Bibliometric Analysis for Papers on Topics Related to Drinking Water

This is a 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). 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.

The analysis was completed using Thomson's Essential Science Indicators (ESI) and Journal Citation Reports (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. The chief indicators of output, or productivity, are journal article publication counts. For influence and impact measures, ESI employs both total citation counts and cites per paper scores. The former reveals gross influence while the latter shows weighted influence, also called impact. JCR 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.

Summary of Analysis

More than one-quarter of the drinking water publications are highly cited papers. A review of the citations indicates that 178 (25.8%) of the drinking water papers qualify as highly cited when using the ESI criteria for the top 10% of highly cited publications. Nineteen (2.8%) of the drinking water papers actually qualify as highly cited when using the criteria for the top 1%, and 2 (0.3%) of these papers qualify as very highly cited when using the criteria for the top 0.1%.

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 9 of the 13 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.

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.

[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.

Seven of the drinking water papers qualify as hot papers. Using the hot paper thresholds established by ESI as a benchmark, 7 hot papers, representing 1% 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 13 of the 22 ESI fields. The distribution of the papers among these 14 fields and the number of citations by field are presented in Table 1.

Table 1. Drinking Water Papers by ESI Fields

No. of Citations	ESI Field	No. of EPA DW Papers	Average Cites/Paper
2,162	Environment/Ecology	193	11.20
1,743	Pharmacology & Toxicology	129	13.51
1,298	Chemistry	77	16.86
1,095	Biology & Biochemistry	109	10.05
818	Engineering	115	7.11
545	Clinical Medicine	22	24.77
422	Immunology	23	18.35
142	Neuroscience & Behavior	10	14.20
79	Mathematics	6	13.17
16	Plant & Animal Science	3	5.33
8	Agricultural Sciences	2	4.00
3	Computer Science	1	3.00
3	Physics	1	3.00
Total = 8,334		**Total = 691**	

There were 178 (25.8% of the papers analyzed) highly cited EPA drinking water papers in eight fields—Biology & Biochemistry, Chemistry, Clinical Medicine, Engineering,

Environment/Ecology, Immunology, Mathematics, and Pharmacology & Toxicology—when using the ESI criteria for the **top 10% of papers**. Table 2 shows the number of EPA drinking water papers that met the **top 10% threshold in ESI**. Nineteen (2.8% of the papers analyzed) of these papers qualified as highly cited when using the ESI criteria for the **top 1% of papers**. These 19 papers covered seven fields—Engineering, Pharmacology & Toxicology, Environment/Ecology, Mathematics, Clinical Medicine, Chemistry, and Biology & Biochemistry. Table 3 shows the number of EPA papers in those seven fields that met the **top 1% threshold in ESI**. There were 2 (0.3% of the papers analyzed) very highly cited EPA drinking water papers in two fields—Engineering and Pharmacology & Toxicology. These two papers met the ESI criteria for the **top 0.1% of papers**.

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

Citations	ESI Field	No. of Papers	Average Cites/Paper	% of EPA Papers in Field
1,584	Environment/Ecology	62	25.55	32.12%
1,054	Chemistry	33	31.94	42.86%
907	Pharmacology &Toxicology	20	45.35	15.50%
669	Engineering	40	16.72	34.78%
427	Biology & Biochemistry	13	32.85	11.93%
387	Clinical Medicine	3	129.00	13.64%
174	Immunology	3	58.00	13.04%
78	Mathematics	4	19.50	66.67%

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

Citations	ESI Field	No. of Papers	Average Cites/Paper	% of EPA Papers in Field
149	Engineering	6	24.83	5.22%
348	Pharmacology & Toxicology	4	87.00	3.10%
190	Environment/Ecology	3	63.33	1.55%
53	Mathematics	3	17.67	50.00%
322	Clinical Medicine	1	322.00	4.55%
98	Chemistry	1	98.00	1.30%

Citations	ESI Field	No. of Papers	Average Cites/Paper	% of EPA Papers in Field
7	Biology & Biochemistry	1	7.00	0.92%

The citations for the 19 highly cited papers are presented in Table 4, and the citations for the 2 very highly cited papers are listed in Table 5.

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

ESI Field	No. of Cites	First Author	Paper
Engineering	5	Plewa MJ	Halonitromethane drinking water disinfection byproducts: chemical characterization and mammalian cell cytotoxicity and genotoxicity. *Environmental Science & Technology* 2004;38(1):62-68.
	13	Miles AM	Comparison of trihalomethanes in tap water and blood. *Environmental Science & Technology* 2002;36(8):1692-1698.
	17	Simpson JM	Microbial source tracking: state of the science. *Environmental Science & Technology* 2002;36(24):5279-5288.
	34	Richardson SD	Identification of new ozone disinfection byproducts in drinking water. *Environmental Science & Technology* 1999;33(19):3368-3377.
	36	Richardson SD	Identification of new drinking water disinfection byproducts formed in the presence of bromide. *Environmental Science & Technology* 1999;33(19):3378-3383.
	44	Ryan JD	Bacteriophage PRD1 and silica colloid transport and recovery in an iron oxide-coated sand aquifer. *Environmental Science & Technology* 1999;33(1):63-73.
Pharmacology & Toxicology	47	Hughes MF	Arsenic toxicity and potential mechanisms of action. *Toxicology Letters* 2002;133(1):1-16.
	71	Thomas DJ	The cellular metabolism and systemic toxicity of arsenic. *Toxicology and Applied Pharmacology* 2001;176(2):127-144.
Pharmacology & Toxicology	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.

ESI Field	No. of Cites	First Author	Paper
	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.
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.
	55	Small J	Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. *Applied and Environmental Microbiology* 2001;67(10):4708-4716.
	110	Waller K	Trihalomethanes in drinking water and spontaneous abortion. *Epidemiology* 1998;9(2):134-140.
Mathematics	6	Lipscomb JC	The impact of cytochrome P450 2E1-dependent metabolic variance on a risk-relevant pharmacokinetic outcome in humans. *Risk Analysis* 2003;23(6):1221-1238.
	9	Teunis PFM	Cryptosporidium dose response studies: Variation between isolates. *Risk Analysis* 2002;22(1):175-183.
	38	Swartout JC	A probabilistic framework for the reference dose (probabilistic RfD). Risk Analysis 1998;18(3):271-282.
Clinical Medicine	322	Dupont HL	The infectivity of Cryptosporidium-parvum in healthy-volunteers. *New England Journal of Medicine* 1995;332(13):855-859.
Chemistry	98	Mass MJ	Methylated trivalent arsenic species are genotoxic. *Chemical Research in Toxicology* 2001;14(4):355-361.
Biology & Biochemistry	7	Vinje J	Development and application of a capsid VP1 (region D) based reverse transcription PCR assay for genotyping of genogroup I and II noroviruses. *Journal of Virological Methods* 2004;116(2):109-117.

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

ESI Field	No. of Cites	First Author	Paper
Engineering	5	Plewa MJ	Halonitromethane drinking water disinfection byproducts: chemical characterization and mammalian cell cytotoxicity and genotoxicity. *Environmental Science & Technology* 2004;38(1):62-68.
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 9 of the 13 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
Biology & Biochemistry	1,095	1,063.04	1.03
Chemistry	1,298	481.71	2.69
Clinical Medicine	548	230.32	2.38
Engineering	818	269.25	3.04
Environment/Ecology	2,162	1045.76	2.07
Immunology	422	308.69	1.37
Mathematics	79	9.91	7.97
Neuroscience & Behavior	142	147.59	0.96
Pharmacology & Toxicology	1,743	1013.34	1.72
Plant & Animal Science	16	16.19	0.99

ESI Field	Total Cites	Expected Cite Rate	Ratio
Agricultural Sciences	8	8.27	0.97
Computer Science	3	0.53	5.66
Physics	3	8.03	0.37

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

EPA DW Papers in that Journal	Journal	Impact Factor (IF)	JCR IF Rank
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
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

EPA DW Papers in that Journal	Journal	Impact Factor (IF)	JCR IF Rank
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

EPA Papers in that Journal	Journal	Immediacy Index (II)	JCR II Rank
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
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, 7 hot papers, representing 1% of the drinking water papers, were identified in five fields—Chemistry, Clinical Medicine,

Environment/ Ecology, Engineering, 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
Chemistry	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.
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.
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.
		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.
Engineering	4	4 cites in November-December 2003	Miles AM, et al. Comparison of trihalomethanes in tap water and blood. *Environmental Science & Technology* 2002;36(8):1692-1698.
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.
		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.

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.