

Scientific research in Bangladesh and a comparison with India and Pakistan

Dilruba Mahbuba^{1,2} Ronald Rousseau^{3,4}

16 July 2008

Abstract

As part of a research program analysing research in Bangladesh we provide a comparison between research indicators related to India, Bangladesh and Pakistan. In this investigation we make use of Web of Science (WoS) data as well as Scopus data (using the SCImago website). Special attention is given to collaboration data and to the evolution of country h-indices. Lists of most-cited (in the WoS) articles are provided. These lists clearly show that India often is a partner in large multinational research groups (receiving a considerable amount of citations in this way), while this is not the case for Pakistan and Bangladesh.

1 Introduction

Many recent articles compare India with China, see e.g. (Arunachalam, 2008; Balaram, 2004; Guan & Ma, 2004). Here we take another point of view comparing a smaller neighbouring country (Bangladesh) with a local giant (India) and another neighbour, namely Pakistan. We review developments in scientific research between 1973 and 2006 in these three South Asian countries Bangladesh, India and Pakistan (in short: the BIP-countries). How do Pakistan and Bangladesh compare with each other and with the Indian giant? How can countries with such a low literacy rate (see Table 1), struggling against

poverty and adverse weather conditions compete in the international scientific fray?

Table 1 Literacy rates of Bangladesh-India-Pakistan

	B	I	P
Total	45.3	64.8	43.9
Male	49.6	75.3	54.8
Female	40.8	53.7	32.1

We intend to clarify some of these questions using basic scientometric tools, measuring the evolution of scientific research in these three countries during the period 1973-2006 (recall that Bangladesh became independent in 1971). The Web of Science and the SCImago Journal & Country Rank database, based on SCOPUS data (<http://www.scimagojr.com/>) are our main sources.

We note that studies related to these countries have been published in the past. Sadana et al. (2004), for instance, note that collaboration on health research across the region (the BIP-countries, including Afghanistan, Nepal, Bhutan, Sri Lanka and the Maldives) must be strengthened. Because of the large number of people involved (India on its own is the second most populous country in the world) and the high disease burden on the population health related research is very important for these countries. We will return to this point further on

¹ Executive Director's Division, Library and Information Services Unit (LISU); ICDDR,B; Mohakhali, Dhaka 1212, Bangladesh; ruba@icddr.org

² Antwerp University (UA), IBW, Venusstraat 35, B-2000, Antwerp, Belgium

³ KHBO (Association K.U.Leuven), Industrial Sciences and Technology, B-8400, Oostende, Belgium; ronald.rousseau@khbo.be

⁴ K.U.Leuven, Steunpunt O&O Indicatoren, Dekenstraat 2, B-3000 Leuven, Belgium

in this article when discussing some special aspects related to Bangladesh. B.M. Gupta et al. (2002) studied collaborations of India with Bangladesh, Pakistan, Sri Lanka and Nepal in the period 1992-1999, based on WoS data. The strongest collaborative linkages are with Bangladesh. India-Bangladesh collaborative articles have also the highest impact. These authors also found that geographical proximity played an important role. Of the 79 institutional collaborative linkages between India and Bangladesh 32 were between neighbouring West Bengal and Bangladesh. The strongest collaboration ties, however, were between ICDDR, B (Dhaka) and NICED (Kolkata).

2 WoS publication data: a comparison between BIP-countries

Figure 1 shows the number of publications per year (all document types) included in the WoS with a least one Bangladeshi or Pakistani address. As India's number of publications is much higher (more than 31,000 publications in the WoS in the year 2006) the numbers shown in Figure 1 are India's number of publications divided by ten. In this way trends can be compared. During the period 1973-2002 the number of publications in Bangladesh and Pakistan increased linearly, be it at a somewhat higher rate for Pakistan than for Bangladesh. Since 2002 the number of WoS publications increased at a much higher rate in both countries, but is, especially for Pakistan really spectacular. A similar phenomenon is true for India (with an increase rate between that of Bangladesh and Pakistan).

3 SCImago data

Table 2 shows, for countries in the Southeast Asian region, the number of citations per document over the period 1996-2006, the h-index over the same period and the number of citations per document for the year 2006. Countries are ranked according to the h-index over the whole period. Data are provided by SCImago, based on Scopus.

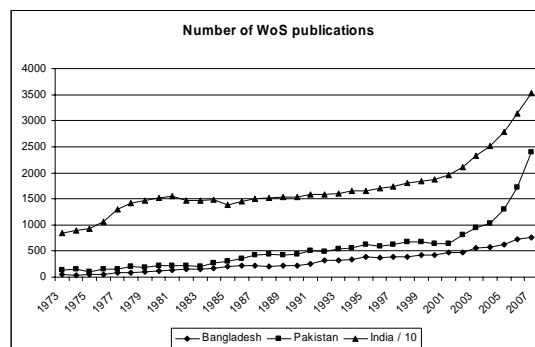


Fig. 1 Bangladesh, Pakistan and India: number of WoS publications

Table 2. Southeast Asian countries with the highest number of publications included in Scopus (according to SCImago Journal & Country Rank).

A	B	C	D	E
1	Japan	7.55	372	0.33
2	South Korea	5.84	161	0.22
3	China	3.14	161	0.12
4	India	3.91	146	0.19
5	Taiwan	5.53	139	0.22
6	Hong Kong	7.07	135	0.29
7	Singapore	6.35	114	0.31
8	Thailand	5.70	84	0.23
9	The Philippines	6.73	58	0.40
10	Malaysia	4.04	56	0.24
11	Indonesia	5.75	52	0.39
12	Vietnam	6.28	52	0.37
13	Pakistan	2.80	45	0.16
14	Bangladesh	3.73	40	0.24

A: Rank; B: Country; C: Citations per document, period 1996-2006; D: h-index, period 1996-2006; E: Citations per document, year 2006

Source URL: <http://www.scimagojr.com>. SCImago Research Group. Copyright 2007. Data Source: Scopus®

For this period Scopus contains 6,447 Bangladeshi documents. These articles received 21,794 citations among which 4462 are country self-citations. The h-index for Bangladesh is 40, 45

H. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin

Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting
Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

This is an Open Access document licensed under the Creative Commons License BY

<http://creativecommons.org/licenses/by/2.0/>

for Pakistan and 146 for India. According to the h-index for this period, Bangladesh occupies the 14th position in the region. Yet, its number of citations per publication over the latest ten years is 3.73. This is higher than Pakistan's and even China's, but not as high as India's. Japan's average number of citations per publications (Japan being the scientific leader in the region), however, is more than twice Bangladesh'.

We note that SCImago Journal & Country Rank, based on SCOPUS data (2006), includes 4 journals from Bangladesh, 22 from Pakistan and 157 from India.

4 Collaboration data

Table 3 shows the countries with which Bangladesh and Pakistan collaborate most over the whole period 1973-2006 (WoS data). Table 4 focuses on recent times showing collaboration data since the year 2000. Remarkably, the top-10 countries for the whole period (Table 3) are almost the same, the only difference being that Australia in the Bangladeshi list is replaced by China in the Pakistani list.

Table 3 Collaboration data: 1973-2006 (WoS)

	Bangladesh	Pakistan
1	USA – 1,648	USA – 1,928
2	UK – 937	UK – 1,547
3	Japan – 842	Germany – 650
4	India – 489	Canada – 362
5	Sweden – 319	Japan – 320
6	Germany – 218	China – 302
7	Australia – 214	Sweden – 245
8	Canada – 156	Italy – 243
9	Italy – 128	India – 217
10	Switzerland - 123	Switzerland - 194

Moreover, Bangladesh and Pakistan have collaborated on 60 articles in the period 2000-2007. Some interesting points: the UK lost its second position in the latest period (Germany in India and Japan in Bangladesh are second). It seems that historic ties slowly lose their grip (Zitt et al., 2000). South Korea has entered the top-10 list in Pakistan and India. France is fifth in India, but not among the top-10 for Bangladesh or Paki-

stan. Similarly, Sweden is fifth in the Bangladesh list, but, in the most recent period, not in the top-10 for the other two countries. It dropped out of the Pakistani list.

Table 4 Collaboration data: 2000-2007 (WoS)

	Bangladesh	Pakistan	India
1	USA – 877	USA – 970	USA – 13,748
2	Japan – 655	UK – 779	Germany – 5,540
3	UK – 482	Germany – 415	UK – 4,516
4	India – 262	Canada – 260	Japan – 3,999
5	Sweden – 167	China – 249	France – 2,928
6	Germany – 156	Japan – 211	Canada – 1,963
7	Australia – 130	Italy – 156	South Korea – 1,947
8	Canada – 86	India – 145	China – 1,819
9	Malaysia – 86	Saudi Arabia – 132	Italy – 1,645
10	Netherlands - 83	South Korea - 121	Australia – 1,516

5 Evolution of country h-indices

In this section we discuss country h-indices. Recall that a country's h-index is defined as the largest natural number (representing a rank) such that the first h publications with at least one (co-)author with an address in this country received each at least h citations. The first h articles in such a ranked list form the country's h-core (Hirsch, 2005; Liu & Rousseau, 2008). Figure 2 shows the yearly h-index of these three countries. By the term 'yearly h-index' we mean the h-index in May 2008 of articles with a least one Bangladesh (Pakistan, India) address and published in the year indicated on the horizontal axis. Such series are called Type 1 in the Liu-Rousseau (2008) classification of time series. It is remarkable that these time series are not decreasing. For Bangladesh we see a platform over a long period: between 1983 and 2002 the h-index stays between 22 and 25, with the only

H. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin

Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting

Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

This is an Open Access document licensed under the Creative Commons License BY

<http://creativecommons.org/licenses/by/2.0/>

exception of year 1995, when it is 27. Pakistan's type I h-index increases slowly from 12 in 1973 to 28 in 2000. A similar behaviour is true for India's h-index: it increases from 56 in 1973 to 86 in 2001. What happens when we normalize these h-indices by dividing by the number of publications? Do we get decreasing curves? Results are shown in Fig. 3. By and large we obtain decreasing curves for Bangladesh and Pakistan. For these countries the increase in the number of publications compensated the shorter time period over which citations could be received. India's curve (h/publications, multiplied by ten), however stays practically flat from 1977 till 2000.

Recall from Table 2 that the 2006 h-indices for the BIP-countries, according to SCImago (based on Scopus) are as follows: India - 146; Pakistan - 45 and Bangladesh - 40. In the WoS, Bangladesh's h-index for the period 1973-2007 is 84, while Pakistan's is 72; India's h-index is 214.

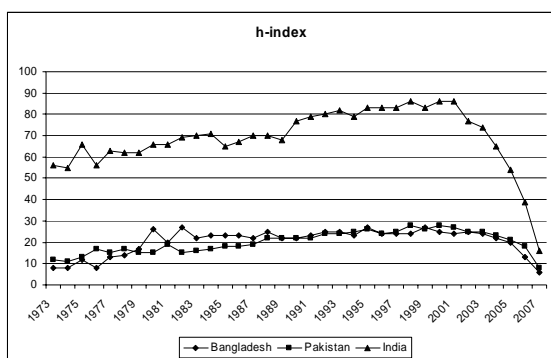


Fig. 2 Country h-index

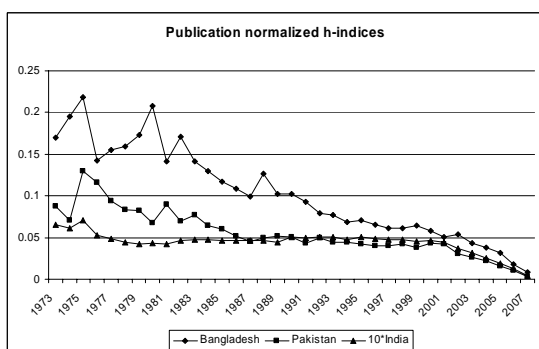


Fig. 3 Normalized country h-index

6 Most-cited articles in the WoS

In this section we list the most-cited articles published by authors of the BIP-countries. At the end of this section we formulate some observations.

A. India: period 1973-1987

The most cited article with an Indian address is:

Disordered electronic systems, written by PA Lee and TV Ramakrishnan, and published in 1985 in *Reviews of Modern Physics*, 57(2), p.287-337, which has been cited 2,909 times. Lee's address, however, was MIT (Ma, USA).

The most-cited article with only Indian addresses is the second most-cited one during this period:

Transparent conductors – a status review, written by KL Chopra, S. Major and DK Pandya, published in 1983 in *Thin Solid Films*, 12(1), p. 1-46. It has been cited 1,095 times.

India: period 1988-1997

Here the most-cited article is a single-authored one (with only an Indian address) and is, like the two previous ones, also a review.

Supramolecular synthons in crystal engineering – a new organic synthesis, written by GR Desiraju and published in 1995 in *Angewandte Chemie – International edition in English*, 34(21), p. 2311-2327. It has been cited 1,718 times.

India: period 1998-2008

The most-cited article with an Indian address is:

Review of particle physics, written by S. Eidelman et al. in 2004 in *Physics Letters B*, 592(1-4), p. 1-1109. This is a long review written by 148 authors, with only one Indian address among a long list of addresses. It has been cited 2,881 times.

The most-cited purely Indian contribution during this period is:

GPI-anchored proteins are organized in submicron domains at the cell surface, published in

1998 in *Nature*, 394(6695), p. 798-801, and cited 610 times.

Note that this publication is of the *article* type, and hence not a review.

B. Bangladesh: period 1973-1987

The most-cited article is:

Detection of entero-toxigenic *Escherichia-coli* by DNA colony hybridization, written by SL Moseley, I. Huq, Arma Alim, M. So, M. Samad-pourmotalebi and S. Falkow, in 1980 in *Journal of Infectious Diseases*, 142(6), 892-898. It has been cited 335 times.

This is indeed an article and not a review. The most-cited purely Bangladesh article turned out to be number 24 of the list.

Interruption of shigellosis by hand washing, written by MU Khan in 1982 in *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 76(2), p. 164-168. It has been cited 130 times.

Bangladesh: period 1988-1997

The most-cited Bangladesh publication during this period is a one-page letter published in *Lancet*:

Large outbreak of clinical cholera due to vibriocholerae non-01 in Bangladesh, written by MJ Albert, AK Siddique, MS Islam, ASG Faruque, M. Ansaruzzaman, SM Faruque and RB Sack, published in 1993 in *Lancet*, 341(8846), p. 704, and cited 237 times.

Bangladesh: period 1998-2008

Again, the most-cited publication is a one-page letter, this time published in the journal *Nature*:

Arsenic poisoning of Bangladesh groundwater, written by R. Nickson, J. McArthur, W. Burgess, KM Ahmed, P. Ravenscoft and M. Rahman, published in 1998 in *Nature*, 395(6700), p. 338. It has been cited already 328 times.

Among the top fifty articles we did not find a single one that had only Bangladeshi addresses.

C. Pakistan: period 1973-1987

The most-cited Pakistani article during this period is:

A heuristic algorithm for the M-machine, N-job flowshop sequencing problem, written by M. Nawaz, EE Enscoree and I. Ham, published in 1983 in *Omega-International Journal of Management Science* and cited 283 times.

This article, however, has also an American address. The most-cited purely Pakistani article for this period is:

Occurrence of mycorrhizas in halophytes, hydrophytes and xerophytes, and of endogone spores in adjacent soils, written by AG Khan and published in 1974 in the *Journal of General Microbiology*, 81(March), p. 7-17. It has been cited 112 times.

Pakistan: period 1988-1997

An article written by PA Underhill and eight co-authors is the most-cited one of this period:

Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography, published in 1997 in *Genome Research*, 7(10), p.996-1005, cited 415 times.

The most-cited purely Pakistani article during this period is:

Breeding for salinity tolerance in plants, by M. Ashraf, published in 1994 in *Critical Reviews in Plant Sciences*, 13(1), p.17-42 and cited 131 times.

Pakistan: period 1998-2007

The most-cited article is again by Underhill and a group of co-authors (20 co-authors this time):

Y chromosome sequence variation and the history of human populations, published in 2000 in *Nature Genetics*, 26(3), p.358-361, and cited 364 times.

The most-cited purely Pakistani article is:

Water purification by electrical discharges by MA Malik, A.Ghaffar and SA Malik, published in 2001, in *Plasma Sources Science & Technol-*

ogy, 10(1), p. 82-91, cited 72 times.

This overview of most-cited publications clearly shows that India participates in huge international groups, while Pakistan and Bangladesh do not. Yet, these countries too depend for their scientific research for a large part on international collaborations, be it on a smaller scale.

7 Bangladesh and Pakistan: dominance of a few centres and of few fields

According to the WoS research in Bangladesh is dominated by two institutes both situated in the country's capital Dhaka: the University of Dhaka (the capital's main university) and the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR,B) publishing 18% and 17% of all Bangladesh' publications. These numbers can be compared to the ones for Pakistan's University of Karachi (16 %) and Quaid I Azam University – including its former name of University of Islamabad (15 %).

According to SCImago (based on Scopus) India published most in the fields of Medicine, Chemistry, Physics and Astronomy and Agricultural and Biological Sciences. Pakistan and Bangladesh publish most in the fields of Medicine and Agricultural and Biological sciences.

8 Special attention to Bangladesh

We performed searches in the WoS on CU=Bangladesh AND TS='some topic'. Results (Table 5) show that Bangladesh pays a lot of attention to diseases, especially diarrhoea and arsenic poisoning and health problems of rural women. Bangladesh has 10,486 publications in the WoS over the whole period 1973-May 2008. For curiosity's sake we also added one typical Western topic, namely sport* but, not surprisingly, did not find a single publication about this topic. We compared with the relative number of recent (i.e. 2007 to May 2008) Belgian publications on these same topics. These percentages are always lower except for the topics 'sport*' and 'men'. Comparing with Pakistan (also data over the whole period 1973-May 2008), it seems that Bangladeshi researchers take more care of

health concerns, in particular of rural women.

Table 5. Topics studied in Bangladesh, and a comparison with Pakistan and Belgium

A	B	C	D
Water	1098(=10.5%)	1508(=7.5%)	4.8%
Health*	863(=8.2%)	888(=4.4%)	5.9%
Rural	824(=7.9%)	220(=1.0%)	0.3%
Diarrh*	807(=7.7%)	172(=0.9%)	0.4%
Cholera*	555(=5.3%)	26(=0.1%)	<0.1%
Women	455(=4.3%)	309(=1.5%)	2.7%
Men	101(=1.0%)	132(=0.7%)	1.2%
Arsen*	284(=2.7%)	40(=0.2%)	0.2%
Pregnan*	201(=1.9%)	193(=0.9%)	0.9%
Jute*	194(=1.8%)	3(<0.1%)	0
Sport*	0	2(<0.1%)	0.3%

A: topic ; B: Results for Bangladesh; C: Results for Pakistan; D: percentage results for Belgium

It is well-known that unification of addresses and institutional names is a big problem (de Bruin & Moed, 1990). When we tried to analyse data from Thomson Scientific's Institutional Citation Report 1981-2007 for ICDDR,B (CD version) we found 53 variants for ICDDR,B. These are shown in the appendix.

9 Conclusion

A (partial) comparison has been provided between research indicators related to India, Bangladesh and Pakistan. Data originate from Thomson Scientific's Web of Science (WoS) as well as from Scopus (using the SCImago website). Especially for Pakistan publication data show a spectacular increase over the latest years. Special attention is given to collaboration data and to the evolution of country h-indices. Lists of most-cited (in the WoS) articles are provided. These lists clearly show that India often is a partner in large multinational research groups, while this is not the case for Pakistan and Bangladesh. Comparing with Pakistan it seems that Bangladeshi researchers take more care of health concerns, in particular of rural women. It is further observed that Pakistan as well as Bangladesh have only a few research centres. Problems with the unification of addresses are

H. Kretschmer & F. Havemann (Eds.): Proceedings of WIS 2008, Berlin

Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting

Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI)

This is an Open Access document licensed under the Creative Commons License BY

<http://creativecommons.org/licenses/by/2.0/>

illustrated.

References

- Arunachalam, S. (2008). The science race continues in Asia. *Current Science*, 94(7), 848-849.
- Balaram, P. (2004). Science and technology proficiency: China and India. *Current Science*, 86(6), 755-756.
- De Bruin, R.E. and H.F. Moed (1990). The unification of addresses in scientific publications. In: *Informetrics 89/90* (L. Egghe & R. Rousseau, eds.). Amsterdam: Elsevier, pp. 65-78.
- Guan, J.C. and N. Ma (2004). A comparative study of research performance in computer science. *Scientometrics*, 61, 339-359.
- Gupta B.M., U.M. Munshi and P.K. Mishra (2002). S&T collaboration of India with other South Asian countries. *Current Science*, 83, 1201-1209.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572.
- Liu YX and R. Rousseau (2008). Definitions of time series in citation analysis with special attention to the h-index. *Journal of Informetrics* (to appear).
- Sadana, R., C. D'Souza, A.A. Hyder and A.M.R. Chowdhury (2004). Importance of health research in South Asia. *BMJ*, 328, 826-830.
- SCImago (2007). SJR — SCImago Journal & Country Rank. Retrieved May 27, 2008, from <http://www.scimagojr.com>
- Zitt, M., E. Bassecouard and Y. Okubo (2000). Shadows of the past in international cooperation: collaboration profiles of the top five producers of science. *Scientometrics*, 47, 627-657.

Appendix

Variant names for ICDDR,B (preferred name for the International Centre for Diarrhoeal Disease

Research, Bangladesh).

1. B Ctr Hlth & Populat Res
2. B Ctr Hlth & Populat Studies
3. B Ctr Hlth Populat Studies
4. Bangladesh Ctr Hlth & Populat Hlth
5. Bangladesh Ctr Hlth & Populat Res
6. Bangladesh ICDDR B; Ctr Hlth & Populat Res
7. Ctr Hlth & Populat Res B
8. ICCDDR B
9. ICDCR B
10. ICDDDE B
11. ICDDR
12. ICDDR B
13. ICDDR B; Ctr Hlth & Populat Res Mohakhali
14. ICDDR B; Ctr Hlth & Populat Res;
15. ICDDR BANGLADESH;
16. ICDDR,B
17. ICDDR; Ctr Hlth & Populat Res
18. ICDDR B
19. ICDDR B
20. ICDDR B Ctr Hlth & Populat Res
21. ICDRRB
22. INT CENT DIARRHOEAL DIS RES
23. INT CTR CONTROL DIARRHOEAL DIS
24. INT CTR DIARRHOEAL DIS RES B
25. INT CTR DIARRHOEAL DIS RES
26. Int Ctr Diarrhea & Dis Res
27. INT CTR DIARRHEA DIS RES BANGLADESH
28. Int Ctr Diarrheal Dis & Res
29. INT CTR DIARRHEAL DIS RES
30. INT CTR DIARRHEAL DIS RES BANGLADESH
31. INT CTR DIARRHOEAL DIS RES
32. INT CTR DIARRHOEAL DIS RES
33. INT CTR DIARRHOEA DIS
34. INT CTR DIARRHOEA DIS RES
35. INT CTR DIARRHOEA DIS RES BANGLADESH
36. INT CTR DIARRHOEAL
37. INT CTR DIARRHOEAL DIS
38. INT CTR DIARRHOEAL DIS & RES
39. INT CTR DIARRHOEAL DIS DIS
40. INT CTR DIARRHOEAL DIS RES
41. Int Ctr Diarrhoeal Dis Res
42. Int Ctr Diarrhoeal Dis Res B
43. INT CTR DIARRHOEAL DIS RES BANGLADESH

44. INT CTR DIARRHOEAL DIS RES
BANGLADESH ICDDR
 45. INT CTR DIARRHOEAL RES
 46. INT CTR DIARRHOEL DIS RES
 47. INT CTR DIARRHOEL DIS RES
BANGLADESH
 48. Int Ctr Diarrhoel Dis Res Ctr
 49. INT CTR DIARROHOEAL DIS RES
 50. Int Ctr Hlth & Pop Res
 51. Int Ctr Hlth & Populat Res
 52. INT DIARRHOEAL DIS RES CTR
 53. INTL CTR DIARRHEAL DIS RES
- Source: Thomson Corporation