



A new database of the references on international clinical practice guidelines: a facility for the evaluation of clinical research

Magnus Eriksson¹ · Annika Billhult^{1,2} · Tommy Billhult¹ · Elena Pallari^{3,6} · Grant Lewison^{4,5}

Received: 29 August 2019 / Published online: 14 December 2019
© The Author(s) 2019

Abstract

Although there are now several bibliographic databases of research publications, such as Google Scholar, Pubmed, Scopus, and the Web of Science (WoS), and some also include counts of citations, there is at present no similarly comprehensive database of the rapidly growing number of clinical practice guidelines (CPGs), with their references, which sometimes number in the hundreds. CPGs have been shown to be useful for the evaluation of clinical (as opposed to basic) biomedical research, which often suffers from relatively low counts of citations in the serial literature. The objectives were to introduce a new citation database, clinical impact[®], and demonstrate how it can be used to evaluate research impact of clinical research publications by exploring the characteristics of CPG citations of two sets of papers, as well as show temporal variation of clinical impact[®] and the WoS. The paper includes the methodology used to retain the data and also the rationale adopted to achieve data quality. The analysis showed that although CPGs tend preferentially to cite papers from their own country, this is not always the case. It also showed that cited papers tend to have a more clinical research level than uncited papers. An analysis of diachronous citations in both clinical impact[®] and the WoS showed that although the WoS citations showed a decreasing trend after a peak at 2–3 years after publication, this was less clear for CPG citations and a longer timescale would be needed to evaluate their impact on these documents.

Keywords Research papers · Clinical practice guidelines · References · Clinical research evaluation · Clinical impact

Introduction

The use of counts of citations to research papers has been a standard method for research evaluation for many years (Kostoff 1998; Ingwersen et al. 2000; Leydesdorff 2008; Moed 2009). Their use was made practicable by the creation of the Science

✉ Magnus Eriksson
magnus.eriksson@minso.se

Extended author information available on the last page of the article

Citation Index by Garfield (1955, 1972), first as large printed volumes, subsequently as CD-ROMs and most recently as the Web of Science (WoS, © Clarivate Analytics). A rival database published by Elsevier, Scopus, was introduced in 2004 (Manafy 2004). It has different properties, and also a different journal coverage, so that citation scores to individual papers will usually also differ (Meho and Yang 2007; Meho and Sugimoto 2009; Bergman 2012; Leydesdorff et al. 2016; Martin-Martin et al. 2018). Although both these databases are proprietary, and are available on subscription to most universities and research institutes in OECD countries, they may be unavailable to some research organisations in other countries, who will, however, be able to obtain citation scores for their papers from a third database, Google Scholar, which is non-proprietary and available in most countries.

There is now a very extensive literature on the theory and practice of citation analysis. The distribution of citation scores has been the subject of extensive research, much of it mathematical (van Raan 2001; Anastadiadis et al. 2010; Ruiz-Castillo 2013; Abramo et al. 2016). It has also been established that the citations to groups of papers, which must strictly be counted in a fixed time window, vary with the parameters of the papers such as the numbers of authors, the numbers of addresses, the numbers of funding acknowledgments, the field of study, and whether the paper is applied (clinical for biomedical research) or basic (Lewison and Dawson 1998; Roe et al. 2010; van Eck et al. 2013). The last of these factors (Research Level, RL) has a big influence, and (with some exceptions) clinical papers tend to be less well cited than basic ones. One effect of this is that when biomedical research proposals are being evaluated by committees of award, clinical work may appear to be less influential, and so have less chance of being funded. However, it may well have more immediate influence on clinical practice.

One way of measuring this influence is through its citation in clinical practice guidelines (CPGs), but as there is no comprehensive dedicated database of such citations, it cannot be easily demonstrated. Citations in CPGs do however appear to some extent in other altmetric services, although they suffer from some issues such as errors in attribution and possible duplicates in potentially between 20 and 50% of the citing CPGs (Tattersall and Carroll 2018) in some cases. It is also unclear how these services handle the different contexts in which research is cited in CPGs, since CPGs also tend to list excluded studies in their reference lists.

CPGs are being published in many countries, and in regions within them. In the UK there are two well-established sets, one from the Scottish Inter-collegiate Guidelines Network (SIGN), which began in 1993 (anon 1997; Miller 2002), and one for use in England and Wales from the National Institute for health and Care Excellence (NICE), which started in 1999 (Cluzeau and Littlejohns 1999; Wailoo et al. 2004). Figure 1 shows the numbers of current guidelines published by these two organisations by year of publication. Both sets retire CPGs and replace them with new ones as new evidence appears.

In Sweden several organisations publish CPGs, both national agencies such as the National Board of Health and Welfare (*Socialstyrelsen*) and the Swedish Medical Products Agency (*Läkemedelsverket*), and regional organizations such as the regional Centres for Health Technology Assessments and the Regional Cancer Centres. The regional organisations are partnerships of the 20 Swedish counties, based on the six Swedish medical regions (Fig. 2). However, the majority of the CPGs are produced by the national agencies (Fig. 3).

Researchers in King's College London (KCL) have been active in the processing of CPGs, and have covered ones from 19 other European countries (Pallari and Lewison 2019). This analysis of CPG references was carried out as part of a European Union

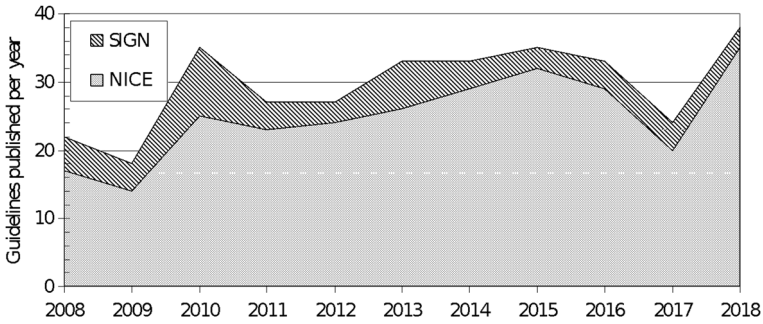
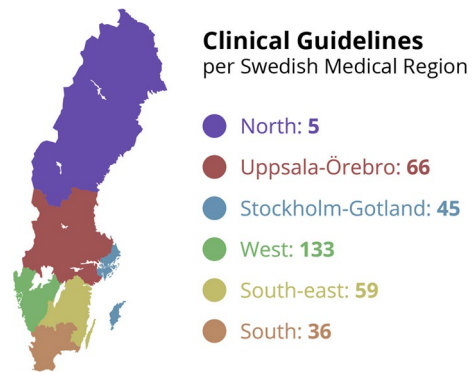


Fig. 1 Numbers of CPGs published each year by NICE and SIGN in the UK, 2008–18

Fig. 2 Clinical practice guidelines published in each Swedish Medical Region (2018)



Guidelines per year

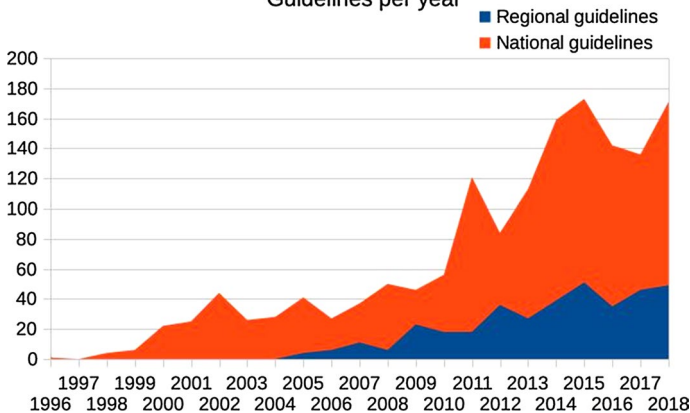


Fig. 3 Number of clinical practice guidelines published in Sweden by year, 1997–2018

contract to map European research outputs and impacts in five non-communicable diseases (NCDs). These were cardiovascular disease and stroke (CARDI), diabetes (DIABE), mental disorders (MENTH), cancer (ONCOL), and respiratory diseases (RESPI).

The references in continental European CPGs, most of which are unsurprisingly from the same subject areas as the guidelines themselves, similarly form a good sample of European research outputs—and also those of other countries, notably the USA (Begum et al. 2016; Pallari et al. 2018a, b).

Early work (Grant 1999; Grant et al. 2000) and by Lewison and colleagues (Lewison and Wilcox-Jay 2003; Lewison 2007; Lewison and Sullivan 2008; Pallari et al. 2018a, b) has shown conclusively that the references are, on average, very clinical; they are heavily cited by CPGs from their own country; they are not particularly recent, although this varies a lot by country; and the individual papers are often highly cited by other papers (in the WoS or Scopus).

Because of the second point, which also pertains to biomedical research papers overall (Bakare and Lewison 2017), evaluation of a set of papers with our new database can serve two distinct purposes: the influence of the selected papers on the national health care provision, and the influence internationally on other health care systems.

Which of these two is more important will depend on the mission statement of the funding body, or of the research performer.

In this paper, we describe the methodology used to create the clinical impact[®] database that allows us to compile a large set of references in CPGs (more than 1,120,000 processed references, of which 450,000 are used in this paper based on their classification as *Included* references) in a quick and accurate manner, and then how we match these to one set of papers published by a hospital in Scotland, and to another set of papers acknowledging the support of a Swedish collecting medical research charity. The latter papers are from the 10 years of 2009–2018, but the Scottish hospital papers are from three individual years, 2000, 2004 and 2008. We will try to identify differences between papers that are cited in CPGs and those that are not, in regard to nationality and Research Level (RL) of the papers. We will also explore the temporal variation of CPG citations and academic citations by looking at the diachronous distribution of citations. The numbers of CPGs from the different countries consist, at the time of this study, of 10,058 CPGs from 28 different countries and international organisations.

Methodology

CPGs processed in the clinical impact[®] database are documents created with the intention to provide guidance to health care practitioners, created by established guideline providers, published electronically online, and containing references to published research. Since CPGs often differ in format, uniform concepts have been established to manage CPGs and the references in the clinical impact[®] database.

- *Guideline document* A single CPG, often composed of several files. The additional files may be supplements or appendices.
- *Guideline provider* An organisation that provides a number of CPGs on their web page.
- *Guideline provider collection* A number of CPGs of the same type found on a Guideline Provider's web page. A Guideline Provider may have several Guideline collections on their pages, e.g. National Practice Guidelines, Screening programs, Health Technology Assessments (HTAs), Enquiry services, or Advice. All Guideline collections are updated annually.

- *Originator* The organisation, or organisations in a collaboration, that developed a specific CPG. The Originator and Guideline provider are often, but not always, the same organisation.

Processing of the CPGs to create the citation database

The clinical impact[®] database was created using Minso Solutions Ref Extractor tool, a proprietary software for identifying references in documents and create citation databases. The data collection process starts by identifying and entering metadata for a Guideline provider collection into the ref extractor tool. metadata added to guidelines provider collections are the providing organisation, its URL, nationality and organization type—Governmental or Non-governmental. The files are grouped and further metadata are added such as guideline provider collection, title, originator, language, date of publication, number of pages, identifiers (DOI or ISBN), and, if applicable, a successor or predecessor tag. Successor and predecessor state the relation between new editions and revisions of individual CPGs, which sometimes take place annually, with similar sets of references. In CPGs that use systematic reviews to assess evidence, the references are usually split into several lists. The references are classified in the clinical impact[®] database as either *Included* references, *Excluded* references, or *Additional* references, based on the type of reference list. *Included* references are the research papers that provide the body of evidence underpinning the recommendations in the CPGs. *Excluded* references have methodological flaws or other characteristics causing the authors of the CPG to deem them irrelevant or regard them as insufficient and not reliable as evidence. *Additional* references are those to which the CPG reader is referred for further information, sometimes even listing literature recommended for patients. Note that only the *Included* references are used for analysis in this paper. In CPGs with only one general reference list, excluded and additional references are generally left out by the CPG authors, and therefore all references in those CPGs are classified as *Included* references. Extracted references are matched against external bibliographic databases for verification and retrieval of publication identifiers such as DOI, PMID, and UT Accession Number. Excluded references and “junk”—strings of characters that resemble references picked up by the extraction process, are not matched to external databases. This classification is likely very important to take into consideration when studies of papers cited in CPGs are being conducted, since the included, excluded, and additional references each account for nearly a third each of all references in the CPGs, but only the included references form the evidence base for the recommendations. The distribution of included, excluded, and additional references in clinical impact[®] may however change as more CPGs are processed. Certain papers are decidedly more cited in CPGs than others, and these tend to be papers on methodology for systematic literature reviews and evidence grading. In fact, among the top 20 cited papers in CPGs, the first clinical research paper appears at number 19, see Table 1.

The external identifiers (DOI, PMID) allow the references extracted from the CPGs to be matched to sets of papers obtained by other means in databases such as Scopus and the WoS. These databases contain more descriptive parameters than Medline, and so papers with particular characteristics resulting from a search in one of these databases can then be matched to the clinical impact[®] database.

At KCL, the processing procedure was somewhat different. The references on the PDF versions of the CPGs were either pasted directly into MS Excel, or into MS Word, where they could be individually numbered (if necessary) and put on a single line of text. References not in journals were removed. They were then copied and pasted into MS Excel, and separated

Table 1 Top 20 cited papers in the clinical impact[®] database

	Cited paper	Citations in CPGs
1	Browman GP, Levine MN, Mohide EA, Hayward RS, Pritchard KI, Gafni A, Laupacis A. The practice guidelines development cycle: a conceptual tool for practice guidelines development and implementation. <i>J Clin Oncol.</i> 1995 Feb;13(2):502–12	246
2	Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. <i>BMC Med Res Methodol.</i> 2007 Feb 15;7:10	202
3	Browman GP, Newnam TE, Mohide EA, Graham ID, Levine MN, Pritchard KI, Evans WK, Maroun JA, Hodson DI, Carey MS, Cowan DH. Progress of clinical oncology guidelines development using the Practice Guidelines Development Cycle: the role of practitioner feedback. <i>J Clin Oncol.</i> 1998 Mar; 16(3):1226–31	165
4	Owens DK, Lohr KN, Atkins D, Treadwell JR, Reston JT, Bass EB, Chang S, Helfand M. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—agency for healthcare research and quality and the effective health-care program. <i>J Clin Epidemiol.</i> 2010 May;63(5):513–23	113
5	Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. <i>BMJ.</i> 2008 Apr 26;336(7650):924–6	85
6	Whiting PF, Ruitjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM; QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. <i>Ann Intern Med.</i> 2011 Oct 18;155(8):529–36	84
7	DerSimonian R, Laird N. Meta-analysis in clinical trials. <i>Control Clin Trials.</i> 1986 Sep;7(3):177–88	75
8	Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Grimshaw J, Hanna SE, Littlejohns P, Makarski J, Zitzelsberger L; AGREE Next Steps Consortium. AGREE II: advancing guideline development, reporting and evaluation in health care. <i>CMAJ.</i> 2010 Dec 14;182(18):E839–42.	75
9	Atkins D, Chang SM, Gartlehner G, Buckley DI, Whitlock EP, Berliner E, Matchar D. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. <i>J Clin Epidemiol.</i> 2011 Nov;64(11):1198–207.	73
10	Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schünemann HJ. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. <i>J Clin Epidemiol.</i> 2011 Apr;64(4):383–94	65
11	Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. <i>BMJ.</i> 2011 Oct 18;343:d5928	64
12	Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. <i>J Clin Epidemiol.</i> 2009 Oct;62(10):1006–12	60

Table 1 (continued)

	Cited paper	Citations in CPGs
13	AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. <i>Qual Saf Health Care.</i> 2003 Feb;12(1):18–23	50
14	Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. <i>Stat Med.</i> 2002 Jun 15;21(11):1539–58.	45
15	Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. <i>BMJ.</i> 2003 Sep 6;327(7414):557–60	39
16	Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. <i>Stat Med.</i> 1998 Dec 30;17(24):2815–34	39
17	Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. <i>BMC Med Res Methodol.</i> 2003 Nov 10;3:25	38
18	Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, Atkins D; Methods Work Group. Third US Preventive Services Task Force. Current methods of the US Preventive Services Task Force: a review of the process. <i>Am J Prev Med.</i> 2001 Apr;20(3 Suppl):21–35	36
19	Heart Outcomes Prevention Evaluation Study Investigators, Yusuf S, Sleight P, Pogue J, Bosch J, Davies R, Dagenais G. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. <i>N Engl J Med.</i> 2000 Jan 20;342(3):145–53	32
20	Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, Jackson RD, Beresford SA, Howard BV, Johnson KC, Kotchen JM, Ockene J; Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. <i>JAMA.</i> 2002 Jul 17;288(3):321–33	32

into components: authors, title, and publication year. The words in the titles were separated with hyphens rather than spaces. Individual search statements were then created, assembled into groups of 20, and then run against the WoS so as to identify these references, whose details were then downloaded into text files. These were converted into MS Excel files by means of a macro (Visual Basic Application program) written by Philip Roe of Evaluametrics Ltd. The spreadsheet included the DOI and PMID for each reference, if available, for matching to the corresponding values for any research papers whose CPG citations were sought.

An example of the use of the new database for research evaluation

In order to illustrate the use of the new database for research evaluation, we provide two worked examples of the information that can be provided for a research performer, namely a Scottish hospital, and for a research funder, namely a Swedish collecting charity. The time periods of the research papers attributable to each differ in these illustrative examples. In practice, we would identify all “their” papers over a fairly long and continuous period, perhaps the previous 10 or 15 years. The relevant papers would be sought in the WoS, and their DOIs and PMIDs would be listed, as well as all other bibliometric parameters, so that an analysis could be made of those that led to more CPG citations, and the information used for management purposes.

The identification of papers from a Scottish hospital, and their classification

Identification of the Scottish hospital papers is a comparatively simple task, and the main hospital name is used in the WoS address field, but with the restriction that the address must also be in Scotland (there is another clinic with the same name in Sri Lanka). The bibliographic data from the identified papers are then downloaded as a series of text files, with full bibliographic data (author names, title, source, document type, addresses, month and year of publication, and funding information). However, the only identifiers needed for cross-matching are the DOI and PMID codes. A few of the WoS papers do not have either of these codes; for these a match can be made on the title. Papers were collected for three publication years, 2000, 2004, and 2008.

We use a series of macros that allow the papers in the list to be characterised by a number of parameters. One classifies them by their research level (RL, on a scale from clinical observation=1.0 to basic research=4.0) based both on the words in their titles (Lewison and Paraje 2004) and the journals in which they have been published. A second macro analyses the addresses on the paper, and allocates fractional scores to each country represented, based on their presence. (For example, a paper with two Scottish addresses and one from Sweden would be classed as UK 0.67, SE 0.33.) Other macros parse the titles and journal names to classify papers within a major disease area, such as cancer, by their disease manifestation (e.g., breast cancer, leukaemia) and research domain or type (e.g., genetics, surgery). These classifications can assist a client to see any significant differences between papers cited in CPGs and those that are not.

The identification of papers from a Swedish collecting charity

Identification of the papers from the Swedish collecting charity is more challenging than the papers from the Scottish hospital. Although the WoS has recorded financial

acknowledgements since late 2008 in the Science Citation Index Expanded (Sirtes 2013; Begum and Lewison 2017; Alvarez-Bornstein et al. 2017), their recording in papers in the Social Sciences Citation Index only began in 2013. However, these are likely to have less influence on CPGs which tend to deal with the technical aspects of diagnosing and treating disease. The funding organisations are listed in a column headed FU, but there are two problems. One is that some funders have not supported the research being described, but have paid for other work by the authors. These acknowledgements need to be removed (Lewison and Sullivan 2015). The second problem is that the names of the funders are given in a large variety of different formats—in the case of the European Commission, several thousand. We therefore adopted the practice, starting in 1993, to give each funder a three part code (Jeschin et al. 1995; Dawson et al. 1998). An example is MRC-GA-UK, where MRC uniquely identifies the UK Medical Research Council, GA that it is a Government Agency (and so independent of ministerial control), and UK the country. This process of coding funders has turned out to be rather complex (Begum and Lewison 2017), but for any given funder our thesaurus (which now contains upwards of 150,000 different names) can show the various formats in which researchers have described it.

We are then able to assemble a search strategy that includes the various formats in which any given funder has been described, often in two different languages, and with various similar terms, such as “association”, “foundation”, “fund” or “group”, with and without the country name. This can then be applied to the WoS for the relevant years in order to identify the papers funded by the organisation. For collecting charities, which may have their own institutes or laboratories, we may also need to seek papers that have an address indicative of their support. For example, in the UK the large charity, Cancer Research UK, has labs in Glasgow (the Beatson Institute), Manchester (formerly the Paterson, since 2013 the Manchester Institute) and Therapeutic Discovery Laboratories in Babraham, Cambridge and London. Papers from these laboratories might, or might not, also have acknowledgements of support from Cancer Research UK.

Papers from the Swedish charity were identified and downloaded to file for the ten publication years, 2009–18. In total, they were more numerous than the papers from the Scottish hospital, and almost all of them were on one disease area, in which the charity specialised.

Diachronous citations

We took the opportunity afforded by the large numbers of citations in the clinical impact[®] database to explore the temporal variation of these two types of citations, viz. here and in the WoS. Diachronous citations are ones given subsequently to a cohort of papers published in a given calendar year. In principle, the number of these grows steadily with time, although in practice the fall-off in the number of citations with time in the WoS may allow an estimate of the total to be expected.

We took the three cohorts of papers from *all* Scottish hospitals in 2000, 2004 and 2008 in order to compare the temporal distributions of diachronous citations to them both by WoS papers and from the clinical impact[®] database. These comparisons were designed to show how the time distributions of references in the WoS and the clinical impact[®] database compared, and hence how long it would take for a given set of papers to be fairly evaluated on the basis of its footprint in the clinical impact[®] database. These three sets of papers were studied purely to show the temporal variation in the numbers

of diachronous citations that they could expect to receive, and so how long it would take for any other samples of research output to be fairly evaluated.

Results

Which papers were cited in clinical practice guidelines?

We generated two sets of papers, one from the Scottish hospital, and the second from the Swedish charity. In the 3 years, 2000, 2004 and 2008, there were 972 articles and reviews in the WoS from the hospital, and in the 10 years 2009–18, there were 1595 papers from the Swedish charity. Nearly all of them had PMIDs (946 and 1575 respectively), and of the remainder, most had DOIs (18 and 19). These identifiers were all matched against the clinical impact[®] database, and matches were found for 59 papers (146 total citations), and 34 papers (58 citations), respectively. Relative to the numbers of citable papers (with PMIDs), the percentages of cited ones were 6.1% and 2.2%, respectively. Two of the Scottish hospital papers were frequently cited in CPGs: one in 20 and another in 18 CPGs.

Because most countries' CPGs tend to over-cite papers from their fellow countrymen and countrywomen, it might be expected that most of the citations of these papers on CPGs would be from NICE and SIGN guidelines, and from Swedish ones, respectively. However, this was not the case. The Scottish individual hospital papers received 38 (26%) of their citations from the UK (25 from NICE and 13 from SIGN CPGs), but also many from Sweden (46), the US and Norway (18 from each). The Swedish charity papers received 30 of their citations (52%) from Swedish CPGs, with 9 each from Finland and from the UK.

Differences between cited and uncited papers from the Scottish hospital

The main difference was that the papers cited in CPGs were much more clinical than the uncited ones. The 49 cited papers had a mean research level, RL p , of 1.07 and the 703 uncited papers have RL $p=1.98$. This difference of 0.91 is large, and the difference between the numbers of clinical, basic, and "both" papers between the two groups is statistically significant with $p\sim 0.23\%$ (based on the Chi squared value from the Poisson distribution with three degrees of freedom). The cited papers were less international, with an average non-UK contribution of 0.095 per paper, compared with 0.142 for the uncited ones. The US contribution per paper was much greater on the latter set (0.035 for the uncited papers compared with 0.009 for the cited ones).

Papers that were classed as "clinical trials", because their titles contained words such as *controlled trial*, *placebo-controlled*, *randomised*, *randomized*, or *phase I, II, or III*, or *phase 1, 2, or 3*, were much more likely to be cited in CPGs. Of the 29 clinical trials papers, 10 were cited in CPGs (20.4%), whereas of the 870 other papers only 19 were so cited (2.2%). This is statistically significant with $p < 0.001$.

There were not enough CPG citations on papers from the Swedish charity to make a comparison between the ones cited by CPGs in the clinical impact[®] database, and the uncited ones, worthwhile.

Fig. 4 Diachronous clinical impact citations. The variation with time in the numbers of diachronous citations per paper for all Scottish hospitals papers in three separate cohorts cited by CPGs in the clinical impact® database

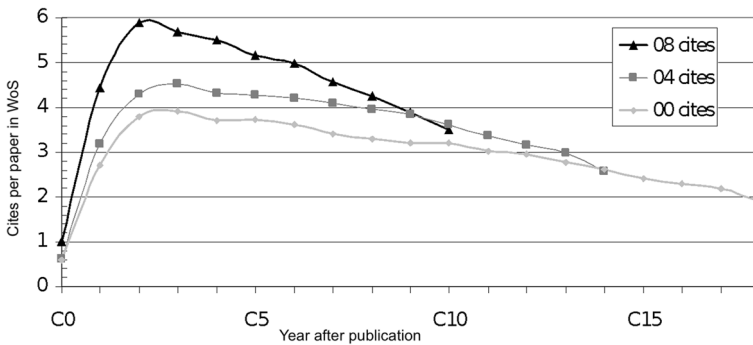
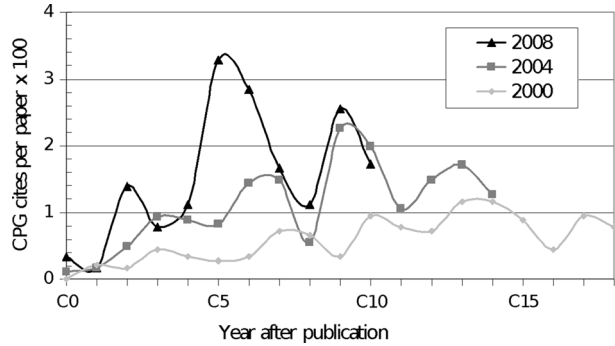


Fig. 5 Diachronous WoS citations. The variation with time in numbers of diachronous citations per paper for all Scottish hospitals papers in three separate cohorts cited by papers in the Web of Science (WoS) database

Diachronous citations

Our selections of papers from the Scottish hospitals and infirmaries in 2000, 2004 and 2008 and in the WoS numbered 1893, 1894 and 1841 respectively. The annual number of diachronous citations in the CPGs processed for the clinical impact® database are shown in Fig. 4, and those in the WoS in Fig. 5 for the years after paper publication. The numbers of citations are, of course, much higher in Fig. 5. The WoS citation curves show a clear peak between years 2 and 3, and a gradual decline thereafter, but the CPG curves are much less smooth, and show a peak citation year about 5 years after publication. It is therefore difficult to gauge when the numbers of CPG citations should be determined as a means to evaluate the importance of the research, but a rather wider citation window seems to be indicated.

Discussion

For the two sets of papers the number of citations in CPGs can appear quite low, but it is in the nature of CPG citations that they are few compared to citations in academic journals. In order to fully capture the impact on CPGs a longer citation window is most likely required. For the papers from the Scottish hospital there was an exception to the previous observation based on the WoS that papers are more cited in their local CPGs. It is likely that this is an artefact from the difference in number of guidelines, and therefore the number of possible citations, between the different countries. It is also true that different guidelines are updated in different ways and in different time intervals, which manifested in one of the highly cited papers from the Scottish hospital. Most of the highest cited papers from the Scottish hospital were in oncology, and were cited in several Guideline Provider Collections, but one of the most highly cited papers stood out. The highly cited paper was cited in a CPG that is updated almost yearly, with a new edition each time. The paper therefore got repeated citations by being cited in a frequently updated CPG. In the clinical impact[®] database this is made visible because the CPGs are marked with the successor-predecessor relation, and no manual review of the guideline texts was necessary. The marking of the CPGs is particularly useful for the detection of successor-predecessor relationships in CPGs in languages with which the reader is not familiar. In future studies this would be possible to manage with clinical impact[®] by selecting a timeframe in which different editions of the same CPGs published within the timeframe have their citations counted once. This would even out the differences in citation numbers that arise from differences in CPG updating frequencies between different CPG publishers.

Papers with a low RL tend to be more cited in CPGs than papers with a higher RL. This difference in citation scores is likely a deliberate result of the search strategies used by the authors of the CPGs, based on their clinical nature. Since papers with a low RL tend to be less cited than basic research in classic bibliometric studies based on academic citations, study of the CPG citations could provide valuable insight on the application and benefit of clinical research, and in particular which research types or domains are having a measurable effect on patient care.

In terms of evaluation, the data can be used to measure impact quantitatively on an organisational level, to report output impact on a research project or funding program, or to identify papers highly cited in CPGs for research impact case studies. The clinical impact[®] data could also be a tool for pharmaceutical companies, contract research organisations, and medical technology companies to track research on their products and services and their impact on Guidelines and recommendations on an international scale. Since these citations have not previously been collated into a searchable database, research performers and funders have been missing valuable information about the impact of clinical research papers, and the clinical impact[®] database is therefore a major advance in the bibliometrics of research evaluation.

The study has some limitations. At present the clinical impact[®] database is necessarily incomplete as it does not yet cover all countries that publish CPGs, which are increasing rapidly in number. As a result the clinical impact[®] database could at the time of this study not be used to examine the impact of the analyzed data sets on CPGs from countries outwith Europe and north America. Second, some of the CPG references are to documents other than articles and reviews in journals, and these may be research outputs that the database cannot evaluate. And third, it is clear that it may take many years for the clinical

impact of research papers on CPGs to become evident, whereas a citation window of (say) 5 years is normally long enough to gauge citation impact in the WoS.

The production of papers from the two organisations in this study will continue, and the existing sets of papers we analyzed will continue to garner citations as more CPGs are processed in clinical impact[®]. As CPGs from more countries are processed, and new matching processes for research output outside journal articles are developed, a follow-up study could examine and address some of the limitations of this study.

Acknowledgements We are grateful to Philip Roe of Evaluametrics Ltd for the Visual Basic Application programs (macros) used to convert the downloaded text files from the WoS into MS Excel spreadsheets, and for their subsequent analysis. At Minso Solutions AB we are happy to acknowledge the work of Tamara Rosini, main developer of the software for the clinical impact[®] database and the ref extractor tool.

Compliance with ethical standards

Conflict of interest This study has no external funding. Magnus Eriksson, Annika Billhult and Tommy Billhult are shareholders of Minso Solutions AB which provides the clinical impact[®] database.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



References

- Abramo, G., D'Angelo, C. A., & Soldatenkova, A. (2016). The dispersion of the citation distribution of top scientists' publications. *Scientometrics*, *109*(3), 1711–1724. <https://doi.org/10.1007/s11192-016-2143-7>.
- Alvarez-Bornstein, B., Morillo, F., & Bordons, M. (2017). Funding acknowledgments in the web of science: Completeness and accuracy of collected data. *Scientometrics*, *112*, 1793–1812. <https://doi.org/10.1007/s11192-017-2453-4>.
- Anastasiadis, A. D., de Albuquerque, M. P., de Albuquerque, M. P., & Mussi, D. B. (2010). Tsallis q-exponential describes the distribution of scientific citations—A new characterization of the impact. *Scientometrics*, *83*(1), 206–218. <https://doi.org/10.1007/s11192-009-0023-0>.
- Anon 1997 The development of national clinical guidelines and integrated care pathways and audit of practices against these standards: A collaborative project by all Scottish clinical genetics services. *Journal of Medical Genetics* *34*(S1): 515.
- Bakare, V., & Lewison, G. (2017). Country over-citation ratios. *Scientometrics*, *113*(2), 1199–1207.
- Begum, M., & Lewison, G. (2017). Web of science research funding information: Methodology for its use in analysis and evaluation. *Journal of Scientometric Research*, *6*(2), 65–73.
- Begum, M., Lewison, G., Wright, J. S. F., Pallari, E., & Sullivan, R. (2016). European non-communicable respiratory disease research, 2002–13: Bibliometric study of outputs and funding. *PLoS ONE*, *11*(4), e0154197. <https://doi.org/10.1371/journal.pone.0154197>.
- Bergman, E. M. L. (2012). Finding citations to social work literature: The relative benefits of using web of science, scopus or google scholar. *The Journal of Academic Librarianship*, *38*(6), 370–379. <https://doi.org/10.1016/j.acalib.2012.08.002>.
- Cluzeau, F. A., & Littlejohns, P. (1999). Apprising clinical practice guidelines in England and Wales: The development of a methodologic framework and its application to policy. *The Joint Commission Journal Quality Improvement*, *25*(10), 514–521.
- Dawson, G., Lucocq, B., Cottrell, R., & Lewison, G. (1998). Mapping the landscape: National Biomedical Research Outputs 1988–95 London: The Wellcome Trust. Policy Report no. 9. ISBN: 1869835956.
- Garfield, E. (1955). Citation indexes for science—New dimension in documentation through association of ideas. *Science*, *122*(3159), 108–111. <https://doi.org/10.1126/science.122.3159.108>.

- Garfield, E. (1972). Citation analysis as a tool in journal evaluation—Journals can be ranked by frequency and impact of citations for science policy studies. *Science*, 178(4060), 471–479. <https://doi.org/10.1126/science.178.4060.471>.
- Grant, J. (1999). Evaluating the outcomes of biomedical research on healthcare. *Research Evaluation*, 8(1), 33–38.
- Grant, J., Cottrell, R., Cluzeau, F., & Fawcett, G. (2000). Evaluating, “payback” on biomedical research from papers cited in clinical guidelines: Applied bibliometric study. *BMJ*, 320(7242), 1107–1111. <https://doi.org/10.1136/bmj.320.7242.1107>.
- Ingwersen, P., Larsen, B., & Wormell, I. (2000). Applying diachronic citation analysis to research program evaluations. In *Web of Knowledge - A Festschrift in Honor of Eugene Garfield : ASIST Monograph Series* (pp. 373–387).
- Jeschin, D., Lewison, G., & Anderson, J. (1995). A bibliometric database for tracking acknowledgements of research funding. In M. E. D. Koenig & A. Bookstein (Eds.), *proceedings of the fifth biennial conference of the international society for scientometrics and informetrics*. Learned Information Inc., Medford NJ, USA. ISBN: 1-57387-010-2.
- Kostoff, R. N. (1998). The use and misuse of citation analysis in research evaluation. *Scientometrics*, 43(1), 27–43.
- Lewison, G. (2007). The references on UK cancer clinical guidelines. In D. Torres-Salinas, & H. F. Moed (Eds.), *proceedings of the 11th biennial conference of the international society for scientometrics and informetrics* (pp. 489–498). Madrid, Spain.
- Lewison, G., & Dawson, G. (1998). The effect of funding on the outputs of biomedical research. *Scientometrics*, 41(1–2), 17–27.
- Lewison, G., & Paraje, G. (2004). The classification of biomedical journals by research level. *Scientometrics*, 60(2), 145–157. <https://doi.org/10.1023/b:scie.0000027677.79173.b8>.
- Lewison, G., & Sullivan, R. (2008). The impact of cancer research: How publications influence UK cancer clinical guidelines. *British Journal of Cancer*, 98(12), 1944–1950. <https://doi.org/10.1038/sj.bjc.6604405>.
- Lewison, G., & Sullivan, R. (2015). Conflicts of interest statements on biomedical papers. *Scientometrics*, 102, 2151–2159. <https://doi.org/10.1007/s11192-014-1507-0>.
- Lewison, G., & Wilcox-Jay, K. (2003). Getting biomedical research into practice: the citations from UK clinical guidelines. In G. Jiang, R. Rousseau, & Y. Wu (Eds.), *proceedings of the ninth biennial conference of the international society for scientometrics and informetrics* (pp. 152–160). Beijing, China.
- Leydesdorff, L. (2008). Caveats for the use of citation indicators in research and journal evaluations. *Journal of the American Society for Information Science and Technology*, 59(2), 278–287.
- Leydesdorff, L., de Moya-Anegón, F., & de Nooy, W. (2016). Aggregated journal–journal citation relations in scopus and web of science matched and compared in terms of networks, maps, and interactive overlays. *Journal of the Association for Information Science and Technology*, 67(9), 2194–2211. <https://doi.org/10.1002/asi.23372>.
- Manafy, M. (2004). Scopus: Elsevier expands the scope of research. *Econtent–Digital Content Strategies and Resources*, 27(11), 9–11.
- Martin-Martin, A., Orduna-Malea, E., Thelwall, M., & Lopez-Cozar, E. D. (2018). Google scholar, web of science, and scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160–1177. <https://doi.org/10.1016/j.joi.2018.09.002>.
- Meho, L. I., & Sugimoto, C. R. (2009). Assessing the scholarly impact of information studies: A tale of two citation databases—Scopus and web of science. *Journal of the American Society Information Science and Technology*, 60(12), 2499–2508. <https://doi.org/10.1002/asi.21165>.
- 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 Information Science Technology*, 58(13), 2105–2125. <https://doi.org/10.1002/asi.20677>.
- Miller, J. (2002). The Scottish intercollegiate guidelines network (SIGN). *The British Journal Diabetes Vascular Disease*, 2(1), 47–49.
- Moed, H. F. (2009). New developments in the use of citation analysis in research evaluation. *Archivum Immunologiae et Therapiae Experimentalis*, 57(1), 13–18.
- Pallari, E., Fox, A. W., & Lewison, G. (2018b). Differential research impact in cancer practice guidelines evidence base: Lessons from ESMO NICE and SIGN. *ESMO OPEN*, 3(1), UNSP e000258. <https://doi.org/10.1136/esmoopen-2017-000258>.
- Pallari, E., & Lewison, G. (2019). 22 How Biomedical Research Can Inform Both Clinicians and the General Public. In W. Glänzel, et al. (Eds.), *Springer handbook of science and technology indicators* (pp. 583–609). Cham: Springer.

- Pallari, R., Lewison, G., Ciani, O., Tarricone, R., Sommariva, S., Begum, M., et al. (2018a). The impacts of diabetes research from 31 European Countries in 2002 to 2013. *Research Evaluation*, 27(3), 270–282. <https://doi.org/10.1093/reseval/rvy006>.
- Roe, P. E., Wentworth, A., Sullivan, R., & Lewison, G. (2010). The anatomy of citations to UK cancer research papers. In *proceedings of the 11th conference on S&T indicators* (pp. 225–226). Leiden.
- Ruiz-Castillo, J. (2013). The role of statistics in establishing the similarity of citation distributions in a static and a dynamic context. *Scientometrics*, 96(1), 173–181. <https://doi.org/10.1007/s11192-013-0954-3>.
- Sirtes, D. (2013). Funding acknowledgements for the German Research Foundation (DFG). The dirty data of the web of science database and how to clean it up. In J. Gerroiz, E. Schiebel, C. Gumpenberger, M. Horlesberger, H. Moed (Eds.), *proceedings of the 14th biennial conference of the international society for scientometrics and informetrics*, (pp. 784–795). Vienna.
- Tattersall, A., & Carroll, C. (2018). What can altmetric.com tell us about policy citations of research an analysis of altmetric.com data for research articles from the university of sheffield. *Frontiers in Research Metrics Analytics*, 2, 9.
- van Eck, N. J., Waltman, L., van Raan, A. F. J., Klautz, R. J. M., & Peul, W. C. (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. *PLoS ONE*, 8(4), e62395.
- van Raan, A. F. J. (2001). Competition amongst scientists for publication status: Toward a model of scientific publication and citation distributions. *Scientometrics*, 51(1), 347–357. <https://doi.org/10.1023/A:1010501820393>.
- Wailoo, A., Roberts, J., & Brazier, J. (2004). Efficiency, equity, and NICE clinical guidelines—Clinical guidelines need a broader view than just the clinical. *BMJ*, 328(7439), 536–537. <https://doi.org/10.1136/bmj.328.7439.536>.

Affiliations

Magnus Eriksson¹  · Annika Billhult^{1,2} · Tommy Billhult¹ · Elena Pallari^{3,6}  · Grant Lewison^{4,5}

Annika Billhult
annika.billhult@minso.se

Tommy Billhult
tommy.billhult@minso.se

Elena Pallari
elena.pallari@kcl.ac.uk

Grant Lewison
grant.lewison@kcl.ac.uk

¹ Minso Solutions AB, Sven Eriksonsplatsen 4, Boras 503 38, Sweden

² University of Boras, Allégatan 1, Boras 503 32, Sweden

³ Centre for Implementation Science, Institute of Psychiatry, Psychology & Neuroscience (IoPPN), Health Service and Population Research Department, King's College London, London SE5 8AF, UK

⁴ School of Cancer and Pharmaceutical Sciences, Division of Cancer Studies, King's College London, Guy's Hospital, London SE1 9RT, UK

⁵ Evaluametrics Ltd, 157 Verulam Road, St Albans AL3 4DW, UK

⁶ MRC Clinical Trials and Methodology, University College London, 90 High Holborn, London WC1V 6LJ, UK