Annals of Library and Information Studies
Vol. 57, Sept 2010, pp. 310-326

Mapping of cholera research in India using HistCite

K Thulasi[1] and Subbiah Arunachalam[2*]

[1]National Centre for Science Information, Indian Institute of Science, Bangalore 560 012, India

[2]Centre for Internet and Society, No. 194, 2nd 'C' Cross, Domlur 2nd Stage, Bangalore 560 071, India
*Author for correspondence: Email: subbiah.arunachalam@gmail.com

Cholera research in India over the past six decades has been mapped using HistCite. The analysis based on data from *Science Citation Index Expanded* reveals not only the significant papers, key players, important institutions and core journals, but also provides a visual representation of evolution of knowledge in the field showing the cognitive links between key papers both from within India and elsewhere. The seminal nature of the early work of S N De is clearly seen.

Introduction

India owes a debt to Eugene Garfield. It was his 1986 essay[1] on Prof. Sambhu Nath De's work with cholera, which paved the way to a more effective strategy for treatment and control,[2-4] that brought to light the grave lacunae in the Indian peer review system. De, whom Prof. Joshua Lederberg had nominated more than once for the Nobel Prize in Physiology or Medicine, was never elected to any of the science or medical academies in India, was never awarded any prize (apart from the Coats medal by the Calcutta University) and died in 1985 unhonoured and unsung in India's scientific circles. The first major national level recognition of De's work came in July 1990, five years after he passed away, when *Current Science* brought out a special issue on S N De and cholera enterotoxin (*Curr Sci*, Vol. 59, No. 13 &14, 623-714), largely thanks to the initiative taken by one of us. After reading Garfield's essay and many of De's papers Balaram,[5] editor of *Current Science*, wrote "That De received no major award in India during his lifetime and our Academies did not see it fit to elect him to their Fellowships must rank as one of the most glaring omissions of our time." He attributed this failure to collective myopia that failed to distinguish men of straw from scientists of substance.'[5]

We have pleasure in dedicating this paper, which in a way is an extension of his 1986 paper on cholera research in India,[1] to Gene Garfield on his 85th birthday. We are happy we have used HistCite, a tool developed by Garfield and colleagues.[6-9]

Cholera remained a neglected disease for long. Although Europe and North America were affected during some earlier pandemics, thanks to the high standards of living and hygiene and advances in public health, these continents have been virtually free from cholera from the beginning of the 20th century. Consequently, there is not much motivation or urgency to do research in cholera in the West, and pharmaceutical companies, which carry out much research in Europe and North America, would obviously not invest huge sums on a disease where the return on investment was not expected to be high. The burden, at least till recently when philanthropic foundations stepped in to support research in neglected diseases, was largely on countries like India where cholera continues to be an important public health problem and what is worse incidence of new cases go hugely underreported.[10]

For this study, we decided to look at cholera research in India in the past more than six decades, the period being largely dictated by the availability of data. The subscription to *Web of Science (WoS)* at the Indian Institute of Science allows searching the literature from 1945 onwards.

Methodology

Our analysis of cholera research is based on a bibliometric tool, viz. HistCite, developed by Garfield and colleagues. HistCite is an analytical and visualization tool[6-8] that enables analysis of a subject

and helps a searcher identify the most significant work on a topic and trace its evolution. It also helps to identify highly productive and highly cited authors in any chosen area of research, top and high impact journals and prominent institutions and most importantly, highly cited papers based on local citation scores (LCS) and global citation scores (GCS). The historiographs, with their graphical representation, help to visualize the historical development of research based on the most highly cited papers both within the field and in all of science as reflected by the entire *Web of Science*. We have analyzed data from *Web of Science* (WoS) covering the period 1945 - 19 Mar 2010. The data includes all the 1,750 papers on cholera research from India (labeled the IP collection) as well as all papers citing these 1,750 publications which form the complete cholera India collection of 10,242 papers spanning 1945 - 19 Mar 2010 (labeled the IPC collection). All publication and citation data for this paper were downloaded during the third week of March 2010 from the *Science Citation Index Expanded* section of *Web of Science* (*WoS*).

Data Collection

We describe data collection in some detail to emphasize the fact that gathering bibliographic data for scientometric analysis is not an easy and straightforward task. We were forced to adopt two different approaches to collect data from *WoS*. For the period 1945-1971, address information was not indexed in *WoS*. Hence to collect publications only from India, we used the following search strategy. We selected the following search terms to describe cholera research by trial and error and after consulting with leading cholera researchers in India such as Dr G B Nair and Dr T Ramamurthy:

cholera or el tor or vibrio cholerae or v.cholerae or v cholerae or vibrio-cholerae or cholerae or O139 or enterotoxin

The search resulted in 240 records worldwide for the period 1945-1954. In order to get publications from India, we analyzed these 240 records based on authors with Indian names [Chatterjee HN, De SN, Soman DW, Chanda NN, Sengupta KP, Bhatnagar SS, Desa J, Divekar PV, Shrivastava DL, Agarwala SC, Bhar B, Bhattacharyya K, Bose SR, Ghosal SC, Ghosh MM, Iyer SN, Jungalwalla N, Lahiri SC, Menon IGK, Menon PS, Murti CRK, Narayanan EK, Rao VKM, Roychandhury PK, Sarkar JK, Subbarow Y, Tribedi BP, Veeraraghavan N]. This retrieved 21 publications from India for 1945 to 1954 based on Indian author names. Incidentally, we later learnt that Yellapragada Subbarow had already moved to USA (and was working at Harvard University and Lederle laboratories), but in those days that was an exception.

We executed the same search strategy for the period 1955-1964 and obtained 439 records for the world as a whole. Refining this data using Indian names we obtained 112 records.

In *WoS*, while analyzing by author, only a maximum of 500 results are shown. Hence instead of searching in one shot 1945-1971, we had to execute several searches for shorter periods (1945-54, 1955-1964, 1965-67, 1968-69 and 1970-1971) to make sure no relevant record was left out. Secondly, from the list of papers, we selected only those with authors that have Indian names. We excluded Muslim and Christian names, as we were not sure about their nationality. However, the number of Christian and Muslim names we encountered was rather small and any omission would not have unduly affected our analysis.

For the time span 1972-2010, for which *WoS* has a more elaborate set of metadata, we combined search terms with country field India retrieving 1,393 records (19 March 2010):

1,393 records; TS=(cholera or el tor or vibrio cholerae or v.cholerae or v cholerae or vibrio-cholerae or cholerae or O139 or enterotoxin) and CU=India.

To these we added all the 31 papers of S N De which are indexed in *SCI Expanded*.[11] All of these put together added to 1,768 papers. As many of S N De's papers (those on cholera) were included twice, HistCite removed the duplicates and we were finally left with 1,750 entries.

The 1,750 papers from India pertaining to cholera (called for convenience the Indian cholera papers collection or IP collection) were cited in all by 10,817 papers. Some of these citing papers might have cited more than one cited paper. While building the collection comprising all Indian papers and citations to them, HistCite software removes the duplicates and the collection of Indian cholera papers + citations to

them (conveniently called IPC collection) has 10,242 papers from the period 1945-2010.

Before we move on, here is a caveat. The Topic Search (TS=) is not ideal because it also searches in the "Keywords Plus" field. This field adds “significant, frequently occurring words in the titles of an article's cited references,” so if one writes a paper on cancer and one cites papers on cholera that paper on cancer would be included as a core paper on cholera. We wonder if there have been any studies to measure the accuracy of this method. Soren Paris, a collaborator of Garfield in HistCite analysis, believes it may not be a serious issue (personal communication).Ideally, one should only search by document title, abstract and keywords. That is not possible in *WoS*, so all publications have to be downloaded and imported into a database and then filtered.

Analysis

An important feature of HistCite analysis is the distinction made between citations from within the initial set of papers being considered, in our case the IP collection or the IPC collection, called local citation score (LCS), and citations from all of the global literature indexed in *Web of Science*, called the global citation score (GCS). If our IP collection would have been only topical (rather than 2-dimensional - relating to cholera and belonging only to Indian authors), then the difference would be easily interpreted: top LCS papers are most important for the topic and more relevant to it, while the top GCS may be only tangentially related to cholera. In our case though, the interpretation of the HistCite results based on the IP collection cannot be straightforward. In the case of the IPC collection (10,242 papers) the situation is different. It constitutes a more representative sample of world literature on cholera, although it will not be as complete a collection as the one we would have come up with if we had started with cholera research in general and not confined to India.

Table 1 lists cholera research papers from India since 1945. Only the 12 most highly cited papers by local citation score (LCS) are shown. Here we are concerned with the number of papers within the 1,750 papers on cholera from India, which have cited a particular paper, and not with the number of times the paper has been cited in the global literature indexed by *WoS*. LCS shows the number of citations from papers within the HistCite collection. In this case we are concerned with cholera research performed in India. In this table as well as in other tables, papers are referred to by a unique number (assigned by HistCite analysis) and we mention only the first author name, although each author in a multi-author paper is given credit for the paper. In Table 2, we list cholera research papers from India in descending order of global citations received by the papers (GCS). We notice that the 12 entries in the two tables are not the same. Some papers have been cited more often by cholera researchers in India as reflected by their high LCS and others are generally cited well by global researchers in cholera as well as other areas. For example, of the three seminal papers on cholera Sambhu Nath De wrote in the 1950s,[2-4] only two (the 1953 paper which became a citation classic and the 1959 *Nature* paper) are found listed in Table 1, but not the 1956 paper. But all three are listed in Table 2. Paper number 1506 [Mayor S, Pagano RE, *Nature Reviews Molecular Cell Biology*, 2007 Aug; 8 (8): 603-612] listed in Table 2 is cited more than 170 times but not once by any Indian cholera researcher and therefore it is not listed in Table1. Papers by Nair *et al*. (No. 812), Finkelstein *et al*. (No. 109) and two papers by Shimada *et al*. (No. 766 and No. 795) are listed in Table 1 as they have large enough LCS but not in Table 2. This is so simply because the threshold of number of global citations we have set for inclusion in Table 2 is much higher than the GCS of these papers. Papers by RB Sack *et al*. (No. 355) and D E Schaefer *et al*. (No. 341) find a place in Table 2 but not in Table 1 as they have not been cited often enough by Indian cholera researchers. It will be worthwhile examining why these two papers published in reputed journals in 1970 and 1971 were not cited more often by Indian cholera researchers (Personal communication from Alexander I Pudovkin, 6 August 2010). Were not many Indian researchers aware of these papers or did they know about the papers but found them not directly relevant to their work?

In Table 3 we list names of authors who have published a large number of papers. We have shown only the top 26. As our study covers a large period - over six decades - some authors might have been

Table 1 — Research papers on Cholera from India as seen from *Web of Science* [1945-2010] (IP collection)arranged according to descending order of local citation scores

[Records: 1,750; Authors: 2,738; Journals: 390; Cited references: 25,900]

#	Rec. Date/Author/Journal	LCS	***GCS***
1	761 Ramamurthy T, et al. Lancet. 1993 Mar 13; 341 (8846): 703-704	148	339
2	21 De SN, Chatterje DN, Journal of Pathology and Bacteriology. 1953; 66 (2): 559-562	131	585
3	812 Nair GB, et al, Journal of Infectious Diseases. 1994 May; 169 (5): 1029-1034	59	109
4	31 Dutta NK, Habbu MK, British Journal of Pharmacology and Chemotherapy. 1955; 10 (2): 153-159	58	216
5	962 Mukhopadhyay AK, et al, Journal of Clinical Microbiology. 1996 Oct; 34 (10): 2537-2543	43	61
6	766 Shimada T, et al, Lancet. 1993 May 22; 341 (8856): 1347-1347	42	116
7	109 Finkelstein RA, Mukerjee S, Proceedings of the Society for Experimental Biology and Medicine. 1963; 112 (2): 355-359	40	121
8	795 Shimada T, et al, Current Microbiology. 1994 Mar; 28 (3): 175-178	39	91
9	994 Sharma C, et al, Journal of Infectious Diseases. 1997 May; 175 (5): 1134-1141	35	53
10	782 Ramamurthy T, et al, Journal of Medical Microbiology. 1993 Oct; 39 (4): 310-317	32	52
11	772 Bhattacharya SK, et al, Journal of Infection. 1993 Jul; 27 (1): 11-15	31	58
12	52 De SN, Nature. 1959; 183 (4674): 1533-1534	29	175

Only top 12 entries have been included. Out of 1,750 papers, only 923 papers have been cited at least once, by anyone, among these 1,750 papers.

Table 2 — Cholera research papers from India (IP collection) arranged according to descending order of global citation scores

[Records: 1,750; Authors: 2,738, Journals: 390; Cited references: 25,900]

#	Rec. Date / Author / Journal	LCS	GCS
1	21 De SN, Chatterje DN, Journal of Pathology and Bacteriology. 1953; 66 (2): 559- 562	131	585
2	761 Ramamurthy T, et al, Lancet. 1993 Mar 13; 341 (8846): 703-704	148	339
3	355 Sack RB, et al, Journal of Infectious Diseases. 1971; 123 (4): 378-&	8	243
4	341 Schafer DE, et al, Proceedings of the National Academy of Sciences of the United States of America. 1970; 67 (2): 851-&	2	242
5	31 Dutta NK, Habbu MK, British Journal of Pharmacology and Chemotherapy. 1955; 10 (2): 153-159	58	216
6	1506 Mayor S, Pagano RE, Nature Reviews Molecular Cell Biology. 2007 Aug; 8 (8): 603-612	0	177
7	52 De SN, Nature. 1959; 183 (4674): 1533-1534	29	175
8	37 De SN, et al, Journal of Pathology and Bacteriology. 1956; 71 (1): 201-&	10	154
9	8 Cameron GR, De SN, Journal of Pathology and Bacteriology. 1949; 61 (3): 375- &	2	147
10	238 Pierce NF, et al, Gastroenterology. 1968; 55 (3): 333-&	10	147
11	325 Banwell JG, et al, Journal of Clinical Investigation. 1970; 49 (1): 183-&	6	146
12	846 Karunasagar I, et al, Aquaculture. 1994 Dec 15; 128 (3-4): 203-209	1	143

Only top 12 entries have been included. Out of 1,750 papers based on global citation scores, only the first 1,367 papers have received at least one citation. Others have not been cited at all.

Table 3 — Cholera-India papers (IP collection): Prolific authors

#	Author	*Recs*	*TLCS*	*TGCS*	#	*Author*	*Recs*	*TLCS*	*TGCS*
1	Nair GB	195	1438	3440	14	Ghose AC	35	174	494
2	Bhattacharya SK	94	846	1871	15	Nandy RK	35	156	365
3	Ramamurthy T	92	644	1567	16	Yamasaki S	35	239	617
4	Takeda Y	89	1010	2353	17	Dutta NK	34	183	730
5	Das J	52	295	698	18	Singh DV	33	113	298
6	Mukerjee S	52	250	635	19	Bhadra RK	32	191	433
7	Pal SC	49	277	630	20	Ghosh AN	32	137	366
8	Ghosh A	48	292	632	21	Takeda T	32	524	1108
9	Mukhopadhyay AK	43	395	850	22	Chatterjee SN	31	90	178
10	Sanyal SC	43	128	384	23	De SN	31	218	1343
11	De SP	41	114	402	24	Ghosh RK	31	201	383
12	Shimada T	36	645	1547	25	Bhattacharya MK	29	245	453
13	Ganguly NK	35	58	220	26	Visweswariah SS	29	157	376

Only authors who have published at least 29 papers are included.

Table 4 — Cholera-India papers (IP collection): Journals often used

#	Journal	Recs	TLCS	TGCS
1	Indian Journal of Medical Research	248	451	1112
2	Journal of Medical Microbiology	62	196	511
3	Lancet	45	358	919
4	FEMS Microbiology Letters	44	148	392
5	Bulletin of The World Health Organization	38	179	732
6	Infection and Immunity	37	141	597
7	Journal of Clinical Microbiology	37	319	1010
8	Indian Journal of Experimental Biology	36	64	111
9	Journal of Infectious Diseases	33	197	1173
10	Indian Journal of Animal Sciences	30	4	20
11	Indian Journal of Biochemistry & Biophysics	28	33	97
12	Current Science	26	32	60
13	Journal of Bacteriology	25	123	541
14	Journal of General Microbiology	25	166	454
15	British Medical Journal	22	71	138
16	Journal of Diarrhoeal Diseases Research	20	35	80
17	Indian Journal of Medical Research Section A-Infectious Diseases	19	47	87
18	Journal of Food Science and Technology-Mysore	18	3	56
19	FEMS Immunology and Medical Microbiology	17	56	151
20	Transactions of the Royal Society of Tropical Medicine and Hygiene	17	21	112
21	Epidemiology and Infection	15	84	199
22	Indian Veterinary Journal	15	4	10
23	Antimicrobial Agents and Chemotherapy	14	73	192
24	Journal of Pathology and Bacteriology	14	171	1116

Only the 24 journals most often used are included

actively publishing for over three decades and some might have just entered the publishing phase or dropped out from the field after a few years of active work. Many at the top of the list are from the National Institute of Cholera and Enteric Diseases, Calcutta. GB Nair, T Ramamurthy and SK Bhattacharya have published the largest number of papers on cholera research. Table 4 lists the 24 journals that have

published the largest number of papers on cholera research from India. *Indian Journal of Medical Research*, *Journal of Medical Microbiology* and *Lancet* are the top three journals. Table 5 lists the prolific institutions working on cholera research in India, arranged in descending order of number of papers published. Two institutions based in Calcutta, viz. National Institute of Cholera and Enteric Diseases and Indian Institute of Chemical Biology are the leading Indian institutions in cholera research. Non-Indian institutions among the top 25 are also shown separately. Table 6 includes top 15 papers from the IPC collection, which include both publications from India and citations to them, arranged in descending order of LCS. Table 7 includes the top 15 papers from the IPC collection, arranged in descending order of global citation score (GCS). Table 8 lists the 26 authors publishing the largest number of papers in the IPC collection including publications and citations. There are nine Indian authors out of 26 authors publishing a large number of papers. Table 9 shows the top 25 journals publishing large number of papers in the IPC collection arranged in descending order of number of papers published. Table 10 lists journals occurring most often in the IPC collection including publication and citations arranged by local citation scores. In both Table 9 and Table 10, *Indian Journal of Medical Research* is the only Indian journal, which is an open access journal. Incidentally, ever since the journal became open access its impact factor is increasing. It was 0.600 in 2004, and increased to 0.869 in 2005, 1.224 in 2006, 1.670 in 2007, 1.883 in 2008 and 1.516 in 2009.

Table 11 lists the top 19 institutions in the IPC collection including publication and citations arranged in descending order of number of papers published. We notice that citations per paper is 13.2 for NICED, India's leading research centre in this area, compared to 24.4 for ICDDR, Dhaka, 52 for the University of Maryland, 52.3 for Johns Hopkins University, 41.8 for the Centre for Disease Control, USA, 58.7 for the University of Texas and 60.3 for Harvard University. These numbers are derived from papers from these institutions found in our IPC collection, and not all of these papers need to be cholera research papers. So, we made a quick study of core cholera research papers from these institutions and found that there is a huge difference in citations per paper between the best of Indian institutions in cholera research and leading US institutions (Table 12).

Table 5 — Cholera-India papers (IP collection) classified by institution

#	Institution	Recs	TLCS	TGCS
1	Natl Inst Cholera & Enter Dis	286	1403	3782
2	Indian Inst Chem Biol*	153	657	1634
3	Christian Med Coll & Hosp	56	131	529
4	Postgrad Inst Med Educ & Res	49	86	278
5	Bose Inst	42	149	449
6	Banaras Hindu Univ	38	122	377
7	Indian Inst Sci	35	122	422
8	Cent Drug Res Inst	35	87	239
9	Inst Microbial Technol	30	231	482
10	Jadavpur Univ	30	38	217
11	Saha Inst Nucl Phys	26	67	150
12	Infect Dis Hosp	23	70	233
13	Indian Council Med Res	23	70	223
14	All India Inst Med Sci	21	21	435
15	Indian Vet Res Inst	19	10	78
16	Unknown **	359	1196	5868

Only the 15 institutions publishing at least 19 papers are shown

* Including Indian Institute of Experimental Medicine
** Unknown – For papers published between 1945-1971 institutional names are not included in *WOS* records. Many of S N De's papers written from Calcutta Medical College fall under this category.

The following non-Indian institutions are also among the top 25 institutions:

#	Institution	Records	TLCS	TGCS
1	Int Ctr Diarrhoeal Dis Res, Dhaka, Bangladesh	52	302	1107
2	Int Med Ctr, Tokyo, Japan	37	403	886
3	Kyoto Univ, Kyoto, Japan	30	532	1345
4	Natl Children's Med Res Ctr, Tokyo, Japan	28	479	1023
5	Natl Inst Infect Dis, Tokyo, Japan	26	254	707
6	Natl Inst Hlth, Tokyo, Japan	21	442	954
7	Univ Maryland, Baltimore, MD, USA	21	131	564

Table 6 — Cholera-India IPC collection (1945-2010) including publications and citations arranged in descending order of local citation scores

[Records: 10,242; Authors: 22,890; Journals: 1,800; Cited references: 302,880]

#	Rec. Date / Author / Journal	LCS	GCS
1	25 De SN, Chatterje DN, Journal of Pathology and Bacteriology. 1953; 66 (2): 559-562	584	585
2	3885 Ramamurthy T, et al, Lancet. 1993 Mar 13; 341 (8846): 703-704	340	340
3	870 Schafer DE, *et al*., Proceedings of the National Academy of Sciences of the United States of America. 1970; 67 (2): 851-&	242	242
4	963 Sack RB, *et al*., Journal of Infectious Diseases. 1971; 123 (4): 378-&	241	243
5	3886 Albert MJ, *et al*., Lancet. 1993 Mar 13; 341 (8846): 704-704	236	255
6	1099 Dean AG, *et al*., Journal of Infectious Diseases. 1972; 125 (4): 407-&	222	1024
7	37 Dutta NK, Habbu MK, British Journal of Pharmacology and Chemotherapy. 1955; 10 (2): 153-159	218	218
8	956 Kimberg DV, *et al*., Journal of Clinical Investigation. 1971; 50 (6): 1218-&	211	588
9	362 Craig JP, Nature. 1965; 207 (4997): 614-&	187	330
10	8352 Mayor S, Pagano RE, Nature Reviews Molecular Cell Biology. 2007 Aug; 8 (8): 603-612	177	177
11	97 De SN, Nature. 1959; 183 (4674): 1533-1534	176	176
12	1394 Guerrant RL, *et al*., Infection and Immunity. 1974; 10 (2): 320-327	168	579
13	5148 Faruque SM, *et al*., Microbiology and Molecular Biology Reviews. 1998 Dec; 62 (4): 1301-	164	259
14	3294 Miller VL, Mekalanos JI, Journal of Bacteriology. 1988 Jun; 170 (6): 2575-2583	159	1299
15	55 De SN, *et al*., Journal of Pathology and Bacteriology. 1956; 71 (1): 201-&	153	154

Only the top 15 publications with a minimum LCS of 153 are shown

Table 7 — Cholera-India IPC collection (1945-2010) arranged in descending order of global citation scores

[Records: 10,242, Authors: 22,890, Journals: 1800, Cited References: 302,880]

#	Rec. Date / Author / Journal	LCS	GCS
1	4357 Tenover FC, *et al*., Journal of Clinical Microbiology. 1995 Sep; 33 (9): 2233- 2239	42	3789
2	4926 Nataro JP, Kaper JB, Clinical Microbiology Reviews. 1998 Jan; 11 (1): 142-+	131	1553
3	329 Miller VL, Mekalanos JJ, Journal of Bacteriology. 1988 Jun; 170 (6): 2575- 2583	159	1299
4	1099 Dean AG, *et al*., Journal of Infectious Diseases. 1972; 125 (4): 407-&	222	1024
5	2681 Fridovich I, Annual Review of Pharmacology and Toxicology. 1983; 23: 239- 257	0	801
6	4784 Finlay BB, Falkow S, Microbiology and Molecular Biology Reviews. 1997 Jun; 61 (2): 136-&	21	746
7	1573 Evans DG, *et al*., Infection and Immunity. 1975; 12 (3): 656-667	86	635
8	956 Kimberg DV, *et al*., Journal of Clinical Investigation. 1971; 50 (6): 1218-&	211	588
9	25 De SN, Chatterje DN, Journal of Pathology and Bacteriology. 1953; 66 (2): 559-562	584	585
10	1394 Guerrant RL, *et al.*, Infection and Immunity. 1974; 10 (2): 320-327	168	579
11	3043 Fridovich I, Advances in Enzymology and Related Areas of Molecular Biology. 1986; 58: 61-97	0	578
12	5039 Paton JC, Paton AW, Clinical Microbiology Reviews. 1998 Jul; 11 (3): 450-+	26	562
13	3343 Karlsson KA, Annual Review of Biochemistry. 1989; 58: 309-350	2	547
14	5513 Wommack KE, Colwell RR, Microbiology and Molecular Biology Reviews. 2000 Mar; 64 (1): 69-+	11	544
15	5634 Lucas KA, ., Pharmacological Reviews. 2000 Sep; 52 (3): 375-413	26	528

Only top 15 publications with a minimum GCS of 528 are shown.

Table 8 — Top 26 authors publishing large number of papers in the IPC collection

#	Author	Recs	TLCS	TGCS	#	Author	Recs	TLCS	TGCS
1	Nair GB	277	4439	4926	14	Shimada T	60	1954	2048
2	Takeda Y	127	2960	3614	15	Das J	59	758	777
3	Bhattacharya SK	121	2077	2306	16	Faruque SM	58	1598	1988
4	Sack RB	120	3446	6872	17	Mukerjee S	58	711	744
5	Ramamurthy T	117	1726	1851	18	Qadri F	58	456	769
6	Sack DA	93	1844	4386	19	Ganguly NK	57	258	322
7	Colwell RR	86	1289	3216	20	Huq A	55	740	1197
8	Albert MJ	84	2034	2580	21	Kaper JB	52	1603	5611
9	Sanyal SC	73	674	1115	22	Holmgren J	51	646	2675
10	Mekalanos JJ	70	2117	5624	23	Levine MM	51	871	2194
11	Yamasaki S	68	826	1173	24	De SP	50	426	478
12	Karunasagar I	60	510	768	25	Pierce NF	50	1615	2782
13	Pal SC	60	656	728	26	Svennerholm AM	50	706	2082

Only top 26 authors with a minimum of 50 papers are shown.

Table 9 — Top 20 journals publishing large number of papers in the IPC collection

#	Journal	Recs	TLCS	TGCS
1	Infection and Immunity	407	5180	15084
2	Indian Journal of Medical Research	384	1402	1963
3	Journal of Infectious Diseases	234	4342	10945
4	Journal of Clinical Microbiology	219	3019	9744
5	Journal of Bacteriology	178	2015	5732
6	Applied and Environmental Microbiology	176	1560	4056
7	Lancet	170	2846	5050
8	Journal of Medical Microbiology	168	939	1870
9	FEMS Microbiology Letters	142	820	1530
10	Zhurnal Mikrobiologii Epidemiologii I Immunobiologii	139	80	286
11	Bulletin of the World Health Organization	108	1335	2046
12	Gastroenterology	93	888	5244
13	Microbiology and Immunology	87	379	854
14	Vaccine	87	319	1198
15	Journal of Biological Chemistry	79	370	2518
16	Journal of Food Protection	77	130	813
17	Journal of Applied Microbiology	75	160	658
18	Molecular Microbiology	68	610	2840
19	Proceedings of the National Academy of Sciences of the United States of America	67	1477	3179
20	Epidemiology and Infection	66	551	953

Only journals that have published at least 66 papers are included in the table.

Table 10 — Journals occurring most often in the IPC collection arranged by local citation score

#	Journal	Recs	TLCS	TGCS
1	Infection and Immunity	407	5180	15084
2	Journal of Infectious Diseases	234	4342	10945
3	Journal of Clinical Microbiology	219	3019	9744
4	Lancet	170	2846	5050
5	Journal of Bacteriology	178	2015	5732
6	Applied and Environmental Microbiology	176	1560	4056
7	Journal of Clinical Investigation	50	1487	4765
8	Proceedings of The National Academy of Sciences of the United States of America	67	1477	3179
9	Journal of Pathology and Bacteriology	30	1403	2177
10	Indian Journal of Medical Research	384	1402	1963
11	Bulletin of The World Health Organization	108	1335	2046
12	New England Journal of Medicine	38	1063	4587
13	Journal of Medical Microbiology	168	939	1870
14	Nature	45	924	2972

Journals that have LCS 900 and above are included

Table 11 — Top 20 institutions in the IPC collection

#	Institution	Recs	TLCS	TGCS	#	Institution	Recs	TLCS	TGCS
1	Natl Inst Cholera & Enter Dis	349	4189	4641	11	Postgrad Inst Med Educ & Res	80	317	421
2	Int Ctr Diarrhoeal Dis Res	328	4720	8009	12	US FDA	70	547	1451
3	Univ Maryland	235	4258	12236	13	Inst Pasteur	69	496	1483
4	Indian Inst Chem Biol	199	1857	2490	14	Tufts Univ	62	784	2518
5	Johns Hopkins Univ	167	2533	8739	15	Banaras Hindu Univ	62	577	868
6	Univ Texas	162	1445	9510	16	Jadavpur Univ	60	264	412
7	Harvard Univ	150	2664	9058	17	Natl Inst Infect Dis	59	1122	1373
8	Ctr Dis Control & Prevention*	115	1457	4806	18	Indian Inst Sci	59	430	819
9	Christian Med Coll & Hosp	87	596	847	19	Univ Adelaide	58	480	1185
10	Kyoto Univ	83	2167	2855	20	Unknown**	1209	14241	35041

* includes Centres for Disease Control
** Unknown – For papers published between 1945-1971 institutional names are not included in *WOS* records.

Historiographs of cholera research in India

HistCite analysis provides some insights, which would otherwise be difficult to obtain. It helps us visualize year-on-year developments in a field. We prepared four historiographs by applying HistCite to the publications on cholera research from India since 1945 and citations to them as seen from *SCI-Expanded*. We have not considered cholera research

Table12 : Top seven institutions publishing cholera papers arranged in descending order of citations per paper [Data from Sci Expanded 1945 - 10 August 2010]

#	Institution	Recs	TGCS	Citation per paper
1	Harvard Univ, USA	755	39450	52.2
2	Johns Hopkins Univ, USA	403	15664	38.8
3	Univ Maryland, USA	750	28563	38
4	Ctr Dis Control & Prevention, USA	283	9137	32.3
5	Univ Texas, USA	624	18752	30
6	Int Ctr Diarrhoeal Dis Res, Dhakka	328	8009	24.4
7	Natl Inst Cholera & Enter Dis, India	349	4641	13.2

publications from the entire world and citations to them. One should keep this distinction in mind while interpreting these historiographs.

Historiograph from the top 50 Cholera India papers (IP collection) by LCS

This historiograph (Fig. 1) was generated from the 50 papers that have been most cited among the 1,750 papers (based on local citations, i.e. from within the 1,750 cholera papers from India) using HistCite. The top 50 LCS papers span from 1953 to 2001 with 124 links, with maximum LCS of 148 and minimum LCS of 15. S N De's paper published in 1953, where he demonstrated an animal model for studying cholera using ligated intestinal loops in rabbits,[2] with LCS 131 and GCS 585 is the earliest significant paper on cholera from India. John Craig found this work truly creative and novel, and noted that it "forever altered our concepts surrounding the pathogenesis of secretory diarrhoea."[12] This paper is linked to six nodes in this historiography and it is at the head of a large citation network of influential papers. A subsequent paper of N K Dutta published in 1955, where he described a method for chemotherapeutic investigation of experimental cholera in infant rabbits, also proved to be important being connected to four other nodes in this historiograph.[13] In the first decade of this historiograph, we find two papers by De (nodes 21 and 52), and two papers by Dutta (nodes 33 and 51). These were followed by four papers by S Mukherjee (nodes 95, 96, 98 and 110) in 1963-1964 and two by Finkelstein during the same two years (nodes 109 and 141). Then there is a 14-year long hiatus, from 1968 to 1982, with very little of noteworthy activity. The year 1993 had been particularly productive for NICED with four highly cited papers (nodes 761, 766, 772 and 782). It was the year when NICED and collaborating institutions came up with the discovery of a novel strain of *Vibrio cholerae* with epidemic potential in southern and eastern India (paper by Ramamurthy T *et al*. in *Lancet*, node 761). This paper has the highest LCS of 148 and GCS of 339 and is cited by 13 other highly cited papers (nodes 130, 172, 772, 782, 795, 812, 848, 904, 962, 965, 994, 1015 and 1029). The third most cited paper is by GB Nair *et al*., 1994, on 'Spread of *Vibrio-cholerae* O139 Bengal in India' with LCS 59 and GCS 109. The latest among the top 50 LCS papers is the one by Singh *et al.* in 2001 on 'Molecular analysis of *Vibrio cholerae* O1, O139, non-O1, and non-O139 strains: Clonal relationships between clinical and environmental isolates' with LCS 18. This implies that active research on cholera is still ongoing.

There is one major network with extensive links spanning the period 1987-2001. We see a lot of active research on isolation and analysis of different cholera strains being reported. Based on top LCS records, we find that papers reporting research on different strains of *Vibrio cholerae* like O139 by G B Nair *et al.*, NON-01 by Shimada *et al.*, isolation of several clinical strains of *Vibrio cholerae*, serotyping scheme of *Vibrio cholerae* by Shimada *et al.*, and virulence patterns of *Vibrio cholerae* strains by T Ramamurthy *et al.* have been among the highly cited papers within cholera-India papers. Some of the noted scientists besides S N De are G B Nair, T Ramamurthy, T Shimada, Y Takeda and S K Bhattacharyya of the

National Institute of Cholera and Enteric Diseases, Calcutta, and N K Dutta of the Haffkine Institute, Bombay. Their researches have had high impact within the cholera-India publications. Among the top 50 LCS papers, a few are not linked to any other paper. We also see a relatively small network with just three nodes and three links that includes papers by SS Visveswariah *et al.* on heat stable enterotoxin that has not been pursued by other leading cholera researchers.

Soren Paris points out, "Interestingly, collections that are based on topic as opposed to a specific researcher often have higher frequency of isolates. On the surface, this means that the papers citing these isolates, are not themselves cited enough to get onto the map. If we increased the size of the map, more connections would appear, but just as likely, other new isolates would appear." (Private communication)

Historiograph from the top 50 Cholera India papers (IP collection) by GCS

The top 50 GCS papers – the most cited cholera research papers from India - span the period from 1949 to 2007 with one major network and several islands. This historiograph (figure not given) shows networks connecting papers that have a high GCS, papers receiving a large number of citations from both within the IP collection and outside. There are only 53 links with GCS ranging between maximum 585 and minimum 52. The major network connects all publications on cholera having top GCS (also having impact in other related areas). Four of De's papers have high GCS. Out of the 585 global citations to node 21 (De's 1953 paper), 131 are from within 1,750 cholera papers from India. The rest could be from non-cholera papers from India, cholera papers from the rest of the world and non-cholera papers from the rest of the world. In retrospect, even more important than this paper is De's 1959 *Nature* paper on the enterotoxicity of bacteria-free culture-filtrate of *Vibrio cholerae*,[4] which went virtually unnoticed in the first five years, but later on went on to accumulate 170 citations - clearly a case of delayed recognition.[1,13] Writing about this paper in which De demonstrated for the first time that cholera bacteria secreted enterotoxin, W E van Heyningen and John R Seal said that it "deserves to go down as a classic in the history of cholera, and, indeed, as later developments have shown, in the history of cellular physiology and biochemistry."[14] Around the same time NK Dutta *et al*. of the Haffkine Institute in Bombay published a paper on the role of cholera toxin in experimental cholera in *Journal of Bacteriology*.[15] One is occasionally confronted with the question "Should De share honours with Dutta for the discovery of the cholera toxin?" For the record, the paper by De in *Nature* was submitted and published earlier than that of Dutta as well as cited more often. Apart from this priority in discovery, with his earlier work on developing a technique for studying the mechanism of action of *V. cholerae* on the intestinal mucous membrane, De had broken new grounds and had brought about a paradigm shift.

We observe, as noticed in Fig. 1, a long hiatus between 1973 and 1990. Then appears the paper by T Ramamurthy *et al.* on novel strains of *Vibrio cholerae* (node 761), which has the second most GCS, viz. 339, among all Indian papers on cholera. Among the top GCS papers, the paper by RB Sack *et al.* on enterotoxigenic *Escherichia coli*, the paper by DE Schafer *et al*. on treatment with cholera toxin, and the paper by NK Dutta *et al*. on experimental cholera in infant rabbits have high GCS. Papers by SN De, T Ramamurthy and NK Dutta have both high LCS and high GCS. We also observe that the paper by S Mayor *et al*. on 'Pathways of clathrin-independent endocytosis' though has not been cited by Indian cholera researchers (LCS = 0), is in the sixth position with GCS 177 and Cameron's paper on pulmonary oedema is cited 148 times but only two times by Indian cholera researchers. Thus we see papers by authors like DE Schafer, S Mayor, and GR Cameron mainly had high impact in other areas of research but single-digit LCS. This is true of most papers by non-Indian, mostly western, authors in the table of papers with high GCS, probably because they are not in the main papers on cholera. The exceptions are two papers by Finkelstein, both of which are on cholera research and he had Indian coauthors (N K Dutta and S Mukherjee).

IPC collection – Historiograph by top 50 LCS

The IPC collection has 10,242 records comprising publications from India on cholera research and all citations indexed in *WoS* to these papers. In the

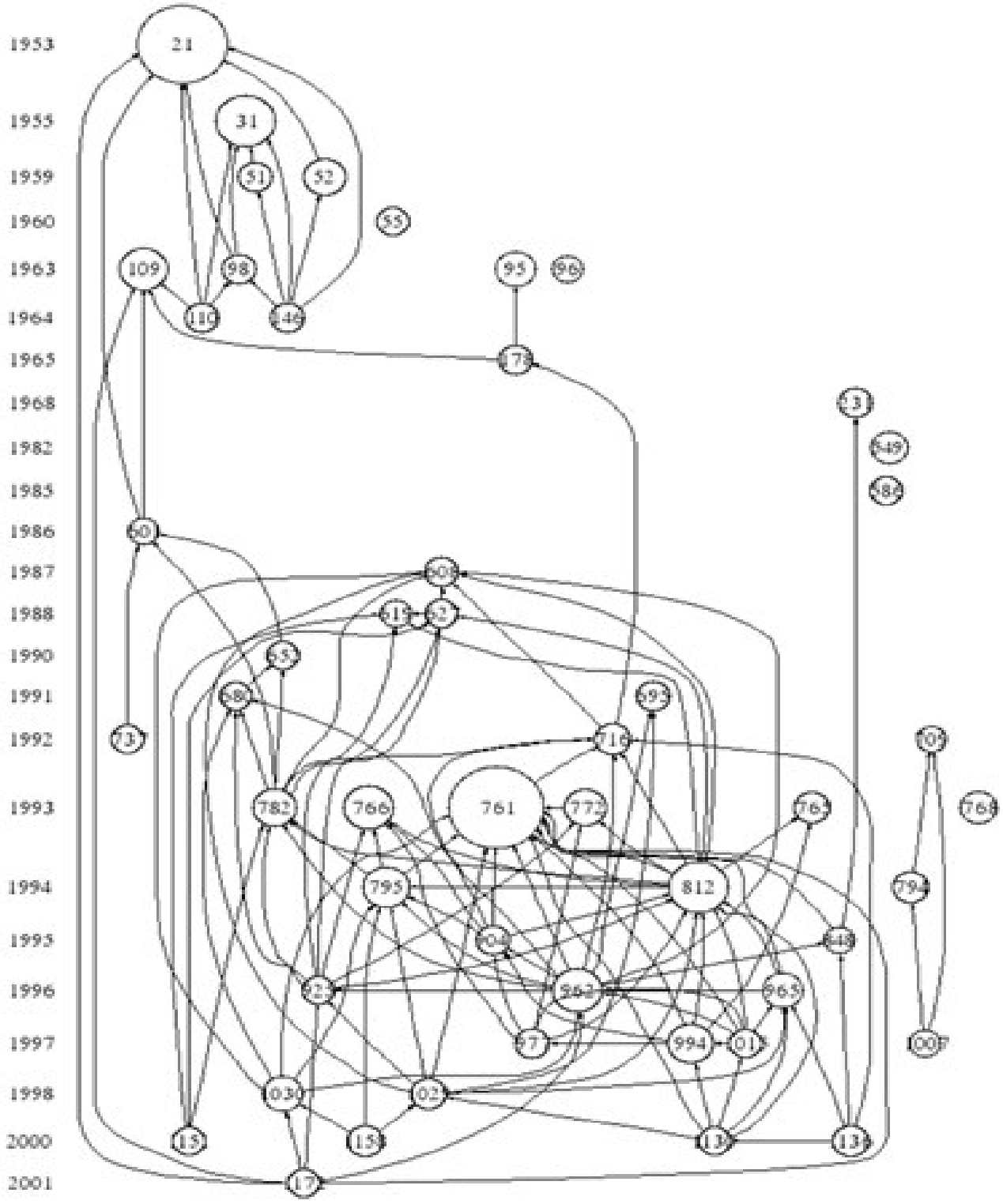

Fig. 1 —- Cholera-India papers (IP collection) - Historiograph by top 50 LCS

Nodes: 50, Links: 124
LCS, top 50; Min: 15, Max: 148 (LCS scaled)

#	Rec.	Bib Inf.	LCS	GCS
1.	21	De SN, 1953, J Pathol Bacteriol, V66, P559	131	585
2.	31	Dutta NK, 1955, Brit J Pharmacol Chemotherapy, V10, P153	58	216
3.	51	Dutta NK, 1959, J Bacteriol, V78, P594	20	95
4.	52	De SN, 1959, Nature, V183, P1533	29	175
5.	55	Bhaskaran K, 1960, J Gen Microbiol, V23, P47	18	68
6.	95	Mukerjee S, 1963, Bull WHO, V28, P333	28	70
7.	96	Mukerjee S, 1963, Bull WHO, V28, P337	16	37
8.	98	Mukerjee S, 1963, Bull WHO, V29, P753	21	43
9.	109	Finkelstein RA, 1963, Proc Soc Exp Biol Med, V112, P355	40	121
10.	110	Mukerjee S, 1964, Brit Med J, V2, P546	21	23
11.	146	Finkelstein RA, 1964, J Infec Dis, V114, P203	21	131
12.	178	Roy C, 1965, Proc Soc Exp Biol Med, V119, P893	20	37
13.	231	Pierce NF, 1968, Brit Med J, V3, P277	21	41
14.	549	Roy NK, 1982, J Gen Microbiol, V128, P1927	26	32
15.	586	Lohia A, 1985, J Bacteriol, V163, P1158	19	24
16.	601	Dattaroy K, 1986, Appl Environ Microbiol, V52, P875	16	43
17.	608	Nair GB, 1987, Appl Environ Microbiol, V53, P1203	20	24
18.	619	Nair GB, 1988, Microbial Ecol, V15, P203	16	37
19.	627	Nair GB, 1988, Appl Environ Microbiol, V54, P3180	19	42
20.	652	Morris JG, 1990, J Clin Invest, V85, P697	18	76
21.	680	Takeda T, 1991, FEMS Microbiol Lett, V80, P23	19	31
22.	695	Shirai H, 1991, J Clin Microbiol, V29, P2517	20	94
23.	709	Visweswariah SS, 1992, Microb Pathog, V12, P209	15	17
24.	716	Ramamurthy T, 1992, Indian J Med Res-A, V95, P125	22	24
25.	737	Sengupta DK, 1992, Infec Immunity, V60, P4848	18	64
26.	761	Ramamurthy T, 1993, Lancet, V341, P703	148	339
27.	765	Bhattacharya MK, 1993, Lancet, V341, P1346	22	44
28.	766	Shimada T, 1993, Lancet, V341, P1347	42	116
29.	768	Chattopadhyay DJ, 1993, J Clin Microbiol, V31, P1579	26	37
30.	772	Bhattacharya SK, 1993, J Infection, V27, P11	31	58
31.	782	Ramamurthy T, 1993, J Med Microbiol, V39, P310	32	52
32.	794	Visweswariah SS, 1994, Eur J Biochem, V219, P727	24	30
33.	795	Shimada T, 1994, Curr Microbiol, V28, P175	39	91
34.	812	Nair GB, 1994, J Infec Dis, V169, P1029	59	109
35.	848	Yamamoto T, 1995, Antimicrob Agents Chemother, V39, P241	17	35
36.	904	Mukhopadhyay AK, 1995, Epidemiol Infect, V115, P427	21	32
37.	925	Saha PK, 1996, J Clin Microbiol, V34, P1114	18	28
38.	962	Mukhopadhyay AK, 1996, J Clin Microbiol, V34, P2537	43	61
39.	965	Mitra R, 1996, Lancet, V348, P1181	27	44
40.	977	Yamasaki S, 1997, Microbiol Immunol, V41, P1	20	40
41.	994	Sharma C, 1997, J Infec Dis, V175, P1134	35	53
42.	1007	Nandi A, 1997, J Cell Biochem, V66, P500	16	27
43.	1015	Sharma C, 1997, J Clin Microbiol, V35, P3348	19	33
44.	1029	Hoshino K, 1998, FEMS Immunol Med Microbiol, V20, P201	22	52
45.	1030	Sharma C, 1998, J Clin Microbiol, V36, P756	28	70
46.	1130	Faruque SM, 2000, FEMS Microbiol Lett, V184, P279	18	25
47.	1136	Garg P, 2000, Epidemiol Infect, V124, P393	22	46
48.	1151	Chakraborty S, 2000, Appl Environ Microbiol, V66, P4022	18	70
49.	1158	Nandi B, 2000, J Clin Microbiol, V38, P4145	24	67
50.	1172	Singh DY, 2001, Appl Environ Microbiol, V67, P910	18	63

Please note that the number representing a particular paper need not be the same in the historiographs generated from the IP and the IPC collections. For example, SN De's 1953 paper is represented by node 21 in Fig. 1 and node 25 in Fig. 2.

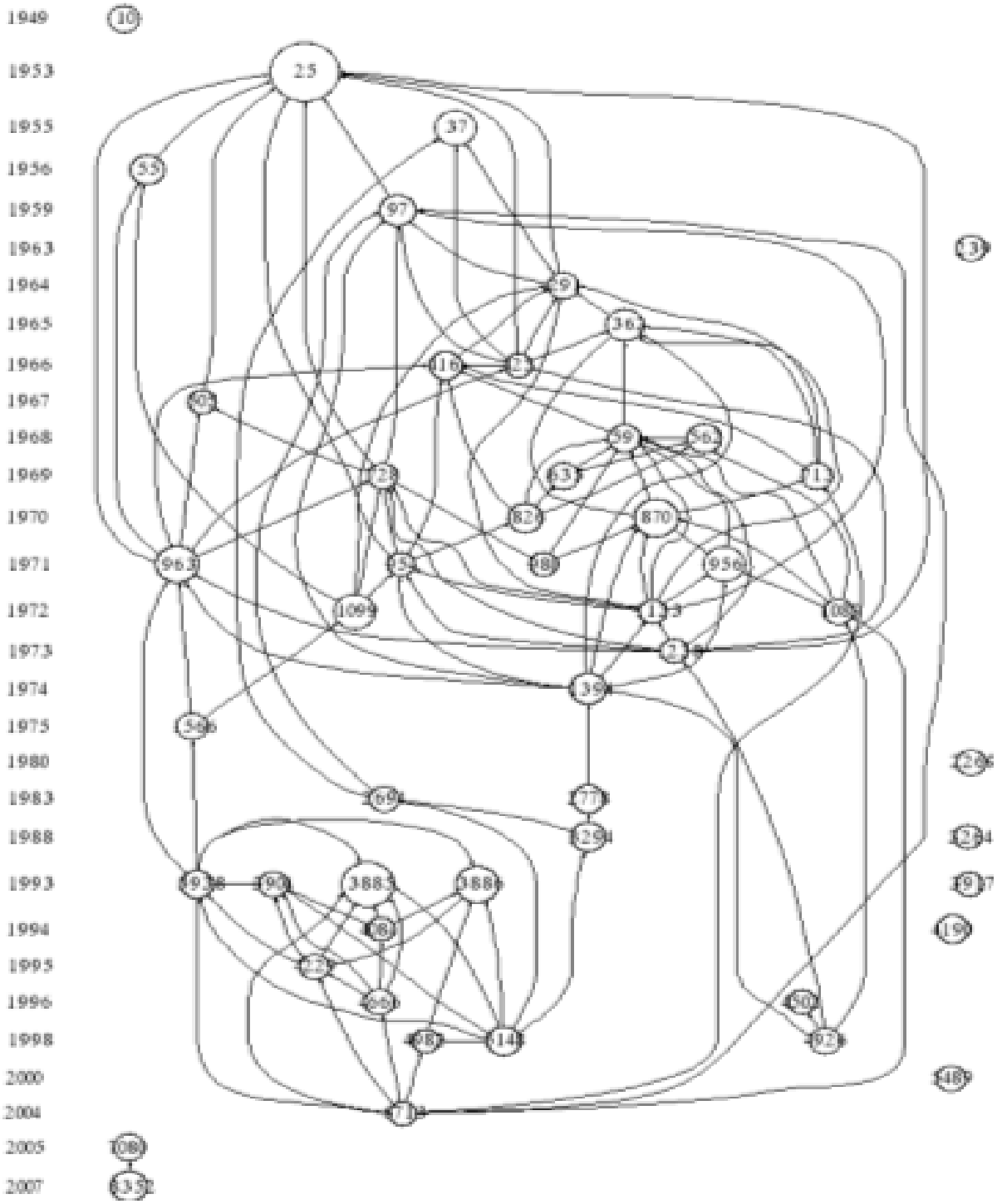

Fig. 2 — Cholera-India-collection (IPC) - Historiograph by top 50 LCS

Nodes: 50, Links: 121

LCS, top 50; Min: 101, Max: 584 (LCS scaled)

#	Rec.	Bibliographic Inf.	LCS	GCS
1.	10	Cameron GR, 1949, J Pathol Bacteriol, V61, P375	129	148
2.	25	De SN, 1953, J Pathol Bacteriol, V66, P559	584	585
3.	37	Dutta NK, 1955, Brit J Pharmacol Chemotherapy, V10, P153	218	216
4.	55	De SN, 1956, J Pathol Bacteriol, V71, P201	153	154
5.	97	De SN, 1959, Nature, V183, P1533	176	175
6.	239	Finkelstein RA, 1963, Proc Soc Exp Biol Med, V112, P355	121	121
7.	291	Finkelstein RA, 1964, J Infec Dis, V114, P203	132	131
8.	362	Craig JP, 1965, Nature, V207, P614	187	330
9.	416	Finkelst.RA, 1966, J Immunol, V96, P440	139	220
10.	425	Kasai GJ, 1966, J Infec Dis, V116, P606	103	149
11.	505	Smith HW, 1967, J Pathol Bacteriol, V93, P531	109	303
12.	562	Pierce NF, 1968, Gastroenterology, V55, P333	147	147
13.	591	Carpente.CC, 1968, J Clin Invest, V47, P1210	140	215
14.	635	Pierce NF, 1969, Ann Intern Med, V70, P1173	128	129
15.	713	Finkelst.RA, 1969, J Exp Med, V130, P185	125	252
16.	728	Gyles CL, 1969, J Infec Dis, V120, P419	115	273
17.	826	Banwell JG, 1970, J Clin Invest, V49, P183	146	146
18.	870	Schafer DE, 1970, Proc Nat Acad Sci USA, V67, P851	242	242
19.	954	Gorbach SL, 1971, J Clin Invest, V50, P881	111	180
20.	956	Kimberg DV, 1971, J Clin Invest, V50, P1218	211	588
21.	963	Sack RB, 1971, J Infec Dis, V123, P378	241	243
22.	980	Field M, 1971, N Engl J Med, V284, P1137	107	226
23.	1089	Field M, 1972, J Clin Invest, V51, P796	115	270
24.	1099	Dean AG, 1972, J Infec Dis, V125, P407	222	1024
25.	1133	Evans DJ, 1972, Nature-New Biol, V236, P137	106	246
26.	1238	Evans DG, 1973, Infec Immunity, V7, P873	110	279
27.	1394	Guerrant RL, 1974, Infec Immunity, V10, P320	168	579
28.	1566	Sack DA, 1975, Infec Immunity, V11, P334	123	450
29.	2269	Blake PA, 1980, Annu Rev Microbiol, V34, P341	119	464
30.	2694	Mekalanos JJ, 1983, Cell, V35, P253	119	300
31.	2778	Mekalanos JJ, 1983, Nature, V306, P551	137	505
32.	3264	Levine MM, 1988, Infec Immunity, V56, P161	101	192
33.	3294	Miller VL, 1988, J Bacteriol, V170, P2575	159	1299
34.	3885	Ramamurthy T, 1993, Lancet, V341, P703	340	339
35.	3886	Albert MJ, 1993, Lancet, V341, P704	236	255
36.	3904	Shimada T, 1993, Lancet, V341, P1347	116	116
37.	3917	Keasler SP, 1993, Lancet, V341, P1661	105	110
38.	3938	Albert MJ, 1993, Lancet, V342, P387	140	163
39.	4084	Nair GB, 1994, J Infec Dis, V169, P1029	109	109
40.	4190	Karunasagar I, 1994, Aquaculture, V128, P203	143	143
41.	4229	Bik EM, 1995, EMBO J, V14, P209	121	166
42.	4502	Savarino SJ, 1996, J Infec Dis, V173, P1019	103	104
43.	4666	Colwell RR, 1996, Science, V274, P2025	119	363
44.	4926	Nataro JP, 1998, Clin Microbiol Rev, V11, P142	131	1553
45.	4982	Karaolis DKR, 1998, Proc Nat Acad Sci USA, V95, P3134	112	207
46.	5148	Faruque SM, 1998, Microbiol Mol Biol Rev, V62, P1301	164	259
47.	5489	Matsumoto C, 2000, J Clin Microbiol, V38, P578	119	119
48.	6713	Sack DA, 2004, Lancet, V363, P223	103	166
49.	7080	Kirkham M, 2005, J Cell Biol, V168, P465	126	126
50.	8352	Mayor S, 2007, Nat Rev Mol Cell Biol, V8, P603	177	177

historiograph based on the top 50 LCS papers - most cited among 10,242 papers within the collection - there is one major network with 121 links (Fig. 2). The top 50 LCS papers span the period 1949-2007 with maximum LCS of 585 and minimum LCS of 101. However, there are only four papers between 1949 and 1959. From 1959 to 2004, we see many links in the network implying a larger number of influential papers on cholera being published. Unlike in Fig. 1 (which is based on publications on cholera from India only), we do not see a large period of inactivity in Fig. 2. Researchers elsewhere in the world were publishing influential papers. The earliest important paper included in the network, is the 1953 paper by S N De on the mechanism of action of *Vibrio cholerae* with highest LCS 584 and is cited by nine other highly cited papers (nodes 55, 97, 291, 425, 505, 728, 954, 963 and 1238). This paper is followed by a paper by T Ramamurthy *et al*. in 1993 on a novel strain of *Vibrio cholerae* with epidemic potential, with an LCS of 340 which is cited by six other highly cited papers (nodes 3938, 4084, 4229, 4666, 5148 and 6713). 'Elevated concentration of adenosine 3'-5'-cyclic monophosphate in intestinal mucosa after treatment with cholera toxin' by DE Schafer *et al*. and 'Enterotoxigenic *Escherichia-coli* isolated from patients with severe cholera-like disease' by RB Sack *et al*. have an LCS of 242 and 241 respectively. Papers with large number of citations have reported research on enterotoxicity of V*ibrio cholerae*, novel strains of *Vibrio cholerae*, experimental studies with cholera toxin, enterotoxigenic *Escherichia coli*, *Vibrio cholerae* non-01, cholera sera, clathrin independent endocytosis, epidemiological studies of *Vibrio cholerae*, and pathogenicity of strains of bacterium *coli* among others. Highly cited and influential papers have been authored by S N De, T Ramamurthy, NK Dutta, DE Schafer, RB Sack, and MJ Albert (names of first authors only are given). We also observe that most of the top LCS papers have the same LCS and GCS, implying their research has impact only on cholera research. AG Dean's paper on a test for *Escherichia-coli* enterotoxin using infant mice has great impact in other areas too, with GCS 1024 and LCS 222.

IPC collection – Historiograph by top 50 GCS

This historiograph (not shown here) has only 50 links with maximum GCS of 3,789 and minimum GCS of 328. This is because there are not many cross-citations within these highly cited papers. There is one network with few links among them and several islands. This is because the papers report research not only on cholera but also on related areas of research. The top 50 GCS publications span the period 1953-2004. JP Craig's paper published in *Nature*, 1965, on toxin in cholera stools is the earliest on the network. Most recent paper on the network is the 2000 paper by J Hacker *et al.* on the pathogenicity islands and the evolution of microbes. Among the top 10 papers by GCS, only S N De figures from India. The 1995 paper by FC Tenovar *et al.* on interpreting chromosomal DNA restriction patterns has the highest GCS of 3,789 followed by the paper by JP Nataro *et al.* on diarrheagenic *Escherichia coli* and the paper by VL Miller *et al.* on a novel suicide vector and its use in construction of insertion mutations with GCS 1,553 and 1,299 respectively. BD Spangler's paper on 'Structure and function of cholera-toxin and the related *Escherichia-coli* heat-labile enterotoxin' is a review on all the previous research done till 1992 as it cites many papers on the network. There are several unconnected nodes that have no impact on the papers in the existing network.

Conclusion

HistCite analysis of cholera research in India since 1945 has thrown some light on how the field has evolved. It has provided a simple visualization of the formation of a cognitive network of interconnected papers and who takes the baton from whom. It has helped us identify research that had impacted only other cholera researchers and research that had a wider impact on science as a whole. Clearly institutions such as NICED, Calcutta, and ICDDR, Dhaka, are doing much valuable research. But HistCite analysis using the larger IPC collection comprising both publications from India and citations to them has shown that papers with high impact come from the West. The average number of citations per cholera research paper is around 13.2 for NICED, India's leading institution for cholera research, as against upwards of 30 for the University of Maryland, Johns Hopkins University, Centre for Disease Control and the University of Texas and above 52 for Harvard University. That is an area Indian cholera researchers need to pay attention.

The HistCite analysis of the 1,750 papers from India on cholera has shown us the value of key papers such as the 1953 paper by S N De and D N Chatterjee on animal model for studying cholera, the 1993 paper on

'novel strain of *Vibrio cholerae'* by T Ramamurthy *et al*. and the 1994 paper on 'spread of *Vibrio-cholerae* O139 Bengal' by G B Nair *et al*. There have been other influential papers by NK Dutta, DE Schafer, RB Sack and MJ Albert. But it has also shown that the arguably more important paper of De (the 1959 *Nature* paper) has not received as many citations as his 1953 paper describing a novel technique. In fact, it turned out to be case of delayed recognition.

We believe such studies by people outside cholera research can at best be a supplement to experts' views, such as the overview on cholera research in India by Das[16] and the article on Sambhu Nath De's legacy by Nair and Narain.[17] One wonders had such a study been performed in the 1960s, would De have received recognition for his outstanding contributions? Or if people were unable to see such brilliance, what could a mere prop like bibliometric analysis have achieved? One is left wondering if standards of peer review have improved in India since the days when De's work was totally ignored. Very few in India had known or heard about Venkatraman Ramakrishnan and his work till the Nobel Committee jolted them from their slumber last October. Once he won the Prize, of course, the whole nation was after him for a few weeks.

Acknowledgment

We are grateful to Alexander Pudovkin, Subbiah Gunasekaran, Soren Paris and Nico Rasters for their critical reading of an earlier version of the manuscript and suggestions for improvement. Eugene Garfield was kind enough to read the manuscript and made valuable suggestions. Part of the work reported was carried out when Arunachalam was a visiting professor at the Institute of Mathematical Sciences, Chennai, in the latter half of 2009.

References

1 Garfield E, Mapping cholera research and the impact of Sambhu Nath De of Calcutta, *Current Contents* 14(1986) 3-11. Reproduced in *Essays of an information scientist*, Vol. 9 (1986), pp.103-111; www.garfield.library.upenn. edu/essays/v9p103y1986.pdf.

2 De S N and Chatterje D N, An experimental study of the mechanism of action of Vibrio cholerae on the intestinal mucous membrane. *Journal of Pathology and Bacteriology*, 66 (1953), 559-562.

3 De SN, Bhattacharya K and Sarkar, JK, A study of the pathogenicity of strains of Bacterium coli from acute and chronic enteritis. *Journal of Pathology and Bacteriology*, 71 (1956), 201-209.

4 De SN, Enterotoxicity of bacteria-free culture-filtrate of *Vibrio cholerae*. *Nature*, 183 (1959) 1533-1534.

5 Balaram P, Sambhu Nath De. *Current Science*, 1990; 59: 627.

6 HistCite, http://www.histcite.com/index.htm.

7 Garfield E, Pudovkin AI and Istomin VS, Why do we need algorithmic historiography? *Journal of the American Society for Information Science and Technology*, 54 (2003) 400–412.

8 Available at: http://garfield.library.upenn.edu/papers/jasist 54(5)400y2003.pdf.

9 Garfield E, Historiographic mapping of knowledge domains literature, *Journal of Information Science*, 30 (2) (2004) 119–145; DOI: 10.1177/0165551504042802. Historiograph Compilation *HistCite* Guide, <http://www.garfield.library.upenn.edu/histcomp/guide.html>.

10 Kanungo S Sah BK, Lopez AL, Sung JS, Paisley AM, Sur D, Clemensb JD & Balakrish Nair G, Cholera research in India: an analysis of reports, 1997-2006, *Bulletin of the World Health Organization*, 88 (2010), 185-191.

11 Arunachalam S and Thulasi K, Remembering Sambhu Nath De: An assessment of the impact of his work on cholera research, *Indian Journal of Medical Research*, 131 (2010) (to be published).

12 Craig JP, S.N.De - An appreciation, *Current Science*, 59 (1990), 637.

13 Dutta, NK and Habbu MK, Experimental cholera in infant rabbits: A method for chemotherapeutic investigation. *British Journal of Pharmacology and Chemotherapy*, 10 (1955), 153-159.

14 van Heyningen WE and Seal JR, *Cholera: the American scientific experience*, *1947 - 1980.* Boulder, CO: Westview Press, 1983. p. 61.

15 Dutta NK, Panse MV, Kulkarni DR, Role of cholera toxin in experimental cholera, *Journal of Bacteriology*, 78 (4) (1959) 594-595.

16 Das J, Cholera research in India - An overview, *Current Science*, 59 (1990), 671-676.

17 Balakrish Nair G and Narain J P, From endotoxin to exotoxin: De's rich legacy to cholera, *Bulletin of the World Health Organization,* 88 (2010) 237–240.