Measuring research impact: *citation metrics in scholarship*

Chinmay Shah

Editor, National Journal of Integrated Research In Medicine
Department of Physiology,
Government Medical College, Bhavnagar,
Bhavnagar. Gujarat. INDIA, cjshah79@yahoo.co.in

CITATION METRICS

- What Types of Data are best for which Purposes?
  - There are no all-purpose indicators
  - Start by identifying the question the results are supposed to answer, then collect data
  - Clearly define
    - Purpose of the evaluation
    - Types of data required
    - How the results will be used

WHAT DO INSTITUTIONS WANT TO FIND OUT FROM CITATION METRICS

- What is the university's research performance?
- Are we competitive compared with our peers?
- How can the university forecast growth?
- Which are our centers of excellence?
- What is our citation ranking?
- What is the influence of our research?
- Which are our most influential papers?
- Which are our top researchers?

Other use

- In the digital age it's increasingly *"**Get cited or perish**"*
- Departments track citation counts for individuals/subgroups
- Noticeable increase in requests for citation metrics from faculty, especially for tenure and promotion & graduate students
- Emphasis on quantitative data & evaluating output to make decisions in academic units
- Citation counts used in tenure & funding decisions by institutions
- To consider a broad range of “usage” and “metrics”

- To discuss what shapes success (or lack thereof)
- Does “value” shift across disciplines?
- connection between ***impact | quality | importance***?
- Rank or visualize impact of publications, scholars & journals
- Discover new research areas and trends by *mapping*
- See which journals are most important, and publish in those journals

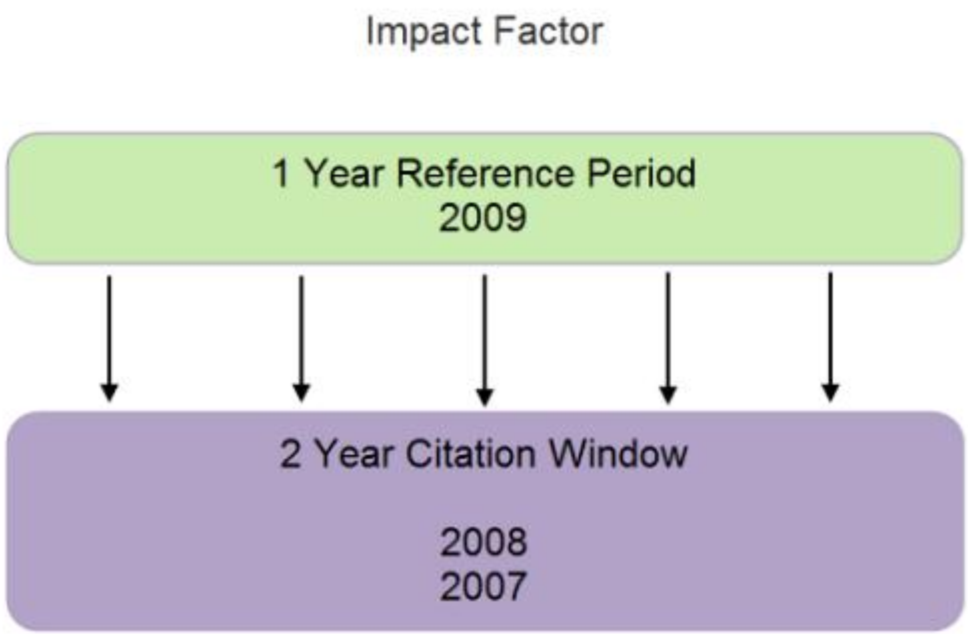
How is citation analysis done?

- citation, Journal & Author Metrics are part of field called *bibliometrics*
- Biometrics s *"discipline of measuring performance of a researcher, collection of articles, journal, research discipline or institution“*. It involves *"application of statistical analyses to study patterns of authorship, publication, and literature use"* (Lancaster 1977)
- **What is being counted**? # papers published per individual, institutional aggregated # of citations, cites per paper, cites per journal, cites per book, cites per book chapter …& so on

Traditional Impact Measures

- *Journal impact factor (Web of Science , Eigenfactor)*
- *Individual scholar impact:* ***h*-index,***(Web of Science, also Google Scholar),* **G-index**
  Individual scholar impact: i10-index (Google Scholar)
- *Individual article impact: Citation count (Web of Science; Google Scholar)*
- *InCites*

Impact Factor



Journal Impact Factor

Cites in 2007 to articles published in:	2006 = 10096	Number of articles published in:	2006 = 352
	2005 = 9958		2005 = 319
	Sum: 20054		Sum: 671

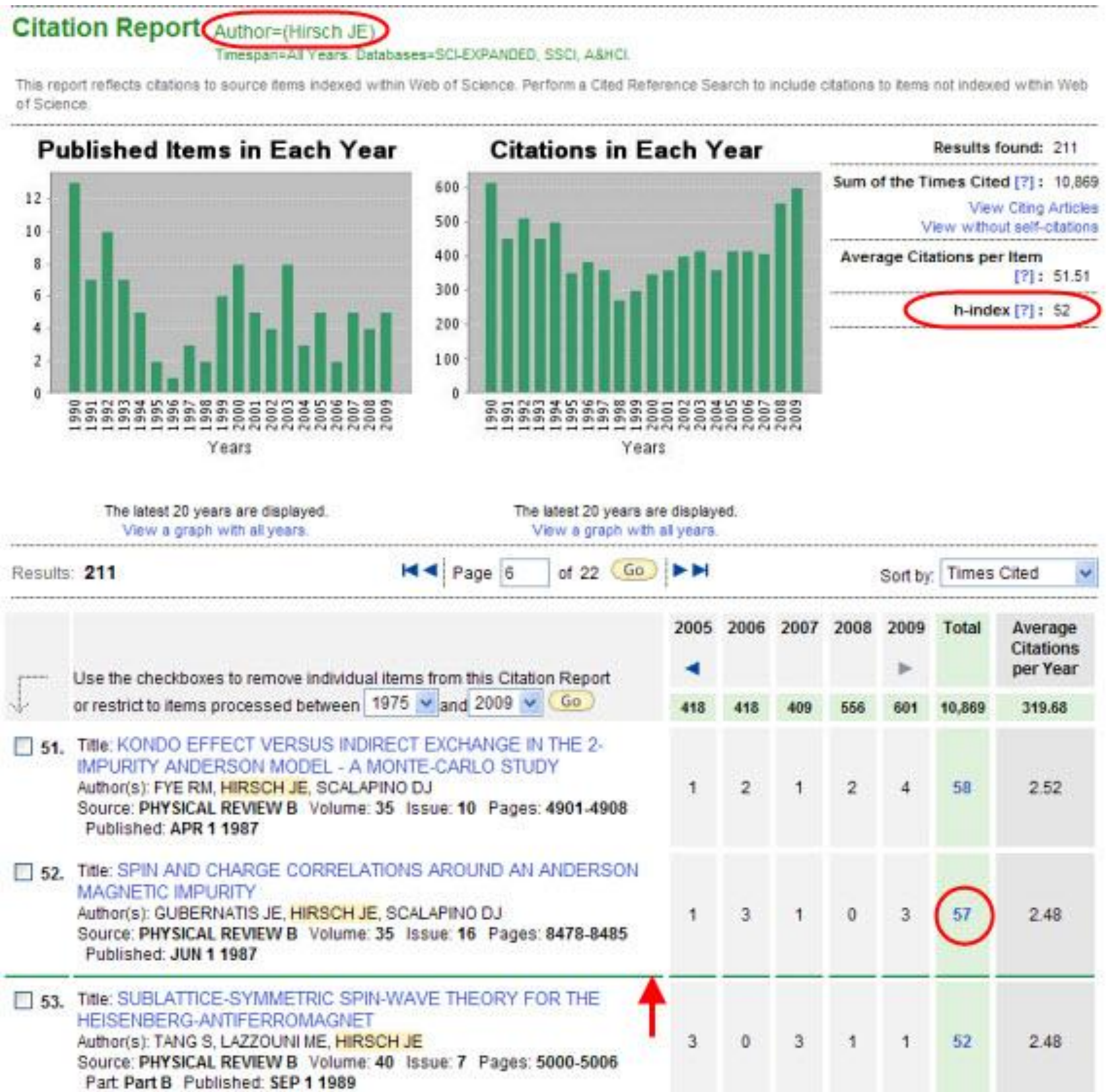
Calculation: $\frac{\text{Cites to recent articles}}{\text{Number of recent articles}} = \frac{20054}{671} = \mathbf{29.887}$

Journal Immediacy Index

Cites in 2007 to articles published in 2007	=	2343
Number of articles published in 2007	=	366

Calculation: $\frac{\text{Cites to current articles}}{\text{Number of current articles}}$ $\frac{2343}{366}$ = **6.402**

The h-index is defined by how many **h** of a researcher's publications (Np) have at least **h** citations each



Variation

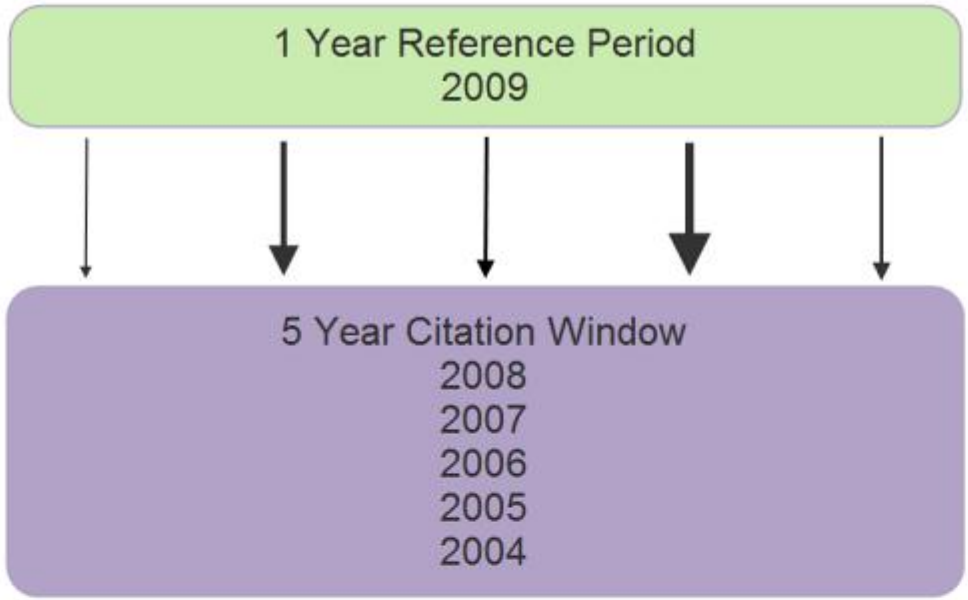
- **The m-index,** introduced by the creator of the h-index, is defined as the h-index divided by the number of years since the researcher’s first publication. The index is meant to normalize the h-index so that early- and late-stage scientists can be compared. The m-index averages periods of high and low productivity throughout a career, which may or may not be reflective of the current situation of the scientist.
- The h-index is relatively unaffected by a small number of exceptionally well-cited articles (eg, reviews). But the case can be made that researchers who have published a landmark paper should get the proper credit for it.

- **The g-index** was developed for this reason. Like the h-index, when a researcher's publications are listed in decreasing order of citations received, the g-index is the largest number such that the top g articles received, in total, at least g^2 citations. Therefore, a few well-cited papers can significantly increase the g-index relative to the corresponding h-index.
- Like the g-index, the **e-index** aims to address the number of "excess" citations above and beyond the h-index . The e-index is defined as the square root of the sum of the "excess" citations in the papers that contributed to the h-index.

Eigenfactor metrics

- While Impact Factor has a one year census period and uses the two previous years for the target window, the Eigenfactor metrics have a one year census period and use the five previous years for the target window

Eigenfactor Metrics



5-Year Journal Impact Factor

Cites in {2007} to items published in:	2006 = 1783	Number of items published in:	2006 = 167
	2005 = 2224		2005 = 190
	2004 = 3092		2004 = 274
	2003 = 2857		2003 = 289
	2002 = 2280		2002 = 195
	Sum: 12236		Sum: 1115

Calculation: Cites to recent items / Number of recent items = 12236 / 1115 = **10.974**

The **i10-index**

- This indicates the number of academic publications an author has written that have at least ten citations from others. It was introduced in July 2011 by Google as part of their work on Google Scholar, a search engine dedicated to academic and related papers

DISCIPLINARITY INDEX

- A measure of the concentration of a set of papers over a set of categories. The index ranges from 0 to 1, where the higher the number, the more concentrated the set. For example, an index of .9 indicates a high level of concentration. This index is based on the Herfindahl index, which is commonly used in economics to assess market share

Metrics: Individual papers

1. Cites: Total citation count for selected paper.

2. Cites2: 2nd generation cite count based on total citations received by the *citing* articles.

3. Expected Citation Rate: An average rate of citation for all the papers of that document type (articles, reviews, letters, etc.), in that journal, for that selected year. This is a metric to evaluate citation counts.

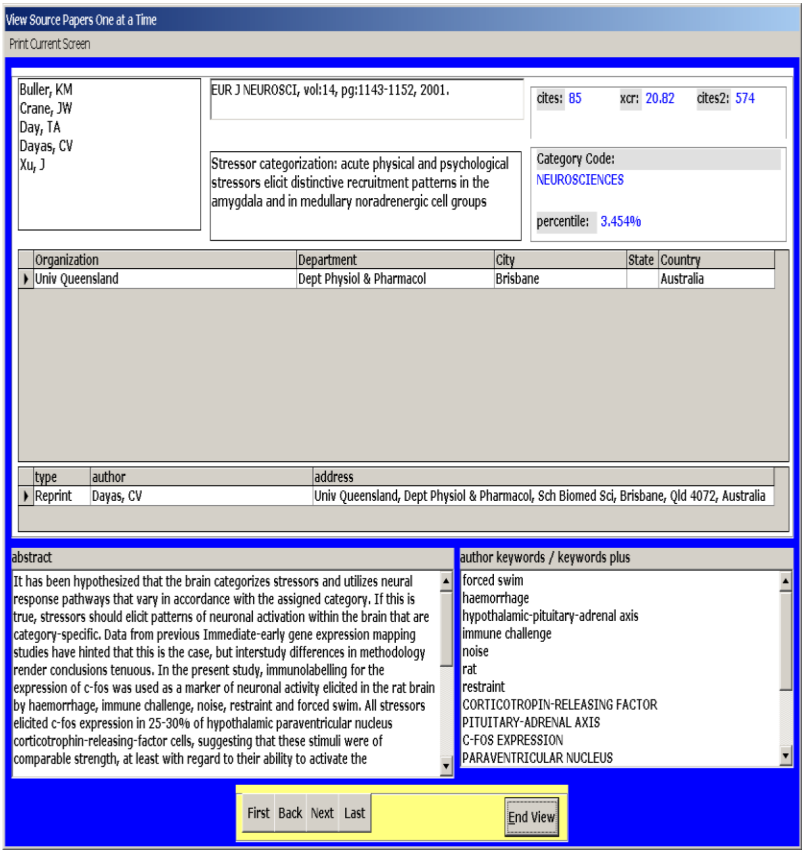
4. Ratio: Ratio of expected cites to actual cites

5. Field: Subject area for the journal in which the paper appeared.

6. %: Percentile position of the paper based on citations in the same field.

Metrics: Individual papers

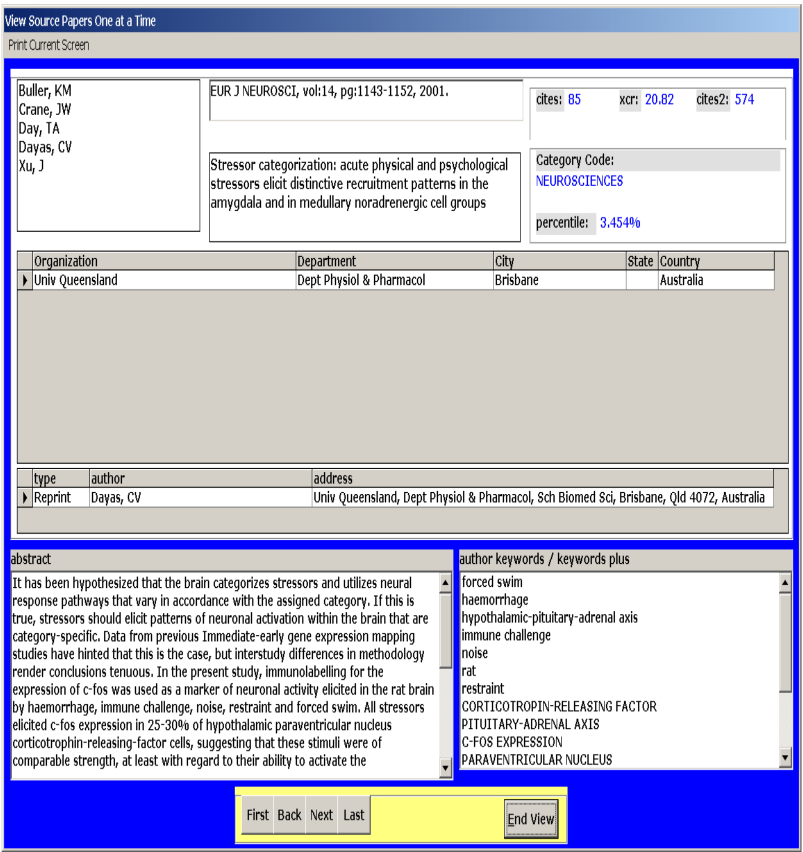
1. Cites: Total citation count for selected paper.



Cites: 85

Metrics: Individual papers

2. Cites2: 2nd generation cite count based on total citations received by the *citing* articles.

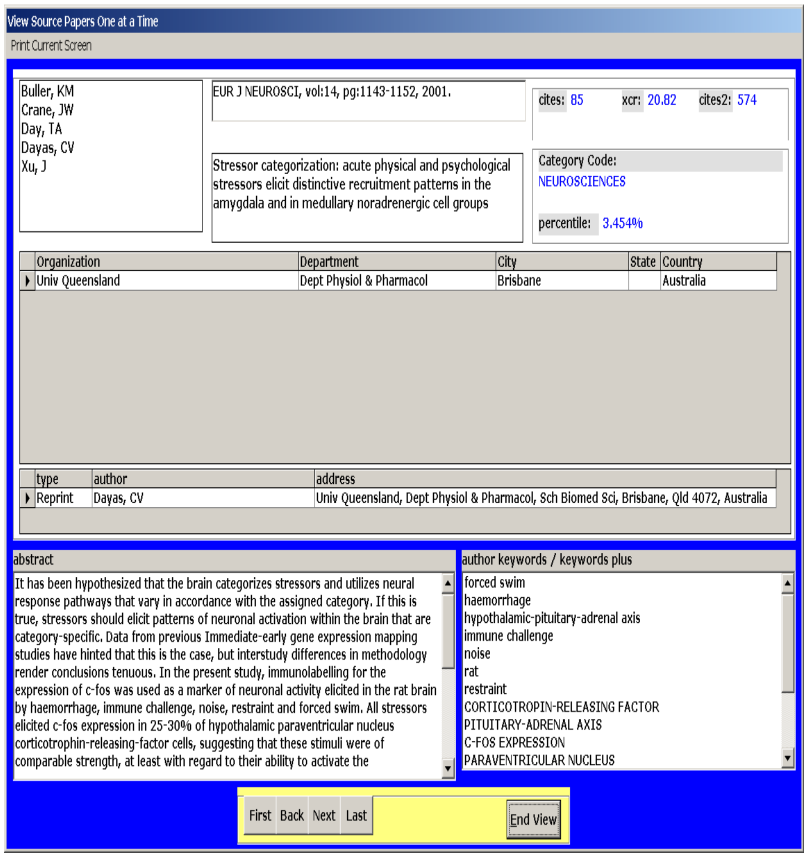


Cites: 85

Cites2: 574

Metrics: Individual papers

3. Expected Citation Rate: An average rate of citation for all the papers of that document type (articles, reviews, letters, etc.), in that journal, for that selected year. This is a metric to compare peer journal papers.



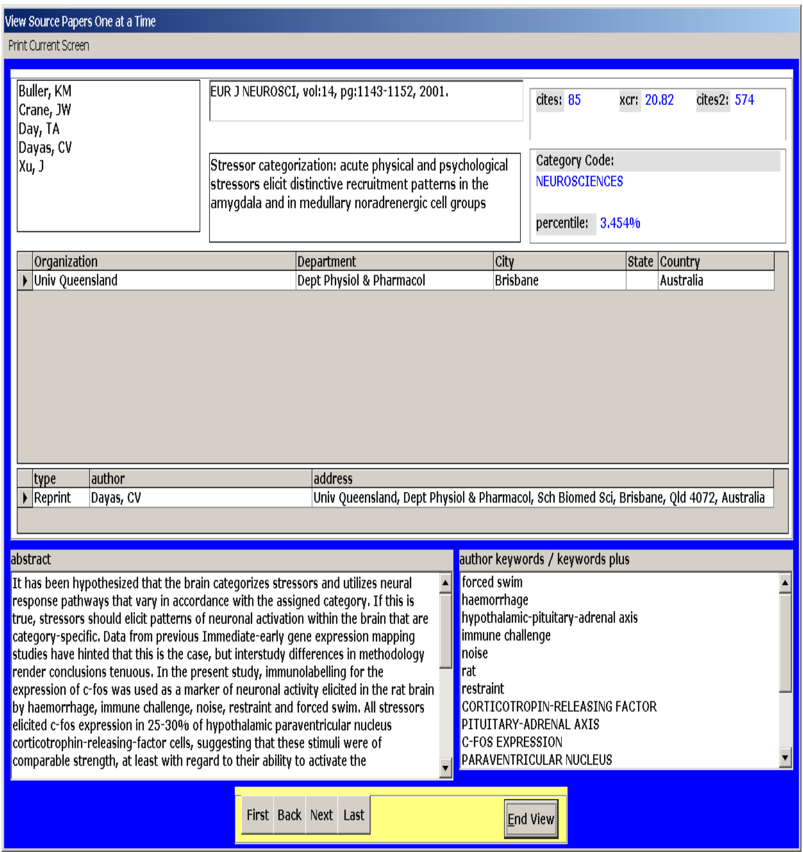
Cites: 85
Cites2: 574

Expected Citation Rate: 20.8

(All *Articles* from *European Journal of Neuroscience* in *2001* received on average 20.8 cites through year-end 2006.)

Metrics: Individual papers

4. Ratio: Ratio of expected cites to actual cites



Cites: 85

Cites2: 574

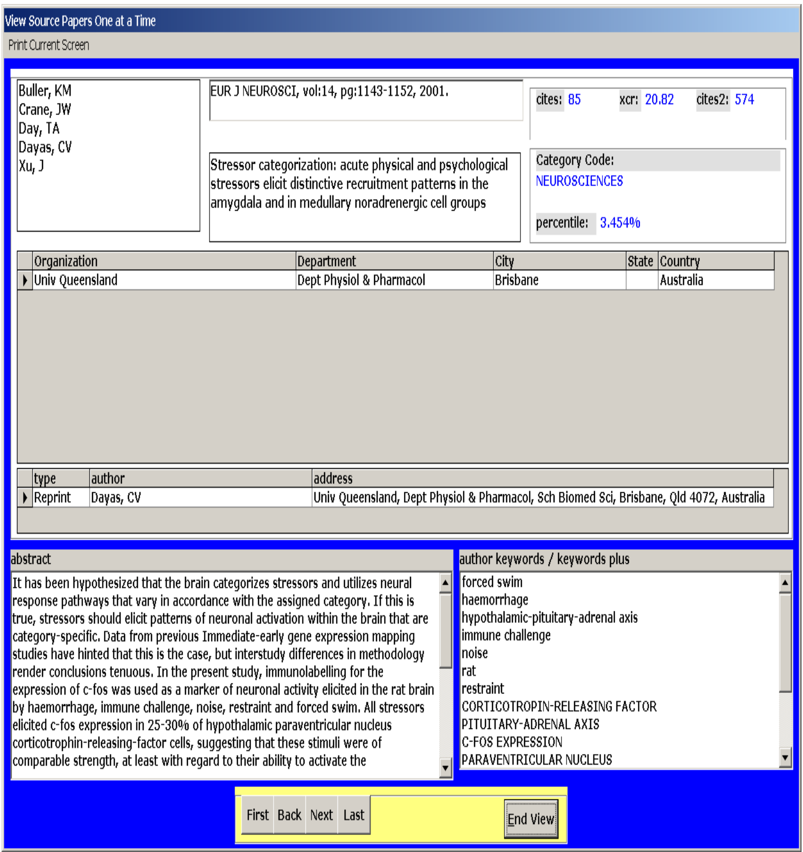
Expected Citation Rate: 20.8

Ratio: 4.1

[85: 20.8 = 4.1]

Metrics: Individual papers

5. Field: Subject area for the journal in which the paper appeared.



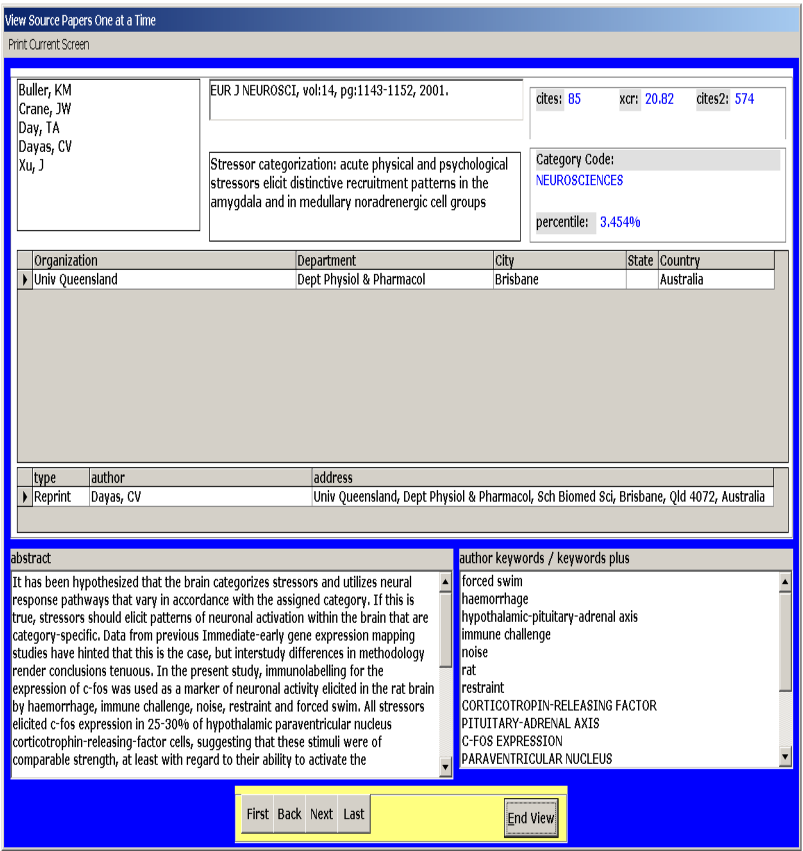
Cites: 85
Cites2: 574
Expected Citation Rate: 20.8
Ratio: 4.1

Field: Neuroscience

[Note: For the multidisciplinary journals *Science*, *Nature* and *PNAS*, all articles and reviews are reassigned based on the primary category to which the article's *citing* and *cited* journals are assigned.]

Metrics: Individual papers

6. Percentile: position of the paper based on citations in the same field and year.



Cites: 85

Cites2: 574

Expected Citation Rate: 20.8

Ratio: 4.1

Field: Neuroscience

Percentile: 3.5%

[The 85 cites to this Neuroscience paper places it in the top 3.5% based on the citation distribution to all papers published in this field in 2001.]

Metrics for groups of papers

1. Total # papers and total # cites: combined numbers for the set

2. Mean times cited: Total cites divided by total papers. [average impact]

3. Median times cited: Midpoint for citations

4. H-Index: Number of papers (*N*) in a given dataset having *N* or more citations.

5. C-Index: Sum of all actual citations divided by sum of all expected citations.

6. Average Percentile: average of the field percentile measures which are based on field and year of publication

7. Disciplinarity: reflects the level of multidisciplinarity in a set of papers, ranging from 0 to 1, where the lower the number, the greater the multidisciplinarity. (Herfindal Index)

H-Index: Number of papers (*N*) in a given dataset having *N* or more citations.

14 papers in this set had 14 or more citations

Summary Metrics

File Sort Options View Options Collection

What You See:

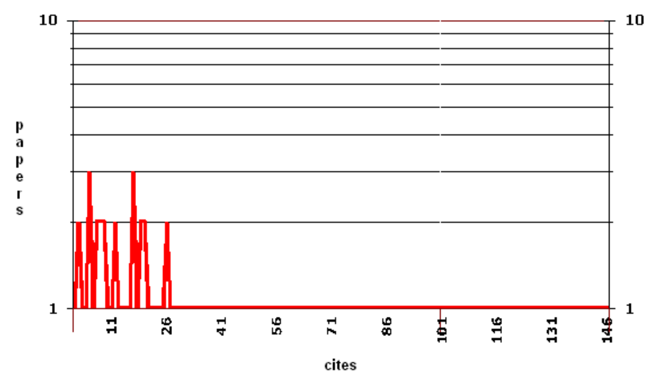
Citation Frequency Distribution: How many papers were cited n number of times. Analysis was run against the collection of source papers called: 'Biotech_Appl_Theor_Microbio'.

Papers	Cites
1	146
1	47
2	26
1	25
1	22
2	20
2	19
3	17
1	15
1	14
2	12
1	11
1	10
2	9
2	8
2	7
3	5
1	4
1	3
2	2
3	0

Total Rows: 21

Metrics for Current Dataset

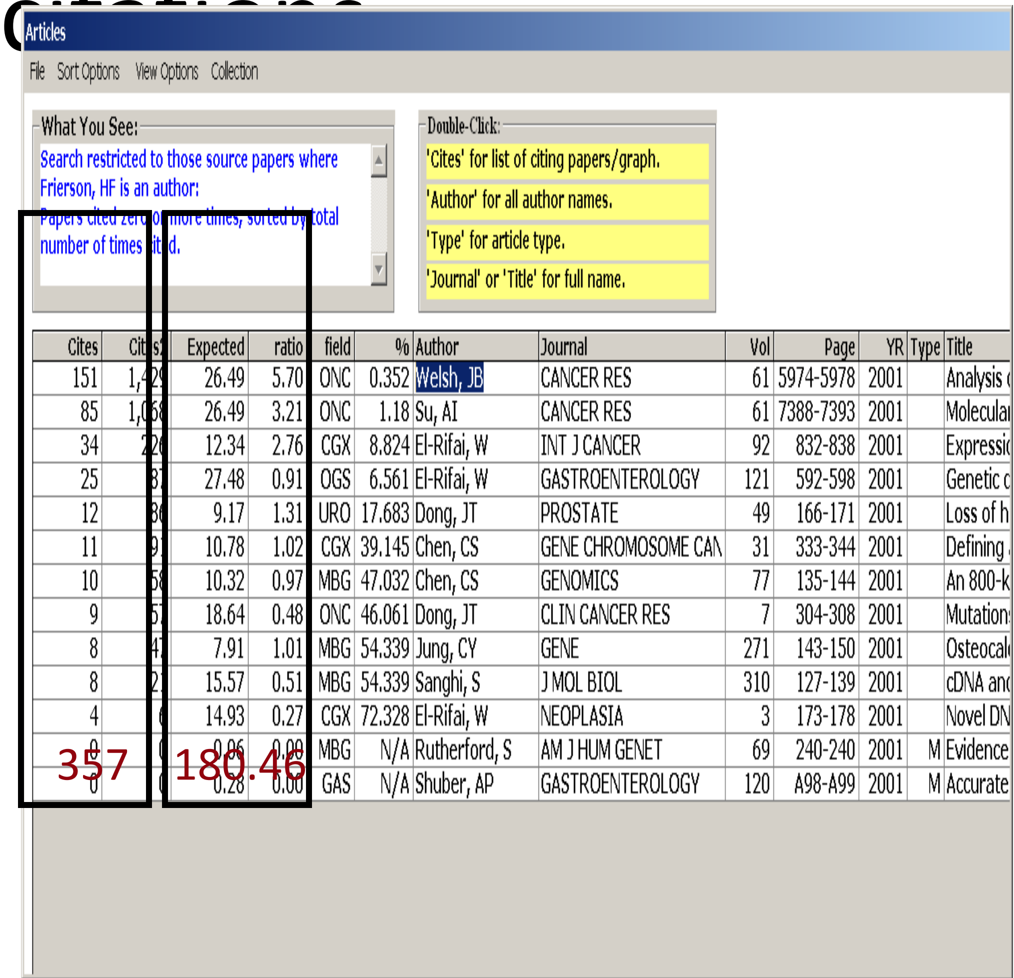
Total Papers = 35, Total Cites = 569
MEAN times Cited = 16.26, MEDIAN times Cited = 11
H-index = 14
C-index = 1.18
Average Percentile = 29.85
Disciplinarity = 0.59



New Analysis...

C-Index: Sum of all actual citations divided by sum of all expected citations

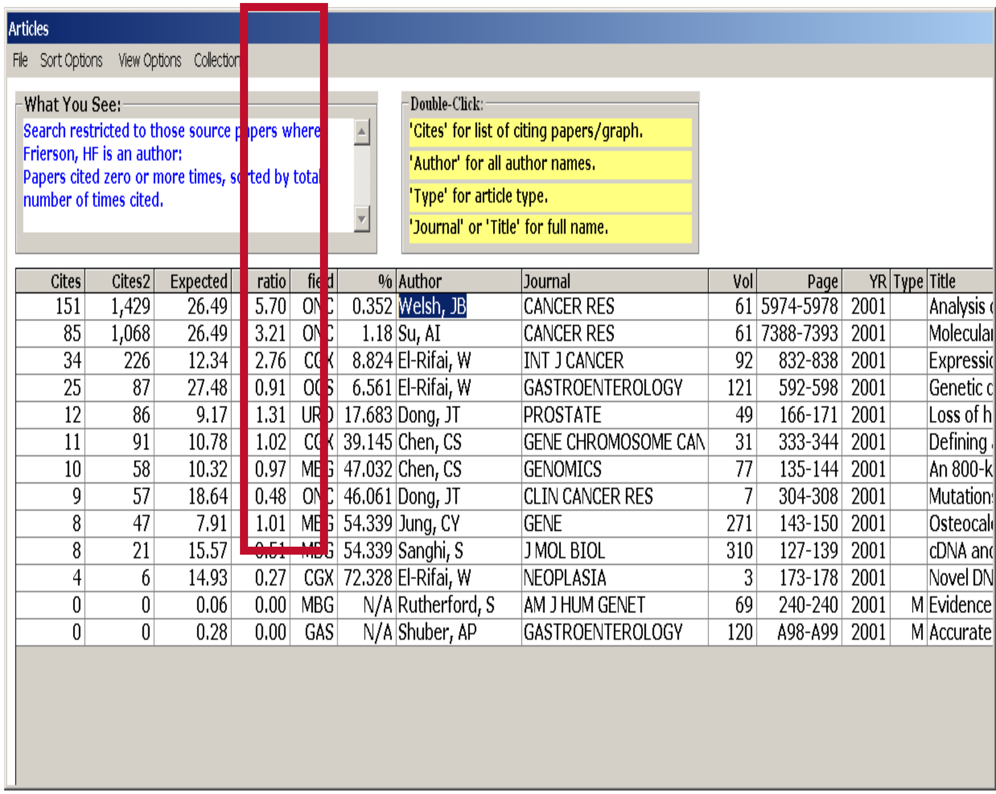
C index = 1.98



Cites	Cit[illegible]s2	Expected	ratio	field	%	Author	Journal	Vol	Page	YR	Type	Title
151	[illegible]	26.49	5.70	ONC	0.352	Welsh, JB	CANCER RES	61	5974-5978	2001		Analysis (
85	[illegible]	26.49	3.21	ONC	1.18	Su, AI	CANCER RES	61	7388-7393	2001		Molecular
34	[illegible]	12.34	2.76	CGX	8.824	El-Rifai, W	INT J CANCER	92	832-838	2001		Expressio
25	[illegible]	27.48	0.91	OGS	6.561	El-Rifai, W	GASTROENTEROLOGY	121	592-598	2001		Genetic c
12	[illegible]	9.17	1.31	URO	17.683	Dong, JT	PROSTATE	49	166-171	2001		Loss of h
11	[illegible]	10.78	1.02	CGX	39.145	Chen, CS	GENE CHROMOSOME CAN	31	333-344	2001		Defining
10	[illegible]	10.32	0.97	MBG	47.032	Chen, CS	GENOMICS	77	135-144	2001		An 800-k
9	[illegible]	18.64	0.48	ONC	46.061	Dong, JT	CLIN CANCER RES	7	304-308	2001		Mutation
8	[illegible]	7.91	1.01	MBG	54.339	Jung, CY	GENE	271	143-150	2001		Osteocal
8	[illegible]	15.57	0.51	MBG	54.339	Sanghi, S	J MOL BIOL	310	127-139	2001		cDNA and
4	[illegible]	14.93	0.27	CGX	72.328	El-Rifai, W	NEOPLASIA	3	173-178	2001		Novel DN
0	[illegible]	0.06	0.00	MBG	N/A	Rutherford, S	AM J HUM GENET	69	240-240	2001	M	Evidence
0	[illegible]	0.28	0.00	GAS	N/A	Shuber, AP	GASTROENTEROLOGY	120	A98-A99	2001	M	Accurate

Average Percentile: average of the field percentile measures which are based on field and year of publication

Average Percentile = 31.62



Cites	Cites2	Expected	ratio	field	%	Author	Journal	Vol	Page	YR	Type	Title
151	1,429	26.49	5.70	ONC	0.352	Welsh, JB	CANCER RES	61	5974-5978	2001		Analysis
85	1,068	26.49	3.21	ONC	1.18	Su, AI	CANCER RES	61	7388-7393	2001		Molecular
34	226	12.34	2.76	CGX	8.824	El-Rifai, W	INT J CANCER	92	832-838	2001		Expressio
25	87	27.48	0.91	OGS	6.561	El-Rifai, W	GASTROENTEROLOGY	121	592-598	2001		Genetic c
12	86	9.17	1.31	URO	17.683	Dong, JT	PROSTATE	49	166-171	2001		Loss of h
11	91	10.78	1.02	CGX	39.145	Chen, CS	GENE CHROMOSOME CAN	31	333-344	2001		Defining
10	58	10.32	0.97	MBG	47.032	Chen, CS	GENOMICS	77	135-144	2001		An 800-k
9	57	18.64	0.48	ONC	46.061	Dong, JT	CLIN CANCER RES	7	304-308	2001		Mutation
8	47	7.91	1.01	MBG	54.339	Jung, CY	GENE	271	143-150	2001		Osteocal
8	21	15.57	0.51	MBG	54.339	Sanghi, S	J MOL BIOL	310	127-139	2001		cDNA and
4	6	14.93	0.27	CGX	72.328	El-Rifai, W	NEOPLASIA	3	173-178	2001		Novel DN
0	0	0.06	0.00	MBG	N/A	Rutherford, S	AM J HUM GENET	69	240-240	2001	M	Evidence
0	0	0.28	0.00	GAS	N/A	Shuber, AP	GASTROENTEROLOGY	120	A98-A99	2001	M	Accurate

Disciplinarity

- Reflects the level of multidisciplinarity in a set of papers, ranging from 0 to 1, where the lower the number, the greater the multidisciplinarity. (Herfindal Index)