



Searching the scientific literature: Implications for quantitative and qualitative reviews

Yelena P. Wu¹, Brandon S. Aylward¹, Michael C. Roberts^{*}, Spencer C. Evans

Clinical Child Psychology Program, University of Kansas, 2010 Dole Human Development Center, 1000 Sunnyside Avenue, Lawrence, KS 66045, USA

HIGHLIGHTS

- ▶ Literature search results vary across engines, rendering different conclusions.
- ▶ Multiple search engines should be used with strategies tailored to each system.
- ▶ Standardized reporting of search results, particularly for reviews, is recommended.

ARTICLE INFO

Article history:

Received 22 February 2012
 Received in revised form 14 June 2012
 Accepted 26 June 2012
 Available online 7 July 2012

Keywords:

Literature review
 Search strategy
 Meta-analysis
 Qualitative review
 Information retrieval

ABSTRACT

Literature reviews are an essential step in the research process and are included in all empirical and review articles. Electronic databases are commonly used to gather this literature. However, several factors can affect the extent to which relevant articles are retrieved, influencing future research and conclusions drawn. The current project examined articles obtained by comparable search strategies in two electronic archives using an exemplar search to illustrate factors that authors should consider when designing their own search strategies. Specifically, literature searches were conducted in PsycINFO and PubMed targeting review articles on two exemplar disorders (bipolar disorder and attention deficit/hyperactivity disorder) and issues of classification and/or differential diagnosis. Articles were coded for relevance and characteristics of article content. The two search engines yielded significantly different proportions of relevant articles overall and by disorder. Keywords differed across search engines for the relevant articles identified. Based on these results, it is recommended that when gathering literature for review papers, multiple search engines should be used, and search syntax and strategies be tailored to the unique capabilities of particular engines. For meta-analyses and systematic reviews, authors may consider reporting the extent to which different archives or sources yielded relevant articles for their particular review.

© 2012 Elsevier Ltd. All rights reserved.

Contents

1. Introduction	553
2. Method	554
2.1. Procedure	554
2.1.1. Coding system	555
2.1.2. Coder training and reliability	555
3. Results	555
4. Discussion	556
Acknowledgements	557
References	557

* Corresponding author at: Clinical Child Psychology Program, University of Kansas, 2010 Dole Human Development, 1000 Sunnyside Avenue, Lawrence, KS 66047, USA. Tel.: +1 785 864 4226; fax: +1 785 864 5024.

E-mail address: mroberts@ku.edu (M.C. Roberts).

¹ Present address: Cincinnati Children's Hospital Medical Center; Behavioral Medicine & Clinical Psychology, MLC 7039, 3333 Burnet Avenue, Cincinnati, OH 45229.

1. Introduction

All empirical publications and review papers, including those in *Clinical Psychology Review* discuss and rely on previously published work. Researchers and clinicians alike routinely use search engines

such as PsycINFO and PubMed to obtain scholarly articles that inform reviews of the literature (e.g., meta-analyses, introduction sections of articles), interpretation of results, future research directions, and clinical practice. Among the various methods that can be used to identify relevant literature, web-based search engines (e.g., PubMed) are perhaps the quickest and most accessible (see Arnold, Bender, & Brown, 2006; Falagas, Pitsoini, Malietzis, & Pappas, 2008, for reviews). In fact, the vast majority of reviews and meta-analyses typically use electronic archives as a key source for identifying relevant literature, in addition to other methods of identifying relevant literature (Lipsey & Wilson, 2001). In the years 2007–2009, for example, 65 meta-analyses were published in three representative journals (*Health Psychology*, *Psychological Bulletin*, *Journal of Consulting and Clinical Psychology*). Of these articles, the search engines PubMed/Medline and/or PsycINFO were used in almost all ($n = 61$; 93.8%) studies to obtain relevant articles.

Although searching of electronic databases has many benefits, including being able to search large archives in a short time, there are a variety of challenges that arise when using these engines. Importantly, these challenges may affect the results of searches if users are not aware of potential differences between search engines and the archives with which they are connected. First, the archives that search engines access may differ in the journals that are included and in the proportion of literature that is relevant for particular research or clinical areas (Gavel & Iselid, 2008; Lohonen, Isohanni, Nieminen, & Miettunen, 2010; Watson & Richardson, 1999a, 1999b). For instance, McDonald, Taylor, and Adams (1999) demonstrated that commonly used electronic databases differ in the psychiatry journals indexed, such that approximately one-third of psychiatry journals are indexed in only one database. Second, journals and their respective fields differ in the terminology used to represent the same idea (e.g., “pediatric” versus “child”) and using the same search terms across engines may lead to disparate results simply because one term is preferred over another within particular literatures. Third, search engines can have different structures, such as unique search capabilities, which can lead to different results. In PsycINFO, for example, users can filter search results by “quantitative study” whereas users cannot apply this filter in PubMed. Fourth, search engines and archives use different index terms that can affect search results. Two of the most commonly used resources in clinical psychology are PubMed and PsycINFO. PubMed, a search engine that draws from MEDLINE in addition to other sources uses the U.S. National Library of Medicine (2011) Medical Subject Headings, (MeSH) which are assigned to individual articles. In contrast, PsycINFO, an archive administered by the American Psychological Association (APA), uses index terms from the *Thesaurus of Psychological Index Terms* (Tuleya, 2007). Although some index terms are common to both MeSH and the *Thesaurus of Psychological Index Terms*, there are numerous terms unique to each system. In addition, there are terms that describe the same construct but that differ between systems (e.g., “Attention Deficit Hyperactivity Disorder” [ADHD] in the *Thesaurus* and “Hyperkinetic Disorder” in MeSH). Indeed, these differences between databases and their respective search engines lead to varying results when parallel searches are run (Arnold et al., 2006; Brettle & Long, 2001; Conn et al., 2003). Due to these differences in archive content, terminology, index terms, and search capabilities between particular databases, some studies have examined ways to tailor and optimize electronic search strategies (Eady, Wilczynski, & Haynes, 2008; Jenuwine & Floyd, 2004).

In sum, the existing literature indicates that search results can differ between databases and that search strategies should account for the capabilities and structures of individual search engines (Kelly & St. Pierre-Hansen, 2008). However, this literature has focused primarily on the search results obtained when the same search terms are used across different databases, or the results of tailored search strategies within single databases (e.g., PsycINFO). Few studies have examined the extent to which search results differ when search strategies are

tailored to the index terms and search capabilities of particular engines. This issue is an important concern to authors of systematic reviews and meta-analyses who want to ensure that the whole literature is comprehensively searched so that accurate conclusions and interpretations can be drawn. Researchers may overlook or underestimate the extent to which archives and search engines function differently, use different algorithms, and include different parts of related literatures. However, it is unknown to what extent the two major search engines for clinical psychology literature (i.e., PsycINFO, PubMed) provide unique or duplicate returns. In addition, no studies to our knowledge have examined potential differences in terms of the content of articles (e.g., whether articles address biological correlates, treatment issues) retrieved from different archives. An examination of these issues would inform both the search strategies that authors use and understanding of potential differences in article content between archives. Furthermore, these issues are particularly relevant for intersecting research fields (e.g., medicine and psychology) and fields contributing to the application of quantitative methods of research in psychology (e.g., mathematics, statistics).

Consistent with prior literature examining broad methodological issues affecting study design and result presentation (e.g., Kratochwill & Levin, 2009; Lane & Sandor, 2009), the primary aim of this review was to investigate a methodology (i.e., literature searching) which can affect research design, results (e.g., meta-analytic results), and interpretation of results. Specifically, we examined potential differences in literature search results when comparing search strategies and archives. Using exemplar topics, comparable searches were conducted in PsycINFO and PubMed to identify systematic differences in the results obtained by the two search engines. The results were examined for (a) differences in the relevance of the articles returned with regard to the topic of investigation (see *Procedure* section), (b) differences in select features of article content, and (c) the extent to which the search engines returned unique versus duplicate results.

Previously published studies on literature search methodology have typically focused on specific content areas (e.g., health care worker burnout, rehabilitation services for individuals with severe mental illness; Arnold et al., 2006; Brettle & Long, 2001) and have used the results of these topic-focused searches to illustrate larger literature search issues. Consistent with this methodology, we chose a particular content area on which to focus in our searches. Specifically, we selected ADHD and bipolar disorder because of the growing literature regarding diagnostic and comorbidity issues for these disorders (e.g., Carlson, 1998; Geller et al., 2002), as well as the substantial literatures addressing broader facets of each disorder. These disorders are also of immediate relevance to the current revision process for the International Classification of Diseases (ICD) and the Diagnostic and Statistical Manual (DSM; First, 2010; Reed, 2010). We therefore focused our searches on ADHD and bipolar disorder and issues of classification and diagnosis. In addition, we chose to focus our literature searches on review and meta-analytic articles, because researchers and practitioners alike rely on reviews to provide summaries of the extant literature and to identify primary sources which may be of interest (Lipsey & Wilson, 2001). Thus, similar to the previously published literature in this area, although our search speaks to a specialty question for a particular domain, the aim of our analysis is to produce information with search implications for other domains and subfields within clinical psychology.

2. Method

2.1. Procedure

Literature searches targeting review articles on bipolar disorder and ADHD and issues of classification and/or differential diagnosis, published in the five years between 2004 and 2008, were conducted in PsycINFO and PubMed on the same day. In order to focus the searches on review articles, limits were placed so that articles from PsycINFO

were retrieved if they were literature reviews, meta-analyses, or systematic reviews, and articles from PubMed were retrieved if they were meta-analyses or reviews. The limits varied due to fundamental differences in the options available in each search engine. The search terms for each engine also differed due to the unique structures of the search engines (see Table 1). Subject terms for the PsycINFO search were selected based on manual review of the PsycINFO thesaurus (Tuleya, 2007) and terms used in the PubMed search were selected based on review of the MeSH headings (National Library of Medicine, 2010). Search results (citations and abstracts) were downloaded into EndNote citation management software.

2.1.1. Coding system

Coders read each abstract to determine whether the article met the current study's criteria for relevance. Articles were coded as relevant if they: (a) were a meta-analysis or a review paper, (b) focused primarily on classification or diagnosis issues, and (c) focused on bipolar disorder and/or ADHD. Relevant articles were coded for their inclusion of several content features. Specifically, the content of relevant articles were coded using six items assessing for age coverage (i.e., children/adolescents [< 18 years old], adults [18 years and older], or both), whether they discussed cultural and sociodemographic issues (e.g., ethnicity, gender, language), reliability and validity of classification/diagnostic systems, biological correlates of the disorders (e.g., genetic variations, neuroanatomical markers), psychosocial correlates of the disorders (e.g., social skills, family stress) and whether treatment/intervention was a primary component of the article (e.g., randomized controlled trials).² Abstracts were coded because oftentimes decisions of inclusion/exclusion for reviews and meta-analyses are based only on information available in the abstract.

2.1.2. Coder training and reliability

Articles were coded by four undergraduate research assistants (two juniors, two seniors) majoring in psychology. Initial training included review of the coding system and practice coding on a subset of articles ($n = 16$) until minimal discrepancies were obtained. Once coders began coding independently, 10% of the articles were coded by two or more individuals, including a master's-level graduate student. Discrepancies between coders were resolved by discussion and, if needed, by consultation with a third party (i.e., the third author). Before discussing any discrepancies, the reliability among coders across all categories (i.e., relevancy of articles and codes related to article content) was good (Cohen's $\kappa = 0.82$). After resolving discrepancies, coders agreed on 100% of the codes. The final, agreed-upon codes were used for analyses.

3. Results

The literature searches yielded 90 unique articles in PsycINFO and 739 in PubMed. In addition, 35 articles were identified in both PsycINFO and PubMed. Of the total, 11 articles were not codable because the abstracts were not in English or an abstract was not available (e.g., the article was a commentary which did not have an abstract). While PubMed returned a greater total number of relevant articles ($n = 137$) than PsycINFO ($n = 37$), there was a significant difference between the proportion of relevant articles identified uniquely by the two search engines ($\chi^2 = 23.3$, $df = 1$, $p < .001$). Specifically, of the articles identified only in PsycINFO, 37 (41.1%) were relevant. In contrast, 137 (18.5%) articles identified only in PubMed were categorized as relevant. Also, 19 (54.3%) of the articles identified in both databases were coded as relevant. See Table 2 for the characteristics of the relevant articles by database.

² This coding system was initially developed by the World Health Organization for analyzing articles relevant to the revision of the International Classification of Diseases under Prapat Sharan and Geoffrey M. Reed, and was tailored for the purposes of the current study.

Table 1
Search Terms.

PsycINFO ^a	PubMed ^b
attention deficit hyperactivity disorder	attention deficit hyperactivity disorder
hyperkinetic disorders	hyperkinetic disorders
bipolar disorder	bipolar disorder
bipolar affective disorder	bipolar affective disorder
taxonomies	classification ^c
psychodiagnostic typologies	diagnosis ^c
diagnosis	diagnosis, differential ^c
psychodiagnosis	
differential diagnosis	

^a Limits placed on the searches resulted in articles being retrieved if: (a) The specified search terms were included in article titles or abstracts, and (b) the article was a literature review, meta-analysis, or systematic review.

^b Limits placed on the searches resulted in articles being retrieved if: (a) The specified search terms were included in article titles or abstracts, and (b) the article was a meta-analysis or review.

^c Terms were entered as MeSH Subheadings.

Out of the relevant articles identified, PubMed and PsycINFO yielded significantly different proportions of articles on ADHD and on bipolar disorder ($\chi^2 = 4.4$, $df = 1$, $p < .05$). Specifically, 73.0% ($n = 27$) of the relevant articles in PsycINFO were on bipolar disorder versus 53.3% ($n = 73$) in PubMed. In contrast, 46.0% ($n = 63$) of relevant articles in PubMed addressed ADHD while 27.0% ($n = 10$) of articles in PsycINFO addressed ADHD. The two search engines also yielded significantly different proportions of articles providing information about the age of study participants ($\chi^2 = 8.1$, $df = 1$, $p < .05$; PubMed, 61.3%, $n = 84$; PsycINFO, 35.1%, $n = 13$). Of the articles in which age information was provided, PsycINFO and PubMed had similar proportions of relevant articles focused on children (ages 17 and younger) and adults (18 years or older). There were no other significant differences between the proportion of relevant articles in PubMed only versus PsycINFO only versus articles identified in both PubMed and PsycINFO in terms of content (i.e., whether cultural or sociodemographic issues, reliability/validity of diagnosis, biological correlates, psychosocial correlates, or treatments/interventions were addressed).

It should be noted that these analyses are based on the proportion of relevant articles obtained in each engine, which is calculated based on the total number of unique articles in each search engine. To compare the search engines based on absolute numbers of articles obtained, PubMed exceeded PsycINFO on the number of relevant articles for each disorder and on each content feature (Table 2).

Table 2
Number of relevant articles by search engine and article content.

	Search engine							
	PsycINFO only		PubMed only		Duplicates		Total	
	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)	<i>n</i>	(%)
Disorder addressed								
BD only ^a	27	(73.0)	73	(53.3)	13	(68.4)	113	(58.6)
ADHD only	10	(27.0)	63	(46.0)	5	(26.3)	78	(40.4)
BD and ADHD	0	(0.0)	1	(0.7)	1	(5.3)	2	(1.0)
Total	37	(100.0)	137	(100.0)	19	(100.0)	193	(100.0)
Article content								
Age information provided	13	(35.1)	84	(61.3)	8	(42.1)	105	(54.4)
Cultural/Sociodemographic	6	(16.2)	25	(18.2)	3	(15.8)	34	(17.6)
Reliability/Validity	8	(21.6)	41	(29.9)	9	(47.4)	58	(30.1)
Biological correlates	8	(21.6)	44	(32.1)	4	(21.1)	56	(29.0)
Psychosocial correlates	15	(40.5)	53	(38.7)	6	(31.6)	74	(38.3)
Treatment issues	19	(51.4)	79	(57.7)	11	(57.9)	109	(56.5)

Note. BD = Bipolar disorder; ADHD = Attention-Deficit/Hyperactivity Disorder; Duplicates = Articles identified in both the PsycINFO and PubMed searches.

A qualitative review of the keywords attached to the relevant articles identified uniquely in PubMed and PsycINFO suggested that certain keywords are used in both databases but that others are used more frequently in one database. For example, both databases included keywords addressing evidence-based practice; however, articles in PsycINFO only used the term “evidence-based practice” ($n=5$) whereas those in PubMed primarily used “evidence-based medicine” ($n=7$ versus $n=1$ using “evidence-based practice”). Other differences between keywords included that PubMed articles used keywords identifying age-related factors (e.g., adolescent, adult, age of onset, age factors) more frequently than PsycINFO articles did (175 occurrences in PubMed versus 12 occurrences in PsycINFO). In addition, PubMed articles used medication keywords (e.g., anticonvulsants, drug therapy) and gene-related keywords (e.g., genetic predisposition to disease, genetics) to a greater extent than those in PsycINFO (Medication keywords: 149 occurrences in PubMed versus 17 occurrences in PsycINFO; gene-related keywords: 28 occurrences in PubMed versus 0 occurrences in PsycINFO). PubMed also included keywords combined with qualifiers from the MeSH database (e.g., bipolar disorder/classification/diagnosis).

4. Discussion

The results of this analysis hold several implications relevant to literature searches, including searches for review papers (e.g., meta-analyses, systematic reviews, qualitative reviews). First, the findings highlight that search engines such as PubMed and PsycINFO may produce different numbers of articles for particular searches, even if search terms have been tailored for the respective archives. Encouragingly, there is some overlap in search results between the search engines examined in the current study, which is consistent with prior studies (e.g., [Brettle & Long, 2001](#); [McDonald et al., 1999](#)). However, there was not as much overlap as one might expect, given that search terms were adapted for the unique characteristics of each search engine and database. These results underscore the importance of conducting literature searches in multiple search engines. In order to gain a comprehensive picture of the literature, investigators should utilize both PubMed and PsycINFO. Although these recommendations might seem self-evident, no previous work appears to have provided empirical support for them.

Clinicians and research investigators may also want to search other archival sources (e.g., Scopus, Web of Science, EBSCO, Megafire, Google Scholar, EMBASE, CINAHL) in order to capture the full extent of the potentially relevant literature; however, the trade-offs between different search engine's sensitivity and specificity should be weighed, depending on the goal of the literature search. For example, high-sensitivity search engines will return a greater number of total articles, both relevant and irrelevant (e.g., PubMed in the current analysis), thereby affording one a greater degree of exhaustiveness in the searches, which may be preferred for authors of meta-analyses and systematic reviews.

Second, the results suggest that some aspects of article content may differ among articles uniquely indexed in particular archives. For example, the search engines used in the current analysis tended to index relevant articles on disorders at different rates. Because differences in content may contribute unique information to the conclusions drawn for review and other empirical articles, users will likely want to make every effort to identify the uniquely relevant literature in multiple databases. Furthermore, the results of the exemplar searches indicate that across both search engines, certain article content areas may be covered more comprehensively than others. For instance, there were fewer articles addressing cultural or sociodemographic issues compared with psychosocial correlates of the disorders.

Third, before conducting searches in multiple databases, authors are encouraged to gain an understanding of the indexing systems or controlled vocabulary and search capabilities (e.g., filters) that these engines utilize ([American Psychological Association, 2012](#); [Arnold et al., 2006](#); [Falagas et al., 2008](#); [U.S. National Library of Medicine, 2011](#)). For instance, there are multiple ways to create search strategies within and across

database search engines, such as by using different subject terms. Authors will also want to acquaint themselves with the terminology and keywords or index terms that are likely to be used in different engines (e.g., terminology that may differ between subfields or journals). One method of identifying keywords unique to a particular database is to examine keywords used within relevant articles. Consultation with a research librarian will likely be helpful when designing search strategies tailored to particular search engines ([Conn et al., 2003](#); [DeLuca et al., 2008](#); [Lipsey & Wilson, 2001](#)).

To assist authors in obtaining the maximum number of relevant search results, improvements in indexing will also be important. A number of authors have outlined problems associated with existing methods of indexing, including the length of time between when articles are published and when they are indexed for different archives and search engines (e.g., [Conn et al., 2003](#); [Lohonen et al., 2010](#)). For example, PubMed includes articles in its archive before index terms are assigned, whereas PsycINFO assigns index terms to articles as they are added to the archive. In addition, improvements to indexing systems may include ensuring that high-use search terms (e.g., randomized controlled trials, systematic reviews) are used as index terms, that index terms are applied consistently and accurately, and that authors are encouraged to choose article keywords that are consistent with the indexing system used by certain databases ([Eady et al., 2008](#)).

There are several limitations of the current analysis that should be noted. The results of one particular literature search were outlined that serves as an illustration of the issues that can be encountered during the literature search process. In addition, the current review focused on meta-analyses and review papers. Results (e.g., percent overlap between articles indexed in databases) may differ with the use of other subject terms or strategies, search engines, literature archives, or with the inclusion of other publication types (e.g., book chapters, empirical articles other than reviews or meta-analyses). In order to maximize researchers' and clinicians' ability to identify the maximum number of relevant references for their needs, future research might investigate how other archives and their search engines, and controlled vocabulary or subject terms identify unique and overlapping search results. Furthermore, examining the degree of overlap of search results across archives may lead to more useful conclusions about how successful meta-analysts and reviewers are in ascertaining all relevant articles to their searches. The capture-recapture method used in other fields (e.g., epidemiology) may lend valuable insights into these endeavors ([Chao, Tsay, Lin, Shau, & Chao, 2001](#)). Finally, we coded the abstracts of relevant articles in order to more closely parallel 'real-world' searches where investigators read article abstracts to determine the relevance of articles to their purposes; however, future studies could also code articles based on the full text and examine article content of non-relevant articles.

In the case of meta-analyses or systematic reviews, where the literature search process is the primary method of “gathering data,” standardized reporting of literature search results would lead to a better understanding of different search engines and their unique strengths for different bodies of literature. Currently, authors of review articles typically note which search engines they used to identify relevant literature, but rarely indicate the proportion of the literature identified by the different search engines. In a manual review of several journals (*Psychological Bulletin*, *Journal of Consulting and Clinical Psychology*, *Health Psychology*; years 2007 through 2009), no meta-analytic articles described the number or percent of articles that came from each search engine or which articles may have been identified by more than one search engine. Publication guidelines for review papers do not yet recommend that authors describe the number of relevant articles identified in each archive or the overlap in search results between search engines (e.g., [Meta-Analysis Reporting Standards \[MARS\], American Psychological Association, 2010](#); [Preferred Reporting Items for Systematic Reviews and Meta-Analyses \(PRISMA\) statement, Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009](#)). However, in order to ensure that literature searches for review articles are as explicitly described (to

ensure replicable results are obtained) as the methods used in empirical studies, review article authors may consider reporting the number of relevant articles identified in each archive and through other methods (e.g., hand-searching).

In conclusion, the current results indicate that, for conducting scientifically based literature reviews and meta-analyses, multiple searches using different archives and search engines are likely necessary for a credible search strategy. For particular literatures, search engines may yield different proportions of relevant articles. In addition, tailoring search strategies to meet the capabilities and indexing systems of different search engines is an essential step in the literature identification process.

Acknowledgements

The authors acknowledge the work of Pratap Sharan and Geoffrey Reed, who developed an earlier version of the coding system used in the current study, and the assistance of Alex Kish, Erin Anderson, and Emilie Durgan with data collection and entry.

References

- American Psychological Association (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.
- American Psychological Association (2012). PsycINFO. Retrieved from <http://www.apa.org/pubs/databases/psycinfo/index.aspx>
- Arnold, S. J., Bender, V. F., & Brown, S. A. (2006). A review and comparison of psychology-related electronic resources. *Journal of Electronic Resources in Medical Libraries*, 3, 61–80. http://dx.doi.org/10.1300/J383v03n03_04.
- Brettell, A. J., & Long, A. F. (2001). Comparison of bibliographic databases for information on the rehabilitation of people with severe mental illness. *Bulletin of the Medical Library Association*, 89, 353–362 (Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/journals/72/>)
- Carlson, G. A. (1998). Mania and ADHD: Comorbidity or confusion. *Journal of Affective Disorders*, 51, 177–187. [http://dx.doi.org/10.1016/S0165-0327\(98\)00179-7](http://dx.doi.org/10.1016/S0165-0327(98)00179-7).
- Chao, A., Tsay, P. K., Lin, S. -H., Shau, W. -Y., & Chao, D. -Y. (2001). The application of capture-recapture models to epidemiological data. *Statistics in Medicine*, 20, 3123–3157. <http://dx.doi.org/10.1002/sim.996>.
- Conn, V. S., Isaramalai, S., Rath, S., Jantarakupt, P., Wadhawan, R., & Dash, Y. (2003). Beyond MEDLINE for literature searches. *Journal of Nursing Scholarship*, 35, 177–182. <http://dx.doi.org/10.1111/j.1547-5069.2003.00177.x>.
- DeLuca, J. B., Mullins, M. M., Lyles, C. M., Crepaz, N., Kay, L., & Thadiparthi, S. (2008). Developing a comprehensive search strategy for evidence based systematic reviews. *Evidence Based Library and Information Practice*, 3, 3–32 (Retrieved from <http://ejournals.library.ualberta.ca/index.php/EBLIP>)
- Eady, A. M., Wilczynski, N. L., & Haynes, R. B. (2008). PsycINFO search strategies identified methodologically sound therapy studies and review articles for use by clinicians and researchers. *Journal of Clinical Epidemiology*, 61, 34–40. <http://dx.doi.org/10.1016/j.jclinepi.2006.09.016>.
- Falagas, M. E., Pitsouni, E. L., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *The Journal of the Federation of American Societies for Experimental Biology*, 22, 338–342. <http://dx.doi.org/10.1096/fj.07-9492LSF>.
- First, M. B. (2010). Clinical utility in the revision of the Diagnostic and Statistical Manual of Mental Disorders (DSM). *Professional Psychology: Research and Practice*, 41, 465–473. <http://dx.doi.org/10.1037/a0021511>.
- Gavel, Y., & Iselid, L. (2008). Web of Science and Scopus: A journal title overlap study. *Online Information Review*, 32, 8–21. <http://dx.doi.org/10.1108/14684520810865958>.
- Geller, B., Zimmerman, B., Williams, M., DelBello, M. P., Bolhofner, K., Craney, J. L., et al. (2002). DSM-IV mania symptoms in a prepubertal and early adolescent bipolar disorder phenotype compared to attention-deficit hyperactive and normal controls. *Journal of Child and Adolescent Psychopharmacology*, 12, 11–25. <http://dx.doi.org/10.1089/10445460252943533>.
- Jenuwine, E. S., & Floyd, J. A. (2004). Comparison of Medical Subject Headings and text-word searches in MEDLINE to retrieve studies on sleep in healthy individuals. *Journal of the Medical Library Association*, 92, 349–353 (Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/issues/195450/>)
- Kelly, L., & St. Pierre-Hansen, N. (2008). So many databases, such little clarity: Searching the literature for the topic aboriginal. *Canadian Family Physician*, 54, 1572–1573 (Retrieved from <http://www.cfp.ca/>)
- Kratochwill, T. R., & Levin, J. R. (2009). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods*, 15, 124–144. <http://dx.doi.org/10.1037/a0017736>.
- Lane, D. M., & Sandor, A. (2009). Designing better graphs by including distributional information and integrating numbers, words, and images. *Psychological Methods*, 14, 239–257. <http://dx.doi.org/10.1037/a0016620>.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.
- Lohonen, J., Isohanni, M., Nieminen, P., & Miettunen, J. (2010). Coverage of the bibliographic databases in mental health research. *Nordic Journal of Psychiatry*, 64, 181–188. <http://dx.doi.org/10.3109/08039480903337378>.
- McDonald, S., Taylor, L., & Adams, C. (1999). Searching the right database: A comparison of four databases for psychiatry journals. *Health Libraries Review*, 16, 151–156. <http://dx.doi.org/10.1046/j.1365-2532.1999.00222.x>.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6, e1000097. <http://dx.doi.org/10.1016/j.jclinepi.2009.06.006>.
- National Library of Medicine (2010). Medical subjects headings (MeSH). Retrieved September 10, 2010, from <http://www.nlm.nih.gov/mesh/>
- Reed, G. M. (2010). Towards ICD-11: Improving the clinical utility of WHO's International Classification of Mental Disorders. *Professional Psychology: Research and Practice*, 41, 457–465. <http://dx.doi.org/10.1037/a0021701>.
- Tuleya, L. G. (2007). *Thesaurus of psychological index terms* (11th ed.). Washington, DC: American Psychological Association.
- U.S. National Library of Medicine (2011). MEDLINE fact sheet. Retrieved from <http://www.nlm.nih.gov/pubs/factsheets/medline.html>
- Watson, R. J. D., & Richardson, P. H. (1999a). Accessing the literature on outcome studies in group psychotherapy: The sensitivity and precision of Medline and PsycINFO bibliographic database searching. *The British Journal of Medical Psychology*, 72, 127–134. <http://dx.doi.org/10.1348/000711299159763>.
- Watson, R. J. D., & Richardson, P. H. (1999b). Identifying randomized controlled trials of cognitive therapy for depression: Comparing the efficiency of EMBASE, Medline, and PsycINFO bibliographic databases. *The British Journal of Medical Psychology*, 72, 535–542. <http://dx.doi.org/10.1348/000711299160220>.