

SCIENTIFIC AUTHORSHIPS AND COLLABORATION NETWORK ANALYSIS ON CHAGAS DISEASE: PAPERS INDEXED IN PUBMED (1940-2009)

Gregorio GONZÁLEZ-ALCAIDE(1), Jinseo PARK(2), Charles HUAMANÍ(3), Joaquín GASCÓN(4) & José Manuel RAMOS(5)

SUMMARY

Chagas disease is a chronic, tropical, parasitic disease, endemic throughout Latin America. The large-scale migration of populations has increased the geographic distribution of the disease and cases have been observed in many other countries around the world. To strengthen the critical mass of knowledge generated in different countries, it is essential to promote cooperative and translational research initiatives. We analyzed authorship of scientific documents on Chagas disease indexed in the *Medline* database from 1940 to 2009. Bibliometrics was used to analyze the evolution of collaboration patterns. A Social Network Analysis was carried out to identify the main research groups in the area by applying clustering methods. We then analyzed 13,989 papers produced by 21,350 authors. Collaboration among authors dramatically increased over the study period, reaching an average of 6.2 authors per paper in the last five-year period. Applying a threshold of collaboration of five or more papers signed in co-authorship, we identified 148 consolidated research groups made up of 1,750 authors. The Chagas disease network identified constitutes a "small world," characterized by a high degree of clustering and a notably high number of Brazilian researchers.

KEYWORDS: Chagas disease; Bibliometrics; Cooperative behavior; Network analysis; Research areas; Research groups.

INTRODUCTION

Chagas disease is a chronic, tropical, parasitic disease, endemic throughout Latin America. It affects 7.7 million people, with around 50,000 new cases each year, and it kills an estimated 12,000 people annually. Early diagnosis of the disease is difficult due to its clinical characteristics, and treatment for chronic strains is not always effective (DE ANDRADE *et al.* 1996; SOSA *et al.* 1998). The parasite was first identified by the Brazilian infectologist Carlos Chagas in 1909, but it was not until the 1960s that disease control programs were first established (ABAD FRANCH *et al.* 2011; DIAS 2009; SILVEIRA 2002). Rural migration to urban areas in the 1960s and the 1970s increased the risk of contracting the disease, while recent large-scale migration has led to the further spread of the disease to a number of other countries worldwide (GASCON *et al.* 2010; MUÑOZ *et al.* 2009; JACKSON *et al.* 2010).

Research on Chagas disease has evolved throughout the last decades due to changes in public health policies and social factors, with an increase in the research areas, as can be seen in the rise in scientific publications, most of them published in international journals (RAMOS *et al.* 2011). For this reason, it is important to identify the main authors responsible for publications on Chagas disease and their collaborative patterns in order to promote cooperative and translational scientific research initiatives.

Social Network Analysis (SNA) is an analytical method based on

the graph theory, which identifies links between individuals or agents in order to analyze the social structures that emerge from those relationships. Proponents of this method postulate that the structure resulting from the network analysis allows better interpretation of the behavior or the attributes of the network's components than the analysis of each individual or agent in isolation (SCOTT 1991; WATTS 2004). Applying SNA to research on scientific communication and dissemination of knowledge through the analysis of co-authorships in publications in one field of knowledge, investigators can identify the groups that are actively participating in research in a given area. Likewise, they can describe the degree of interaction and the influence of different agents or groups within the research community (GONZÁLEZ ALCAIDE *et al.* 2010; NEWMAN 2001). The use of SNA has seen considerable development in the past several years (MELIN & PERSSON 1996), complementing and deepening the initial approaches taken since the 1960s in the field of bibliometrics (first by PRICE and then by CRANE) in relation to the concept of "invisible colleges" and the analysis of citation and co-citation networks (BARABASI *et al.* 2002; CRANE 1969; MULLINS 1980; PRICE 1965; PRICE 1966).

The objective of this study was to analyze the authorship of scientific manuscripts on Chagas disease published in scientific journals indexed in the *Medline* database from 1940 to 2009 and to develop a social network analysis applied to co-authorship of scientific papers about American trypanosomiasis.

(1) Department of History of Science and Documentation Department, Universidad de Valencia, Valencia, Spain. E-mail: gregorio.gonzalez@uv.es

(2) Korea Institute of Science & Technology Information (KISTI). Deajeon, South Korea. E-mail: jayoujin@hanmail.net

(3) Sociedad Científica de San Fernando, Universidad Nacional Mayor de San Marcos, Lima, Perú. E-mail: huamani_ca@hotmail.com

(4) Barcelona Center for International Health Research (CRESIB, Hospital Clinic-Universitat de Barcelona), Barcelona, Spain. E-mail: jgascón@clinic.ub.es

(5) Department of Internal Medicine, Hospital General Universitario de Alicante, Alicante, Spain. E-mail: jramosrincon@yahoo.es

Correspondence to: Gregorio González-Alcaide, Faculty of Medicine and Odontology, University of Valencia. 15, Blasco Ibáñez Avenue. 46020 – Valencia, Spain. E-mail: gregorio.gonzalez@uv.es

MATERIAL AND METHODS

A literature review was carried out using the *Medline* database, with the Medical Subject Headings (MeSH) "Chagas disease" and "*Trypanosoma cruzi*." Complementing these generic terms, the MeSH thesaurus included specific descriptors, synonyms and variant spellings in the search process. The query terms were also entered in the "Title" and "Abstract" fields. The study period was from 1940 to 2009. The PubMed platform (<http://www.ncbi.nlm.nih.gov/pubmed>) was accessed on 25 May 2010.

The information obtained from the registers was put into a database using Microsoft Access. A standardization process was carried out to consolidate variations of author names. The criterion followed in this process was the occurrence of the institutional signature associated with the variations of names and surnames.

The following bibliometric indicators were computed: number of papers published, number of authors, number of signatures (or number of contributions in all papers), and authors per paper or collaboration index (mean number of signatories per paper).

A regression analysis was used to explore the trend of publications and signatures (dependent variable) over year-long periods (independent variable). The number of publications was treated in two ways in the regression analysis: on its natural scale to explore the linear trend, and after a natural-log transformation to explore the percentage of change per year and to determine whether the rise in publications followed an exponential trend.

The most productive authors were also identified, determining (for the top 80 authors [≥ 50 papers]) the total number of publications (articles, reviews, editorials, notes and other published documents), the number of articles, the percentage of papers signed by researchers as first and last author, the number of their total collaborators (the number of authors who have co-authored at least one paper), the number of their collaborators with ≥ 5 papers and ≥ 10 papers, (that is, their consolidated collaborators, with whom they regularly publish or with whom they have disseminated a considerable portion of their scientific output).

A SNA was carried out. The networks consist of nodes and links: nodes represent authors, while links connect nodes in the form of co-authorships. All co-authorships were identified for each document. Subsequently, the number of co-authorships was recalculated in order to determine the collaboration intensity, as many co-authorships are redundant in a large collection of documents. This information was introduced into an algorithm to identify clusters of authors or research groups. A cluster was defined as the presence of at least two authors who were linked to one another by a number of co-authorships equal to or greater than 5. The algorithm was then iterated for a number of co-authorships equal to or greater than 10. These thresholds were applied to concentrate the analysis on the most intense links, thus enabling an appropriate visualization and representation of the network.

Pajek software was used for constructing the author's networks and network analysis. Fucherteman-Reingold and Kamada-Kawai algorithms were applied for node spatial distribution.

RESULTS

A total of 13,989 papers produced by 21,350 authors were analyzed. We identified 1,008 authors who published more than nine papers (4.7%); 6,623 authors who published between two and nine papers (31%) and 13,719 authors who only published one paper (64.3%).

AUTHORSHIP

Figure 1 shows the evolution of authors per paper collaboration index by five-year period, which has dramatically increased over the study period, with an average of less than three authors per paper from 1940 to 1979, 3.7 from 1980 to 1989, 4.6 from 1990 to 1999, and 5.9 from 2000 to 2009. In the last five-year period, the number of authors per paper has reached an average of 6.2 authors per paper.

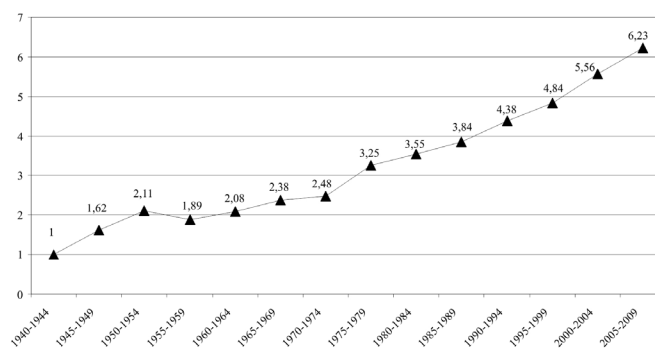


Fig. 1 - Distribution of authors per paper index on Chagas disease by five-year period.

Figure 2 shows the evolution of the number of articles and the number of signatures by five-year period. The number of articles corresponds to a straight-line fit ($R^2=0.945$); nevertheless, the number of signatures follows an exponential fit ($R^2=0.951$).

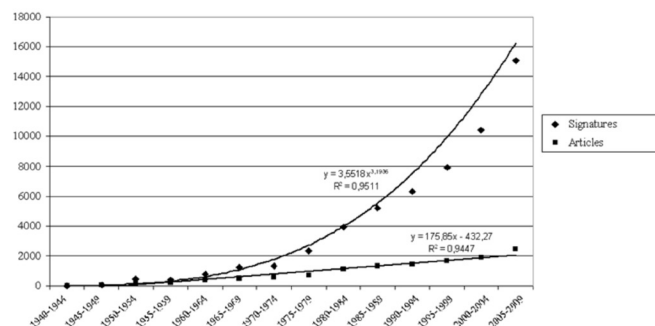


Fig. 2 - Distribution of signatures and published articles on Chagas disease by five-year period.

Table 1 ranks the 80 most productive authors (≥ 50 papers) and their collaborative patterns. The main author was W. de Souza, a Brazilian researcher ($n = 186$ documents), followed by F. Kierszenbaum ($n = 138$) and R. Docampo ($n = 134$), both Latin American investigators working in the USA.

With regard to author order of signatures, the most productive authors tend to hold the final position. Fifty-eight authors (72.5%) signed

Table I
Most productive authors (>50 works) and collaboration patterns on Chagas disease in the *Medline* database (1940-2009)

Author (country)	Number of documents	Number of articles	% of papers as first author	% of papers as last author	Number of collaborators	Number of collaborators (≥ 5 papers)	Number of collaborators (≥10 papers)
Souza, Wanderley de (Brazil)	186	170	8.6	44.62	304	24	6
Kierszenbaum, Felipe (USA)	138	131	30.43	52.17	49	9	4
Docampo, Roberto (USA)	134	122	20.89	39.55	184	24	7
Segura, Elsa L (Argentina)	129	117	10.85	53.49	242	27	11
Cazzulo, Juan José (Argentina)	124	114	10.48	40.32	168	23	7
Chiari, Egler (Brazil)	123	120	6.5	27.64	254	26	9
Tanowitz, Herbert B (USA)	115	105	17.39	33.91	235	30	19
Frasch, Alberto Carlos (Argentina)	112	106	8.93	40.18	326	23	10
Dias, João Carlos Pinto (Brazil)	111	97	21.62	22.52	173	12	1
Brener, Z (Brazil)	108	100	12.04	50	108	15	5
Amato Neto, Vicente (Brazil)	107	99	24.3	9.34	178	18	6
Miles, Michael A (UK)	105	98	16.19	42.86	213	17	5
Levin, Mariano J (Argentina)	98	92	7.14	37.75	325	29	6
Schenone, H (Chile)	96	91	44.79	30.21	170	12	5
Colli, Walter (Brazil)	93	85	1.07	48.39	142	11	4
Tibayrenc, Michel (France)	90	87	16.67	51.11	189	14	4
Tarleton, Rick L (USA)	87	79	19.54	52.87	192	12	1
Rassi, Anis (Brazil)	86	70	24.42	19.77	166	7	1
Urbina, Julio A (Venezuela)	84	74	27.38	25	169	22	3
González Cappa, Stella M (Argentina)	84	81	20.24	46.43	141	19	4
Stoppani, Andrés OM (Argentina)	84	81	3.57	85.71	74	10	4
Wittner, Murray (USA)	84	78	2.38	30.95	158	21	10
Andrade, Sonia G (Brazil)	81	80	40.74	33.33	97	6	3
Romanha, Alvaro J (Brazil)	80	75	3.75	36.25	219	14	5
Zingales, Bianca (Brazil)	77	75	19.48	31.17	196	17	5
Schenkman, Sergio (Brazil)	75	69	18.67	48	159	14	3
Gazzinelli, Ricardo T (Brazil)	75	65	16	32	197	24	6
Teixeira, Antonio RL (Brazil)	75	69	32	41.33	146	12	2
Gürtler, Ricardo E (Argentina)	75	75	30.67	30.67	90	19	8
Mady, Charles (Brazil)	73	64	20.55	13.7	178	19	9
Lopes, Edison Reis (Brazil)	73	65	27.4	19.18	165	7	3
Brenière, Simone Frédérique (France)	72	71	31.94	37.5	167	16	7
Solari, Aldo (Chile)	71	69	12.68	50.7	156	15	6
Ouaisi, M Ali (France)	70	67	27.14	37.14	146	17	6
Castro, Solange Lisboa de (Brazil)	70	59	14.28	37.14	160	15	3
Araújo Jorge, Tania C (Brazil)	70	62	18.57	37.14	156	21	4
Pileggi, F (Brazil)	69	67	1.45	85.51	143	26	9
Coura, José Rodrigues (Brazil)	69	61	30.43	42.03	151	9	3
Yoshida, Nobuko (Brazil)	68	62	13.23	42.65	109	13	3
Goldenberg, Samuel (Brazil)	67	66	2.98	49.25	217	12	4
Kuhn, RE (USA)	67	66	10.45	76.12	41	5	1
Meirelles, Maria de Nazareth L (Brazil)	66	61	24.24	36.36	128	7	5
Apt, Werner (Chile)	66	62	39.39	9.09	114	11	5
Marsden, PD (Brazil)	65	60	24.61	33.85	66	9	1

Table I
Most productive authors (>50 works) and collaboration patterns on Chagas disease in the *Medline* database (1940-2009) (cont.)

Author (country)	Number of documents	Number of articles	% of papers as first author	% of papers as last author	Number of collaborators	Number of collaborators (≥ 5 papers)	Number of collaborators (≥10 papers)
Brun, Reto (Switzerland)	65	64	3.08	15.38	199	8	4
Carlier, Yves (Belgium)	64	63	7.81	53.12	143	20	4
Contreras, María del Carmen (Chile)	63	60	12.7	6.35	154	12	5
Luquetti, Alejandro O (Brazil)	62	56	8.06	8.06	186	12	1
Mortara, Renato A (Brazil)	62	59	12.9	32.26	180	12	2
Scharfstein, Julio (Brazil)	61	50	19.67	31.15	180	12	2
Alves, Maria Júlia Manso (Brazil)	61	54	14.75	27.87	119	7	2
Barretto, MP (Brazil)	61	60	37.7	50.82	19	7	3
Chapadeiro, E (Brazil)	61	58	19.67	32.79	103	8	3
Kirchhoff, Louis V (USA)	59	52	18.64	25.42	141	6	1
Silveira, José Franco da (Brazil)	59	55	5.08	49.15	240	16	8
Rojas, A (Chile)	58	56	1.72	31.03	91	11	4
Rocha, Manoel Otávio da Costa (Brazil)	58	47	6.9	41.38	133	15	5
Jansen, Ana Maria (Brazil)	58	54	8.62	62.07	138	10	3
Silva, João Santana (Brazil)	58	52	13.79	46.55	186	10	5
Corrêa Oliveira, Rodrigo (Brazil)	57	50	1.75	28.07	174	21	7
Barnabé, Christian (France)	56	55	12.5	7.14	150	7	2
Lana, Marta de (Brazil)	56	56	14.28	21.43	127	16	8
Bestetti, Reinaldo B (Brazil)	56	33	53.57	12.5	60	7	3
Weiss, Louis M (USA)	55	49	0	1.82	121	25	14
Lederkremer, Rosa M de (Argentina)	55	51	29.09	49.09	60	10	3
Tafari, Washington L (Brazil)	54	51	18.52	29.63	91	12	6
Albuquerque, Sérgio (Brazil)	54	52	1.85	22.22	193	4	0
Bellotti, G (Brazil)	53	51	9.43	3.77	122	19	8
Villalta, Fernando (USA)	53	51	56.6	22.64	73	7	3
Dantas, Roberto Oliveira (Brazil)	53	50	37.73	15.09	54	6	4
Cerecetto, Hugo (Uruguay)	52	50	11.54	28.85	163	23	10
Souto Padrón, Thaïs (Brazil)	52	50	30.77	17.31	107	2	1
Ribeiro, Antônio Luiz Pinho (Brazil)	52	42	28.85	25	90	11	4
López, Manuel C (Spain)	52	52	0	44.23	111	13	5
Guhl, Felipe (Colombia)	52	47	26.92	23.08	132	6	4
Prata, Aluizio (Brazil)	51	47	3.92	35.29	95	5	1
Marin Neto, José Antônio (Brazil)	51	37	31.37	29.41	103	11	3
Ianni, Bárbara M (Brazil)	51	45	13.72	1.96	132	16	8
Ribeiro, RD (Brazil)	50	50	46	12	36	6	3
Basombrío, Miguel Angel (Argentina)	50	49	32	42	142	5	1

more papers in the last position than in the first position. The top ten authors have a much higher number of papers as last author. The main investigators signing as first author of their manuscripts were R.B. Bestetti (53.6% of his publications), R.D. Ribeiro (46%), H. Schenone (44.8%) and S.G. Andrade (40.7%).

The most productive authors have established collaborative links with a great number of authors. However, those with whom they regularly

publish are relatively small in number (≥ 5 papers in co-authorship between 2% and 37% of their collaborators, mean 10.2%; and ≥ 10 papers in co-authorship between 0% and 8% of their collaborators, mean 3.6%).

CO-AUTHORSHIP COLLABORATION NETWORK

Applying a threshold or intensity of collaboration of five or more papers signed in co-authorship, we identified 148 research groups made

up of 1,750 authors (72.4% of the authors who have published five or more papers). The largest group consisted of 1,141 authors. Figure 3 shows the co-authorship collaboration network according to the Fuchertman-Reingold algorithm. The most productive authors (≥ 50 papers) are highlighted with their names and nationalities. Most of them are Brazilian ($n = 46$), with a smaller contribution made by Argentinean ($n = 9$), American ($n = 9$), Chilean ($n = 5$) and European scientists ($n = 8$). Colombia, Uruguay and Venezuela, with one researcher each, complete the roster.

Applying a threshold of 10 or more papers signed in co-authorship, we identified 116 research groups (clusters) made up of 585 authors (58% of the authors who have published 10 or more papers). Many of these investigators published their first paper prior to 1990 (64.3%, $n = 376$), and most of them (75%, $n = 282$) have continued publishing since the turn of the century. Figure 4, drawn with the Kamada-Kawai algorithm, represents groups with 2–5 authors, whereas Figure 5 groups consist of 6–9 authors, Figure 6 groups consist of 10–38 authors, and Figure 7 (the two largest groups) consists of 57 and 58 authors, respectively.

DISCUSSION

The present study identifies and characterizes the collaboration among the scientific community with regard to the study of Chagas disease over a long period of time (70 years), coinciding with the consolidation and expansion of Big Science from the mid-twentieth century to the present.

The study has several limitations: The *Medline* database does not cover all publications and document types on Chagas disease worldwide, and the co-authorship network analysis carried out has a descriptive nature; however, further empirical research is needed to evaluate network performance.

Although *Medline* is not an exhaustive source, previous studies have shown that its use provides a satisfactory resource for high quality and relevant biomedical papers (ROLLIN *et al.* 2010). The main MeSH used for indexing Chagas disease literature, after “*Trypanosoma cruzi*” and “Chagas disease” were “Chagas cardiomyopathy” (10.2%), “molecular sequence data” (7.8%), and “insect vector” (7.4%); the main qualifiers were “immunology” (24.6%), “parasitology” (22.1%), “metabolism” (20.5%), and “genetics” (16.3%) (RAMOS *et al.* 2011).

The diachronic evolution of the number of publications observed corresponds with a linear growth model. The considerable amount of scientific production, together with the large number of “big producer” authors identified (≥ 10 documents), reflects the significant development as well as the maturity of the research in this area and in the larger field of tropical diseases as a whole (FALAGAS *et al.* 2006). On the other hand, it is important to note some differences as well: although pathologies like Chagas disease, malaria and schistosomiasis have seen an increase in scientific research (COSTA LIMA *et al.* 1985; GARG *et al.* 2009; GLOVER & BOWEN 2004; ZHENG *et al.* 2009), investigations into other diseases, such as leprosy, have experienced the opposite trend since the turn of the century (SCHOONBAERT & DEMEDTS 2008). In the case of Chagas disease, a number of factors have contributed to the surge in research initiatives apart from its high prevalence and extensive geographical, endemic spread, including the globalization of the illness due to migration and its resurgence as a co-infection associated with HIV (RAMOS JUNIOR *et al.* 2010).

The evolution in the number of co-authors who participate in the papers on this topic has seen exponential growth, a fact consistent with a more general trend of increased collaboration and especially co-authorship throughout the entire biomedical sector (VANZ & STUMPF 2010). However, the rise in collaboration also entails certain negative aspects, including unjustified hyperauthorship as well as other ethically

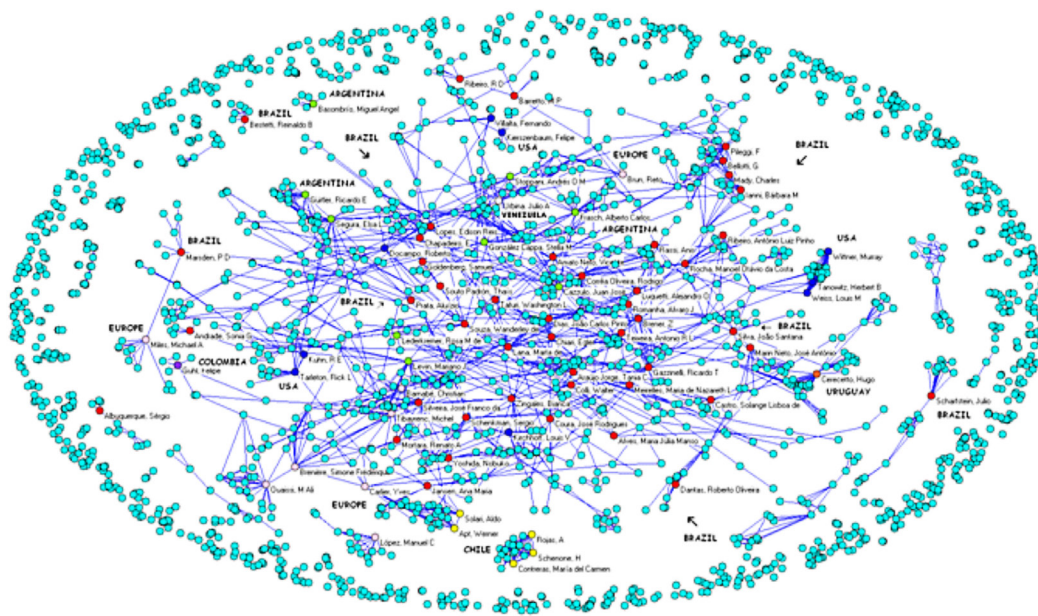


Fig. 3 - Author's network (≥ 5 co-authorships) on Chagas disease in the *Medline* database (1940-2009).

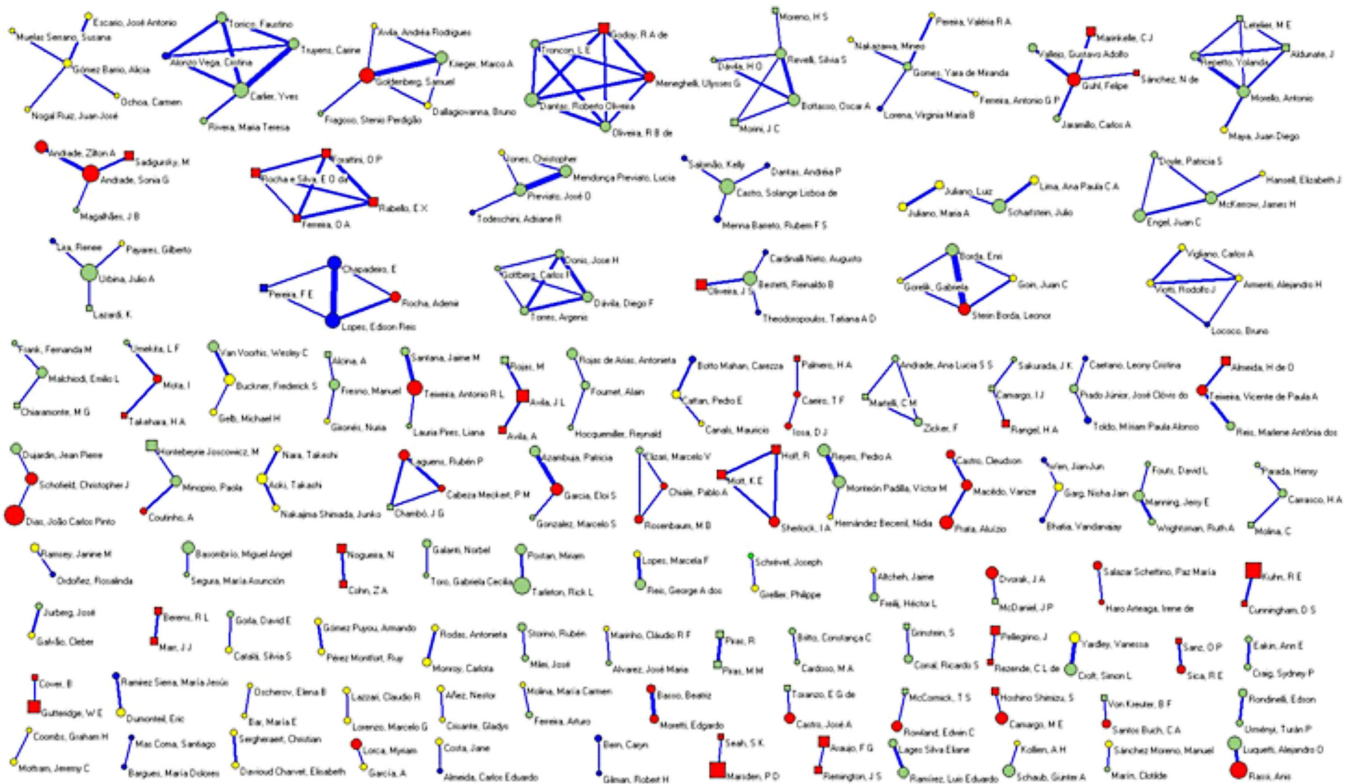


Fig. 4 - Main research groups (2-5 authors and ≥ 10 co-authorship) on Chagas disease in the *Medline* database (1940-2009). Key of figures 4 to 7: color of nodes (red first document published before the 1980s decade, green first document published in the 1980s decade, yellow first document published in the 1990s decade, blue first document published in the 2000s decade); shape of nodes (square last document published before the 2000s decade, circular at least one paper published in the 2000s decade); size of nodes (number of documents published), link thickness (number of co-authorships).

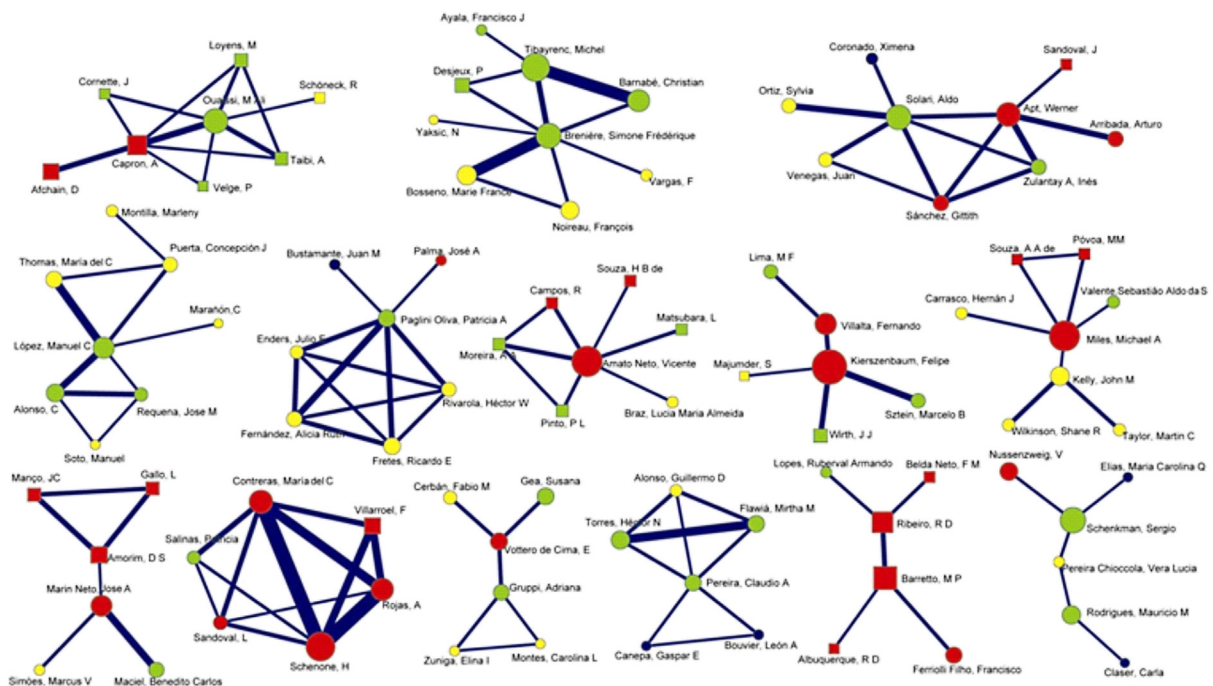


Fig. 5 - Main research groups (6-9 authors and ≥ 10 co-authorship) on Chagas disease in the *Medline* database (1940-2009).

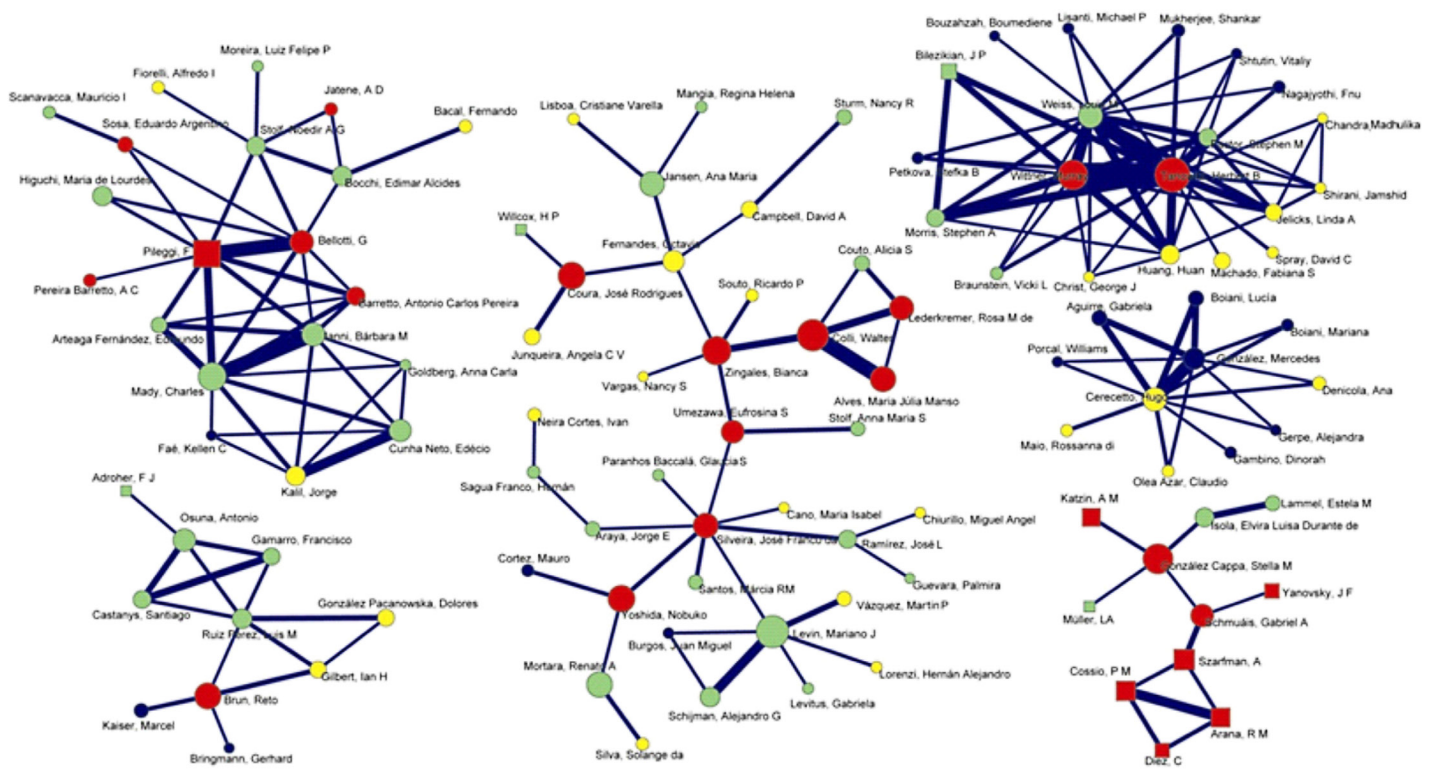


Fig. 6 - Main research groups (10-38 authors and ≥ 10 co-authorships) on Chagas disease in the *Medline* database (1940-2009).

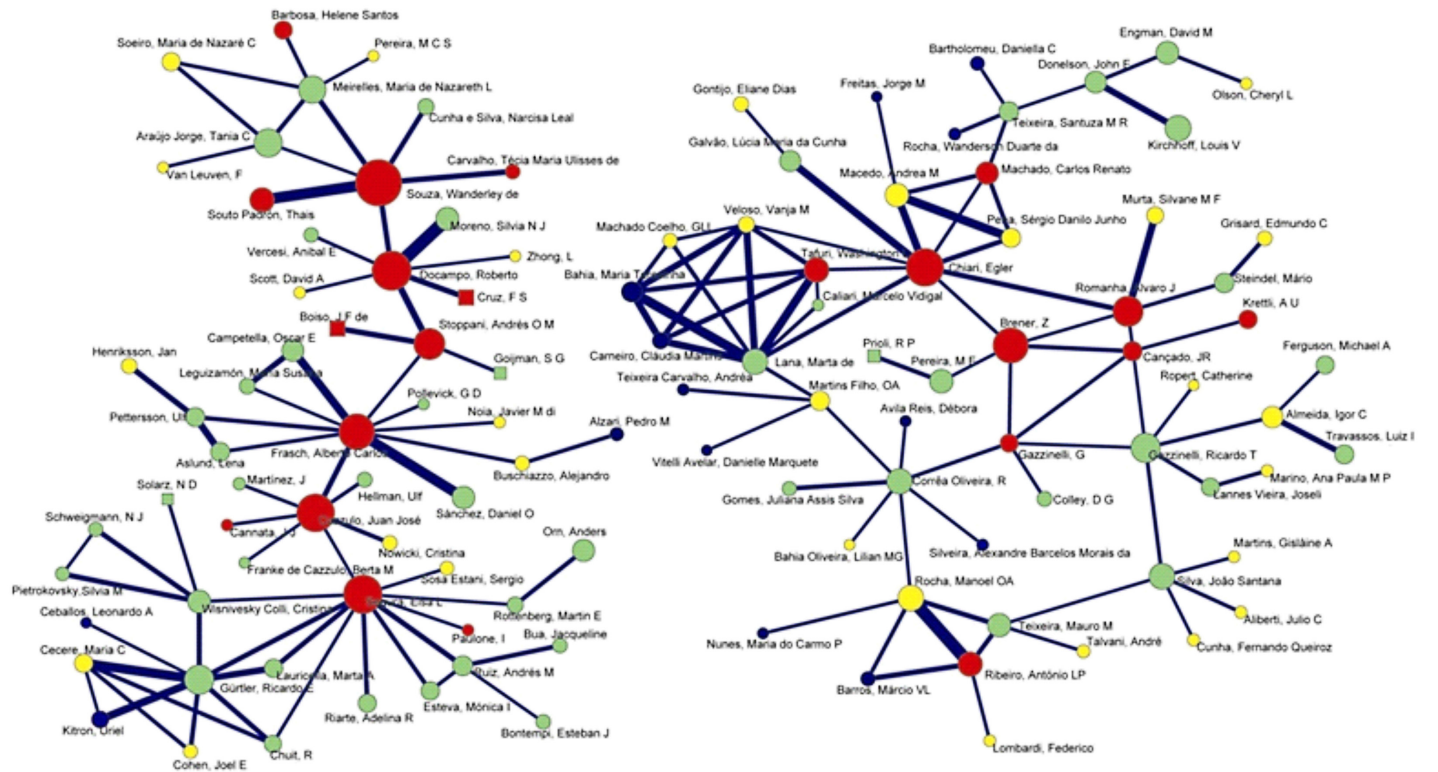


Fig. 7 - Main research groups (57-58 authors and ≥ 10 co-authorships) on Chagas disease in the *Medline* database (1940-2009).

reprehensible behavior that undermines scientific credibility; these hazards should be averted through the necessary corrective measures (GARCÍA *et al.* 2010).

In general, as in other topics, the most productive authors are those who demonstrate a high degree of collaboration (BALES *et al.* 2008; BALES *et al.* 2011). In this sense, it is important to note that although these researchers (identified as those who have published over 49 papers) have established collaborative links with a great number of authors, their stable and consolidated collaborators—that is, those with whom they regularly publish or have disseminated a considerable portion of their scientific output—are a relatively small in number. Again, this characteristic of scientific cooperation extends to other disciplines and areas of knowledge (GONZÁLEZ ALCAIDE *et al.* 2008).

Furthermore, with regard to author order, the most productive authors tend to hold the final position, reserved for senior authors who direct the research; this position has also been correlated to other variables, such as the age or the academic/professional prestige of the authors. The first position, on the other hand, is occupied by junior authors who have assumed responsibility for performing the work (COSTAS & BORDONS 2011).

The analysis of co-authorships in scientific publications is a good method to characterize the patterns of collaboration in an area of knowledge because publishing the results of a study is in fact, together with the generation of knowledge and promotion of translational activities (above all, the genuine impact towards public health measures), the ultimate purpose of the investigators' work. Some of the main features in scientific co-authorship networks include the following:

- Most researchers are connected either directly or through intermediaries, constituting the so-called "giant component" that draws together the majority of the authors (approximately 80%–90% of those belonging to the network). NEWMAN (2004a) determined that the giant component comprised between 82% and 92% of authors in different scientific disciplines; in the case of the co-authorship network on Chagas disease, this figure was 84.1%.

- Scientific co-authorship networks display a high degree of clustering, in which cohesive local communities, sub-networks, or subgroups are commonplace. Their members are more closely linked by mutual interrelationships than the members of the network as a whole, so the quantification of the frequency, threshold or intensity of the collaborations constituted the criteria utilized in the present study to identify the existing research groups in the area of Chagas disease. Because a greater number of co-authorships show a stronger and more consolidated social connection and a higher degree of interaction and collaboration, this criterion was considered more appropriate in identifying research subgroups than other clustering methodologies (CHANG & CHEN 2011).

- Scientific communities are commonly known as a "small world," in which the average "distance" or degree of separation between any two authors varies logarithmically according to the size of the community. NEWMAN (2001) found that the average distance between any two authors in a large scientific network was six other authors. The average distance in the case of the network on Chagas disease was somewhat

smaller (4.7), indicating a more centralized community with notable interaction among its component agents. Milgram's notion of a "small world," where people are connected by short path lengths, and which is associated with the phrase "six degrees of separation", has recently been revised by KLEINFELD (2002). This researcher emphasized that the importance of social capital (the ability to make personal connections) and variables such as social status were more significant than the distance between the individuals. WATTS & STROGATZ developed a mathematical model in order to explain the "small world" phenomenon, identifying a few random connectors that reduced the distance between nodes in the large networks. Nevertheless, with regard to the distance in the co-authorship network of scientists analyzed, it should be stressed that linkage to others through more than one researcher does not necessarily entail interaction (even the co-authors of a single paper might have contributed to the study without interacting with their colleagues). Furthermore, studies have pointed out that social interaction occurs more often through a short chain of contacts. Therefore, at an individual level, other researchers placed farther away than two or three intermediates may not be useful for authors seeking collaboration (KLEINFELD 2002; WATTS & STROGATZ 1998).

- Finally, it is important to highlight the fact that in scientific networks, many of the distributions of a random variable and its frequency follow a power law distribution, in line with the pattern of the so-called "scale-free networks" (BARABASI & ALBERT 1999). For example, a few nodes are much more closely connected than the rest; these nodes represent a reduced number of central or influential authors, while the other nodes constitute peripheral authors. A small number of authors known as the "best connected scientists" exercise the important role of intermediaries, ensuring a rapid connection (i.e., a reduced number of steps) between different authors and groups making up the network. If these authors were to be eliminated, significant portions of the network would be disconnected, and the network itself would be weakened by structural holes. Hence, these scientists help to assure accessibility between authors, cohesion among them and dissemination of knowledge throughout the network (NEWMAN 2004b). One significant quality prevalent among these scientists was described by GOH and colleagues: normally, these central, influential authors do not collaborate with other authors of similar characteristics (GOH *et al.* 2003).

The Chagas disease network identified in the study has a high number of research subgroups, characterized by their markedly national nature. This fact can be explained by the interest taken by individual researchers in Chagas disease and Trypanosomiasis in different countries (due in large part to the broad geographical dissemination of the disease, which has spilled over from Latin America into the United States, Canada and Europe), resulting in a myriad of papers (ABAD FRANCH *et al.* 2011; RAMOS *et al.* 2011). The higher number of Brazilian researchers can be attributed to the greater size and development of this country's scientific system, which has become the principal scientific reference in South America. Also at play is the special attention paid to this disease by initiatives such as the Brazilian Network of Healthcare and Studies in Co-infection of Trypanosomiasis, HIV and Other Immunosuppressive Diseases (Rede Brasileira de atenção e estudos em co-infecção *Trypanosoma cruzi*/HIV [e outras condições de imunossupressão]) (ALMEIDA *et al.* 2009; COSTA LIMA *et al.* 1985; RAMOS JUNIOR *et al.* 2010). Chile and Argentina (the latter has the second highest prevalence of the disease in the region) are well behind Brazil in terms

of research production. The United States and some European countries have also become referential for their research into this topic, in part because globalization and migratory flows have extended the reach of the infection and increased the importance of tropical diseases for the public health systems in these countries. As a result, these countries, which are in the vanguard of scientific excellence and development, have increased their collaborations with scientific publications in Latin America and the Caribbean. This cooperation is aided by geographical proximity, in the case of the United States, and by cultural, linguistic, and historical links, in the case of the countries in Europe (KEISER & UTZINGER 2005; FALAGAS *et al.* 2006; WILLIAMS *et al.* 2008).

For Chagas disease as well as other widespread endemic diseases, it is essential to promote the development of transnational initiatives in order to consolidate the research potential and critical mass of knowledge generated by research centers and groups in different countries. For this reason, a good reference is the afore-mentioned “Rede brasileira de atenção e estudos em co-infecção *Trypanosoma cruzi*/HIV (e outras condições de imunossupressão)”, founded in 2006 in order to coordinate and organize the Brazilian healthcare network and to integrate established international groups into it. These efforts serve to promote synergies that contribute to an international network that improves care and strengthens the development of strategic studies in the area (ALMEIDA *et al.* 2009; RAMOS JUNIOR *et al.* 2010). In addition, it would be necessary to foster research in countries with high prevalence of the disease but where no notably productive authors have been identified, such as in Mexico, Bolivia, Guatemala, Ecuador, Honduras, El Salvador and Paraguay (RAMOS *et al.* 2011). This is also true for countries like Venezuela or Colombia, where only one key researcher has been identified. Indeed, effective disease control depends on research that responds to population’s needs as identified in disease control programs tailored to each individual country (SILVEIRA 2002).

The most productive and collaborative subgroups of the co-authorship networks—that is, the core collaborators that constitute the engine of research in the area—were identified by applying a threshold of ten or more collaborative papers to their members. Many of these investigators have been continuously active in the field since the 1980s, and 79.8% of the authors with more than ten publications also published in the first decade of the twenty-first century. These individuals occupy central positions in the research groups in which they participate, bringing together the rest of the researchers around them. Previous studies have demonstrated that scientific communities and research groups endure longer if they have a strong, active nucleus of members, suggesting that a stable core of investigators is fundamental to ensuring the long-term sustainability of the groups. More importantly, these authors play an important role in terms of cultivating the scientific community and integrating new members (WU *et al.* 2009; BARABASI *et al.* 2002).

The importance of strong links within the scientific co-authorship networks is illustrated by the fact that the authors who establish and consolidate relationships with just one other author or with a small group of authors can develop their work more efficiently and with a higher degree of citation than authors who collaborate sporadically with a greater number of co-authors. (ABBASI *et al.* 2010; KRACKHARDT 1992). Moreover, the authors who establish non-redundant collaborative relationships with other authors (that is, relationships with authors who are not already linked among themselves) have an advantage over

other authors in terms of efficiency and degree of citation (ABBASI & ALTMANN 2011). This idea has already been developed by Burt in relation to the theory of “structural holes” (BURT 1995; BURT 2004).

RESUMEN

Redes de colaboración y autorías científicas sobre la enfermedad de Chagas: análisis de las publicaciones indexadas en PubMed (1940-2009)

La enfermedad de Chagas es una enfermedad parasitaria tropical, endémica en muchos países y regiones de América, si bien, los movimientos de población han incrementado su distribución geográfica y se han constatado casos en muchos países del mundo. En este sentido, resulta fundamental promover iniciativas de investigación cooperativas y transnacionales, con el propósito de aunar la masa crítica de conocimiento generada en los diferentes países. Se estudian las publicaciones científicas sobre la enfermedad de Chagas recogidas en la base de datos *Medline* entre 1940 y 2009. Mediante indicadores bibliométricos se han analizado los patrones de colaboración y se ha efectuado un Análisis de Redes Sociales para identificar los principales grupos de investigación. Se han analizado 13.989 documentos publicados por 21.350 autores. La evolución de la colaboración ha experimentado un notable crecimiento, alcanzando un promedio de 6,2 autores por trabajo en el último quinquenio. Aplicando un umbral de colaboración de 5 o más trabajos firmados en coautoría, se han identificado 148 grupos de investigación conformados por 1.750 autores. La red de Chagas identificada conforma un ‘mundo pequeño’ con un elevado grado de agrupamiento, destacando el elevado número de investigadores brasileños.

REFERENCES

1. Abad Franch F, Celeste Vega M, Rolón MS, Santos WS, Rojas de Arias A. Community participation in Chagas disease vector surveillance: systematic review. *PLoS Negl Trop Dis*. 2011;5:e1207.
2. Abbasi A, Altmann J, Hwang J. Evaluating scholars based on their academic collaboration activities: two indices, the RC-index and the CC-index, for quantifying collaboration activities of researchers and scientific communities. *Scientometrics*. 2010;83:1-13.
3. Abbasi A, Altmann J. On the correlation between research performance and Social Network Analysis measures applied to research collaboration networks. *IEEE, Proceedings of the 44th Hawaii International Conference on System Sciences*. 2011. [Conference on the Internet]. [cited 2011 Nov 30] Available from: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5718491>
4. Almeida EA, Ramos Junior AN, Correia D, Shikanai Yasuda MA. Rede brasileira de atenção e estudos na co-infecção *Trypanosoma cruzi*/HIV e em outras condições de imunossupressão. *Rev Soc Bras Med Trop*. 2009;42:605-8.
5. Bales ME, Johnson SB, Weng C. Social network analysis of interdisciplinarity in obesity research. *AMIA Annu Symp Proc*. 2008;40:870.
6. Bales ME, Johnson SB, Keeling JW, Carley KM, Kunkel F, Merrill JA. Evolution of coauthorship in Public Health services and systems research. *Am J Prev Med*. 2011;41:112-7.
7. Barabasi AL, Albert R. Emergence of scaling in random networks. *Science*. 1999;286:509-12.
8. Barabasi AL, Jeong H, Neda Z, Ravasz E, Schubert A, Vicsek T. Evolution of the social network of scientific collaborations. *Physica A*. 2002;311:590-614.

9. Burt RS. Structural holes: the social structure of competition. Cambridge: Harvard University Press; 1995.
10. Burt RS. Structural holes and good ideas. *Am J Sociol*. 2004;110:349-99.
11. Chang YF, Chen CM. Classification and visualization of the social science network by the minimum span clustering method. *J Am Soc Inf Sci Technol*. 2011;62:2404-13.
12. Costa Lima JA, Schmitt Rosa CM, Piegas MH, Peixinho A, Schmidt A, Briquet de Lemos AA, *et al*. Analysis of scientific information published in Brazil in 5 years on Chagas disease, schistosomiasis, malaria, leishmaniasis and filariasis. *Educ Med Salud*. 1985;19:209-26.
13. Costas R, Bordons M. Do age and professional rank influence the order of authorship in scientific publications? Some evidence a micro-level perspective. *Scientometrics*. 2011;88:145-61.
14. Crane D. Social structure in a group of scientists: a test of the "invisible college" hypothesis. *Am Sociol Rev*. 1969;34:335-53.
15. De Andrade AL, Zicker F, De Oliveira RM, Almeida Silva S, Luquetti A, Travassos LR, *et al*. Randomised trial of efficacy of benzimidazole in treatment of early *Trypanosoma cruzi* infection. *Lancet*. 1996;348:1407-13.
16. Dias JCP. Elimination of Chagas disease transmission: perspectives. *Mem Inst Oswaldo Cruz*. 2009;104(Suppl 1):41-5.
17. Falagas ME, Karavasiou AI, Bliiziotis IA. A bibliometric analysis of global trends of research productivity in tropical medicine. *Acta Trop*. 2006;99:155-9.
18. García CC, Martruccelli CRN, Rossilho MMF, Denardin OVP. Authorship for scientific papers: the new challenges. *Rev Bras Cir Cardiovasc*. 2010;25:559-67.
19. Garg KC, Kumar S, Madhavi Y, Bahl M. Bibliometrics of global malaria vaccine research. *Health Info Libr J*. 2009;26:22-31.
20. Gascon J, Bern C, Pinazo MJ. Chagas disease in Spain, the United States and other non-endemic countries. *Acta Trop*. 2010;115:22-7.
21. Glover SW, Bowen SL. Bibliometric analysis of research published in Tropical Medicine and International Health 1996-2003. *Trop Med Int Health*. 2004;9:1327-30.
22. Goh KI, Oh E, Kahng B, Kim D. Betweenness centrality correlation in social networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2003;67(1 Part 2):017101.
23. González-Alcaide G, Valderrama-Zurián JC, Ramos-Rincón JM. Producción científica, colaboración y ámbitos de investigación en enfermedades infecciosas y microbiología clínica (2003-2007). *Enferm Infecc Microbiol Clin*. 2010;28:509-16.
24. González-Alcaide G, Alonso-Arroyo A, González de Dios J, Sempere AP, Valderrama-Zurián JC, Alexandre-Benavent R. Redes de coautoría y colaboración institucional en Revista de Neurología. *Rev Neurol*. 2008;46:642-51.
25. Jackson Y, Gétaz L, Wolff H, Holst M, Mauris A, Tardin A, *et al*. Prevalence, clinical staging and risk for blood-borne transmission of Chagas disease among Latin American migrants in Geneva, Switzerland. *PLoS Negl Trop Dis*. 2010;4:e592.
26. Keiser J, Utzinger J. Trends in the core literature on tropical medicine: a bibliometric analysis from 1952-2002. *Scientometrics*. 2005;62: 351-65.
27. Kleinfield JS. The small world problem. *Society*. 2002;39:61-6.
28. Krackhardt D. The strength of strong ties: the importance of philos in organizations. In: Nohria N, Eccles R, editors. *Networks and organizations: structure, form, and action*. Boston: Harvard Business School Press; 1992. p. 216-39.
29. Melin G, Persson O. Studying research collaboration using coauthorships. *Scientometrics*. 1996;36:363-77.
30. Mullins NC. *Social networks among biological scientists*. New York: Arno Press; 1980.
31. Muñoz J, Gómez i Prat J, Gállego M, Gimeno F, Treviño B, López-Chejade P, *et al*. Clinical profile of *Trypanosoma cruzi* infection in a non-endemic setting: immigration and Chagas disease in Barcelona (Spain). *Acta Trop*. 2009;111:51-5.
32. Newman ME. The structure of scientific collaboration networks. *Proc Natl Acad Sci USA*. 2001;98:404-9.
33. Newman ME. Coauthorship networks and patterns of scientific collaboration. *Proc Natl Acad Sci USA*. 2004a;101(Suppl 1):5200-5.
34. Newman ME. Who is the best connected scientist? A study of scientific coauthorship networks. In: Ben-Naim E, Frauenfelder H, Toroczkai Z, editors. *Complex Networks*. Berlin: Springer; 2004b. p. 337-70.
35. Price DJS. Networks of scientific papers. *Science*. 1965;149:510-5.
36. Price DJS, Beaver D. Collaboration in an invisible college. *Am Psychol*. 1966;21: 1011-8.
37. Ramos JM, González Alcaide G, Gascón J, Gutiérrez F. Mapping of Chagas disease research: analysis of publications in the period between 1940 and 2009. *Rev Soc Bras Med Trop* 2011;44:708-16.
38. Ramos Júnior AN, Correia D, Almeida EA, Shikanai-Yasuda MA. History, current issues and future of the Brazilian network for attending and studying *Trypanosoma cruzi* HIV coinfection. *J Infect Dev Ctries*. 2010;4:682-8.
39. Rollin L, Darmoni S, Caillard JF, Gehanno JF. Searching for high-quality articles about intervention studies in occupational health: what is really missed when using only the *Medline* database? *Scand J Work Environ Health*. 2010;36:484-7.
40. Scott J. *Social network analysis: a handbook*. London: Sage Publications; 1991.
41. Silveira AC, editor. *O controle da doença de Chagas nos países do cone sul da América: história de uma iniciativa internacional 1991-2001*. Brasília: Organização Pan-americana da Saúde; 2002. p. 15-43.
42. Schoonbaert D, Demedts V. Analysis of the leprosy literature indexed in *Medline* (1950-2007). *Lepr Rev*. 2008;79:387-400.
43. Sosa Estani S, Segura EL, Ruiz AM, Velazquez E, Porcel BM, Yamposis C. Efficacy of chemotherapy with benzimidazole in children in the indeterminate phase of Chagas' disease. *Am J Trop Med Hyg*. 1998;59:526-9.
44. Vanz SAS, Stumpf IRC. Colaboração científica: revisão teórico-conceitual. *Perspect Ciên Inf*. 2010;15:42-55.
45. Watts DJ, Strogatz SH. Collective dynamics of 'Small-World' networks. *Nature*. 1998;393:440-2.
46. Watts DJ. The "new" Science of networks. *Annu Rev Sociol*. 2004;30:243-70.
47. Williams JR, Bórquez A, Basáñez MG. Hispanic Latin America, Spain and the Spanish-speaking Caribbean: a rich source of reference material for public health, epidemiology and tropical medicine. *Emerg Themes Epidemiol*. 2008;5: 17.
48. Wu B, Zhao F, Yang S, Suo L, Tian H. Characterizing the evolution of collaboration network. In: *SWSM 09: Proceeding of the 2nd ACM workshop on Social Web Search and Mining*. Hong Kong: SWSM'09; 2009. p. 33-40.
49. Zheng HC, Yan L, Cui L, Guan YF, Takano Y. Mapping the history and current situation of research on John Cunningham virus: a bibliometric analysis. *BMC Infect Dis*. 2009;9:28.

Received: 10 January 2012

Accepted: 10 April 2012