

Health Research Networks Based on National CV Platforms in Brazil and Uruguay

Cecilia Tomassini^{1,*}, Claudia Cohanoff¹, Sofia Robaina¹, Jesús P Mena-Chalco²

¹Research Commission, Universidad de la República, (UdelaR) Uruguay, Montevideo, URUGUAY.

²Universidade Federal do ABC, (UFABC), BRAZIL.

ABSTRACT

Collaborations are an essential part of the research process. In the health area, these involve a great diversity of actors in various scientific and health subsystems. The study of collaborations has been developed mostly from the analysis of co-authorship in articles indexed in international platforms. However, these sources present some limitations to capture the production of knowledge in Latin American countries. This paper seeks to diversify the sources of information and the units of analysis for the study of collaborations in health research, by exploring data from two original and little-used sources of information in national and public Curriculum Vitae (CV) platforms: LattesCv (Brazil) and CVUy (Uruguay). Based on a wide sample of research projects extracted from CVs, networks for knowledge production are analyzed at micro (researchers) and meso (institutions) levels in each country. This preliminary analysis allows us not only to generate evidence on the nature and evolution of health research networks, but also to evaluate the advantages and limitations of CVs as a new source for the study of collaborative networks.

Keywords: Health research, Collaboration networks, Curriculum vitae, Brazil, Uruguay.

Correspondence

Cecilia Tomassini

Research Commission, Universidad de la República, (UdelaR) Uruguay. Jackson 1303, Montevideo, URUGUAY.
Email: ctomassini@csic.edu.uy

Received: 31-10-2020

Revised: 20-02-2021

Accepted: 12-03-2021

DOI: 10.5530/jscires.10.1s.25

INTRODUCTION

In this study we understand health research from a broad perspective, where the production of knowledge is one of the functions that drives the Health Innovation System (HIS).^[1] Collaborations are a constitutive and central part of the systemic vision of innovation and have been studied in multiple ways. In general, this definition confronts the linear approach of innovation, based on the idea that there is an interactive process which takes place among actors and institutions.^[2] In other words, the acknowledgment that innovation can hardly exist without diversity in connections.^[3] Latin American countries have very different and unequal HIS^[4] leading to different capacities to mobilize collaborations. These capacities are particularly important when it comes to promoting collaborations with the health subsystem for solving local problems.^[5]

The study of collaboration has been based on bibliometric data from articles indexed in international databases. However, the use of these data presents a series of geographic and language biases that limit the analysis for Latin American countries. This

paper seeks to explore new sources of information and units of analysis for the study of collaborations in health research. To that end, an exploratory analysis of collaborations in research projects contained in the Curriculum Vitae (CV) of researchers is carried out, for the case of two Latin American countries that have public curriculum platforms: Brazil (LattesCv-platform) and Uruguay (CVuy-platform).

Social Network Analysis (SNA) was applied to the data contained in more than 1,448 CVs in the case of Uruguay and 12,252 in the case of Brazil. We explore to what extent the data on collaborations built from CVs allow us to understand the evolution of collaboration networks between researchers (micro level) and between HIS institutions (meso level). The analysis focuses on the evolution of collaboration networks between the year 2000 and 2015. This period allows us to contextualize the collaboration networks in a special situation for the promotion and investment in STI in Latin America, from a scenario of economic crisis and little STI investment to a scenario of recovery and larger investment.

We are particularly interested in answering four exploratory questions in each country: (i). How does the size of collaboration networks in health research projects evolve over the period? (ii) How does the cohesion and connectivity between network actors evolve? (iii). Who are the main actors in the network and what role do they play? (iv) How do networks between actors in different HIS subsystems

Copyright

© The Author(s). 2021 This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

evolve? Of particular interest is the analysis of the nature of interinstitutional collaborations that bring together actors focused on the production of knowledge with external actors from other subsystems, especially health services providers. By answering these questions, we hope not only to generate evidence on the functioning of health research networks in each country, but also to evaluate the advantages and limitations of the use of CVs as a new source for the study of collaborative networks.

Background

Health research and health innovation systems

Despite growing recognition of the research as a tool to improve healthcare and healthcare systems, there is great difficulty in defining what health research is and how it could effectively provide solutions.^[6] In this study, we understand that health research involves a broad range of areas (biomedicine, social sciences, management and administration, public policy, engineering and medical technologies, among others) and when it is focused on problem solving it requires a collaborative process between different actors.

This collaborative process occurs within cross-sector systems which involve a diversity of agents in different domains of action, or subsystems.^[1] Combined, these subsystems constitute the HIS. Albuquerque and Cassiolatto^[6] identify at least four types of HIS agents: (i) universities and research institutes, which bring together the scientific community and are the center of knowledge and technologies that make up the system. (ii) Healthcare services (hospitals, clinics, medical centers, etc.), which are mainly focused on their healthcare function but at the same time show a strong interaction with universities and the industrial sector. (iii) Health policy and regulatory institutions, including government agencies but also professional associations that play a decisive role in selecting agendas and financing. (iv) The industrial sector (service and supply companies, laboratories, etc.).

According to Consoli and Mina (1), innovation in HIS is nourished by knowledge transfer, especially from scientific research and clinical practice. This knowledge is highly dispersed within the system, and among disciplines, technology fields, geographical locations and institutions. Therefore, collaboration is the “gateway” for knowledge generation aimed at solving health problems.

The Social Network Analysis (SNA) is complementary to the innovation system approach, since it allows to identify the system’s main actors, their strategic positions, the information flows and their evolution in time. Several studies have used this tool to identify the distinctive features of the HIS at national, regional, international or sector level. For example, in the late 1990s, Hicks and Katz (1996)^[7] point out the strategic role

of hospitals within the HIS in the United Kingdom. Consoli and Mina (1) also highlight the role of hospitals, specifically university and research hospitals due to their double function and their potential role as “led users”.^[8] In the case of Brazil, previous research also shows the growing participation of university hospitals in the production of scientific knowledge, combined with their other functions.^[9]

The application of the innovation system approach in Latin American countries has shown that this region - in addition to local and contextual potentialities - shows common barriers related to the development of STI. Some of the main barriers highlighted by Arocena and Sutz^[3] are: (i) low levels of investment in infrastructure, (ii) lack of knowledge demand, (iii) poor promotion of endogenous knowledge and opportunities to apply it, (iv) difficulties to reconcile knowledge supply and demand, and (v) lack of stable linkage between STI institutions and productive actors at a national level. As pointed out by several authors these barriers also apply to the case of HIS in the Latin-American region.^[10-13]

Investment in STI in Latin America has been historically low, none of the countries exceeding 1% of GDP in Research and Development (R&D) investment (Ricyt, 2020). Brazil has been an exception experiencing a progressive increase in R&D investment between 2000 and 2013 up to approximately 2% of GDP, which since 2014 has dropped to 1%.^[14] Uruguay shows a much lower level of R&D investment. Despite an increase in 2006, it has never exceeded 0,5% of GDP.^[15]

Even though a country’s research capacity depends on multiple factors, human resources are crucial, particularly regarding scientists with a PhD training. Brazil has 6.9 doctors per 10,000 inhabitants^[16] and, given the implementation of successive national plans and the main role played by the scholarship system, it stands out in the region for its highly mature postgraduate system. Postgraduate program in the health sciences sector is remarkably diverse with most doctoral students graduating in this field.^[17] Uruguay’s postgraduate program offer, on the other hand, is more recent and its scholarship system has seen a significant boost as of 2008. The total number of Uruguayan doctors is 4.5 per 10,000 inhabitants.^[18]

Despite its diversity, Latin American countries share the common challenge of applying national capacities towards promoting interactive learning focused on solving problems.^[19] To achieve this, collaboration and exchange between diverse areas of knowledge (formal and tacit) are key. In the health sector, this is particularly important, not only because knowledge is highly distributed, but also because progress in improving health - for example, the development of new diagnoses and treatments - requires close collaboration

between basic medical research, clinical research, product development and healthcare.^[1]

Collaborations for knowledge production

Scientific collaboration studies have a long tradition in social sciences and scientometrics. Laudel (2002)^[20] defines it as “a system of research activities by several actors related in a functional way and coordinated to attain a research goal corresponding with these actors’ research goals or interests” (18, p. 5).

In recent decades, several authors have highlighted the emergence of new forms of knowledge that leave behind isolated paths to increasingly substitute them for collaborations among different kinds of actors and disciplines.^[21] Some of the main reasons for collaboration include the need to jointly afford the increasingly high costs of frontier research, the increase in interconnectivity between countries, the development of communication technologies, the need for addressing complex issues integrating knowledge.^[22] Particularly in the health sector, from the late 1990s, a strong impulse for collaboration began at a global level based on the belief that it could help reduce disparities between countries. This impulse was mainly boosted by: i) scientific developments such as genomics and statistical techniques, ii) technological requirements necessary for analyzing large data sets, iii) and major global funding initiatives.^[23]

In addition to its advantages, collaboration also implies disadvantages, for example, in terms of time and resources.^[22,24] And they may involve different costs, Boschma (2005)^[25] points out that the levels of proximity or distance between collaborating actors can speed up or block innovative knowledge.

Historically, the study of scientific collaborations has been operationalized based on the co-authorship of articles. In Brazil, co-authorship networks have been widely used to analyze collaboration on research in specific diseases,^[26,27] or with a focus on strategic health research institutions^[28] among others. No records of this type of studies were found in Uruguay.

Co-authorship analyses have advantages in terms of reliability and replicability; however, collaboration and co-authorship are not synonymous.^[20,22] Not all contributors appear as authors and not all authors actually participate in the collaboration. In addition, bibliographic datasets used for this bibliometric analysis show a series of geographic and language biases.^[29] In recent years, several studies seek to diversify the sources of information to evaluate STI activities avoiding the aforementioned biases, for example using webometrics.^[30]

This research proposes the use of a new source of information in order to complement the analyses on scientific collaborations at a national level. For this purpose, we will focus on data extracted from CV platforms with a focus on research projects. We define research projects as the process for the production of basic, applied or experimental knowledge that is guided by objectives, limited in time and carried out with specific resources. Projects are developed within institutional frameworks with their own missions, involve several disciplinary, multidisciplinary, formal and tacit knowledge, and constitute workspaces for information exchange between several actors, which are involved with different levels of commitment, collaboration and conflict. By proposing the use of this unit of analysis we are focusing on a stage of knowledge production different from the one usually analyzed with data from published articles. If we think of knowledge production as a process, then data from research projects and data from articles could complement each other. The former captures the beginnings of knowledge production, which, depending on their success, use different dissemination channels or are not necessarily disseminated. While the latter capture a portion of the production of knowledge that is disseminated in a format that is primarily of interest to the academic community.

Methods and sources

CVs as sources of information to study STI activities

Different studies have used CVs as a source of data to evaluate national STI activities and capacities, for example, researchers’ academic careers,^[31,32] international mobility,^[33] dynamics of academic collaboration in general^[34] and of co-authorship networks in particular,^[35] as an assessment tool of STI activities,^[36] among others. There are multiple advantages of using a CV as source: its use is nearly universal, information has a relatively standardized format, in most cases access it is easily available, and longitudinal data is provided on the performance of individuals in different areas of their professional life.^[36] But it is not a source without disadvantages.^[37] Information contained in a CV is of variable quality and completeness, and emphasis may be put on different aspects depending on the individual specialization and the intended goal as the CV was created. Truthfulness of the information is also variable because it is self-declared and may change over the course of time. In addition to the fact that it may not be updated with the latest information (data truncation). Finally, it requires a great effort of standardizing the information. The existence of standardized CVs platforms managed by public STI agencies is a key factor in gaining access to quality information.

The availability of platforms for public access to CVs is the main justification for analyzing the cases of Uruguay and Brazil. The CV formats in both countries were created in a similar way, in fact the CV in Uruguay is based on the Brazilian LattesCv.

In both cases, CVs are used to evaluate applicants to programs from various STI agencies. The CVUy platform in Uruguay, is managed by the National Research and Innovation Agency (ANII). It is a standardized and automated system launched in 2008 and currently has more than 12,200 records. In the case of Brazil, there are two large databases, one at research group level and their members, the Directory of Research Groups (DRG), and one at individual level, the mentioned Lattes-CV. The latter was launched by the National Council for Scientific and Technological Development (CNPq) in 1999 and, in 2007, it exceeded one million registered researchers.^[38]

Data retrieval processes in both platforms followed a common logic. Both extractions were based on obtaining information on collaboration from the integration of research teams projects declared in CVs. Variables regarding projects applied to both datasets are: title, summary, starting year, source of financing, number of members; and on individuals or project members: name, ID in CV, name of main institution, type of institution, area and knowledge discipline.

However, each platform allows different definitions of the population under analysis. In the case of Brazil, the interface between the DGP and LattesCv data allows to identify the principal investigators, or group leaders, from a broad definition of health, according to the application sector. In this case the population under analysis were 12,252 leaders of research groups with application in the human health sector, who registered 154,879 research projects in their CVs throughout the period 2000-2015. In the case of Uruguay, the CVuy does not allow to distinguish the leading of research groups, nor does it allow to define health research according to the sector of application. So, the population under analysis were defined from the main area of performance that the researcher declared and the main area of the research projects: Medical Sciences and Health (MSH). In this case, the population under study were 1,448 researchers who registered 3,183 in their CVs throughout the period.

The database building is divided into four common stages:

1. Retrieval of CVs: i) in Brazil it was carried out with the scriptLattes^[39] program based on the collection and summarization of all LattesCV in HTML format ii) in Uruguay, upon a formal request of the information, all the CVUys were obtained in HTML format, and were then organized in a structured way.
2. Consolidation and cleanup in two phases: i) integration of project data and individual researchers' data in a database; ii) removal of duplicate information and definition of time frame (projects started between 2000 and 2015).
3. Disambiguation of textual data (names of institutions) and coding in categories. For disambiguation, VantagePoint®

software was used for Lattes-CV, and OpenRefine for the CVUy.

4. Consistency check based on the analysis of descriptive statistics.

The criteria applied made it possible to obtain two large database comprising researchers from a wide variety of institutions and knowledge areas. Although the differences in the unit of analysis do not allow us the comparison between countries, we are able to evaluate the specificities of the collaboration networks in health in each country.

Research collaboration networks construction

Collaboration networks are analyzed at micro and meso levels. Figure 1 summarizes the network creation procedure. In the first kind of network, nodes are individuals who participate in research projects and have registered their CVs in the system. In the second kind of network, nodes are the institutions where researchers belong. In both networks, links are established when a researcher's name appears in a project's

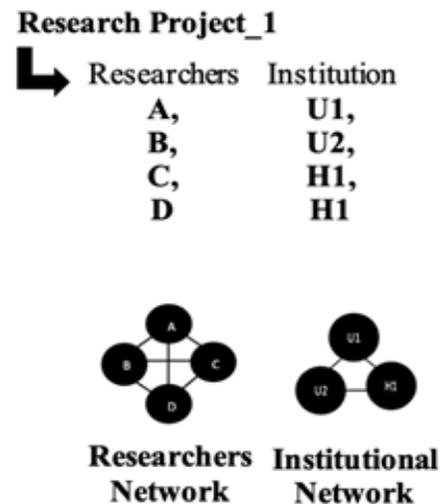


Figure 1: Network construction at a micro and meso level.

team and/or when two researchers register the same project in their CV. These are non-directed networks where we assume reciprocity in collaborations.

For each country networks are analyzed by comparing the evolution of metrics at a node and structure level in four periods (2000-2003; 2004-2007; 2008-2011; 2012-2015).¹ This division into periods allows us to observe how the evolution of the networks is inserted in a particular economic and investment context in ITS in the region. Although each country presents its peculiarities, this is a period characterized

1. For details on the calculation of each indicator, see Table S1 in this article's supplementary material.

by a strong economic crisis and low STI investment at the beginning of the 2000s, followed by a period of recovery and improvements in STI investments and a subsequent period of stagnation approximately as of 2012.

The main indicators and metrics used to answer the research questions are:

1. Networks size evolution: using indicators of the size of the network we can identify periods of increasing or decreasing collaborations among researchers and institutions.

2. Networks cohesion and connectivity: By analyzing how Density (D) and Average Degree Centrality (ADC) evolves we explore the network cohesion. In addition, since in research practices it is usual to work between relatively small groups, it is interesting to observe the structure of communities based on the Clustering Coefficient (CCo). A network connectivity analysis is conducted based on the evolution of its Degree Centrality (DC), the size of the Largest Component (LC) and the Average Path Length (APL) of the LC.

3. Actors centrality: Based on measures such as Degree Centrality (DC), Betweenness Centrality (BC); Closeness Centrality (CC), we can identify who are the main actors in the network and what role do they play?

4. Subnetwork of interinstitutional collaborations: To analyze the collaborations between different components of the SIH, we focused on the main characteristics of subnetworks, on the basis of four categories: (i). STI institutions, including mainly researchers in universities and research institutes; (ii). Healthcare services, mainly researchers in medical centers, hospitals, clinics and healthcare centers. Following the recommendations of Hicks and Katz (1996), university hospitals are not coded under the university category but as a different kind of institution. (iii). Companies, including researchers and professionals in companies that develop or provide technology or services for the health sector; and finally, (iv). S&T and health support institutions, governmental agencies and several non-profit organizations.

RESULTS AND DISCUSSION

Collaboration networks in health research in Uruguay

During the period of analysis, the development of Uruguay's research system experiences at least three relevant changes: (i) Strengthening of research programs and creation of a new institutional framework; (ii) Strengthening of postgraduate programs and expansion of scholarship funding; (iii). Increase in investment in STI activities, though with fluctuations. Although between 2005–2010 the public component assigned to STI activities at a national level increased by a factor of five, a drop was observed in the following years. [40] In 2008, the budget allocated to university research also

increased.[41] Over this period, capacity building in the health sector shows a positive development. According to the analysis of the extracted CVs, there is a sustained increase in the number of PhD graduates in 2000–2015, as well as an increase in the proportion of those who graduate in the country to the detriment of those who do so abroad.² The analysis of the CV data allows us to observe that collaboration networks in research projects seem to accompany Uruguay's maturation process regarding its research promotion system. They report growth throughout the period as well as an increase in connectivity among researchers (Figure 2). At the

2. For an analysis on Uruguay's health PhD base on CVuy data see supplementary material.

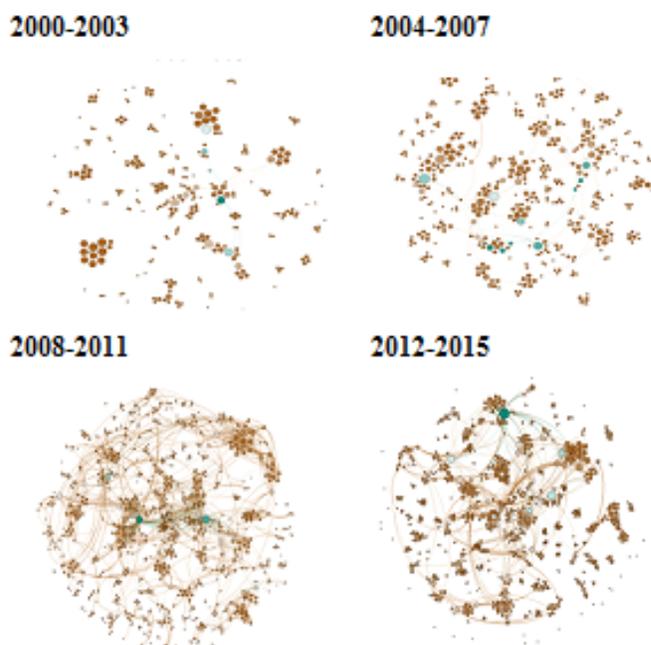


Figure 2: Evolution of project networks in health research – Uruguay – CVUy.*

*Nodes are researchers registered in the CVUy and links are shared research projects. The size of the nodes illustrates the DC and the color illustrates the BC (darker shades of green mean a greater DC).

beginning of the series, there are 275 participating researchers in the network and by the end of the period the number of researchers increases by a factor of four. Connections between researchers also increase, from more than 300 to more than 2,000 (Table 1).

Density measures indicate a low global network cohesion, that means, very few of all possible relationships between network actors actually take place. However, it is relevant to note that the analyzed networks show an increasing level of connectivity over the course of time. The ADC, that is, the average of nodes adjacent to each researcher in the network, ranges from 3 to 6 depending on the period. There is also

Except for the MSP, which is the main national health policies authority, the rest are public STI institutions, three schools (FMED, FCIEN and FQ), two research institutes (IIBCE and IP_MV) and one training program (PEDECIBA). All show incremental growth in their centrality measures (DC, BC, CC) throughout the period.⁵ These institutions play a major role in information flows and exchanges related to the production of health knowledge at a national level. FMED shows the highest indicators of centrality. It maintains direct links with most institutions in the network (DC 134) and acts as mediator in collaborations between many of the network's node pairs (BC 0.46). In addition, it has a strategic position so as to easily receive information released by other nodes (CC 0.69). The rest of the institutions mentioned above also have high degrees of DC and CC, but lower degrees of intermediation. In terms of geographic distribution, a high centralization of collaborations is observed in the capital of the country. However, the establishment of the Regional University Center in the north of the country (CENUR_Lit. Norte) seems to be a great step to change this. Said institution shows high measures of centrality since its creation towards the present and seems to be acting as a bridge between leading health institutions in the capital and the north of the country.

At the network's periphery, there are several institutions that play different roles in the production of knowledge in health. Interinstitutional subnetworks show that the LC

of the public university, but no further information can be disaggregated in the data extracted from some CVuy.

5. The evolution of the centrality measures for the main institutions is detailed in the supplementary material.

1. Nat. STI Inst. and Companies (N:46; E:95) E:71)



3. Nat. STI Inst. and Healthcare Services (N:60; E:115)

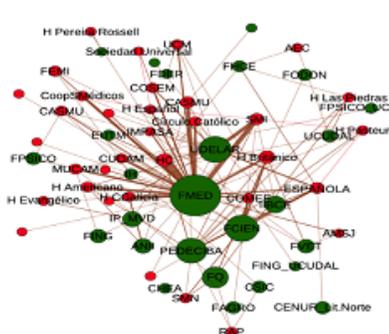


Figure 4: Subnetwork of interinstitutional collaborations (LC) – Uruguay – CVUy.*

of collaboration between STI institutions and companies comprises 46 nodes and only 9% of the links of the global network (Figure 4). Centrality measures of the companies are low and their links with STI institutions are not constant in time except for some of them, such as the Uruguayan Center for Molecular Imaging (CUDIM) (non-estate public corporation focused on diagnosis, research, training and applications in the health sector), Atgen (private laboratory that was created as the first spin-off incubated in FCIEN-UDELAR) and Celsius (private laboratory acquired by the Dermur Pharma group). Although these data need further research, the positive evolution shown by collaborations between UDELAR, the Pasteur Institute of Montevideo and private sector company ATGen is an essential precedent to explain the joint development of the first national test to detect SARS-CoV-2 that made possible the country's current success at expanding its diagnostic capacity to cope with the pandemic.

Furthermore, the collaboration network with the government sector shows a LC with 49 nodes and 12% of the links of the global network. Except for the MSP and the Ministry of Social Development (MIDES), there are no systematic links sustained throughout time and their centrality measures are low. Finally, the interinstitutional subnetwork with healthcare services comprises an array of public and private actors, including clinics and hospitals. This subnetwork has the highest LC comprising 60 institutions and 15% of the links of the global network. Two hospitals have the highest centrality measures: the university hospital (HC-UDELAR) and the British Hospital. In contrast to the previous examples, several institutions in this subnetwork establish stronger and longer lasting links in every period, in particular several hospitals and healthcare services. Connections with hospitals, particularly universities, have an important specificity for collaboration networks in health. They can act as institutions that demand and spread knowledge and technologies, but also as institutions that produce knowledge and capacities.

Collaboration networks in health research in Brazil

Brazil stands out in the region for the development of an institutional framework for the promotion of STIs in health, as well as for its strong scientific community and the ability to develop a research system that can effectively contribute to improve the quality of life of its population.^[44,45] During the analyzed period, the STI promotion system in Brazil undergoes considerable changes regarding its institutional framework, infrastructure and budget. It should be noted that during the analyzed period, several efforts were made to decentralize and expand the higher education and postgraduate offer, including the establishment of new federal universities

who graduate in the country are considerably more than those who graduate abroad.⁶

The collaboration network analyzed for the case of Brazil comprises a total of 6,459 research group leaders with application in the health sector conducting research projects during the analyzed period (Table 2). Every period shows a rise in the network's size (Figure 5), that is, a growing number of leaders and their collaborations. The same as in Uruguay's collaboration networks, the highest network growth is observed in 2008-2011. The significant decrease in the network's size in the last period stems from a lack of update of the CV data due to the closeness in time of the extraction date.

In this case, the network's structure shows low cohesion and high fragmentation. Density measures and connectivity

6. For an analysis on Brazil's PhD training in health base on LattesCv see supplementary material.

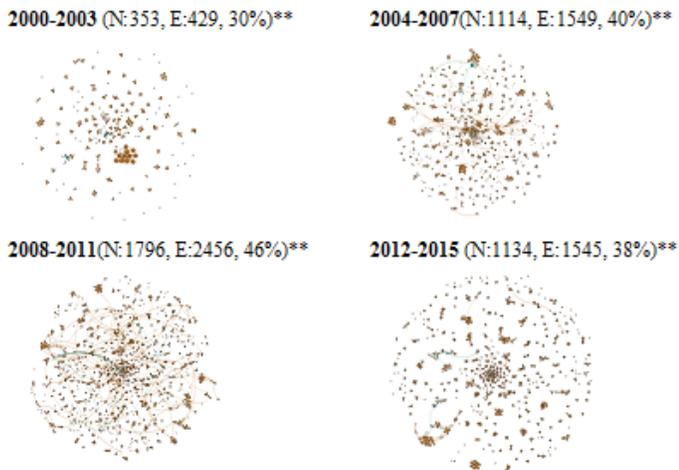


Figure 5: Evolution of project networks between group leaders – Brazil-LattesCv.*

*Size of the network can be visualized in image. (N: Nodes, E: Edges, % of nodes in total network.)

and the expansion of national and international scholarship programs for postgraduate studies.

In the health research area, a rise in budget is observed, especially from 2003 to 2006, followed by a drop and subsequent stagnation.^[44] Research groups focused on health sciences account for approximately 5,609, that is, 16% of all groups nationwide in 2014.^[16] On the other hand, groups with direct application in human health gather a great diversity of disciplines and represent more than 10,800 until 2015.^[46] According to data on the academic training of these group leaders, the number of PhD graduates grows continuously until 2008, followed by a subsequent drop. In addition, those

Table 2: Metrics of collaboration networks 2000 to 2015 –Brazil-LattesCv.*

	2000 - 2003	2004 - 2007	2008 - 2011	2012 - 2015	2000 - 2015
Number of projects	21,238	39,398	54,695	39,548	154,879
Number of nodes	1196	2761	3880	2959	6459
Number of edges	943	2613	3854	2697	8365
Average Degree	1,577	1,893	1,987	1,823	2.59
Density	0.001	0.001	0,001	0,001	0
Largest Component (LC)	17	150	423	74	3595
Average Path Length (APL) of the LC	3	6	14	5	11
Proportion of the LC	0.01	0.05	0.11	0.03	0.56
Clustering Coefficient (CCo)	0.665	0.639	0.604	0.649	0.464
STI Institution Nodes**	1129 (94%)	2592 (94%)	3639 (94%)	2767 (94%)	6031 (93%)
Health Services Nodes**	36(3%)	115(4%)	152(4%)	107(4%)	253(4%)

*Nodes with DC ≤ 1 **Researcher's main institution in LattesCv

Source: own elaboration based on Lattes-CV(CNPq)

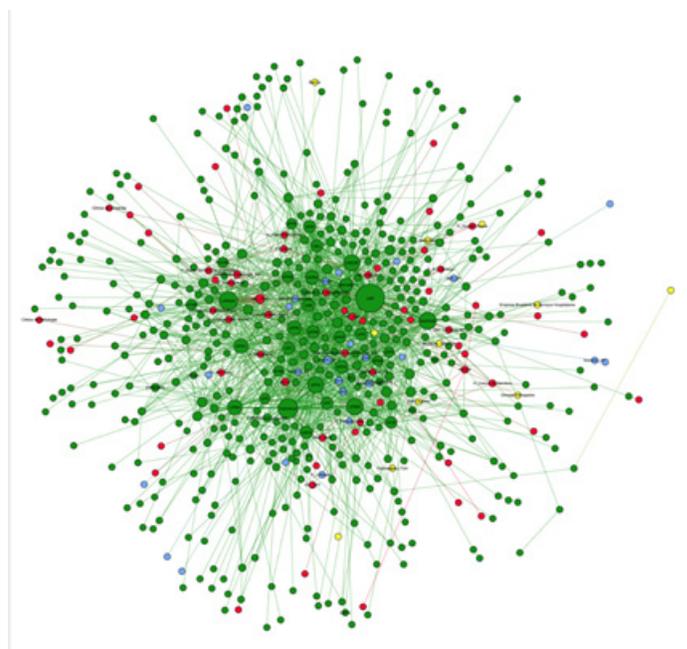


Figure 6: Network of institutional collaborations (2000-2015) LC –Brazil-LattesCv *(N:577, E:1858)

are very low, the networks' average degree does not exceed two connections in none of the periods. Nonetheless, an improvement in connectivity towards the present is observed. The proportion of leaders with 0 grade decreases as the proportion of leaders with grade 1 or higher increases.⁷ In other words, towards the present, more and more research group leaders establish collaborations in projects with other leaders. The improvement in network connectivity is confirmed by LC growth towards the third period going up to 10% of the total network. The network shows a high clustering coefficient in all periods. On the other hand, the distance between pairs of nodes varies substantially in the periods depending on the size of the LC (Table 2). If we consider the entire analysis period, the distance between any two pairs of nodes in the network is high, 11 steps away. This characteristic could be explained by the fact that the network only captures a part of collaboration between most established researchers, and should be complemented in the future with information of others researchers.

Most are researchers at STI institutions (93%). Nevertheless, the network also includes a group of leaders (4%) that belong to health system institutions, such as hospitals, clinics, non-profit organizations and government agencies.

Institutional collaboration networks in research projects increase in size throughout the periods and comprise a total of 1,074 institutions (Figure 6). As can be seen in Figure 6, these networks are mostly composed by national STI institutions, especially federal universities. Six institutions show the highest centrality measures (DC, BC and CC) in 2000–2015: University of São Paulo (USP), Oswaldo Cruz Foundation (FIOCRUZ), Federal University of Rio Grande do Sul (UFRGS), Federal University of São Paulo (UNIFESP), Federal University of Minas Gerais (UFMG) and Federal University of Rio de Janeiro (UFRJ).⁸ Their strategic position in the network makes them the institutions with the greatest capacity to direct collaborations and spread knowledge in health in Brazil. Most of them are public universities located in the south and southeast region, except for FIOCRUZ, which is a public research institute under the Ministry of Health (MS) based in different regions of the country. It should also be highlighted the relevant role played by other federal universities located in the north and northeast region. These institutions, which show increasing centrality measures throughout the period, act as a bridge between their nearby nodes and the rest of the network, showing their strategic position bringing together regions with less resources with those with greater resources. The Federal University of Bahia (UFBA) stands out for its role connecting several institutions

based in Bahia with the rest of the network and accounting for high BC indicators. Geographic location is a key factor in understanding Brazil's dynamics in health research, particularly due to its high concentration in the southern and southeastern regions. Nonetheless, some slight improvements in capacity decentralization are observed over time in the north and northeast regions.^[47,48]

Interinstitutional subnetworks gathering leaders from institutions with different functions collaborating for the health research system comprise only 188 nodes and 202 edges, that is, a small portion (18%) of the global network (Figure 7). However, growth is observed in these subnetworks from the first periods of the series to the present. Group leaders in

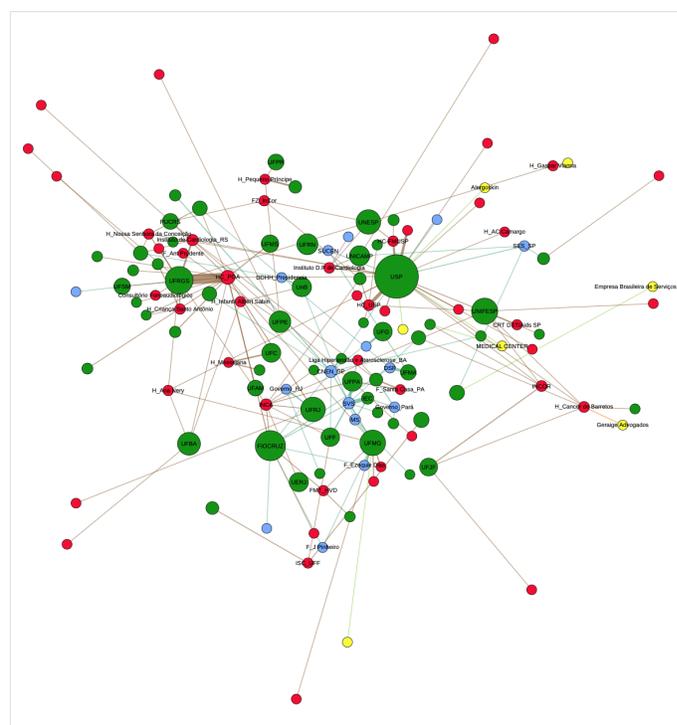


Figure 7: Subnetwork of interinstitutional collaborations (2000-2015) LC – Brazil- LattesCv.* (N:130, E:170)

hospitals, such as the Clinic Hospital of POA, the Barretos Cancer Hospital or the Clinic Hospital of USP play a leading role in these networks. Particularly university hospitals are key to collaborations that have grown over time. Collaboration between leaders from different government agencies in the health sector is also observed, especially in the Secretariat of Health Surveillance (SVS), MS and the Secretariat of Health of São Paulo (SES_SP).

CONCLUSION

The collaboration networks between health scientists have been analyzed by using data of research projects from their CVs over a 16-year period (2000–2015). We found that data

7. The evolution of the DC and the LC is detailed in the supplementary material.

8. The evolution of the centrality measures for the main institutions is detailed in the supplementary material.

from both CV platforms can be used for network analysis and provide significant evidence to understand the dynamics of collaboration at national level. In both countries, the networks analyzed at micro level grow in size and connectivity from the middle of the period. In general terms, the evolution of the networks seems to converge with the institutional and budget strengthening in STI systems observed in both countries.

The project data in both platforms also allowed us to analyze the inter-institutional collaborations of the different HIS actors, based on coding the types of institutions. This analysis confirms the leading role played by public's research institutions in these two Latin American countries, especially public universities and health research institutes. Furthermore, the data shows that both countries have a diversified base of collaborations between STI institutions, hospitals and government agencies. The presence of diverse actors that represent the supply and demands of knowledge could help guide research towards meeting the demands of the health system. The expert users located in hospitals and government agencies are not just sources of information, but they are also well aware of local problems and their possible solutions.^[8] Although, the data shows that these interinstitutional subnetworks grow over time, they continue to be a small portion of the global network in both countries and the participation of several health subsystem actors is not constant over time. In that sense, the lack of connection between HIS subsystems, health policies and the population's health needs, continues to be a problem.^[12]

On the other hand, the analyzed data captures only a minor role of the business and industrial sector in health research projects, showing a limitation on CVs data collected. Particularly, it should be noted that data collected regarding group leaders' institutions is not a good source of information for capturing interactions with the industrial or business sector in Brazil. However, other dimensions contained in the DGP platforms have shown greater potential illustrating typical dynamics in the university-private sector interaction showing the relevance of collaborations between public health STI institutions, national laboratories and multinational companies.^[49] In the future, it will be necessary to explore other variables within the Lattes-CV in order to obtain longitudinal information on company participation in research projects, for example their participation in research projects funding.

Although this preliminary analysis provides new sources of information as well as substantial evidence to understand the dynamics of collaborations for the knowledge production, it is also limited by the design of CVs, especially because of the little attention that the CVs manager agencies put in collecting the co-participation of non-academic actors. If the focus of scientific evaluation is mainly on rewarding bibliographic productivity based on published articles, then information

on collaborations with non-academic actors will be little considered in the design of CVs. Nevertheless, to strengthen the HIS, progress should be made in expanding the role of both health and STI subsystems, increasing participation of actors involved in guaranteeing public health. This requires reducing the costs of collaboration and rewarding these activities. This is a problem that concerns scientific evaluation systems in our region and in the world. As Kickbusch^[50] points out, the challenge is to find performance indicators that generate mutual benefits to the actors involved, so that collaborations can be sustained over time and thus improve the quality of health research and services.

In the future, it seems necessary to reinforce at least three lines of analysis to expand the use of these sources: (i) explore the potentiality of CVs data to analyses causal explanations and determining factors for collaboration networks and their structure, (ii) diversify the empirical evidence and units of analysis to explore types of collaborations. The CVs are a rich source of information that can be used to integrate collaborations in technical, bibliographic, teaching materials, among others, providing a more comprehensive overview of the various forms of knowledge and technologies production in developing countries. (iii) Explore the qualitative data contained in CVs, for example, applying text mining techniques in project summaries or research lines, with the aim of better understanding what research topics and problems make up the agendas in the different knowledge areas, institutions and country regions.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Consoli D, Mina A. An Evolutionary Perspective on Health Innovation Systems. *Journal of Evolutionary Economics*. 2009;19(2):297-319.
2. Edquist C. System of Innovation Approaches-their Emergence and Characteristics. In: *System of Innovation Technologies, Institutions and Organizations*. Edquist Charles. London: Printer; 1997.
3. Arocena R, Sutz J. Subdesarrollo e innovación. Madrid: Cambridge University Press; 2003;232.
4. Alger J, Becerra-Posada F, Kennedy A, Martinelli E, Cuervo LG. Sistemas nacionales de investigación para la salud en América Latina: Una revisión de 14 países. *Revista Panamericana de Salud Publica*. 2009;26:447-57.
5. Soares MC, Cassiolato JE. Health Innovation Systems, Equity and Development. Rio de Janeiro, RJ, Brasil: e-papers; 2015.
6. Albuquerque ED, Cassiolato JE. As especificidades do sistema de inovação do setor saúde. *Brazilian Journal of Political Economy*. 2002;22(4):701-19.
7. Hicks D, Katz JS. Hospitals: The hidden research system. *Science and Public Policy*. 1996;23(5):297-304.
8. Hippel VE. The dominant role of users in the scientific instrument innovation process. *Research Policy*. 1976;5(3):212-39.
9. Araujo KMD, Mourão PAS, Leta J. Balance between education- and research-oriented publications from a Brazilian University Hospital. *Brazilian Journal of Medical and Biological Research*. 2005;38(9):1285-91.
10. Gadelha CAG, Maldonado J, Vargas MA, Barbosa P, Costa LS. A Dinâmica inovativa do sistema produtivo da saúde. Rio de Janeiro, RJ, Brasil: FIOCRUZ; 2012.
11. Soares MC, Cassiolato J. Health Innovation Systems, Equity and Development. Rio de Janeiro, RJ, Brasil: e-papers; 2015.
12. Morel CM, Acharya T, Broun D, Dangi A, Elias C, Ganguly NK, *et al.* Health

- Innovation Networks to Help Developing Countries Address Neglected Diseases. *Science*. 2005;309(5733):401-4.
13. Natera JM, Tomassini C, Vera-Cruz AO. Policy analysis and knowledge application for building a healthy health innovation system in developing countries. *Innovation and Development*. 2019;9(2):159-68.
 14. Koeller P, Viotti R, Rauhen A. Dispêndios do governo federal em c & t e p&d: esforços e perspectivas recentes. *Radar Tecnologia, Produção e Comércio Exterior*. 2016;48:13-8.
 15. RICYT. Indicadores comparativos. ricyt.org. 2020. [cited 2020 Oct 20]. Available from: <http://www.ricyt.org/category/indicadores/>
 16. CNPq. Séries Históricas até 2015. cnpq.br. 2014 [cited 2020 Oct 16]. Available from: <http://www.cnpq.br/web/guest/series-historicas>
 17. Centro de Gestão e Estudos Estratégicos. Descentralização do fomento à ciência, tecnologia e inovação no Brasil. Brasília, DF: Centro de Gestão e Estudos Estratégicos; 2010.
 18. Méndez L, Pellegrino A, Robaina S, Vigorito A. Primer censo de personas uruguayas e inmigrantes con título de doctorado. Informe de Resultados. Programa de Población, FCS; 2019. [cited 2020 Sep 19]. Available from: <https://www.colibri.udelar.edu.uy/jspui/bitstream/20.500.12008/22319/1/DT%20UM-PP%2003.pdf>
 19. Arocena R, Sutz J. Weak knowledge demand in the South: Learning divides and innovation policies. *Science and Public Policy*. 2010;37(8):571-82.
 20. Laudel G. What do we measure by co-authorships?. *Research Evaluation*. 2002;11(1):3-15.
 21. Gibbons M, Limoges C, Nowotny H, Schwartzman S, Seot P, Trow M. The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies. SAGE; 1994;196.
 22. Katz JS, Martin BR. What is Research Collaboration?. *Research Policy*. 1997;26(1):1-18.
 23. Parker M, Kingori P. Good and Bad Research Collaborations: Researchers' Views on Science and Ethics in Global Health Research. *PLoS One*. 2016;11(10):1-19.
 24. Beaver D. Reflections on scientific collaboration (and its study): Past, present and future. *Scientometrics*. 2001;52(3):365-77.
 25. Boschma R. Proximity and Innovation: A Critical Assessment. *Regional Studies*. 2005;39(1):61-74.
 26. Morel CM, Serruya SJ, Penna GO, Guimarães R. Co-authorship Network Analysis: A Powerful Tool for Strategic Planning of Research, Development and Capacity Building Programs on Neglected Diseases. *PLoS Negl Trop Dis*. 2009;3(8):e501.
 27. Vasconcellos AG, Morel CM. Enabling Policy Planning and Innovation Management through Patent Information and Co-Authorship Network Analyses: A Study of Tuberculosis in Brazil. *PLoS One*. 2012;7(10):e45569.
 28. Fonseca BPF, Fernandes E, Fonseca MVA. Collaboration in science and technology organizations of the public sector: A network perspective. *Science and Public Policy*. 2017;44(1):37-49.
 29. Chavarro, D, Rafols, I, Tang, P. To what extent is inclusion in the Web of Science an indicator of journal 'quality'?, *Research Evaluation*. 2018;27(2): 106-118. Report No.: ID 2990653. Available from: <https://papers.ssrn.com/abstract=2990653>
 30. Lang PB, Gouveia FC, Leta J. Cooperation in health: A cluster analysis of 190 research institutions. *Journal of Scientometric Research*. 2013;2(3):22-30.
 31. Sabatier M, Carrere M, Mangematin V. Profiles of Academic Activities and Careers: Does Gender Matter? An Analysis Based on French Life Scientist CVs. *The Journal of Technology Transfer*. 2006;31(3):311-24.
 32. Dietz JS, Chompalov I, Bozeman B, Lane EO, Park J. Using the Curriculum Vita to Study the Career Paths of Scientists and Engineers: An Exploratory Assessment. *Scientometrics*. 2000;49(3):419-42.
 33. Dietz JS, Chompalov I, Bozeman B, Lane EO, Park J. Using the Curriculum Vita to Study the Career Paths of Scientists and Engineers: An Exploratory Assessment. *Scientometrics*. 2000;49(3):419-42.
 34. Lee S, Bozeman B. The Impact of Research Collaboration on Scientific Productivity. *Social Studies of Science*. 2005;35(5):673-702.
 35. Mena-Chalco JP, Digiampietri LA, Lopes FM, Cesar RM. Brazilian bibliometric coauthorship networks: Brazilian Bibliometric Coauthorship Networks. *Journal of the Association for Information Science and Technology*. 2014;65(7):1424-45.
 36. Bozeman B, Dietz JS, Gaughan M. Scientific and technical human capital: An alternative model for research evaluation. *International Journal of Technology Management*. 2001;22(7/8):716.
 37. Cañibano C, Bozeman B. Curriculum vitae method in science policy and research evaluation: The state-of-the-art. *Research Evaluation*. 2009;18(2):86-94.
 38. Digiampietri LA, Mena-Chalco JP, Pérez-Alcázar JJ, Tuesta EF, Delgado KV, Mugnaini R, *et al.* Minerando e caracterizando dados de Currículos Lattes. *Anais*; 2012.
 39. Mena-Chalco JP, Marcondes R, Marcondes R. Script Lattes: An open-source knowledge extraction system from the Lattes platform. *Journal of the Brazilian Computer Society*. 2009;15(4):31-9.
 40. ANII. Indicadores De Ciencia, Tecnología E Innovación En Uruguay 2016. Montevideo: Agencia Nacional de Investigación e Innovación (ANII); 2016;55. [cited 2020 Oct 29] Available from: <https://www.anii.org.uy/upcms/files/listado-documentos/documentos/bolet-n-de-indicadores-de-cti-uruguay-2016-1-.pdf>
 41. UDELAR. Propuesta al país 2020-2024. Plan estratégico de desarrollo de la Universidad de la República. Montevideo: Universidad de la República; 2020;373. [cited 2020 Oct 29] Available from: https://udelar.edu.uy/portal/wp-content/uploads/sites/48/2020/09/Presupuesto_2020-2024.pdf
 42. Newman MEJ. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*. 2004;101(suppl 1):5200-5.
 43. Granovetter MS. The Strength of Weak Ties. *American Journal of Sociology*. 1973;78(6):1360-80.
 44. Guimarães R, Souza LEPFD, Santos LMP, Serruya SJ. Não há saúde sem pesquisa: Avanços no Brasil de 2003 a 2010. *Revista Baiana de Saúde Pública*. 2012;36(1):55-65.
 45. Morel C. Health research and the millennium development goals: Global challenges and opportunities, national solutions and policies. *Ciência and Saúde Coletiva*. 2004;9(2):261-70.
 46. Tomassini C. Interações entre o sistema de ciência, tecnologia e inovação e o sistema de saúde no Brasil (2000-2014). Uma análise a partir da perspectiva do conhecimento e a inovação para a inclusão social: Universidade Federal do Rio de Janeiro; 2017.
 47. Chiarini T, Oliveira VP, do Couto e Silva Neto FC. Spatial distribution of scientific activities: An exploratory analysis of Brazil, 2000-10. *Science and Public Policy*. 2014;41(5):625-40.
 48. Tomassini C, Bianchi C, Couto Soares MC. Health-related knowledge production in Brazil: Regional interaction networks and priority setting. *Innovation and Development*. 2019;9(2):187-204.
 49. Gadelha CA, Vargas MA, Alves NG. Pesquisa translacional e sistemas de inovação na saúde: Implicações para o segmento biofarmacêutico. *Saúde em Debate*. 2019;43(spe2):133-46.
 50. Kickbusch I. Policy Innovation for Health. Springer Science and Business Media; 2008;209.

PhD training in health in Uruguay and Brazil

In the understanding that PhD training constitutes a critical stage in building and developing R&D capacities, it was interesting to approach the evolution and geographical location of health researchers' PhD training in both countries. The information contained in the CVs was analyzed for this purpose and complement the network analysis.

The case of Uruguay

Among researchers participating in health research networks, there is a sustained increase in the number of PhD graduates in 1984–2015, as well as an increase in the proportion of those who graduate in the country to the detriment of those who do so abroad (Figure S1).

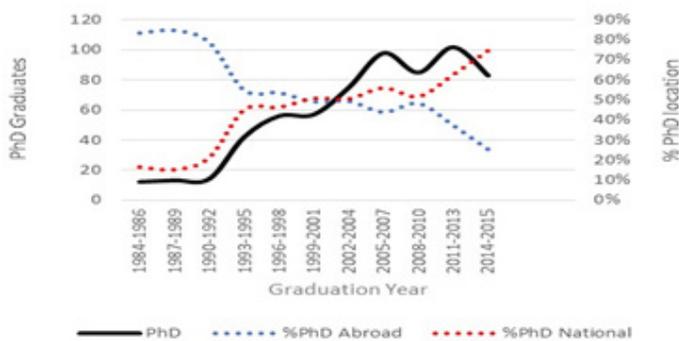


Figure S1: Evolution of PhD graduation of health researchers in Uruguay, 1984-2015.

Until 1992, the proportion of graduates abroad was significantly higher, which is associated with emigration during the dictatorial regime (1975–1985) of a significant contingent of scientists and young people who then continued their training abroad. As of 1986, the proportion of those who graduate in the country increases to become a majority from the beginning of the 2000s until the end of the period. This relocation of PhD studies among health researchers can be associated with the establishment of several new institutions that: i) open work and academic opportunities, ii) offer the possibility to obtain a PhD degree in the country, and iii) support PhD training in general and in the health area in particular. The top five countries chosen by those who graduate abroad in 2000–2015 are: Spain, France, Argentina Brazil and the United States. In 2012–2015, half of the health PhDs graduate in the region, which may be associated with the growing number of postgraduate programs offered also at a regional level.

The case of Brazil

According to the analyzed information, PhD qualifications of group leaders in the health area grew continuously until 2008 (Figure S2). The effect of the drop observed towards the

Table S1: Description of ARS metrics and indicators.

Indicator	Definition and calculation formula	Notation in the article
Number of nodes	Total number of nodes in the network	n
Number of ties	Total number of links between nodes	t
Density	Measures how close the network is to being complete. A complete network means that all possible connections take place between all nodes, in which case the density is 1. Number of effective ties in the network, expressed as a proportion of the number possible. In a network with undirected links, the density is $D = 2 \frac{t}{n(n-1)}$	D
Average Degree Centrality	Average number of ties that each node in the network has. $ADC = \frac{t}{n}$ Where t is the total number of links; n is the total number of nodes	ADC
Clustering Coefficient	Average of the individual clustering coefficient $CCo = \frac{\sum C_{coi}}{n}$ Where Ccoi is the individual clustering coefficient (i.e., the density of ties among nodes connected to a given node); n is the number of nodes in the network	CCo
Largest Component	It refers to the largest group of nodes that are all connected to each other, directly or indirectly.	LC
Degree Centrality	Number of links that a node has, or number of adjacent nodes	DC
Betweenness Centrality	It is a measure of how often a given node falls along the shortest path between two other nodes. $BC_j = \sum_{i < k} \frac{g_{ijk}}{g_{ik}}$ where g _{ijk} is the number of geodesic paths connecting i and k through j, and g _{ik} is the total number of geodesic paths connecting i and k.	BC
Closeness Centrality	It is a measure of how close a node is to all the others in the network, higher CC measurements mean greater closeness to all other nodes. Total distance (in the graph) of a given node from all other nodes. $CC = \frac{1}{\sum_y d(y,x)}$ Where d(y,x) is the distance between nodes y and x.	CC

present is explained by the date of data extraction in 2015. The vast majority of group leaders are consolidated researchers with completed doctorates, so it is more difficult to find recent doctorates in the sample.

A significant majority of those who graduate in 1984–2015 obtain their degree in the country and not abroad. The number of those graduating in the country increases throughout the period except for the period from 1993 to 1995. This may be related to Brazil’s major economic crisis in the 1980s that led to larger emigration flows in general during that period, which is when those graduating in said triennium begin their PhD studies.

An increasing number of graduates and their growing nationalization is associated with a highly mature national postgraduate system, which has been continuously expanding since the 1970s. Most leaders of health research groups that obtain their PhD degrees abroad choose the United States and Canada as their study destinations, as well as the following European countries: Spain, France, Portugal, Great Britain and Germany. Brazilians, unlike what is observed in Uruguay,

do not choose countries in the region to carry out their PhD studies. On the contrary, a high-quality and varied PhD program offer makes Brazil a leading host country of students in Latin America.

The analyzed information shows a growing nationalization of PhD training both in Brazil and Uruguay. It is possible to propose a hypothesis that identifies the development of PhD institutionalization in both countries, and the search and demand for greater specialization, as the cause for a migration shift due to PhD training.

Evolution of connectivity of individual collaboration networks in health research projects

The graphs show the evolution of the DC and LG analyzed for each country in the article (Figure S3, S4, S5, S6).

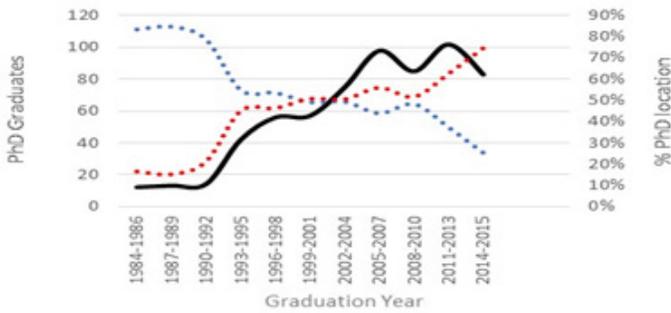


Figure S2: Evolution of PhD graduation of health researchers in Brazil, 1984-2015.

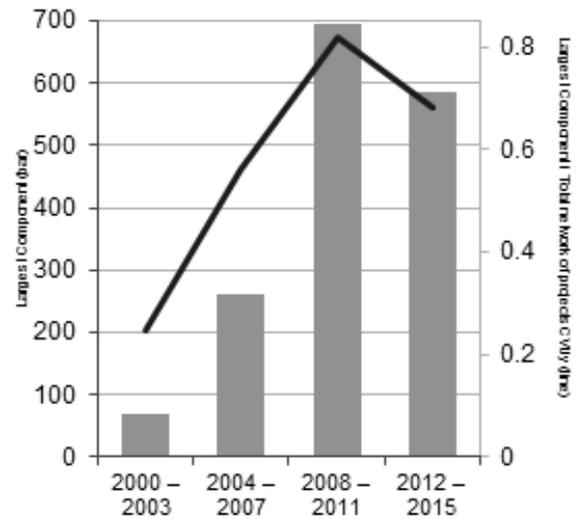


Figure S4: Evolution of Largest Component in the health project network-CVUy

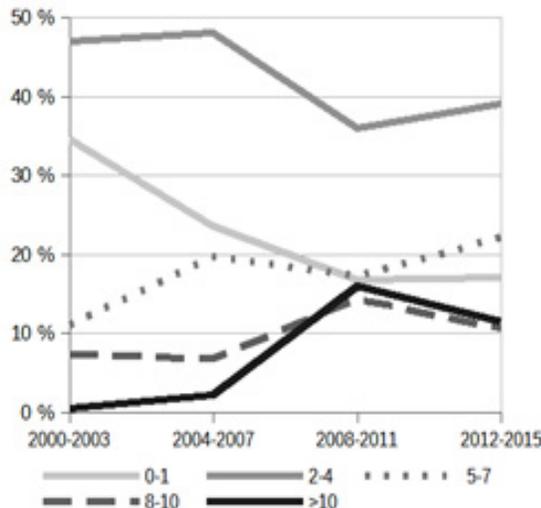


Figure S3: Evolution of connectivity: Degree Centrality (DC) by sections-CVUy.

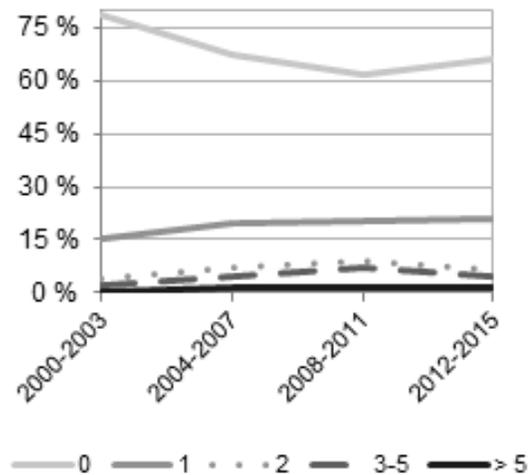


Figure S5: Evolution of connectivity: Degree Centrality (DC) by sections Lattes-CV.

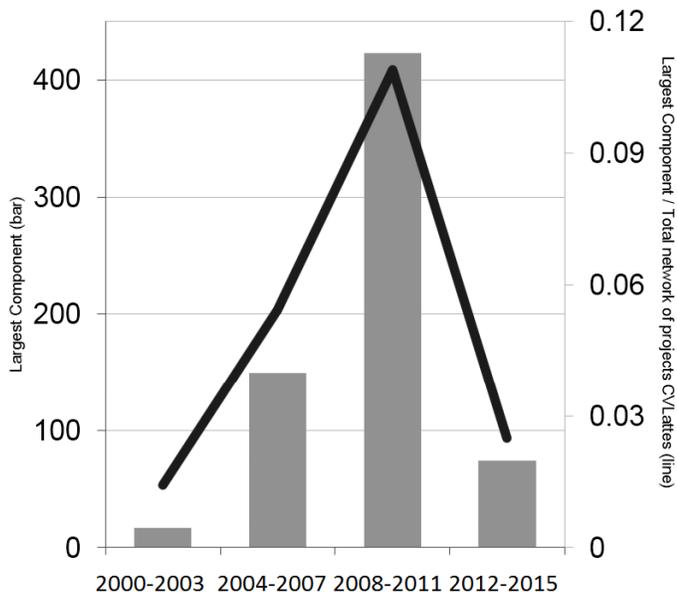


Figure S6: Evolution of Largest Component in the health project network Lattes-CV.