Bibliometric Analysis for Papers on Topics Related to Drinking Water

This is a revised bibliometric analysis of the papers prepared by intramural and extramural researchers of the U.S. Environmental Protection Agency (EPA) on topics related to drinking water (DW). This analysis was revised in May 2007 because the journals were initially categorized into the fields used by Thomson Scientific's *Essential Science Indicators (ESI)* using information provided in Thomson's *Journal Citation Reports (JCR)*; for this revised analysis, the journals were categorized into *ESI* fields using the journal category list by *ESI* that is available on the Internet at http://in-cites.com/journal-list/index.html. The Journal List for ESI was made available in 2006 and the current list contains all of the 12,734 journals covered for *ESI* up to December 31, 2006. This list is updated bimonthly by Thomson. This revised bibliometric analysis will allow comparison of the results of this 2005 analysis to those of the analysis performed in 2007.

For this analysis, 691 papers were reviewed. These 691 papers, published from 1994 to 2005, were cited 8,334 times in the journals covered by Thomson's Web of Science.[1] Of these 691 papers, 567 (82%) have been cited at least once in a journal. Searches of Thomson Scientific's *Web of Science* were conducted to obtain times cited data for the drinking water journal publications. The analysis was completed using Thomson's *ESI* and *JCR* as benchmarks. *ESI* provides access to a unique and comprehensive compilation of essential science performance statistics and science trends data derived from Thomson's databases. For this analysis, the *ESI* highly cited papers thresholds as well as the hot papers thresholds were used to assess the influence and impact of the drinking water papers. *JCR* is a recognized authority for evaluating journals. It presents quantifiable statistical data that provide a systematic, objective way to evaluate the world's leading journals and their impact and influence in the global research community. The two key measures used in this analysis to assess the journals in which the EPA drinking water papers are published are the Impact Factor and Immediacy Index. The Impact Factor is a measure of the frequency with which the "average article" in a journal has been cited in a particular year. The Impact Factor helps evaluate a journal's relative importance, especially when compared to other journals in the same field. The Immediacy Index is a measure of how quickly the "average article" in a journal is cited. This index indicates how often articles published in a journal are cited within the same year and it is useful in comparing how quickly journals are cited.

The report includes a summary of the results of the bibliometric analysis, an analysis of the 691 drinking water research papers analyzed by *ESI* field (e.g., chemistry, microbiology, pharmacology & toxicology), and an analysis of the journals in which the drinking water papers were published.

[1] Thomson's *Web of Science* provides access to current and retrospective multidisciplinary information from approximately 8,700 of the most prestigious, high impact research journals in the world. *Web of Science* also provides cited reference searching.

Summary of Analysis

Nearly one-fifth of the drinking water publications are highly cited papers. A review of the citations indicates that 135 (19.54%) of the drinking water papers qualify as highly cited when using the *ESI* criteria for the top 10% of highly cited publications. Eight (1.16%) of the drinking water papers actually qualify as highly cited when using the criteria for the top 1%, and 1 (0.14%) of these papers qualify as very highly cited when using the criteria for the top 0.1%. None of the drinking water papers met the criteria for the top 0.01%.

The drinking water papers are more highly cited than the average paper. Using the *ESI* average citation rates for papers published by field as the benchmark, in 12 of the 16 fields in which the EPA drinking water papers were published, the ratio of actual to expected cites is greater than 1, indicating that the drinking water papers are more highly cited than the average papers in those fields. For the 16 fields combined, the ratio of actual to expected cites is 1.63, indicating that the drinking water papers are more highly cited than the average paper.

Nearly one-third of the drinking water papers are published in very high impact journals. Two-hundred two (202) of 691 papers were published in the top 10% of journals ranked by *JCR* Impact Factor, representing 29% of EPA's drinking water papers. Sixteen percent (110 out of the 691 papers) of the drinking water papers are published in the top 10% of journals ranked by *JCR* Immediacy Factor.

Six of the drinking water papers qualify as hot papers. Using the hot paper thresholds established by *ESI* as a benchmark, 6 hot papers, representing 0.87% of the drinking water papers, were identified in the analysis.

The author self-citation rate is below average. Four-hundred sixty-eight (468) of the 8,334 cites are author self-cites. This 5.6% author self-citation rate is below the accepted range of 10-30% author self-citation rate.

Highly Cited Drinking Water Publications

The 691 drinking water papers reviewed for this analysis covered 16 of the 22 *ESI* fields. The distribution of the papers among these 16 fields and the number of citations by field are presented in Table 1.

Table 1. Drinking Water Papers by *ESI* Fields

***ESI* Field**	**No. of Citations**	**No. of EPA DW Papers**	**Average Cites/Paper**
Pharmacology & Toxicology	1,969	128	15.38
Environment/Ecology	1,408	176	8.00
Clinical Medicine	1,407	90	15.63
Chemistry	950	63	15.08

ESI Field	No. of Citations	No. of EPA DW Papers	Average Cites/Paper
Microbiology	883	68	12.98
Molecular Biology & Genetics	522	35	14.91
Immunology	390	19	20.53
Engineering	306	50	6.12
Biology & Biochemistry	173	21	8.24
Neuroscience & Behavior	132	9	14.67
Social Sciences, general	82	7	11.71
Agricultural Sciences	72	16	4.50
Plant & Animal Science	20	3	6.67
Physics	16	1	16.00
Computer Science	3	1	3.00
Geosciences	1	4	0.25
Totals	**8,334**	**691**	**12.06**

There were 135 (19.54% of the papers analyzed) highly cited EPA drinking water papers in 11 fields—Agricultural Sciences, Biology & Biochemistry, Chemistry, Clinical Medicine, Engineering, Environment/Ecology, Immunology, Microbiology, Pharmacology & Toxicology, Physics, and Social Sciences—when using the *ESI* criteria for the **top 10% of papers**. Table 2 shows the 135 EPA drinking water papers that met the **top 10% threshold in *ESI***. Eight (1.16% of the papers analyzed) of these papers qualified as highly cited when using the *ESI* criteria for the **top 1% of papers**. These 8 papers covered three fields—Clinical Medicine, Environment/Ecology, and Pharmacology & Toxicology. Table 3 shows the 8 EPA papers in those 3 fields that met the **top 1% threshold in *ESI***. There was 1 (0.14% of the papers analyzed) very highly cited EPA drinking water paper in the field of Pharmacology & Toxicology. This paper met the *ESI* criteria for the **top 0.1% of papers**. None of the papers met the **top 0.01% threshold in *ESI***.

Table 2. Number of Highly Cited Drinking Water Papers by Field (top 10%)

ESI Field	Citations	No. of Papers	Average Cites/Paper	% of EPA Papers in Field
Pharmacology &Toxicology	1,228	26	47.23	20.31%
Environment/Ecology	872	43	20.28	24.43%
Chemistry	717	24	29.88	38.10%
Clinical Medicine	705	9	78.33	10.00%
Microbiology	424	9	47.11	13.24%
Engineering	219	13	16.85	26.00%
Immunology	174	3	58.00	15.79%
Biology & Biochemistry	81	2	40.50	9.52%
Social Sciences, general	78	4	19.50	57.14%
Agricultural Sciences	19	1	19.00	6.25%
Physics	16	1	16.00	100.00%
Totals	**4,533**	**135**	**33.58**	**19.54%**

Table 3. Number of Highly Cited Drinking Water Papers by Field (top 1%)

ESI Field	Citations	No. of Papers	Average Cites/Paper	% of EPA Papers in Field
Pharmacology & Toxicology	446	5	89.20	3.91%
Clinical Medicine	319	1	319.00	1.11%
Environment/Ecology	30	2	15.00	1.14%
Totals	**795**	**8**	**99.38**	**1.16%**

The citations for the 8 highly cited papers (top 1%) are presented in Table 4, and the citations for the 1 very highly cited paper is provided in Table 5.

Table 4. Highly Cited Drinking Water Papers (top 1%)

ESI Field	No. of Cites	First Author	Paper
Pharmacology & Toxicology	116	Kitchin KT	Recent advances in arsenic carcinogenesis: modes of action, animal model systems, and methylated arsenic metabolites. *Toxicology and Applied Pharmacology* 2001;172(3):249-261.
	114	Styblo M	Comparative toxicity of trivalent and pentavalent inorganic and methylated arsenicals in rat and human cells. *Archives of Toxicology* 2000;74(6):289-299.
	98	Mass MJ	Methylated trivalent arsenic species are genotoxic. *Chemical Research in Toxicology* 2001;14(4):355-361.
	71	Thomas DJ	The cellular metabolism and systemic toxicity of arsenic. *Toxicology and Applied Pharmacology* 2001;176(2):127-144.
	47	Hughes MF	Arsenic toxicity and potential mechanisms of action. *Toxicology Letters* 2002;133(1):1-16.
Clinical Medicine	319	Dupont HL	The infectivity of Cryptosporidium-parvum in healthy-volunteers. *New England Journal of Medicine* 1995;332(13):855-859.
Environment/ Ecology	25	Styblo M	The role of biomethylation in toxicity and carcinogenicity of arsenic: A research update. *Environmental Health Perspectives* 2002;110(Suppl 5):767-771.
	5	Plewa MJ	Halonitromethane drinking water disinfection byproducts: chemical characterization and mammalian cell cytotoxicity and genotoxicity. *Environmental Science & Technology* 2004;38(1):62-68.

Table 5. Very Highly Cited Drinking Water Paper (Top 0.1%)

ESI Field	No. of Cites	First Author	Paper
Pharmacology & Toxicology	116	Kitchin KT	Recent advances in arsenic carcinogenesis: modes of action, animal model systems, and methylated arsenic metabolites. *Toxicology and Applied Pharmacology* 2001;172(3):249-261.

Ratio of Actual Cites to Expected Citation Rates

The expected citation rate is the average number of cites that a paper published in the same journal in the same year and of the same document type (article, review, editorial, etc.) has received from the year of publication to the present. Using the *ESI* average citation rates for papers published by field as the benchmark, in 12 of the 16 fields in which the EPA drinking water papers were published, the ratio of actual to expected cites is greater than 1, indicating that the EPA papers are more highly cited than the average papers in those fields (see Table 6).

Table 6. Ratio of Average Cites to Expected Cites for Drinking Water Papers by Field

***ESI* Field**	**Total Cites**	**Expected Cite Rate**	**Ratio**
Agricultural Sciences	72	56.69	1.27
Biology & Biochemistry	173	181.03	0.96
Chemistry	950	427.19	2.22
Clinical Medicine	1,407	814.94	1.73
Computer Science	3	0.53	5.66
Engineering	306	130.57	2.34
Environment/Ecology	1,408	881.24	1.60
Geosciences	1	2.90	0.34
Immunology	390	273.18	1.43
Microbiology	883	585.39	1.51
Molecular Biology & Genetics	522	611.08	0.85
Neuroscience & Behavior	132	140.15	0.94
Pharmacology & Toxicology	1,969	968.94	2.03
Physics	16	4.13	3.87
Plant & Animal Science	20	14.85	1.35
Social Sciences, general	82	18.28	4.48
Totals	**8,334**	**5111.09**	**1.63**

***JCR* Benchmarks**

The Impact Factor is a well known metric in citation analysis. It is a measure of the frequency with which the *average article* in a journal has been cited in a particular year. The Impact Factor helps evaluate a journal's relative importance, especially when compared to others in the same field. The Impact Factor is calculated by dividing the number of citations in the current year to articles published in the 2 previous years by the total number of articles published in the 2 previous years.

Table 7 indicates the number of drinking water papers published in the top 10% of journals, based on the *JCR* Impact Factor. Two-hundred two (202) of 691 papers were published in the top 10% of journals, representing 29% of EPA's drinking water papers.

Table 7. Drinking Water Papers in Top 10% of Journals by *JCR* Impact Factor

EPA DW Papers in that Journal	Journal	Impact Factor (IF)	*JCR* IF Rank
37	Environmental Science & Technology	3.592	487
30	Applied and Environmental Microbiology	3.820	418
24	Environmental Health Perspectives	3.408	538
18	Analytical Chemistry	5.250	248
10	Journal of Analytical Atomic Spectrometry	3.200	605
9	Chemical Research in Toxicology	3.332	555
9	Journal of Infectious Diseases	4.481	311
8	Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis	3.433	530
7	Infection and Immunity	3.875	403
6	Carcinogenesis	4.663	292
5	Epidemiology	4.220	350
4	Journal of Virology	5.225	251
4	Journal of Clinical Microbiology	3.489	519
3	Mutation Research-Reviews in Mutation Research	5.783	210
3	International Journal of Epidemiology	3.289	575
2	Emerging Infectious Diseases	5.340	240
2	American Journal of Epidemiology	4.486	310

EPA DW Papers in that Journal	Journal	Impact Factor (IF)	*JCR* IF Rank
2	TrAC–Trends in Analytical Chemistry	3.539	502
1	New England Journal of Medicine	34.833	5
1	Nature Medicine	30.550	9
1	Chemical Reviews	21.036	23
1	Lancet	18.316	28
1	Cancer Research	8.649	105
1	Mass Spectrometry Reviews	7.364	143
1	FASEB Journal	7.172	149
1	Bioinformatics	6.701	168
1	Nucleic Acids Research	6.575	171
1	Journal of Biological Chemistry	6.482	179
1	Free Radical Biology and Medicine	5.063	260
1	Drug Discovery Today	4.943	271
1	Mutation Research-DNA Repair	3.987	386
1	Drug Metabolism and Disposition	3.652	462
1	Methods	3.622	469
1	Mental Retardation and Developmental Disabilities Research Reviews	3.479	522
1	American Journal of Public Health	3.363	551
1	Journal of the American Society for Mass Spectrometry	3.321	563
1	Journal of Nutrition	3.321	563
Total = 202			

Immediacy Index

The journal Immediacy Index is a measure of how quickly the *average article* in a journal is cited. It indicates how often articles published in a journal are cited within the year they are published. The Immediacy Index is calculated by dividing the number of citations to articles published in a given year by the number of articles published in that year.

Table 8 indicates the number of EPA drinking water papers published in the top 10% of journals, based on the *JCR* Immediacy Index. One-hundred ten (110) of the 691 papers analyzed appear in the top 10% of journals, representing 16% of EPA's drinking water papers.

Table 8. Drinking Water Papers in Top 10% of Journals by *JCR* Immediacy Index

EPA Papers in that Journal	Journal	Immediacy Index (II)	*JCR* II Rank
24	Environmental Health Perspectives	0.869	304
18	Analytical Chemistry	0.657	493
9	Journal of Infectious Diseases	0.889	287
8	Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis	0.721	420
7	Infection and Immunity	0.624	544
6	Carcinogenesis	0.775	379
5	Epidemiology	0.938	264
4	Journal of Virology	1.124	188
3	International Journal of Epidemiology	1.376	131
2	American Journal of Tropical Medicine and Hygiene	1.024	216
2	Emerging Infectious Diseases	1.007	225
2	American Journal of Epidemiology	0.908	281
2	Journal of Applied Toxicology	0.759	391
1	Free Radical Biology and Medicine	0.712	432
1	New England Journal of Medicine	11.719	2
1	Nature Medicine	6.749	5
1	Lancet	5.826	10
1	Chemical Reviews	2.955	40
1	Drug Discovery Today	1.882	86
1	Nucleic Acids Research	1.370	133
1	FASEB Journal	1.247	154
1	Journal of Biological Chemistry	1.231	160
1	ATLA-Alternatives to Laboratory Animals	0.964	247

EPA Papers in that Journal	Journal	Immediacy Index (II)	*JCR* II Rank
1	Cancer Research	0.935	268
1	Drug Metabolism and Disposition	0.791	368
1	Mental Retardation and Developmental Disabilities Research Reviews	0.788	371
1	Bioinformatics	0.736	408
1	American Journal of Public Health	0.682	465
1	Journal of Nutrition	0.647	507
1	Methods	0.596	577
1	Infection Control and Hospital Epidemiology	0.590	586
Total = 110			

Hot Papers

ESI establishes citation thresholds for hot papers, which are selected from the highly cited papers in different fields, but the time frame for citing and cited papers is much shorter—papers must be cited within 2 years of publication and the citations must occur in a 2-month time period. Papers are assigned to 2-month periods and thresholds are set for each period and field to select 0.1% of papers. There were no hot papers identified for the current 2-month period (i.e., January-February 2005), but there were a number of hot papers identified from previous periods.

Using the hot paper thresholds established by *ESI* as a benchmark, 6 hot papers, representing 0.87% of the drinking water papers, were identified in four fields—Clinical Medicine, Environment/ Ecology, Microbiology, and Pharmacology & Toxicology. The hot papers are listed in Table 9.

Table 9. Hot Papers Identified Using *ESI* Thresholds

Field	*ESI* Hot Papers Threshold	No. of Cites in 2-Month Period	Paper
Clinical Medicine	12	12 cites in September-October 1996	Dupont HL, et al. The infectivity of Cryptosporidium parvum in healthy volunteers. *New England Journal of Medicine* 1995;332(13):855-859.

Field	*ESI* Hot Papers Threshold	No. of Cites in 2-Month Period	Paper
Environment/ Ecology	8	10 cites in July-August 2004	Styblo M, et al. The role of biomethylation in toxicity and carcinogenicity of arsenic: a research update. *Environmental Health Perspectives* 2002;110(Suppl 5):767-771.
Microbiology	14	16 cites in April-May 2003	Small J, et al. Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. *Applied and Environmental Microbiology* 2001;67(10):4708-4716.
Pharmacology & Toxicology	8	8 cites in July-August 2002	Styblo M, et al. Comparative toxicity of trivalent and pentavalent inorganic and methylated arsenicals in rat and human cells. *Archives of Toxicology* 2000;74(6):289-299.
	8	10 cites in November-December 2002	Kitchin KT. Recent advances in arsenic carcinogenesis: modes of action, animal model systems, and methylated arsenic metabolites. *Toxicology and Applied Pharmacology* 2001;173(3):249-261.
	8	11 cites in August-September 2004	Nesnow S, et al. DNA damage induced by methylated trivalent arsenicals is mediated by reactive oxygen species. *Chemical Research in Toxicology* 2002;15(12):1627-1634.

Author Self-Citation

Self-citations are journal article references to articles from that same author (i.e., the first author). Because higher author self-citation rates can inflate the number of citations, the author self-citation rate was calculated for the drinking water papers. Of the 8,334 total cites, 468 are author self-cites—a 5.6% author self-citation rate. Garfield and Sher[2] found that authors working in research-based disciplines tend to cite themselves on the average of 20% of the time. MacRoberts and MacRoberts[3] claim that approximately 10% to 30% of all the citations listed fall into the category of author self-citation. Therefore, the 5.6% self-cite rate for the drinking water papers is below the range for author self-citation.

[2] Garfield E, Sher IH. New factors in the evaluation of scientific literature through citation indexing. *American Documentation* 1963;18(July):195-201.

[3] MacRoberts MH, MacRoberts BR. Problems of citation analysis: a critical review. *Journal of the American Society of Information Science* 1989;40(5):342-349.