

# **Network analysis to support public health: evolution of collaboration among leishmaniasis researchers**

Ricardo B. Sampaio<sup>1,2</sup>, Bruna P.F. Fonseca<sup>1</sup>, Ashwin Bahulkar<sup>3</sup>, Boleslaw K. Szymanski<sup>3,4</sup>

Affiliations:

<sup>1</sup>National Institute for Science and Technology on Innovation on Neglected Diseases (INCT/IDN), Center for Technological Development in Health (CDTS), Oswaldo Cruz Foundation (Fiocruz), Rio de Janeiro, RJ, Brazil

<sup>2</sup>Diretoria Regional de Brasilia, Fiocruz, Brasilia, DF, Brazil

<sup>3</sup>Computer Science, Rensselaer Polytechnic Institute, Troy, NY, USA

<sup>4</sup>Spółeczna Akademia Nauk, Łódź, Poland

Financial support: This paper was supported by CNPq, CAPES and FAPERJ (INCT-IDN #573642/2008-7) and by ARL (NS CTA Agreement #W911NF-09-2-0053)

## **Abstract**

Databases on scientific publications are a well-known source for complex network analysis. The present work focuses on tracking evolution of collaboration amongst researchers on leishmaniasis, a neglected disease associated with poverty and very common in Brazil, India and many other countries in Latin America, Asia and Africa. Using SCOPUS and PubMed databases we have identified clusters of publications resulting from research areas and collaboration between countries. Based on the collaboration patterns, areas of research and their evolution over the past 35 years, we combined different methods in order to understand evolution in science. The methods took into consideration descriptive network analysis combined with lexical analysis of publications, and the collaboration patterns represented by links in network structure. The methods used country of the authors' publications, MeSH terms, and the collaboration patterns in seven five-year period collaboration network and publication networks snapshots as attributes. The results show that network analysis metrics can bring evidences of evolution of collaboration between different research groups within a specific research area and that those areas have subnetworks that influence collaboration structures and focus.

## **1 Introduction**

Collaborative networks are useful means to address the complexity inherent in health research. Collaboration is a critical process for innovation development in this area and researcher networks have been considered an integral element for innovative performance in the health sector (Guler & Nerkar 2012). The multifactorial, dynamic, and nonlinear character of disease causation have been reflected in an increased emphasis on transdisciplinary, translational, and network-based research in health (Barabási 2007). In addition to collaboration and networking, there is also a need for greater efficiency in health research, which is an imperative factor for reducing the time between new scientific discoveries and patient benefits (Wilson et al. 2011).

Leishmaniasis has been a challenging neglected tropical diseases (NTD) due to its complex epidemiology and ecology, the lack of simple, easily-applied tools for case management and the paucity of current incidence data, which often results in a failure

on the part of policy-makers to recognize its importance (Alvar et al. 2006; Bern et al. 2008).

The disease occurs on five continents and is considered endemic in 98 countries and three territories, most of which are low and middle-income countries (LMIC) (Alvar et al. 2012). Despite the fact that Brazil and India are among the top three most scientifically productive countries in leishmaniasis research (Ramos et al. 2013) and that their researchers are highly engaged in collaborative networks (González-Alcaide et al. 2013), these countries still account for 80-90% of cases in their specific regions (Alvar et al. 2012). This apparent gap in translating science discoveries to public health products and practices exposes an urgent need to develop new methods to support research and science-related policies and give new directives for scientific investments.

As the results of scientific work in leishmaniasis do not seem to directly affect the reduction of cases in endemic areas, even if scientific development has been nurtured in those regions, understanding the evolution and dynamics of science and collaboration in this theme is key and the driving force for this study. We aim to understand how the global leishmaniasis research network has evolved over time, in both time and space, and generate evidence that could ultimately inform the prioritization of research, financial investments and health policy.

We investigate the following three issues:

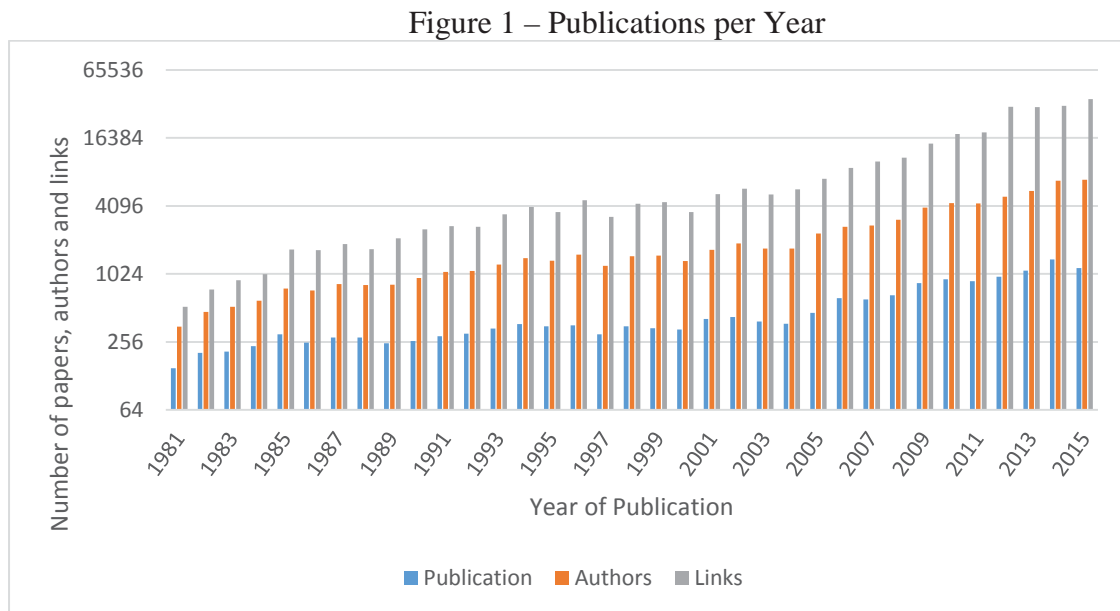
1. Scientific networks are an important way to address needs and challenges of LMIC dealing with NTDs (Carlos M. Morel et al. 2005; Carlos M. Morel et al. 2007). Important questions we attempt to answer include: Are the most endemic countries significantly producing knowledge to address leishmaniasis challenges? What are the patterns of collaboration in leishmaniasis regarding endemic and non endemic countries?
2. Governments and funding agencies are increasingly supporting collaborative research networks to encourage translation of research results into practice, including treatment for NTD (McKew & Pilon 2013; D'Andreta et al. 2013). Is scientific research in leishmaniasis addressing public health needs? How are different countries directing their scientific research efforts? Do they overlap or complement each other?
3. In the initial stages of medical research, when basic scientific research and pre-clinical phase are conducted, broad international collaboration may be highly beneficial. However, in the later phases of clinical trials, broad involvement of local actors who have an intimate understanding of their own health systems and challenges can increase the focus on national research priorities, enhance capacity through bringing together researchers with differing disciplinary skills, and facilitate longer-term trust-based networks (Bennett et al. 2011). In this sense, can we identify collaboration patterns and model the leishmaniasis research network using features intrinsic to the network itself and their members?

This paper focuses on tracking evolution of research topics and collaborations amongst researchers working on leishmaniasis by using scientific publication data from Scopus and PubMed databases from the past 35 years (1981-2015). Three complementary analyses were made: i) Descriptive and statistical analysis of researchers collaboration networks, characterizing scientists based on Medical Subject Headings (MeSH) taxonomy and their countries of affiliations; ii) Evolution of research topics as tracked by lexical analysis of titles and classification of publications; and iii) Link prediction algorithms to confirm existence or absence of persistent patterns in different time periods to identify evolving collaborations. From these analyses, we

extrapolate the potential factors that can influence the global and local collaboration structure in different stages of leishmaniasis research.

## 2 Data

Data collection strategy was based on retrieving scientific publications on leishmaniasis from Scopus database (Elsevier). Queries were directed to the title, abstract and keywords of publications (leish\*) and retrieved 16,927 documents. Filters for Subject Areas on Health Sciences (based on Scopus classification) and articles were further included. The search period comprised 35 years (1981 to 2015), which were later divided into seven periods of five years each. A steady growth of publications, authors and collaboration links (edges) between researchers in the area can be clearly observed (Figure 1).



Next step was to incorporate MeSH terms into the dataset. MeSH is the National Library of Medicine's controlled vocabulary thesaurus, used for indexing articles of biomedical journals available in the PubMed database. It consists of sets of general descriptors (headings) and qualifiers (subheadings) organized in a hierarchical structure that permits categorizing articles at various levels of specificity.

Of the 16,927 publications retrieved from Scopus, 13,806 articles were found in PubMed, (81,5%). The combination of information from these two databases allowed us to use country attributes retrieved from author affiliations available in Scopus and the classification of publications available from PubMed. Using this data, we created three networks. The first is the co-authorship network between countries, based on the professional affiliation of researchers authoring the publications. The second is an information network in which nodes are classification terms (MeSH terms and words from titles and abstracts of publications) with links joining terms that co-occur in the same publications. The third is the co-authorship network between researchers, with links between them indicating joint publications.

The total number of unique authors of the articles was 42,893. Calculations for link prediction and percentage of relevance for network structure included only authors that had three or more publications over the entire period. Disambiguation of author names was performed using the VantagePoint software (Search Technologies Inc.).

### 3 Methods

To build our understanding of the dynamics and evolution of collaboration patterns within leishmaniasis research and evolution of research topics, we used two different methods, as described below.

#### 3.1 Descriptive Analysis of Network Data

##### 3.1.1 Co-authorship Network between countries

Researchers' affiliation data was used to build collaboration networks between countries. In these networks, nodes represent countries, and two or more countries were connected if their members shared the authorship of one or more papers. Visualization of the network graphs were produced with the open-source software Gephi (Bastian et al., 2009). Network community structure was evaluated based on the number of ties within each community.

Degree centrality was used to identify the most collaborative countries in the network. This measure reflects the number of a node's direct connections (Freeman 1979). Countries with high degree centrality are usually focal points of collaboration in the network.

Degree centrality is calculated as the number of direct links a particular node has with other nodes. If a node represents a country, the number of other countries, with which it collaborates, will be its degree centrality. Weighted degree centrality takes into consideration not only the number of nodes to which the given node is connected, but also the number of times each pair of nodes collaborated. If node A collaborated twice with node B and 3 times with node C its degree centrality is 2 but its weighted degree centrality is 5.

We also used betweenness centrality to analyze bridging countries for knowledge transfer and access. Betweenness centrality is based on the extent to which a particular node lies between other pairs of nodes in a network, connecting them (Freeman 1979). Nodes that are often on the shortest path between other nodes are deemed 'central' because they control the flow of information in the network by connecting different groups. They are also the ones whose removal from the network will disrupt communications between other nodes the most.

Betweenness centrality is calculated using the shortest paths between all distinct pairs of network nodes. Thus, given node A, counting all of the shortest paths between any two nodes in the network that pass through node A will result in its betweenness centrality.

Both measures were calculated on each one of the seven five-year period networks snapshots in order to observe their evolution over time.

##### 3.1.2 Lexical Analysis of the information network

Lexical analysis was performed with the software IRAMUTEQ (Interface R for Multidimensional Analysis of Texts and Questionnaires), based on a methodology developed by Max Reinert in the 1980s (Reinert 1983; Reinert 1986). The IRAMUTEQ has been used in sociology, psychology, and political and health sciences for the quantitative analysis of text or textual statistical data (Guarnaccia et al. 2015; Rodrigues et al. 2015).

In order to apply the lexical analysis, four different text corpora were built comprising title, abstract, MeSH and MeSH subheads of the retrieved publications. For each corpus, a classification was incorporated based on the information available for the text. Then, term frequencies have been computed and two-dimensional Cartesian

map visualization of terms was drawn based on component analysis of the word distributions. Each of the four sets was categorized according to the five-year periods and by countries of authors' affiliation.

### 3.2 Link Prediction

#### 3.2.1 Co-authorship network between researchers

We use machine learning-based link prediction to predict collaborations between researchers in one five-year period based on the data from the preceding five-year period. The innovative purpose of using link prediction here is to see if new collaborations can be predicted from trends seen in the creation of old collaborations and therefore to find when the patterns of collaboration are stable and when they evolve. Good predictability in a given time period indicates a stable pattern of collaboration, while low predictability indicates evolving patterns of collaboration.

As in the information network, the original scientific collaboration network was divided into seven network snapshots, each of which covered different five-year periods. Each instance of data contains collaborations that happened during the particular timespan. Nodes (researchers) in each network have several attributes, such as country of affiliation, MeSH keywords attributed to their papers, and the number of papers published during that particular timespan for each MeSH keyword. We use these node attributes and the network structure to predict links in the future. The properties of the seven collaborative networks are listed in Table 1. It is worth noting that from timespan to timespan both number of nodes and average degree are increasing (with one exception of the average degree drop in 1996-2000 period compared to 1991-1995 period). This demonstrates that both the number of researchers and the extent of collaboration were growing over the years.

Table 1: Basic properties of the seven collaborative networks created from seven time spans of the entire publication data.

Snapshot of the network	Number of nodes	Number of edges	Average node degree
1981-1985	670	1170	1.75
1986-1990	1114	2714	2.44
<i>1991-1995</i>	1626	4610	<i>2.84</i>
<i>1996-2000</i>	2007	5638	<i>2.81</i>
2001-2005	2617	9896	3.78
2006-2010	4117	18185	4.42
2011-2015	4832	33549	6.94

#### 3.2.2 Link Prediction Tasks

The goal of traditional link prediction application is to foresee future collaborations among researchers. Here our goal is to use link prediction to identify periods of evolution of collaborations during which collaboration patterns changed. To this end, we compared the performance of link prediction on edges formed between researchers who were new to the field to the one based on edges between experienced researchers. The drastic drop in the performance of link prediction from one period to another indicated a change in the way new researchers collaborate

To use machine learning for link prediction, we used data on links formed in the past. For example, to find which new links are likely to form in the timespan 1991-1995



from 1986-1990 network snapshot, we need to learn from the links formed in this snapshot. So, we look at the new links formed there based on yet earlier 1981-1985 network snapshot, and use these links as training data to predict links in the timespan 1991-1995. Consequently, for each network snapshot for which we execute link prediction, we use two immediately preceding network snapshots. Data from the earliest timespan serves as the training data, those from the middle timespan is used for validation and the latest one enables us to evaluate predictions. It is the information about the nodes and their connections in the middle timespan that is used to predict which new connections are likely to be formed in the latest timespan.

Starting from the earliest period, the 1981-1985 timespan, the network corresponding to the subsequent period was used to validate the machine learning model and the links were predicted for the network snapshot for the 1991-1995 period. We processed data in each period using SVM, Linear Regression and k-NN machine learning algorithms. We repeated this process for subsequent four later periods ending with the five sets of results.

### 3.2.3 Features Used in Link Prediction

We used several cues from the Scopus data to predict new links. The following features were used.

**Network Features - The Number of Common Neighbors:** We used the number of common neighbors between a pair of nodes as a feature because positive correlation between link formation and the number of common neighbors was observed in the literature.

**Individual Traits - Areas of Shared Research:** For every researcher, we know the number of papers published in each area in the particular timespan data from which the given network snapshot has been created. We used as a feature the number of common areas of research in which a pair of nodes has published in the corresponding timespan.

**Individual Traits - Country of Affiliation:** For every researcher, we know the country in which the researcher was based. The feature had the value 1 if both researchers were from the same country and 0 otherwise.

**Individual Traits - Recency and Strength of Collaboration:** Representing collaboration via an unweighted edge misses the temporally evolving nature of collaboration. To address this concern we introduced the weights for collaboration edges that represent its strength and recency. We assumed that strong and recent collaboration among common neighbors of a node increases chances of those neighbors promoting collaboration of the node shared by others. Moreover, a large number of strong common neighbors are indicative of a node belonging to the same research community as its common neighbors, thus making collaboration more feasible. The opposite is likely to happen in the case of a few common neighbors or with weak or aging collaboration.

One feature we used was the *strength* of a link between two nodes which measured the number of papers that those nodes published together. Another was the *recency* of the link which decreased with the age of joint publication by the factor  $f$  (Bahulkar et al. 2016) defined here as

$$(1) \quad f = f_d^k$$

where  $f_d=0.9$  and  $k$  denote the age of the publication in years, with articles published in the last year of the timespan period having  $k=0$ . Clearly, the link strength is always higher than the link recency.

### 3.2.4 Processing Data

Link prediction considered any viable pair of nodes and predicted on the basis of their features if the link is likely to be formed. However, the number of potential edges in a network was growing with the square of the number of nodes, so this approach would be computationally infeasible. To address this concern, we adopted commonly used heuristic that new edges arise between nodes that are at most three hops away from each other. Thus, only pairs of nodes satisfying this condition were considered viable for having an edge.

First step in processing data was using the training validation data to find for each of the four algorithms tested its best parameters for the prediction run. Then, we ran the machine learning algorithms on testing data to produce prediction. We reported results separately for pairs of which both nodes existed only in the testing data (*new pairs*) and for pairs for which at least one node existed in the validation data (*old pairs*). The results of link prediction are discussed in the next section.

## 4 Results

### 4.1 Descriptive Analysis of Network Data

#### 4.1.1 Publication and Collaboration

Figure 2 – Publications of top seven countries per period

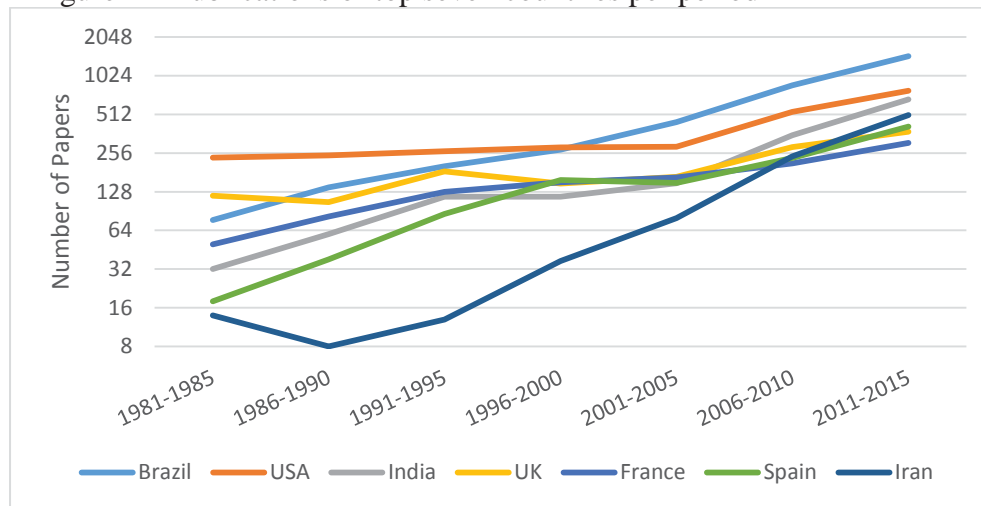
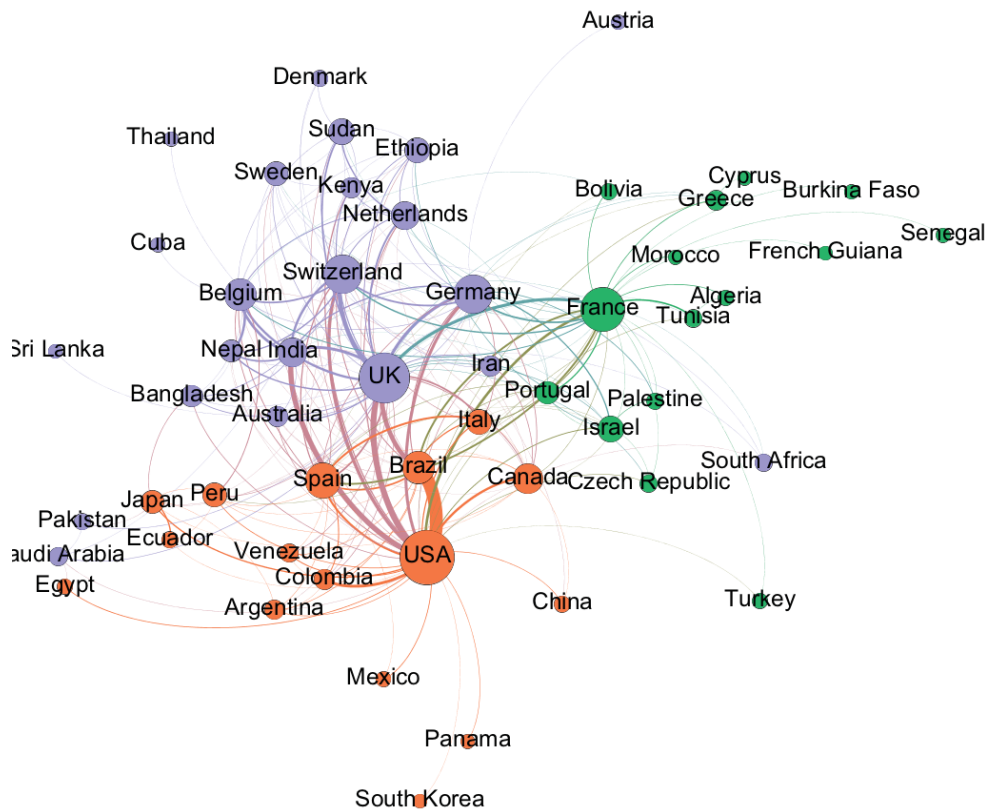


Figure 2 shows the top seven countries in leishmaniasis research, according to the number of papers published from 1981-2015. Brazil and India were among the most scientifically productive countries, but they reached this position after year 2000. Since then, Brazil held the first place in number of publications and India continued to rise from sixth (6<sup>th</sup>) place in the period of 2001 to 2005 to third (3<sup>rd</sup>) place from 2006 until 2015. This finding is in line with the increased financial support for NTD research in these two countries. Another country that showed an improving position on number of publications is Iran, not amongst the top 10 until 2001, but seventh (7<sup>th</sup>) overall and fourth (4<sup>th</sup>) in the last period.

The number of publications is just part of the picture. The use of network analysis for those periods revealed a different perspective in terms of influence over the area.

As scientific research is increasingly collaborative, network analysis methods were used to map countries' collaborations, their contributions and influences. The global network of leishmaniasis research is composed of 127 countries, reflecting the solid international collaborative research efforts for disease control (Figure 3). The top three most collaborative countries, according to their weighted degree centrality, were USA, France and UK, as represented by the larger nodes. During the 35-year period evaluated, these countries have collaborated with 111, 109 and 106 partner countries, respectively.

Figure 3 – Country collaborative network of leishmaniasis research (1981-2015).



Country links were mapped based on the affiliations of the authors of scientific papers. Each node represents one country and two countries were considered connected if their authors shared the authorship of a paper. The size of the nodes is proportional to their weighted degree centrality. The thickness of the links indicates the frequency of collaboration between two nodes. The node color indicates the cluster in which the country has been placed according to its collaboration pattern.

Degree centrality is a proxy for collaboration and not always a measure of the volume of publications. Although Brazil and India had a high number of publications in leishmaniasis, these countries lagged behind in terms of collaborative research.

Analysis of the network's community structure revealed three different clusters: cluster 1, including Brazil, Spain and USA (high degree centrality countries), cluster 2 including UK, Switzerland and Germany, and cluster 3, including France, Israel and Portugal. These clusters group countries that frequently collaborate with each other in leishmaniasis research.



Cluster 1 included several Spanish-speaking countries, suggesting that language proximity may have played an important role in establishing collaborations. Brazil and USA were the countries that most frequently collaborate, with 362 articles published in co-authorship in the period evaluated. India and Brazil, although scientifically productive in this area, were not included in the same cluster, indicating that these countries did not collaborate often.

Table 2 shows the number of publications per country and the degree centrality (Deg), weighted degree centrality (W.deg), betweenness centrality (Bet) and a ratio between W.deg and publications (Wdeg/Pub) for the top seven countries. This ratio is a measure of the percentage of international collaborations these countries had in relation to their entire number of publications. Because one publication might have authors from more than two countries, this result was just an estimate of the number of collaborations per paper each country had.

Table 2: Publication and centrality measures

1981-1985						2001-2005					
Country	Pub	Deg	Wdeg	Bet	Wdeg/Pub	Country	Pub	Deg	Wdeg	Bet	Wdeg/Pub
Brazil	77	5	11	0,015	14%	Brazil	445	27	128	0,0412	29%
USA	236	13	34	0,071	14%	USA	286	50	219	0,262	77%
India	32	3	3	0,000	9%	India	149	11	36	0,041	24%
UK	119	14	43	0,075	36%	UK	167	46	197	0,171	118%
France	50	7	10	0,027	20%	France	165	42	133	0,151	81%
Spain	18	2	2	0,000	11%	Spain	150	24	46	0,054	31%
Iran	14	1	1	0,000	7%	Iran	80	9	27	0,008	34%
1986-1990						2006-2010					
Country	Pub	Deg	Wdeg	Bet	Wdeg/Pub	Country	Pub	Deg	Wdeg	Bet	Wdeg/Pub
Brazil	138	9	29	0,007	21%	Brazil	858	30	206	0,019	24%
USA	246	27	67	0,162	27%	USA	534	64	516	0,152	97%
India	60	0	0	0,000	0%	India	351	35	186	0,039	53%
UK	106	18	37	0,050	35%	UK	285	50	436	0,075	153%
France	82	10	19	0,017	23%	France	213	25	60	0,006	28%
Spain	38	2	2	0,000	5%	Spain	234	36	175	0,022	75%
Iran	8	0	0	0,000	0%	Iran	241	58	250	0,148	104%
1991-1995						2011-2015					
Country	Pub	Deg	Wdeg	Bet	Wdeg/Pub	Country	Pub	Deg	Wdeg	Bet	Wdeg/Pub
Brazil	202	10	41	0,013	20%	Brazil	1448	99	678	0,023	47%
USA	265	26	95	0,124	36%	USA	779	102	1058	0,044	136%
India	117	8	14	0,002	12%	India	669	92	625	0,009	93%
UK	184	27	86	0,129	47%	UK	375	97	869	0,014	232%
France	128	20	45	0,067	35%	France	307	103	647	0,043	211%
Spain	86	5	9	0,002	10%	Spain	411	100	598	0,017	145%
Iran	13	2	2	0	15%	Iran	505	86	243	0,001	48%
1996-2000						1981-2015					
Country	Pub	Deg	Wdeg	Bet	Wdeg/Pub	Country	Pub	Deg	Wdeg	Bet	Wdeg/Pub
Brazil	272	17	78	0,041	29%	Brazil	3440	101	1171	0,017	34%
USA	283	37	186	0,152	66%	USA	2630	111	2175	0,059	83%
India	117	11	46	0,005	39%	India	1496	95	910	0,023	61%
UK	148	40	135	0,199	91%	UK	1384	106	1803	0,031	130%

France	152	26	65	0,128	43%	France	1097	109	1169	0,055	107%
Spain	158	13	31	0,010	20%	Spain	1095	103	863	0,017	79%
Iran	37	5	12	0,008	32%	Iran	898	88	345	0,001	38%

This descriptive analysis shows that the number of collaborations (degree and weighted degree) was increasing in a much faster ratio than the number of publications, indicating that international collaborations were becoming more common with time. Another important result shown in Table 2 is the centrality of the top five countries. While Brazil, India and Iran had a similar pattern of high number of publications and low centralities, USA, UK and France did not have the largest number of publications, but held an important place in terms of number of international collaborations (degree) and fulfilling the gap or the bridge between different nations (betweenness). Those are important features of countries that could play a role in influencing research activities in specific areas (Leydesdorff and Wagner 2009).

The last result analyzed in Table 2 takes into consideration the ratio between weighted degree and the number of publications per country. The countries with higher ratios had a greater ability to cooperate with other countries. At this point, geographical distance and social differences such as language have not been taken into consideration. One of the results on this line of analysis showed that this ratio has grown extensively from 1981 to 2015, with more international collaborations per paper. Another interesting characteristic is that countries such as USA, Brazil and India, the top three countries in terms of publications, have lower ratios than other countries. In contrast, the European countries, especially UK and France, have high number of collaborations.

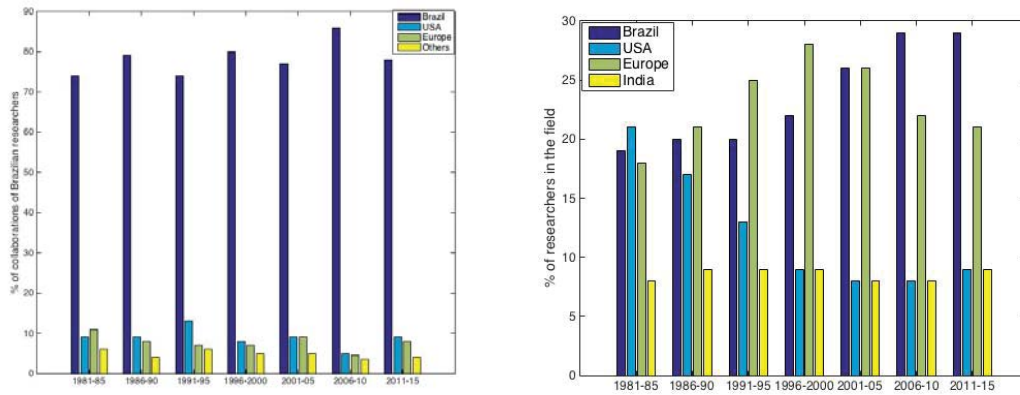
Table 3: Percentage of researchers belonging to a particular country

% nodes	1981-1985	1986-1090	1991-1995	1996-2000	2001-2005	2006-2010	2011-2015
Brazil	18.8	20.1	19.8	22.3	25.9	28.9	29.4
India	8	8.9	9.2	8.9	7.8	8	8.7
US	21.1	17.3	12.9	9	8	8	9
UK	6.1	5.1	5.4	4.4	5.2	4.1	3.7
France	5.8	6.4	7.4	8	6	4.3	4.2
Spain	2.1	4.8	6.5	8.2	7.4	5.4	6
Iran	1	0.6	1	2.4	4.2	5.5	6.5
Germany	1	2	2.1	2.5	2.3	2.9	2.5
Italy	1.4	2.9	4.3	4	4.5	3.3	2.7

The evaluation of the percentage of researchers from each country who contributed in each of the network snapshots shows that there was a gradual increase in the percentage of researchers from Brazil, while the percentage of researchers from several other countries, especially the USA, has declined over time. Table 3 supports this observation. In 1981-1985, Brazil had 18.8% of all researchers, while in 2011-2015, it increased its share to 29.4%. On the other hand, the USA had 21.1% of all researchers in 1981-1985, and it went down to 9% in 2011-2015. Since the total number of researchers itself has grown over years, this points to a huge increase in the number of researchers in Brazil from 1996-2000 onwards, which we will later talk about in the section on link prediction results. As shown in Figure 4, the Brazilian researchers kept

the patterns of collaborations unchanged over the years despite their growing share of active researchers in the area.

Figure 4 – Patterns of collaborations and percentage of researchers in the field over the seven five year periods



We measured how many researchers from each country were represented by nodes with high degrees. These researchers were usually associated with prominent positions, in which they might have had better access to information and resources and hence might have been able to influence the scientific system. We took the top 10% of the nodes sorted by their degrees, identified their countries of origin and presented the results in Table 4. Brazil had the 18.8 % of the total nodes in 1981-1985 and about 42% of the top degree nodes. Conversely, in 2011-2015, Brazil had 29.4 % of the total nodes and only 24.5% of the top degree nodes. This drop could be related to the significant increase in the overall number of researchers starting from the 1996-2000 period. Most of the researchers beginning in that period would be young researchers, early in their careers, so their number of collaborations might have been low. On the other hand, Indian researchers gradually increased their leadership, reaching from 1.5% in the first period to 12.2% in the last. This indicated a gradual change happening in the leadership.

Table 4: Percentage of researchers represented by high degree nodes

% of top degree nodes	1981-1985	1986-1990	1991-1995	1996-2000	2001-2005	2006-2010	2011-2015
Brazil	42	46.8	23.4	34	26.5	31.2	24.5
India	1.5	0.9	0	5.5	7.3	10.8	12.2
US	13.4	15.3	16.7	6	10.3	4.6	19
UK	12	4.5	8.6	5	12.6	7	4.8
France	12	7.2	12.4	6.5	5.3	4.7	4.5
Spain	0	0	1.8	0	6.9	5.1	4.3
Iran	0	0	0	5	1.5	4.4	3.2
Germany	0	0	0	0.5	2.2	2.6	1.4
Italy	0	0	1.3	4.5	2.3	3.2	1.9

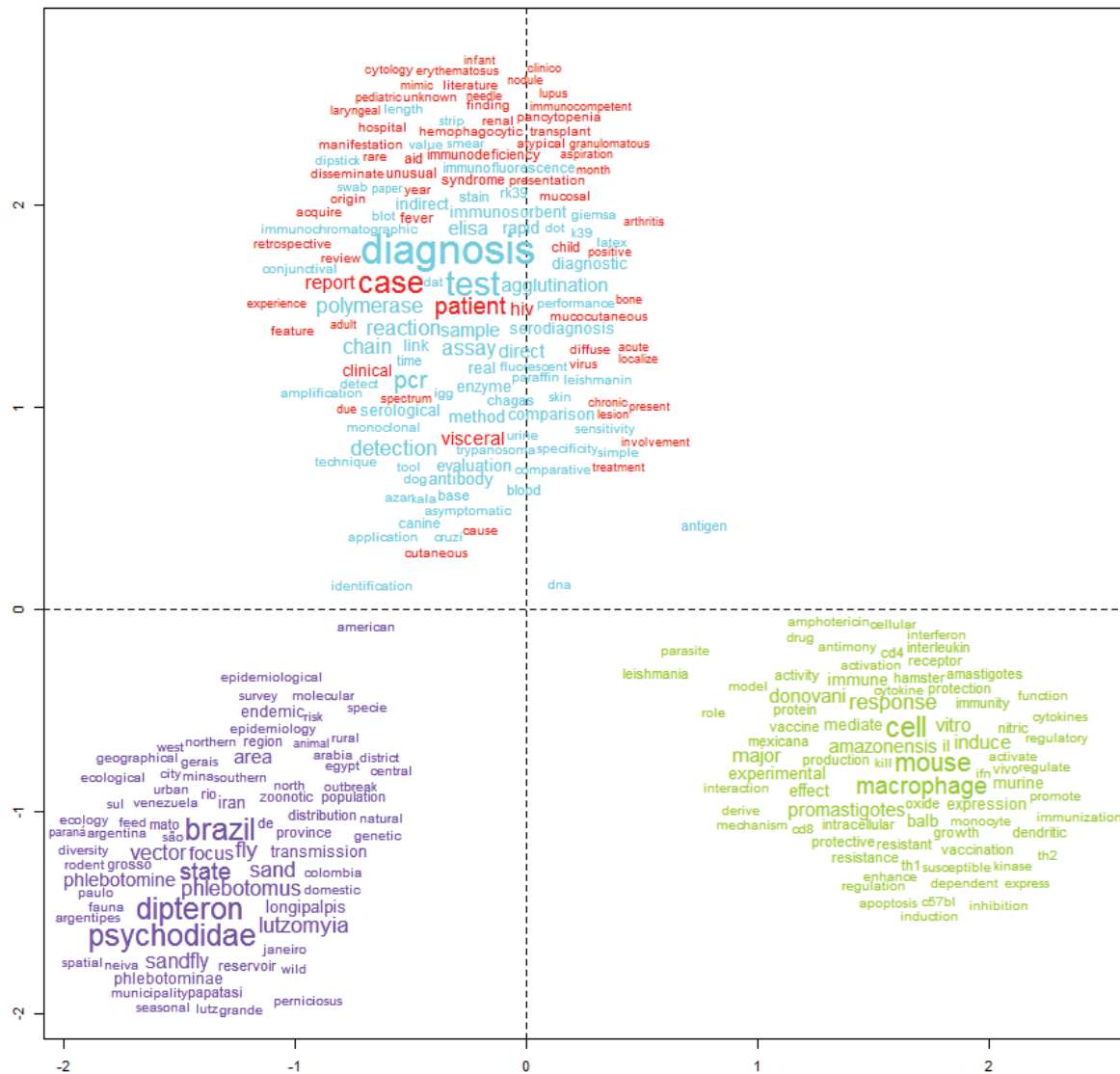
#### 4.1.2 Lexical Analysis

Lexical analysis was applied to title, abstracts, MeSH and subheadings from the top seven countries. Figure 5 shows major research themes from all 35 years captured by article titles. Three clusters in four different colors were identified by Reinert's

method, which classifies the words based on their co-occurrence and morphological analysis.

A closer look at the clusters showed a significant concentration in the top cluster on diagnosis of the disease, where research had a strong focus on the patient. On the bottom left, the words “psychodidae”, “dipeteron”, and “sandfly” showed a clear research pattern dealing with the vector and the dissemination of the disease. On the right bottom corner “cell”, “mouse” and “macrophage” showed a tendency for basic and molecular research. Analyzing how the research themes might change over countries, we found that over the seven countries, France and Spain overlapped with the top cluster, Brazil and Iran strongly overlapped with the bottom left of classification, and USA, UK and India had strongest affinity with the bottom right cluster.

Figure 5 – Classification of the article title words



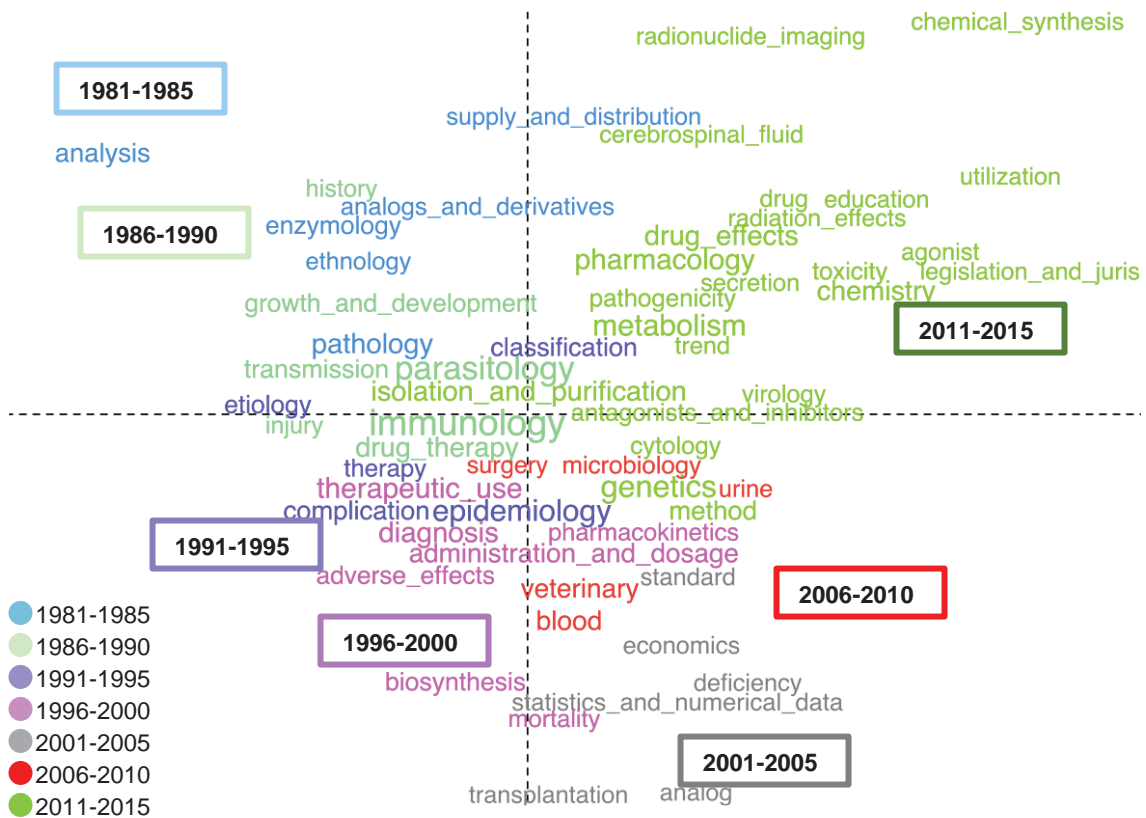
Similar patterns were observed in the analysis of all four textual corpuses. For space consideration, only the results of the subheading and MeSH titles over the time periods and countries were shown.

For the evolution of research over time we show on Figure 6 the MeSH subheads for each five-year period (similar patterns aroused from analysis of title words and abstracts). The subheads are colored according to the five-year period index on the

bottom left of the figure. The colored boxes with the time periods are placed according to the position on the Cartesian figure of the words most common to each period. The component analysis disperses the words according to their proximity to each of the time periods being analyzed. If a word is closer to the classification center, it is more common to other periods. If it is located on the periphery, it is more related to a specific period.

The displacement of the seven five-year periods follows a “u” shape from top left of the figure to the top right. “Immunology”, for example, has a specific color (light green), related to 1986-1990 period, but as it is placed in the classification center, so it might be common to other periods. “Chemical synthesis” on the other hand, is located at top right corner of the classification (dark green), thus, it will be almost exclusive to the 2011-2015 period. Some of the periods are closely related to each other as in the case of 1991-1995 to 1996-2000, which denotes smaller changes in the research topics compared to the other periods.

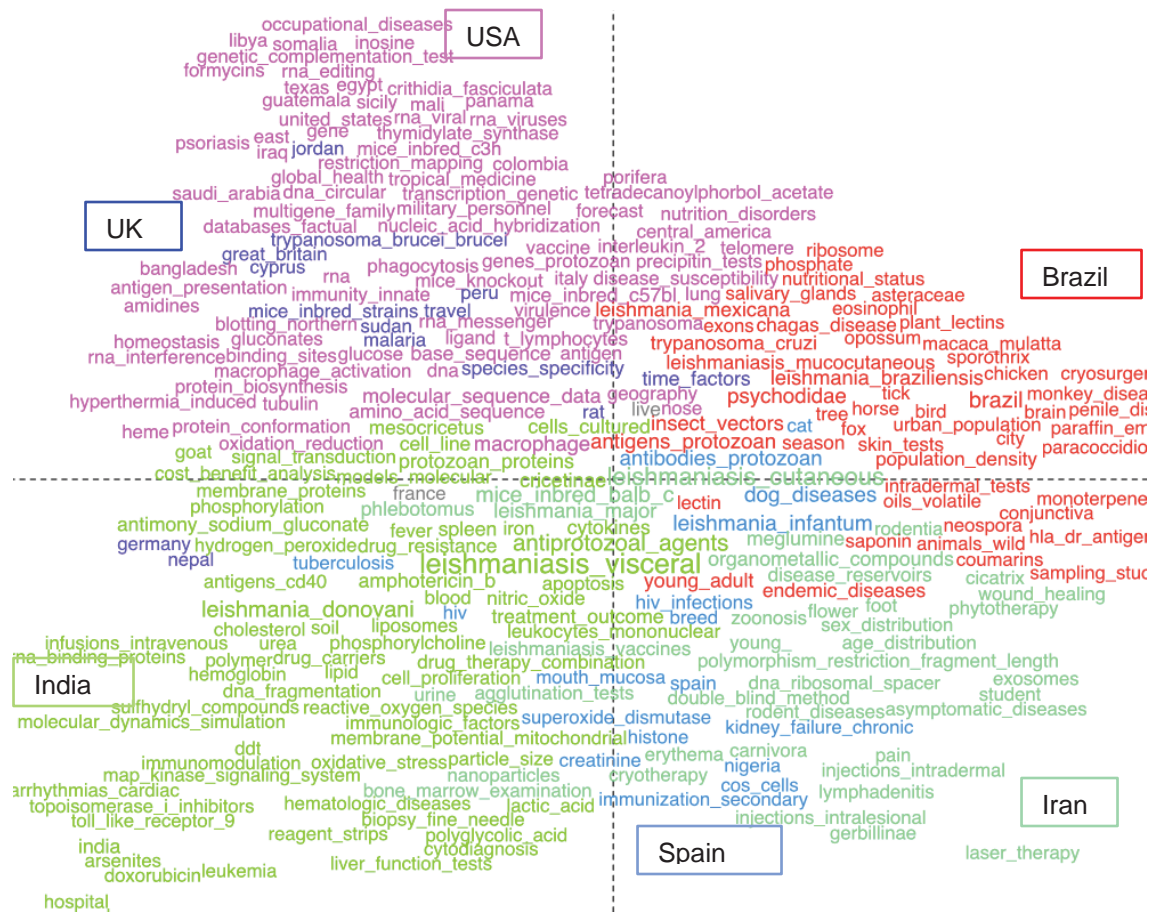
Figure 6 – Subhead Component Analysis distribution based on periods



This analysis sheds the light on the evolution of research over the years and can be summarized as follows. At the earliest decade, in the 1980’s, the research concentrated on disease pathology and transmission, with basic research focusing directly on enzymology. In the next decade, the 1990’s, research became more directly related to leishmaniasis diagnosis, epidemiology and treatment schemes. In the first decade of the 2000’s, the importance of the animal host was evidenced and the economic impacts of the disease also were investigated. In the past five years (2011-2015) drug-related research has again become the center of research focus.

Figure 7 – MeSH terms distributed by country





Country research profiles are summarized in Figure 7. Different leishmania species are closely related to the countries in which they are the primary cause of disease, like *Leishmania donovani* in India, *Leishmania braziliensis* in Brazil, and *Leishmania infantum* in Spain. Research themes between countries also differ, although limits are less clearly defined. Among other themes, USA scientists were concerned with leishmaniasis as an occupational disease, Brazilian researchers seemed to be interested in plant lectins and the Iranian scientific community had a particular interest in laser therapy as a treatment option. Common research themes between all countries included basic immunology and cell biology-related terms (cytokines, apoptosis, antigens etc.)

## 4.2 Link Prediction

In this section, we look at the results from link prediction and observe how the patterns of collaborations change over time. We started by creating features described in Section 3.2, for link prediction task. We used the machine learning algorithms listed in Section 3.2 to run over the data. We measured the accuracy, which represents the percentage of predicted and non-predicted edges that correctly identified in the test data, and recall, which measures the percentage of new edges in test data that are predicted by each of the classification tasks.

Initially, we look at predictability for each of the 5-year consecutive periods between 1991 and 2015. SVM usually gives the best recall, however at the price of a lower accuracy. Using k-NN or Linear Regression yields lower recall but higher accuracy. Since our main objective was to measure the ability of machine learning

techniques to predict links, the results in Table 6 demonstrate that SVM was the best method to meet this objective.

Despite the large changes observed between networks created for each time period (for example, only about 50% of the nodes and even a smaller percentage of the edges in the later network were present in the earlier network), patterns of collaboration were stable for the nodes that existed in both testing and training periods, as evidenced by the high predictability of the links, seen in Table 6.

Table 6. Results of link prediction using machine learning methods

Dataset		SVM		Linear Regression		k-NN	
Earliest-Middle	Latest	Accuracy	Recall	Accuracy	Recall	Accuracy	Recall
1981-1990	1991-1995	<b>80.6</b>	<b>73.5</b>	83.6	67.8	92.6	23.1
1986-1995	1996-2000	<b>79.6</b>	<b>82.5</b>	87.5	65.1	94.5	16.7
1991-2000	2001-2005	<b>82.9</b>	<b>80.6</b>	86.8	72.9	94.6	23.3
1996-2005	2006-2010	<b>83.1</b>	<b>76.6</b>	90.8	61.3	87.8	39.1
2005-2010	2011-2015	<b>82.5</b>	<b>83.5</b>	86.1	73.5	80.5	45.1

We then observe how the quality of the results changed when we extended the past timespans on which predictions were based and at the same time considered the large gap between the training and testing data. The results are shown in Table 7.

Table 7. Predictions based on at least two past timespans

Dataset		SVM		Linear Regression	
Middle	Latest	Accuracy	Recall	Accuracy	Recall
1986-1995	2001-2005	81.7	61.5	74.6	52.6
1986-1995	2006-2010	74.4	64.2	85.5	50.5
1991-2000	2006-2010	82.2	68.2	86.8	60.2
1986-1995	2011-2015	81.8	41.3	88.4	28.7
1991-2000	2011-2015	65.4	73.5	87.3	50.4
1996-2005	2011-2015	83.1	72.8	76.1	74.5

It seemed that using an extended timespan of 10 years has not affected the results much. However, the quality of the results decreased as the gap between training and testing data increased (e.g., using 1986-1995 period to predict 2011-2015 snapshot, so 15 years gap). This was particularly affecting the recall. An explanation for the relatively low recall values in this case was that only about 50% of the nodes in any snapshot of the network occurred in the previous snapshot. So, over the 15 year gap, which spanned four snapshots, only 6% of nodes in the predicted period existed in the training data. The higher level conclusion that can be drawn is that, with time, even established researchers modify their lists of collaborations in time spans longer than five years.

To further measure the effect of extended period on prediction, in Table 8, we used networks aggregated over 15 years. We found that the prediction rates improved significantly when a large timespan was used to predict networks in the distant future. This showed that larger timespans capture relationships and their attributes well and help us make better predictions.

Table 8. Predictions based on three past timespans

Dataset		SVM		Linear Regression	
Middle	Latest	Accuracy	Recall	Accuracy	Recall
1986-2000	2011-2015	68.5	74.5	80.1	63.7
1986-2000	2006-2010	69.7	78.6	78.5	75.5

Finally, to assess changes in the collaboration patterns in different periods, we measured the difference in recall values for edges formed between nodes that appeared for the first time, which we refer to as Recall for Edges from New Nodes, abbreviated as Recall-New, versus edges formed between nodes that existed in the previous time period (training data), which we refer to as Recall for Edges from Old Nodes, abbreviated as Recall-Old.

Table 9. Recall for edges from new nodes vs. that for edges from old nodes.

Middle network	Latest Network	Recall-New	Recall-Old
1986-1990	1991-1995	98.1	73.5
1991-1995	1996-2000	85.2	82.1
1996-2000	2001-2005	35.5	82.3
2001-2005	2006-2010	20.4	78.6
2006-2010	2011-2015	48.3	85.3

We observed that, until the year 2000, the prediction of edges formed from new nodes yielded a strong recall, quite surprisingly even higher than the persisting nodes. This indicates the period in which collaboration was evolving slowly, it was led by established researchers. However, from the year 2000 onwards, recall for the new edges dropped significantly below half of the recall of persisting nodes in years 2001-2005, and to nearly a quarter of the old node recall in years 2006-2010. Hence, the roles of new and old nodes changed completely, with the established researchers continuing moderate rate of collaboration change while the newly joining researchers engaging in different patterns. Interestingly, in the following period of years 2011-2015 the evolution of collaboration patterns started to stabilize as the recall for edges between researchers entering the leishmaniasis research field raised to over half of the recall of the old edges, that represent collaboration of established researchers. This is natural, since, as the new patterns develop and are present in the training data, they are captured by the machine learning based link prediction, and the recall even for new nodes recovers as they engage in the newly established patterns.

## 5 Discussion

The increase in the number of publications on leishmaniasis over the 1981-2015 period reflects not only the efforts in estimating the prevalence of leishmaniasis seen in recent years (WHO, 2010), but also the greater social awareness by funding agencies, including the Gates Foundation and the Wellcome Trust, on this disease (G-Finder, 2014). The inclusion of leishmaniasis in the World Health Organization (WHO) health agenda as an initiative to control the disease in endemic countries could also have influenced this trend (Alvar et al., 2012).

The analysis of scientific productivity in specific research areas is an important measure of a country's investment in research and development (R&D) efforts. Traditionally, high-income countries have accounted for most scientific publications produced in all areas of science (UNESCO, 2001), but here we show that specifically for leishmaniasis this is not always the case. Brazil and India have been increasing research activities in this area (Al-Mutawakel et al. 2010; Ramos et al., 2013), leading scientific productivity and engaging in collaborative activities (González-Alcaide et al., 2013). In these two countries research output is associated with disease endemicity. It is critical for the developing world to promote, through research and publications, those areas of concern that are having a proportionally greater scientific and social impact upon them (Holmgren e Schnitzer 2004) and putting a great research effort into leishmaniasis is a way of tackling their own health needs.

In addition, we show that this increase in leishmaniasis-related research mainly occurred after the 2000's. In Brazil, this may be related to the increase in funding for leishmaniasis from the government through the creation of programs to address neglected diseases (Brasil 2012) or specifically leishmaniasis (Sampaio et al. 2015). The more pronounced role of the Ministry of Health in defining and supporting research priorities since 2003 could also have played a role in this shift. Indian government has also invested in disease elimination (Singh et al. 2016) and given priority to investment in new drug development for leishmaniasis (Kettler e Modi 2001). Social network analysis (SNA) has been applied to understand collaboration networks in NTD and to generate evidence to guide policy-planning efforts in Brazil, Canada and Germany (Carlos Medicis Morel et al. 2009; Vasconcellos & Morel 2012; Phillips et al. 2013; Bender et al. 2015). The overall increase in leishmaniasis research collaboration reflects to some extent the global increase in scientific collaboration (Adams, 2012) and the concern, in an ever globalized world, to safeguard one's own population against introduced tropical diseases (Guerrant & Blackwood 1999).

In this analysis, the presence of USA, France and UK as most collaborative, and therefore most central in the research network, reflects their scientific commitment with global health issues and the increasing trend of high-income countries to conduct research on diseases that used to be restricted to developing countries. The collaboration between these countries with low and middle income countries (LMIC) would be important for research capacity strengthening, especially in the fundamental research and early stages of the medical research specific to the disease.

During the 35 years over which we analyzed the data, the three traditionally leading countries were joined by the three newcomers, Brazil, India and Iran, which devoted significant resources to successfully increase number of researchers and publications in these areas, especially in the last decade. Although Brazil and India were the most scientifically productive countries followed by Iran, they were not the most collaborative. It has been shown that these newcomer countries' researchers have been highly engaged in collaborative activities (González-Alcaide et al. 2013), but these partnerships might have been established with national researchers. Although collaboration has been positively associated with scientific productivity (Lee & Bozeman 2005), studies in LMIC have shown otherwise (Duque et al. 2005).

This can signal the limited role for international collaboration on such diseases, which is essential in the early stages of research focusing on basic research and pre-trials, versus the later stages that involve clinical trials, diagnosis and cures, and therefore need to rely on involvement of the local agents, such as doctors practicing in the areas in which the diseases is endemic. Additionally, it is common to national research programs to emphasize local research investments following policies to

support and foster networks. Research networks should be encouraged in these countries through internationally-oriented calls for proposals or mobility grants. This could help address potential knowledge-gaps and synergies beyond national borders.

Moreover, despite concentrating cases in their specific regions, Brazil and India do not collaborate often and are members of different clusters in the network, indicating a low cooperation between endemic countries. Collaboration between these countries should be encouraged as it would provide access to local knowledge and better understanding of disease transmission, diagnostics and morbidity dynamics in different settings.

India has increased its percentage of influential scientists in the leishmaniasis research network over the years. This might have been due to keeping steady percentage of overall researchers in the field at just below 10% over the entire 35 year period, which resulted in India having percentage of experienced researchers growing over the years.

Lexical analysis provided an overview of the disease research trends through the years. The research themes in the 1980's with their focus on understanding disease pathology and transmission are compatible with early studies of leishmaniasis. The next decade research was clearly concentrated on disease management, diagnosis and treatment methods, and understanding the disease epidemiology. The analyses of the past 15 years showed the relevance of the animal host, a concern with economic impacts of the disease and a search for new forms of treatment. Analysis of the evolution of trends in leishmaniasis research showed that research efforts have recently evolved to a more drug-oriented research, to address an important public health need. The number of available drugs for patient treatment is limited and even these are either exorbitantly priced, have toxic side effects or prove ineffective due to the emergence of resistant strains. However, other important needs might be lacking the necessary scientific or financial investment, such as vector control and vaccine development. Existing vector control methods are not very efficient and yet to this date only three vaccine candidates have gone for clinical trial (Srivastava et al. 2016).

Country thematic profiles reflect interests of their researchers. The specific interest of the USA scientific community in which leishmaniasis is an occupational disease reflects their concern with potential exposure of military personnel acting in the areas where the disease is endemic (Weina et al. 2004). Brazilian researchers' interest in plant lectins shows their response to the need to investigate alternative therapeutic approaches and to develop vaccines (Carvalho & Ferreira 2001; Souza et al. 2013). Several controlled trials evaluating treatments for cutaneous leishmaniasis were carried out in Iran (Khatami et al. 2007), including laser therapy (Asilian et al. 2004), which can explain the particular interest of the Iranian scientific community on this subject.

Country research profiles can provide input for diagnosing the activity of the scientific community of a country. The identification of country-specific themes provides information that goes beyond the construction of bare rankings. Each of the countries evaluated have slightly different motivations or research interests. This provides a window for collaboration and knowledge-sharing between selected countries, according to their specialties/interests, in order to tackle specific leishmaniasis challenges.

Finally, link prediction technique proved to be a very important tool for identifying collaboration pattern changes, which signal evolving collaborations and changing focus of research. This technique was also essential to enable predicting



possible outcomes and the results produced with this research (Kuzmin et al. 2016), thus it can help understanding how and why collaboration takes place.

Link prediction could successfully model the evolution of the leishmaniasis research network. Such predictions could be useful in suggesting unrealized collaborations and thus help to build and maintain strong research teams. In addition, by analyzing the features used for the prediction, we can use this knowledge as a basis for specifying vocabularies for expert description. In the periods just before 2000 the collaboration amongst researchers on the field had a high predictable response for new edges. After that, we believe the newcomers (Brazil, India and Iran) unbalanced the collaboration patterns, with a high percentage of within country collaboration, and new areas of research mainly concerned with specific interests. The prediction technique, which had a high response rate, had a poorer response for edges on new nodes after 2000.

The described above change of collaboration patterns correlates well with growing number of local researchers involved in Brazil and India, and the number of collaborations in which the researchers were involved, demonstrated in earlier sections. Correlation is not equivalent to causation, but the three elements of the change, increased publications in some countries, increased collaboration patterns, and the different center of focus of research in different countries demonstrate the essence of the change observed at these many levels. Clearly, changing focus may trigger change of collaboration patterns, by for example shifted some collaboration from international to national ones. It may also require large teams for certain topics than the ones researched in the past. We plan to investigate such interdependencies of collaboration, and research focus in future work, based on evidence of changes uncovered in this paper.

## References

Adams J. Collaborations: The rise of research networks. *Nature*. 2012; 490(7420): 335-336.

Al-Mutawakel, K. C., Scutaru A. Shami, M. Sakr, D. A. Groneberg and D Quarcoo. Scientometric analysis of the world-wide research efforts concerning Leishmaniasis Parasites & Vectors 2010 3:14.

Alvar, Jorge, Iván D. Vélez, Caryn Bern, Mercé Herrero, Philippe Desjeux, Jorge Cano, Jean Jannin, Margriet den Boer, e the WHO Leishmaniasis Control Team. 2012. “Leishmaniasis Worldwide and Global Estimates of Its Incidence”. *PLOS ONE* 7 (5): e35671. doi:10.1371/journal.pone.0035671.

Alvar, Jorge, Sergio Yactayo, e Caryn Bern. 2006. “Leishmaniasis and Poverty”. *Trends in Parasitology* 22 (12): 552–57. doi:10.1016/j.pt.2006.09.004.

Asilian, A., A. Sharif, G. Faghihi, Sh. Enshaeieh, F. Shariati, e A. H. Siadat. 2004. “Evaluation of CO2 Laser Efficacy in the Treatment of Cutaneous Leishmaniasis”. *International Journal of Dermatology* 43 (10): 736–38. doi:10.1111/j.1365-4632.2004.02349.x.

Bahulkar, Ashwin, Boleslaw K. Szymanski, Omar Lizardo, Yuxiao Dong, Yang Yang, e Nitesh V. Chawla. 2016. “Analysis of Link Formation, Persistence and Dissolution in NetSense Data”. *Applied Network Science* 1: 11.

Barabási, Albert-László. 2007. “Network Medicine — From Obesity to the ‘Diseasome’”. *New England Journal of Medicine* 357 (4): 404–7. doi:10.1056/NEJMe078114.

Bastian, Mathieu, Sebastien Heymann, e Mathieu Jacomy. 2009. “Gephi: An Open Source Software for Exploring and Manipulating Networks”. In *Third International AAAI Conference on Weblogs and Social Media*.

<http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>.

Bender, Max Ernst, Suzanne Edwards, Peter von Philipsborn, Fridolin Steinbeis, Thomas Keil, e Peter Tinnemann. 2015. "Using Co-Authorship Networks to Map and Analyse Global Neglected Tropical Disease Research with an Affiliation to Germany". *PLoS Neglected Tropical Diseases* 9 (12): e0004182. doi:10.1371/journal.pntd.0004182.

Bennett, Sara, Irene Akua Agyepong, Kabir Sheikh, Kara Hanson, Freddie Ssenooba, e Lucy Gilson. 2011. "Building the Field of Health Policy and Systems Research: An Agenda for Action". *PLOS Med* 8 (8): e1001081. doi:10.1371/journal.pmed.1001081.

Bern, Caryn, James H. Maguire, e Jorge Alvar. 2008. "Complexities of Assessing the Disease Burden Attributable to Leishmaniasis". *PLoS Neglected Tropical Diseases* 2 (10): e313. doi:10.1371/journal.pntd.0000313.

Brasil. 2012. *Estratégia Nacional de Ciência, Tecnologia e Inovação 2012 – 2015*. Balanço das Atividades Estruturantes. Brasília, DF: Ministério da Ciência, Tecnologia e Inovação.

Carvalho, P. B. de, e E. I. Ferreira. 2001. "Leishmaniasis Phytotherapy. Nature's Leadership against an Ancient Disease". *Fitoterapia* 72 (6): 599–618.

D'Andreta, Daniela, Harry Scarbrough, e Sarah Evans. 2013. "The Enactment of Knowledge Translation: A Study of the Collaborations for Leadership in Applied Health Research and Care Initiative within the English National Health Service". *Journal of Health Services Research & Policy* 18 (3 Suppl): 40–52. doi:10.1177/1355819613499902.

Duque, Ricardo B., Marcus Ynalvez, R. Sooryamoorthy, Paul Mbatia, Dan-Bright S. Dzorgbo, e Wesley Shrum. 2005. "Collaboration Paradox: Scientific Productivity, the Internet, and Problems of Research in Developing Areas". *Social Studies of Science* 35 (5): 755–85. doi:10.1177/0306312705053048.

Freeman, Linton C. 1979. "Centrality in social networks conceptual clarification". *Social networks* 1 (3): 215–239.

González-Alcaide, Gregorio, Charles Huamaní, Jinseo Park, e José Manuel Ramos. 2013. "Evolution of Coauthorship Networks: Worldwide Scientific Production on Leishmaniasis". *Revista Da Sociedade Brasileira De Medicina Tropical* 46 (6): 719–27. doi:10.1590/0037-8682-0207-2013.

Guarnaccia, Cinzia, Francesca Giannone, Giorgio Falgares, Aldo Ozino Caligaris, e Edith Sales-Wuillemin. 2015. "Differences in Social Representation of Blood Donation between Donors and Non-Donors: An Empirical Study". *Blood Transfusion = Trasfusione Del Sangue*, novembro, 1–7. doi:10.2450/2015.0048-15.

Guerrant, R. L., e B. L. Blackwood. 1999. "Threats to Global Health and Survival: The Growing Crises of Tropical Infectious Diseases--Our 'unfinished Agenda'". *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 28 (5): 966–86. doi:10.1086/514765.

Guler, Isin, e Atul Nerkar. 2012. "The Impact of Global and Local Cohesion on Innovation in the Pharmaceutical Industry". *Strategic Management Journal* 33 (5): 535–49. doi:10.1002/smj.957.

Holmgren, Milena, e Stefan A. Schnitzer. 2004. "Science on the Rise in Developing Countries". *PLOS Biology* 2 (1): e1. doi:10.1371/journal.pbio.0020001.

Kettler, Hannah E., e Rajiv Modi. 2001. "Building local research and development capacity for the prevention and cure of neglected diseases: the case of India". *Bulletin of the World Health Organization* 79 (8): 742–47. doi:10.1590/S0042-96862001000800010.

Khatami, Alireza, Alireza Firooz, Farzam Gorouhi, e Yahya Dowlati. 2007. “Treatment of Acute Old World Cutaneous Leishmaniasis: A Systematic Review of the Randomized Controlled Trials”. *Journal of the American Academy of Dermatology* 57 (2): 335.e1-29. doi:10.1016/j.jaad.2007.01.016.

Kuzmin, Konstantin, Xiaoyan Lu, Partha Sarathi Mukherjee, Juntao Zhuang, Chris Gaiteri, e Boleslaw K. Szymanski. 2016. “Supporting novel biomedical research via multilayer collaboration networks”. *arXiv:1610.09253 [cs, q-bio]*, outubro. <http://arxiv.org/abs/1610.09253>.

Lee, Sooho, e Barry Bozeman. 2005. “The Impact of Research Collaboration on Scientific Productivity”. *Social Studies of Science* 35 (5): 673–702. doi:10.1177/0306312705052359.

Leydesdorff, Loet, e Caroline Wagner. 2009. “International Collaboration in Science and the Formation of a Core Group”. *arXiv:0911.1438 [physics]*, novembro. <http://arxiv.org/abs/0911.1438>.

McKew, John C., e Andre M. Pilon. 2013. “NIH TRND Program: Successes in Preclinical Therapeutic Development”. *Trends in Pharmacological Sciences* 34 (2): 87–89. doi:10.1016/j.tips.2012.10.001.

Morel, Carlos M., Tara Acharya, Denis Broun, Ajit Dangi, Christopher Elias, N. K. Ganguly, Charles A. Gardner, et al. 2005. “Health Innovation Networks to Help Developing Countries Address Neglected Diseases”. *Science* 309 (5733): 401–4. doi:10.1126/science.1115538.

Morel, Carlos M., José R. Carvalheiro, Carmen N. P. Romero, Eduardo A. Costa, e Paulo M. Buss. 2007. “The Road to Recovery”. *Nature* 449 (7159): 180–82. doi:10.1038/449180a.

Morel, Carlos Medicis, Suzanne Jacob Serruya, Gerson Oliveira Penna, e Reinaldo Guimarães. 2009. “Co-authorship Network Analysis: A Powerful Tool for Strategic Planning of Research, Development and Capacity Building Programs on Neglected Diseases”. *PLoS Negl Trop Dis* 3 (8): e501. doi:10.1371/journal.pntd.0000501.

Phillips, Kaye, Jillian Clare Kohler, Peter Pennefather, Halla Thorsteinsdottir, e Joseph Wong. 2013. “Canada’s Neglected Tropical Disease Research Network: Who’s in the Core—Who’s on the Periphery?” *PLOS Negl Trop Dis* 7 (12): e2568. doi:10.1371/journal.pntd.0002568.

Ramos, Jose M, Gregorio Gonzalez-Alcaide, e Maxima Bolanos-Pizarro. 2013. “Bibliometric analysis of leishmaniasis research in Medline (1945-2010)”. *Parasites & Vectors* 6 (março): 55. doi:10.1186/1756-3305-6-55.

Reinert, M. 1983. “Une méthode de classification descendante hiérarchique : application à l’analyse lexicale par contexte”. *Cahiers de l’analyse des données* 8 (2): 187–98.

———. 1986. “Un logiciel d’analyse lexicale”. *Cahiers de l’analyse des données* 11 (4): 471–81.

Rodrigues, Mayara Muniz Dias, Maria Adelaide Silva Paredes Moreira, Eliane de Sousa Leite, Saemmy Grasiely Estrela de Albuquerque, Ronaldo Bezerra de Queiroz, Antonia de Oliveira Silva, Jacira dos Santos Oliveira, et al. 2015. “Activities Offered to the Elderly in Primary Health Care: Social Representations of Elderly and Professionals”. *International Archives of Medicine* 8 (0). doi:10.3823/1853.

Sampaio, Ricardo Barros, Flávia Tavares Silva Elias, Celina Roitman, Ricardo de Godoi Mattos Ferreira, Carlos Medici Morel, Manoel BARRAL Neto, Edgar Marcelino Carvalho, Wagner de Jesus Martins, e Gerson de Oliveira Penna. 2015. “Mobilização para um Programa de Pesquisa Translacional em Leishmanioses: uma

Solução para Saúde Pública.” *Tempus Actas de Saúde Coletiva* 9 (3): 249–67. doi:10.18569/tempus.v9i3.1800.

Singh, Om Prakash, Epcó Hasker, Marleen Boelaert, e Shyam Sundar. 2016. “Elimination of Visceral Leishmaniasis on the Indian Subcontinent”. *The Lancet Infectious Diseases* 16 (12): e304–9. doi:10.1016/S1473-3099(16)30140-2.

Souza, Maria A., Fernanda C. Carvalho, Luciana P. Ruas, Rafael Ricci-Azevedo, e Maria Cristina Roque-Barreira. 2013. “The Immunomodulatory Effect of Plant Lectins: A Review with Emphasis on ArtinM Properties”. *Glycoconjugate Journal* 30 (7): 641–57. doi:10.1007/s10719-012-9464-4.

Srivastava, Saumya, Prem Shankar, Jyotsna Mishra, e Sarman Singh. 2016. “Possibilities and challenges for developing a successful vaccine for leishmaniasis”. *Parasites & Vectors* 9: 277. doi:10.1186/s13071-016-1553-y.

Vasconcellos, Alexandre Guimarães, e Carlos Medicis Morel. 2012. “Enabling Policy Planning and Innovation Management through Patent Information and Co-Authorship Network Analyses: A Study of Tuberculosis in Brazil”. *PLoS ONE* 7 (10): e45569. doi:10.1371/journal.pone.0045569.

Weina, Peter J., Ronald C. Neafie, Glenn Wortmann, Mark Polhemus, e Naomi E. Aronson. 2004. “Old World Leishmaniasis: An Emerging Infection among Deployed US Military and Civilian Workers”. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 39 (11): 1674–80. doi:10.1086/425747.

Wilson, Katherine, Teresa Brady, e Catherine Lesesne. 2011. “An Organizing Framework for Translation in Public Health: The Knowledge to Action Framework”. *Preventing Chronic Disease: Public Health Research, Practice and Policy* 8 (2). [http://www.cdc.gov/pcd/issues/2011/mar/10\\_0012.htm](http://www.cdc.gov/pcd/issues/2011/mar/10_0012.htm).