hepatocarcinogenesis has generated strong interest in preventive screening research and implementation. In addition, there is an urgent and expanding role for novel molecular therapies like sorafenib in the management of this chemo-resistant tumor. The promise of precision medicine in HCC has focused much research on genomic profiling to identify prognostic markers and targets for drug therapy.
With increasing emphasis on evidence-based medicine, recent years have seen a rise in the number of research publications[5]. Bibliometric methods of analyzing research meta-data can provide useful information to funding agencies, federal policymakers and individual researchers and clinicians. Citation analysis can also identify which areas should be prioritized by future research. For physicians, a list of citation classics enables them to efficiently navigate landmark developments in their field.

Despite criticism, citation rate and impact factor metrics are widely used to assess the scholarly impact of a scientific manuscript[6, 4]. Many medical and surgical specialties have utilized bibliometric techniques to identify their citation classics and characterize research trends, including several gastrointestinal malignancies[5-11]. However to our knowledge no such analysis has yet been published on HCC. The objective of our study was to present the 100 top cited articles and to gain insight on the evolution of HCC research.

METHODS

In January 2018, two independent researchers (NA and RS) searched the Scopus database (www.scopus.com) for citations of articles published on HCC. As compared to PubMed, Web of Science or Google Scholar, Scopus is known to have a greater coverage where scientific publications are concerned[12]. In order to broaden the results, all article types, non-human studies and studies without abstract availability were included. No time restriction was applied and all the journals in the database were included in the search. Articles with their main focus on HCC, epidemiology, carcinogenesis, management, surgery and chemotherapy were included.

Our primary search phrases included “liver cancer”, “hepatocellular cancer”, “hepatic cancer”, “hepatoocarcinogenesis” and “HCC”. All these terms were also searched using “carcinoma”, “neoplasms”, “tumor” and “malignancy” and their plural forms. All individual key phrases were searched for in the title, keywords and abstract.

The returned dataset was strictly related to HCC. Abstracts of all articles were thoroughly analyzed for relevance to the topic according to the inclusion and exclusion criteria, independently by both reviewers. Abstracts were obtained from other search engines if they were not available on Scopus and similarly analyzed for relevance. The top 100 selected articles were ordered using “cited by” option. A final list was made after comparing both reviewers’ lists. A difference of 5.5% was observed between the lists. Articles were only included after consensus and agreement of both reviewers.

The final list was exported to Microsoft Excel 2016. The list was analyzed for the following: article title, author, author affiliations, number of authors, journal, total citations, year of publication, citations per year, institution and country of origin of the first author. H index and affiliations of authors were obtained from their Scopus profiles. We calculated the median and inter-quartile range (IQR) for total number of citations and citations per year, as our data were not expected to be normally distributed. Thomson Reuters’ Journal Citation Reports 2017 was used to obtain the journal’s impact factor. Tables and charts were made using Microsoft Excel 2016 (Microsoft Corp., Redmond, WA, USA). Pearson product-moment correlation test was applied to determine the correlation between journal impact factor and the number of papers of that journal in the top 100 list, using IBM SPSS (Statistical Package for the Social Sciences; SPSS Inc., Chicago, IL, USA) version 22. A p value of <0.05 was considered significant. No ethics approval was required for this study as all data are publicly available.

RESULTS

1. Citation Count and Citations per Year

Table 1 shows the top 100 cited articles in the field of HCC, along with their total citations, citation per year and rank according to total citations and citation per year. The citation count ranged from 618-4,833 for these 100 articles, with median of 934.5 (IQR = 582.5). The citations of all the articles sum up to 125,615, with approximately 3% (n = 3762) being self-citations. The number of citations per year ranged from 14.6 to 579.4. The median number of citations per year was 62.8 (IQR = 65.0).

2. Year of Publication, Origin and Authorship

The top ten articles were published after 1996. The 100 most cited articles were published between 1954 and 2012. During this 58 year span, the largest number of these articles (n = 29) were published in the 5-year period from 2006 to 2010, followed by 2001-2005 (n = 24) as shown in Figure 1. During 2006-2010, 5 articles out of the 29 published discussed sorafenib. In contrast, out of the 40 articles published during 1993-2003, 19 articles were based on HCC treatment; 10 articles from the aforementioned 19 were about chemoembolization and radiofrequency ablation, whereas 5 explored liver transplantation and resection. The remaining 4 covered other miscellaneous aspects of HCC treatment.

Twenty-one different countries contributed to the 100 most cited articles. Near half of the articles originated from USA (n = 49), followed by Italy and Spain with an equal number of articles (n = 19). Thirty-one articles had authors from more than one country. One article from the list had unidentified authorship nationality (Figure 2).

A total of 808 authors contributed to these 100 articles. The number of authors for each article ranged from 1 to 49, with each paper having a median of 7 authors (IQR = 7.25). Table 2 lists all the authors with 5 or more publications in top 100 cited articles along with their H-index. Jordi Bruix coauthored the greatest number of articles (n = 19) followed by Josep M Llovet (n = 17), and Hashem B El-Serag (n = 7).

3. Journal and Institutional Affiliations

The 100 most cited articles were published in 28 different journals, with top 5 producing nearly half of the articles (n = 49). These five journals included (1) Hepatology (2) New England Journal of Medicine (3) Gastroenterology (4) The Lancet, and (5) Nature. Hepatology published the most articles (n = 19), but New England Journal of Medicine generated the largest number of total citations (21,276) with 15 papers. The median impact factor for these 28 journals was 15.2 (range = 2.19-72.4). Thirty-five articles in our list were published in journals having impact factors of 30.0 or above. We found no significant correlation between the impact factor of a journal and the number of articles of that journal in the top 100 list (p = 0.146). Journals with more than 2 articles in top 100 are shown in Table 3.

![Figure 1](image_url)  
*Figure 1 Number of publications in each 5-year interval.*
Table 1 Top 100 cited articles, their total citations, citations per year and rank according to citations per year.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>El-Serag HB, Rudolph KL. Hepatocellular Carcinoma: Epidemiology and Molecular Carcinogenesis. Gastroenterology. 2007; 132(7): 2557-76.</td>
</tr>
<tr>
<td>11</td>
<td>Edmondson HA, Steiner PF. Primary carcinoma of the liver. A study of 100 cases among 48,900 necropsies. Cancer. 1954; 7(3): 462-503.</td>
</tr>
</tbody>
</table>

Asghar N et al. Citation classics in hepatocellular carcinoma


A number of institutions contributed to the top 100 articles. University of Barcelona contributed the greatest number of articles (n = 13), followed by Baylor College of Medicine (n = 6). Based on first author institutional affiliations, institutions which contributed more than 2 articles in our list are shown in Table 4.

### 4. Type of Article and Topical Distribution

We divided the top 100 cited articles into 3 categories, according to the type of study: original articles (n = 75), review/expert opinion (n = 20) and official guidelines (n = 5). Figure 3 shows the temporal trend of publication by type of article. Original articles were further subdivided according to article’s main focus into 5 categories: genetics and pathophysiology (n = 27), management (n = 24), epidemiology/natural history (n = 19), prognostication (n = 5) and surveillance (n = 2).

Of the top 40 most cited articles, 35% (n = 14) were focused on management, while this figure rose to 70% (n = 7) among the top 10 articles. This pattern is also reflected in the citations-per-year ranking, indicating that the exceptionally high citation counts of management articles are not attributable to earlier publication.

### DISCUSSION

**Research Activity and Trends**

As shown in Figures 1 and 3, the time period from 1991 to 2010 saw the publication of a large majority of the 100 most-cited articles in the list (n = 82); furthermore, the distribution of these articles through this period reflects an exponential increase in research activity. This pattern of literature growth has been identified by a variety of bibliometric analyses in the past—notably, by studies analyzing the literature on clinically relevant topics such as diabetes[19] or chikungunya[20], which suggests the modern significance of HCC as a disease entity to be dealt with in clinical practice. Additionally, as suggested by previous studies,[11] the increasing number of publications since 1991 may be also be linked to the establishment of the World Wide Web. Fur-
ther analysis of the articles published in this interval pinpoints peak research activity to be in the period from 2006 to 2010. The specific subject matter of these articles provides valuable insight into potential breakthroughs in the field — most importantly, the emergence of sorafenib as a novel therapeutic agent for HCC.

During the 2006-2010 periods alone, 5 articles (out of the 29 published during these years) were seen to explore the therapeutic potential of sorafenib. Collectively, these randomized controlled trials (RCTs) and experimental studies served to highlight the significant survival benefit and acceptable tolerance to sorafenib, which may be responsible for a shift in treatment strategies. While sorafenib-based studies were found to predominate in the 2006-2010 period, the majority of articles focusing on liver transplantation or resection were published from 1993 to 2001, whereas articles exploring chemoeembolization and radiofrequency ablation for unresectable tumors were published mainly from 1995 to 2003. Overall, this signifies a growing interest in applying targeted therapeutics to the field of HCC.

Additionally, analysis of the top-cited articles also shows a sudden decline in research activity from 2011 to 2015 after the exponential increase. This may simply be attributed to the temporal disadvantage of newer articles, given that they have had little time to garner citations. The introduction of a ranking by citations-per-year (Table 1) reduces this bias considerably and emphasizes the significance of recently-published articles higher up in this ranking. On the other hand, the past 5-6 years have also seen the publication and frequent updating of reviews and clinical guidelines in favor of original articles.

Content
Survey of the findings and rationale of the top 10 most-cited articles allows clear identification of the gaps in literature bridged by these studies, and puts their contribution to the field of HCC into perspective. Excluding clinical guidelines, 2 of the top 10 most-cited articles provide clinical evidence of the therapeutic benefit of sorafenib in advanced HCC, while another study highlights liver transplantation as an effective treatment. Apart from clinical studies, 3 of the highest-cited articles review the epidemiological aspects of HCC; proposing strategies for prevention and prognostication in view of known risk factors, defining molecular pathways, and underlining the shifting incidence of HCC from older to younger patient populations over time. Among pathophysiology-based articles, the relationship between HCC incidence and hepatitis B or C seems to be of particular interest, as well as the role of genetic alterations such as p53 mutations. More importantly, the increasing number of citations garnered by management-based articles in recent years may align with attempts being made towards the development and assessment of curative procedures for HCC. Although the initial percentage of articles discussing management in the top 100 list is 24%, this increases to 35% in the top 40 most-cited articles and 70% in the top 10 most-cited articles. This pattern is also reflected in the citations-per-year ranking, indicating that the exceptionally high citation counts of management articles are not attributable to earlier publication. In comparison with bibliometric analyses of other hepatic pathologies, the literature exploring Budd-Chiari syndrome shows a similar focus on clinical reports[15], while studies on non-alcoholic fatty liver disease favor pathogenesis and epidemiology[14].

Journals
Table 3 summarizes the journals that most frequently contained the top-cited articles, along with their respective impact factors. The highest numbers of articles were found to be published in the specialty journal Hepatology. Overall, however, our analysis shows a fairly even distribution of articles to general versus specialty journals (52% vs. 48% respectively). Moreover, the broad readership and high impact factors of journals such as New England Journal of Medicine, The Lancet and Nature indicate the importance of these studies and their results to the scientific community as a whole. Such high-impact journals (impact factor ≥ 30) represent 35% of all publications in the list, while the median impact factor for journals in our list was 15.2. This figure is higher than that reported for citation classics in other fields such as gastric cancer[19] (median impact factor 9.28), reinforcing the relative impact of HCC in academic research.

Unlike several fields such as cardiovascular[17], diabetes[19] and thrombolytics[17] research, our analysis did not identify a correlation between journal impact factor and the number of top-cited articles published in that journal. This may be related to the near-equal contribution of specialty journals among the top 100 cited articles, as these journals tend to have narrower readership—indicating the need for and importance of sub-specialized expertise in the field of HCC.

Origins and Authors
As denoted in Figure 2, USA has contributed the highest number of top-cited articles to HCC literature by a large margin, followed by Western European countries. These findings are unsurprising, given the number of established research institutes in the US and the widespread availability of resources. Furthermore, 31% of articles were found to be multi-origin, thus implying open exchange of information across borders and solidifying the role of such international exchange in producing high-quality research. Apart from this, the overwhelming disease burden of HCC in sub-Saharan Africa and Eastern Asia[19] strongly encourages a global effort to increase the productivity of these countries so as to evaluate these populations more thoroughly.

In terms of authorship, it is interesting to note that the author and institute with the highest number of publications originate from Spain, rather than USA. The recognition of individual authors helps to identify significant influences in the field of research and potentially cements the increased value of their work. However, only 6 authors were found to contribute more than 5 articles to the top 100 list, which may indicate a lack of prominent thought leaders despite the high overall productivity of the field.

Limitations
As a citation classic, this analysis is subject to various types of bias that may interfere with accurate representation by citation count. This includes institutional bias, language bias, powerful person bias, and in-house bias[16]. We calculated self-citations among the total number of citations, and found them to be around 3%, indicating that bias resulting from self-citations is minimal in our analysis. Furthermore, we utilized only one database in our study to collect references, which may have resulted in the omission of articles not recognized by Scopus. Finally, despite the production of a separate list by citations-per-year, high-quality articles published in 2015 or later require a certain number of years to garner citation counts reflecting their impact, due to which any such recent articles would likely be excluded from our list.

ACKNOWLEDGEMENTS

No conflict of commercial, personal, political, intellectual, or religious interests.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
REFERENCES


*Peer Reviewer:* Cerwenka Herwig