



**UNIVERSIDADE DE BRASÍLIA**  
**FACULDADE DE CIÊNCIA DA INFORMAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

**RICARDO BARROS SAMPAIO**

**AS ESTRUTURAS GLOBAIS E REGIONAIS DO CAMPO DE  
PESQUISA, DESENVOLVIMENTO E INOVAÇÃO DAS DOENÇAS  
NEGLIGENCIADAS LEISHMANIOSE E TUBERCULOSE SOB A ÓTICA  
DAS REDES COMPLEXAS**

Brasília-DF

2015



**UNIVERSIDADE DE BRASÍLIA**  
**FACULDADE DE CIÊNCIA DA INFORMAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

**RICARDO BARROS SAMPAIO**

**AS ESTRUTURAS GLOBAIS E REGIONAIS DO CAMPO DE  
PESQUISA, DESENVOLVIMENTO E INOVAÇÃO DAS DOENÇAS  
NEGLIGENCIADAS LEISHMANIOSE E TUBERCULOSE SOB A ÓTICA  
DAS REDES COMPLEXAS**

Tese submetida ao Programa de Pós-Graduação em Ciência da Informação da Faculdade de Ciência da Informação da Universidade de Brasília como requisito para a obtenção do título de Doutor em Ciência da Informação.

**Orientador: Prof. Dr. Jorge Henrique Cabral  
Fernandes**

Brasília-DF

2015

© 2015 Autor: Ricardo Barros Sampaio. Qualquer parte desta publicação pode ser reproduzida, desde que citada a fonte.

Sampaio, Ricardo Barros

As estruturas globais e regionais do campo de pesquisa, desenvolvimento e inovação das doenças negligenciadas leishmaniose e tuberculose sob a ótica das redes complexas / Ricardo Barros Sampaio – Brasília: 2015. 216 f.; Ilustrado; 30 cm.

Tese (doutorado) – Universidade de Brasília, Faculdade de Ciência da Informação, Programa de Pós-Graduação em Ciência da Informação, 2015.

Inclui bibliografia.

Orientação: Jorge Henrique Cabral Fernandes.

1. Análise de redes complexas. 2. Análise de redes sociais. 3. Coautoria. 4. Clusterização. 5. Doenças negligenciadas. 6. Leishmaniose. 7. Tuberculose.



### FOLHA DE APROVAÇÃO

**Título:** *"As estruturas globais e regionais do campo de pesquisa, desenvolvimento e inovação das doenças negligenciadas Leishmaniose e Tuberculose sob a ótica das Redes Complexas"*

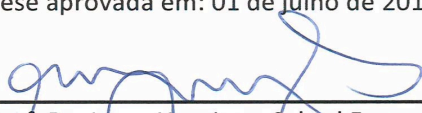
**Autor (a):** Ricardo Barros Sampaio

**Área de concentração:** Transferência da Informação

**Linha de pesquisa:** Gestão da Informação

Tese submetida à Comissão Examinadora designada pelo Colegiado do Programa de Pós-graduação em Ciência da Informação da Faculdade em Ciência da Informação da Universidade de Brasília como requisito parcial para obtenção do título de **Doutor** em Ciência da Informação.

Tese aprovada em: 01 de julho de 2015.



---

Prof. Dr. Jorge Henrique Cabral Fernandes  
Presidente (UnB/PPGCINF)

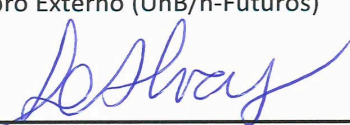
---

Prof. Dr. Carlos Medicis Morel  
Membro Externo (UNESP)



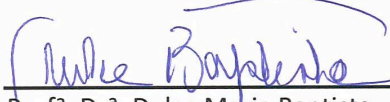
---

Prof. Dr. Isaac Roitman  
Membro Externo (UnB/n-Futuros)



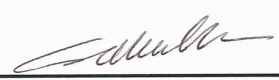
---

Prof.ª Dr.ª Lillian Maria Araújo de Rezende Alvares  
Membro Interno (UnB/PPGCINF)



---

Prof.ª Dr.ª Dulce Maria Baptista  
Membro Interno (UnB/PPGCINF)



---

Prof.ª Dr.ª Suzana Pinheiro Machado Mueller  
Suplente (UnB/PPGCINF)

**DEDICATÓRIA**

À NILMAR, A URIAS e A JOSÉ SAMPAIO  
*Meus pais, que me ajudaram a ser o que sou e a  
chegar onde estou.*

À CINTHIA,  
*Minha esposa, que sempre esteve ao meu lado, nos  
momentos fáceis e difíceis, no Brasil e no mundo,  
que me fará chegar onde estarei amanhã.*

## AGRADECIMENTOS

*Gostaria de agradecer ao Professor Dr. Jorge Henrique Cabral Fernandes que me apoiou no caminho da pesquisa durante todo o processo de doutoramento, indo além dos conceitos e técnicas acadêmicas,*

*À professora Brigitte Gay, minha coorientadora na França durante o estágio de pesquisa, que me apresentou a várias das ferramentas e métodos que foram de extrema importância para a realização dos trabalhos realizados,*

*Aos professores, coordenadores, alunos companheiros de pesquisa e a equipe de apoio do Programa de Pós-Graduação da Faculdade de Ciência da Informação da Universidade de Brasília, que sempre me incetivaram,*

*À banca avaliadora, Prof. Dr. Isaac Roitman com seu profundo conhecimento sobre a pesquisa, Profa. Dra. Lillian Maria Araújo de Rezende Álvares que tem me apoiado em diferentes atividades acadêmicas, Profa. Dra. Dulce Maria Baptista que esteve desde a minha entrada no programa me acompanhando e a Profa. Dra. Suzana Pinheiro Machado Mueller que apesar do tempo pôde fazer uma grande contribuição,*

*À Fundação Oswaldo Cruz, que por meio do Dr. Wagener Martins me proporcionou um arcabouço de informações para aplicar na minha pesquisa, À Dra. Celina Roitman que ajudou na finalização do trabalho e ao Prof. Dr. Carlos Medicis Morel que mesmo à distância incentivou à pesquisa. Aos demais companheiros de trabalho, Bruna Fonseca, Marcelo Jesus Marcus Vinicius, Waldir Campelo e tantos outros da Fiocruz que participaram das várias atividades de pesquisa relacionadas aos trabalhos apresentados,*

*Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro para o desenvolvimento da pesquisa,*

*E por último a toda minha família que sempre esteve presente nos momentos de alegria e entusiasmos durante esta trajetória.*

## RESUMO

Na era da informação, do conhecimento e das mídias digitais, de que forma os padrões de relacionamento interpessoais registrados geram informações sobre a estrutura global e regional da pesquisa em doenças negligenciadas? Este estudo investiga as redes de coautoria de cientistas que trabalham em leishmaniose e em tuberculose, duas das principais doenças negligenciadas no contexto brasileiro e mundial, em busca de revelar de que forma vêm ocorrendo mudanças no universo de participação das pesquisas. Foi identificada a ascensão de pesquisadores de países como Brasil, Índia, China e África do Sul, que se tornaram não apenas relevantes, mas em certas áreas chegam a dominar a estrutura de colaboração em rede e a produtividade no campo observado. Os resultados apresentados possibilitam às agências de financiamento avaliar seu papel em relação aos objetivos de construção e desenvolvimento da capacidade científica, bem como a consistência de parcerias científicas em todo o mundo. Eles também permitem a avaliação dos mecanismos que apoiam e incentivam a investigação em países endêmicos. O estudo também realiza uma análise exploratória acerca dos resultados da aplicação de métodos de análise de redes complexas no universo analisado. Nesse sentido, a pesquisa avalia a criação de instrumentos que apoiem a tomada de decisão em saúde pública, buscando demonstrar como a utilização desses métodos, amparados pelo uso apropriado de tecnologias para tratamento de dados, apresenta alternativas promissoras na avaliação da ciência e das dinâmicas de redes de ciência e tecnologia. O estudo utilizou diferentes bases para a recuperação dos dados, a saber: o PubMed, a Web of Science e o SciELO. Foi analisada a dinâmica das redes de coautoria em publicações científicas além das palavras-chave, das revistas científicas e outros elementos que compõem os chamados metadados das publicações científicas. O resultado final é um conjunto de métodos que podem apoiar o estudo das comunidades científicas e dos grupos de pesquisa com base em seu comportamento específico quanto à comunicação e ao relacionamento entre seus pares.

**Palavras-chave:** Análise de redes complexas. Análise de redes sociais. Coautoria. Clusterização. Doenças negligenciadas. Leishmaniose. Tuberculose.

## ABSTRACT

In the age of information, knowledge and digital media, how the registered interpersonal relationship patterns generate information on global and regional structure of the research in neglected diseases? This study investigates the networks of co-authorship of scientists working on leishmaniasis and tuberculosis, two of the major neglected diseases in the Brazilian and global context, seeking to reveal how changes are taking place in the universe of shared research. It has been identified the rise of researchers from countries such as Brazil, India, China and South Africa, which became not only relevant, but in certain areas come to dominate the networked collaborative structure and productivity in the observed field. The presented results allow the funding agencies to assess its role in relation to construction and development goals of scientific capacity as well as the consistency of scientific partnerships worldwide. They also allow the evaluation of the mechanisms that support and encourage research in endemic countries. The study also conducts an exploratory analysis of the results of the application analysis of complex networks methods in the universe analyzed. In this sense, the research evaluates the development of tools to support public health decisions demonstrating how the use of these methods, supported by appropriate use of technologies for data processing, offers promising alternatives in the evaluation of science and network dynamics on science and technology. The study used different databases for data retrieval, namely: PubMed, Web of Science and SciELO. The dynamics of co-authorship networks in scientific publications was analyzed in addition to keywords, scientific journals and other elements that make up the so-called metadata of scientific publications. The end result is a set of methods that can support the study of scientific communities and research groups based on their specific behavior as communication and the relationship among their peers.

**Keywords:** Analysis of complex networks. Social network analysis. Co-authoring. Clustering. Diseases neglected. Leishmaniasis. Tuberculosis.



## Lista de figuras

<b>FIGURA 1.</b>	<b><i>Status da endemidade da leishmaniose visceral no mundo em 2012</i></b> .....	<b>22</b>
<b>FIGURA 2.</b>	<b><i>Status da endemidade da leishmaniose tegumentar no mundo em 2012</i></b> .....	<b>22</b>
<b>FIGURA 3.</b>	<b><i>Incidência de novos casos de tuberculose estimados no ano de 2012</i></b> .....	<b>24</b>
<b>FIGURA 4.</b>	<b><i>Mapa conceitual da análise de redes sobre a produção científica e tecnológica</i></b> .....	<b>26</b>
<b>FIGURA 5.</b>	<b><i>Rede de amizade entre alunos de uma escola de ensino fundamental</i></b> ....	<b>30</b>
<b>FIGURA 6.</b>	<b><i>Número de publicações na área de colaboração científica por ano</i></b> .....	<b>35</b>
<b>FIGURA 7.</b>	<b><i>Número de publicações por país</i></b> .....	<b>36</b>
<b>FIGURA 8.</b>	<b><i>Número de publicações por autor</i></b> .....	<b>36</b>
<b>FIGURA 9.</b>	<b><i>Número de publicações em coautoria por ano</i></b> .....	<b>37</b>
<b>FIGURA 10.</b>	<b><i>Rede de colaboração em pesquisa sobre colaboração científica</i></b> .....	<b>38</b>
<b>FIGURA 11.</b>	<b><i>Número de publicações em coautoria por ano</i></b> .....	<b>40</b>
<b>FIGURA 12.</b>	<b><i>Mapa da internet</i></b> .....	<b>66</b>
<b>FIGURA 13.</b>	<b><i>Interação proteína-proteína</i></b> .....	<b>69</b>
<b>FIGURA 14.</b>	<b><i>Circuito neural de uma minhoca</i></b> .....	<b>69</b>
<b>FIGURA 15.</b>	<b><i>Estrutura de redes Mundo Pequeno, Livre de Escala e Aleatória</i></b> .....	<b>72</b>
<b>FIGURA 16.</b>	<b><i>Rede centro-periferia</i></b> .....	<b>73</b>
<b>FIGURA 17.</b>	<b><i>Número de publicações científicas por ano (gráfico externo). Número de publicações científicas por autor (gráfico interno)</i></b> .....	<b>80</b>
<b>FIGURA 18.</b>	<b><i>Número de publicações científicas sobre leishmaniose por país, por período</i></b> .....	<b>83</b>
<b>FIGURA 19.</b>	<b><i>Mapa-múndi da quantidade de publicações científicas sobre leishmaniose por país</i></b> .....	<b>84</b>
<b>FIGURA 20.</b>	<b><i>Número de publicações científicas sobre tuberculose por ano</i></b> .....	<b>87</b>
<b>FIGURA 21.</b>	<b><i>Número de publicações científicas sobre tuberculose por país, por período</i></b> .....	<b>88</b>
<b>FIGURA 22.</b>	<b><i>Curva de distribuição de grau dos autores na rede sobre leishmaniose</i></b>	<b>95</b>
<b>FIGURA 23.</b>	<b><i>Correlação entre centralidade de grau e centralidade de intermediação</i></b>	<b>96</b>
<b>FIGURA 24.</b>	<b><i>Curva de distribuição de grau dos autores na rede sobre tuberculose</i></b> ....	<b>98</b>
<b>FIGURA 25.</b>	<b><i>Rede de relações entre hubs</i></b> .....	<b>102</b>
<b>FIGURA 26.</b>	<b><i>Rede de relacionamentos entre grupos sobre leishmaniose</i></b> .....	<b>103</b>

<b>FIGURA 27.</b>	<b>Rede de relações entre hubs .....</b>	<b>105</b>
<b>FIGURA 28.</b>	<b>Rede de relacionamentos entre grupos de pesquisa sobre tuberculose .</b>	<b>106</b>
<b>FIGURA 29.</b>	<b>Rede leishmaniose – período 1985-1988.....</b>	<b>110</b>
<b>FIGURA 30.</b>	<b>Rede leishmaniose – período 2005-2008.....</b>	<b>111</b>
<b>FIGURA 31.</b>	<b>Evolução da entrada e saída de pesquisadores em leishmania .....</b>	<b>112</b>
<b>FIGURA 32.</b>	<b>Evolução da rede de leishmania dentro e entre comunidades .....</b>	<b>113</b>
<b>FIGURA 33.</b>	<b>Interações entre os hubs dos grupos 1 e 3.....</b>	<b>115</b>
<b>FIGURA 34.</b>	<b>Rede de colaboração em pesquisa sobre leishmaniose dos pesquisadores afiliados a instituições brasileiras com os parceiros internacionais.....</b>	<b>127</b>
<b>FIGURA 35.</b>	<b>Rede de colaboração em pesquisa sobre leishmaniose dos pesquisadores afiliados a instituições brasileiras.....</b>	<b>128</b>
<b>FIGURA 36.</b>	<b>Rede de colaboração dos pesquisadores da Fiocruz .....</b>	<b>129</b>
<b>FIGURA 37.</b>	<b>Rede de colaboração do Brasil com outros países.....</b>	<b>131</b>
<b>FIGURA 38.</b>	<b>Rede de colaboração da Fiocruz com os demais países .....</b>	<b>132</b>
<b>FIGURA 39.</b>	<b>Rede de relacionamentos de pesquisa sobre leishmaniose, por grupos .</b>	<b>133</b>
<b>FIGURA 40.</b>	<b>Rede de relacionamentos entre grupos de pesquisa sobre leishmaniose</b>	<b>134</b>
<b>FIGURA 41.</b>	<b>Relações dos grupos com as categorias .....</b>	<b>136</b>
<b>FIGURA 42.</b>	<b>Rede de colaboração em pesquisa sobre tuberculose de pesquisadores afiliados a instituições brasileiras.....</b>	<b>142</b>
<b>FIGURA 43.</b>	<b>Rede de colaboração em pesquisa sobre tuberculose de pesquisadores afiliados a instituições brasileiras.....</b>	<b>144</b>
<b>FIGURA 44.</b>	<b>Rede de colaboração em pesquisa sobre tuberculose de pesquisadores afiliados à Fiocruz.....</b>	<b>145</b>
<b>FIGURA 45.</b>	<b>Rede de relacionamentos entre grupos de pesquisa sobre tuberculose .</b>	<b>149</b>
<b>FIGURA 46.</b>	<b>Rede de relacionamentos entre grupos de pesquisa sobre tuberculose, separada por grupos .....</b>	<b>150</b>
<b>FIGURA 47.</b>	<b>Rede de tuberculose e pesquisadores Fiocruz – grupos .....</b>	<b>151</b>
<b>FIGURA 48.</b>	<b>Rede de relacionamentos dos pesquisadores Fiocruz – grupos .....</b>	<b>152</b>
<b>FIGURA 49.</b>	<b>Relações dos grupos com as categorias .....</b>	<b>153</b>
<b>FIGURA 50.</b>	<b>Relações do grupo 1 com as categorias .....</b>	<b>155</b>
<b>FIGURA 51.</b>	<b>Relações do grupo localizado fora do componente gigante com as categorias .....</b>	<b>155</b>
<b>FIGURA 52.</b>	<b>Relações do grupo 13 com as categorias. ....</b>	<b>156</b>
<b>FIGURA 53.</b>	<b>Relações do grupo 12 com as categorias .....</b>	<b>157</b>

<b>FIGURA 54.</b>	<b>Relações da categoria <i>Immunology</i> com os grupos .....</b>	<b>158</b>
<b>FIGURA 55.</b>	<b>Relações da categoria <i>Plant Sciences</i> com os grupos .....</b>	<b>158</b>
<b>FIGURA 56.</b>	<b>Rede de dois modos mudando para uma rede de um modo .....</b>	<b>166</b>
<b>FIGURA 57.</b>	<b>Número de publicações em leishmaniose por ano.....</b>	<b>167</b>
<b>FIGURA 58.</b>	<b>Número de publicações em tuberculose por ano.....</b>	<b>168</b>
<b>FIGURA 59.</b>	<b>Número de títulos indexados no SciELO por país .....</b>	<b>1697</b>
<b>FIGURA 60.</b>	<b>Número de publicações em leishmaniose por país.....</b>	<b>170</b>
<b>FIGURA 61.</b>	<b>Número de publicações em tuberculose por país.....</b>	<b>170</b>
<b>FIGURA 62.</b>	<b>Número de publicações em leishmaniose por revista .....</b>	<b>172</b>
<b>FIGURA 63.</b>	<b>Número de publicações em tuberculose por revista .....</b>	<b>173</b>
<b>FIGURA 64.</b>	<b>Número de publicações em leishmaniose associadas às áreas de conhecimento .....</b>	<b>174</b>
<b>FIGURA 65.</b>	<b>Número de publicações em tuberculose associadas às áreas de conhecimento .....</b>	<b>174</b>
<b>FIGURA 66.</b>	<b>Rede de colaboração em leishmaniose .....</b>	<b>178</b>
<b>FIGURA 67.</b>	<b>Pesquisadores em leishmaniose com maior relevância.....</b>	<b>179</b>
<b>FIGURA 68.</b>	<b>Rede de colaboração em tuberculose .....</b>	<b>180</b>
<b>FIGURA 69.</b>	<b>Pesquisadores em tuberculose com maior relevância.....</b>	<b>182</b>
<b>FIGURA 70.</b>	<b>Rede de dois modos em leishmaniose. Pesquisadores e revistas.....</b>	<b>183</b>
<b>FIGURA 71.</b>	<b>Rede de dois modos em tuberculose. Pesquisadores e revistas.....</b>	<b>184</b>
<b>FIGURA 72.</b>	<b>Revistas em leishmaniose sem relacionamentos expressivos com outras revistas .....</b>	<b>185</b>
<b>FIGURA 73.</b>	<b>Revistas em tuberculose sem relacionamentos expressivos com outras revistas .....</b>	<b>186</b>
<b>FIGURA 74.</b>	<b>Rede de revistas em leishmaniose.....</b>	<b>187</b>
<b>FIGURA 75.</b>	<b>Rede de revistas em tuberculose.....</b>	<b>189</b>
<b>FIGURA 76.</b>	<b>Rede de relações entre hubs .....</b>	<b>194</b>
<b>FIGURA 77.</b>	<b>Número de publicações de entidades brasileiras em áreas específicas na Web of Science .....</b>	<b>198</b>

## Lista de tabelas

<b>TABELA 1.</b>	<b>Revistas científicas com maior número de publicações sobre leishmaniose .....</b>	<b>81</b>
<b>TABELA 2.</b>	<b>Número de MeSH por período em pesquisas sobre leishmaniose .....</b>	<b>85</b>
<b>TABELA 3.</b>	<b>Número de publicações por ano das vinte revistas mais bem cotadas .....</b>	<b>89</b>
<b>TABELA 4.</b>	<b>Número de MeSH por ano das cem palavras mais frequentes .....</b>	<b>90</b>
<b>TABELA 5.</b>	<b>Estatísticas agregadas básicas para a rede de colaboração em leishmaniose .....</b>	<b>94</b>
<b>TABELA 6.</b>	<b>Estatísticas agregadas básicas para a rede de colaboração em tuberculose .....</b>	<b>97</b>
<b>TABELA 7.</b>	<b>Comunidades e distribuição dos hubs e dos nós do <i>k-core</i>.....</b>	<b>100</b>
<b>TABELA 8.</b>	<b>Comunidades e distribuição dos hubs e dos nós do <i>k-core</i>.....</b>	<b>104</b>
<b>TABELA 9.</b>	<b>Autores, publicações e relacionamentos de pesquisa sobre leishmaniose .....</b>	<b>107</b>
<b>TABELA 10.</b>	<b>Medidas de rede: graus médio e máximo, coeficiente de agregação e distância média da rede sobre leishmaniose e de uma rede randômica de tamanho similar (nós e linhas).....</b>	<b>108</b>
<b>TABELA 11.</b>	<b>Número de publicações por ano dos 15 pesquisadores com mais de 50 publicações nos triênios avaliados.....</b>	<b>122</b>
<b>TABELA 12.</b>	<b>Avaliação temporal das categorias da Web of Science.....</b>	<b>123</b>
<b>TABELA 13.</b>	<b>Número de colaborações do Brasil com outros países.....</b>	<b>124</b>
<b>TABELA 14.</b>	<b>Instituições com publicações em leishmaniose no Brasil.....</b>	<b>124</b>
<b>TABELA 15.</b>	<b>Indicadores da rede de colaboração em pesquisa sobre leishmaniose ...</b>	<b>125</b>
<b>TABELA 16.</b>	<b>Países que mais colaboram com instituições brasileiras e com a Fiocruz no âmbito da rede de colaboração em pesquisa sobre leishmaniose.....</b>	<b>130</b>
<b>TABELA 17.</b>	<b>Indicadores da rede de colaboração em pesquisa sobre leishmaniose ...</b>	<b>135</b>
<b>TABELA 18.</b>	<b>Indicadores da rede de colaboração em pesquisa sobre leishmaniose ...</b>	<b>137</b>
<b>TABELA 19.</b>	<b>Número de publicações por ano dos 25 pesquisadores com mais publicações no período 2003-2013.....</b>	<b>140</b>
<b>TABELA 20.</b>	<b>Avaliação temporal das categorias da Web of Science.....</b>	<b>141</b>
<b>TABELA 21.</b>	<b>Indicadores da rede de colaboração em pesquisa sobre tuberculose.....</b>	<b>146</b>
<b>TABELA 22.</b>	<b>Países que mais colaboram com os cientistas brasileiros e da Fiocruz no âmbito da rede de colaboração em pesquisa sobre tuberculose.....</b>	<b>146</b>

<b>TABELA 23. Comunidades e relações existentes na rede de pesquisa em tuberculose .....</b>	<b>147</b>
<b>TABELA 24. Percentual de títulos e publicações por país .....</b>	<b>171</b>
<b>TABELA 25. Percentual de títulos e publicações por país .....</b>	<b>176</b>
<b>TABELA 26. Dados redes de dois modos.....</b>	<b>177</b>
<b>TABELA 27. Dados redes de dois modos.....</b>	<b>177</b>
<b>TABELA 28. Percentual de títulos e publicações por país.....</b>	<b>200</b>

## Lista de siglas

- ARS – Análise de Redes Sociais
- BMC – Bioorganic & Medicinal Chemistry
- Brics – Brasil, Rússia, Índia, China, África do Sul
- CI – Ciência da Informação
- DNA – Deoxyribonucleic Acid
- DNDi – Drugs for Neglected Diseases Initiative
- DST – Doenças Sexualmente Transmissíveis
- ECI – Escola de Ciência da Informação
- ER – Erdős Rény
- HIV – Human Immunodeficiency Virus
- HPV – Human Papiloma Virus
- IP – Internet Protocol
- LIS – Library and Information Science
- LN – Leishmaniose
- NCBI – National Center for Biotechnology Information
- NLM – National Library of Medicine
- OMS – Organização Mundial da Saúde
- OMS – Organização Mundial da Saúde
- Opas – Organização Panamericana de Saúde
- PD&I – Pesquisa, Desenvolvimento e Inovação
- PPI - Protein Protein Interaction
- PUC – Pontifícia Universidade Católica
- SciELO – Scientific Electronic Library Online
- SUS – Sistema Único de Saúde
- TB – Tuberculose
- UFMG – Universidade Federal de Minas Gerais
- UFPB – Universidade Federal da Paraíba
- UnB – Universidade de Brasília
- USP – Universidade de São Paulo
- WHO – World Health Organization
- WoS – Web of Science
- WWW – World Wide Web

## SUMÁRIO

<b>PARTE I – INTRODUÇÃO .....</b>	<b>17</b>
<b>1 Introdução 17</b>	
<b>1.1 Justificativa.....</b>	<b>18</b>
1.1.1 Doenças negligenciadas .....	20
1.1.2 Leishmaniose.....	21
1.1.3 Tuberculose.....	23
<b>1.2 A análise de redes no apoio à avaliação e à gestão do conhecimento .....</b>	<b>24</b>
<b>1.3 Objetivos da pesquisa .....</b>	<b>27</b>
1.3.1 Objetivos geral e específicos .....	27
1.3.2 Pressupostos.....	28
<b>2 Referencial teórico.....</b>	<b>29</b>
<b>2.1 Ciência de redes.....</b>	<b>29</b>
<b>2.2 Redes de coautoria .....</b>	<b>33</b>
<b>2.3 Análise da colaboração científica por meio de redes de coautoria.....</b>	<b>34</b>
2.3.1 Dados bibliométricos .....	34
2.3.2 Rede de colaboração.....	38
<b>2.4 Colaboração Científica.....</b>	<b>41</b>
2.4.1 Estudos empíricos .....	41
2.4.2 Metodologias.....	43
2.4.3 Redes de coautoria no Brasil .....	46
2.4.4 Redes de coautoria na saúde.....	50
<b>3 Metodologia.....</b>	<b>53</b>
<b>3.1 Modelo exploratório de análise de redes.....</b>	<b>53</b>
3.1.1 Definição da rede .....	55
3.1.2 Manipulação da rede .....	55
3.1.3 Determinação das características estruturais.....	56
3.1.4 Inspeção visual.....	57
<b>3.2 Coleta e tratamento dos dados .....</b>	<b>58</b>
3.2.1 PubMed .....	58
3.2.2 Web of Science .....	59
3.2.3 SciELO .....	60
<b>3.3 Elementos e medidas de redes .....</b>	<b>61</b>
3.3.1 Elementos.....	61
3.3.2 Medidas de rede .....	62
<b>3.4 Tipos e estrutura das redes .....</b>	<b>65</b>
3.4.1 Tipos de redes .....	65
3.4.2 Estrutura das redes.....	70
<b>3.5 Organização dos capítulos de resultados e discussão .....</b>	<b>74</b>
<b>PARTE II - RESULTADOS .....</b>	<b>76</b>
<b>4 Resultados da pesquisa na base PubMed .....</b>	<b>76</b>
<b>4.1 Resultados bibliométricos .....</b>	<b>79</b>
4.1.1 Dados bibliométricos – leishmaniose .....	79
4.1.2 Dados bibliométricos – tuberculose .....	85

<b>4.2</b>	<b>Organização hierárquica da rede .....</b>	<b>93</b>
4.2.1	Medidas da rede leishmaniose.....	94
4.2.2	Medidas da rede tuberculose.....	96
<b>4.3</b>	<b>Formação de grupos e suas relações .....</b>	<b>98</b>
4.3.1	Rede leishmaniose.....	99
4.3.2	Rede tuberculose .....	104
<b>4.4</b>	<b>Evolução temporal das redes e a mudança em suas estruturas.....</b>	<b>107</b>
<b>4.5</b>	<b>Interação entre comunidades e geografia .....</b>	<b>112</b>
<b>5</b>	<b>Resultados da pesquisa na Web of Science .....</b>	<b>117</b>
<b>5.1</b>	<b>Rede de pesquisadores em leishmaniose .....</b>	<b>119</b>
5.1.1	Análise bibliométrica das publicações em leishmaniose .....	121
5.1.2	Análise da rede científica de leishmaniose .....	125
5.1.3	Rede de colaboração internacional.....	130
5.1.4	Subgrupos de pesquisa em leishmaniose.....	132
5.1.5	Grandes áreas do conhecimento e caracterização dos grupos .....	136
<b>5.2</b>	<b>Rede de pesquisadores em tuberculose.....</b>	<b>138</b>
5.2.1	Análise bibliométrica das publicações em tuberculose .....	139
5.2.2	Rede de colaboração em pesquisa sobre tuberculose .....	142
5.2.3	Subgrupos de pesquisa em tuberculose.....	147
5.2.4	Grandes áreas do conhecimento e caracterização dos grupos .....	152
<b>5.3</b>	<b>Descrição das categorias Web of Science .....</b>	<b>159</b>
<b>6</b>	<b>Resultados da pesquisa no SciELO.....</b>	<b>165</b>
<b>6.1</b>	<b>Definição e manipulação das redes.....</b>	<b>165</b>
6.1.1	Dados das redes.....	167
6.1.2	Coleções, revistas e áreas do SciELO .....	168
<b>6.2</b>	<b>Determinação de características estruturais e a inspeção visual das redes .....</b>	<b>175</b>
6.2.1	Dados das redes.....	175
6.2.2	Redes de colaboração.....	178
6.2.3	Redes de afiliação .....	182
<b>PARTE III – DISCUSSÃO .....</b>		<b>191</b>
<b>7</b>	<b>Discussão dos resultados .....</b>	<b>191</b>
7.1	Discussão dos resultados PubMed .....	193
7.2	Discussão dos resultados Web of Science .....	196
7.3	Discussão dos resultados SciELO .....	200
<b>8</b>	<b>Conclusão .....</b>	<b>203</b>
<b>8.1</b>	<b>Pergunta de pesquisa, objetivos e hipóteses .....</b>	<b>204</b>
8.1.1	Objetivos.....	205
8.1.2	Pressupostos.....	206
<b>8.2</b>	<b>Trabalhos futuros .....</b>	<b>207</b>
<b>Referências .....</b>		<b>208</b>



## PARTE I – INTRODUÇÃO

### 1 Introdução

Pelo fato de estarem ligadas à pobreza, até recentemente as doenças negligenciadas eram tradicionalmente deixadas em segundo plano nas agendas internacionais de saúde. Além disso, essas doenças apresentaram, durante muitos anos, pouco incentivo para a indústria quanto aos investimentos em pesquisa, desenvolvimento e inovação (PD & I), ficando fora do mercado farmacêutico (STEVENS, 2008). No entanto, alguns dos países afetados por essas doenças fazem parte das principais economias emergentes. Com isso, os governos desses países, com o apoio de fundações não governamentais, têm mudado esse quadro, proporcionando fundos substanciais destinados a programas de pesquisa e tratamento de doenças negligenciadas, especialmente na última década (BRITO, 2013).

Após um longo período de descaso, o campo de pesquisa das doenças negligenciadas, que afetam um sexto da população do mundo (WHO, 2014), está em busca de novas formas de se fazer e entender as ciências médicas e biológicas. Graças ao apoio e ao financiamento de diferentes entidades, tanto governamentais como não governamentais, elevaram-se os investimentos e a aplicação de novas políticas de PD & I sobre essas doenças, revertendo, em alguns casos, a situação crítica na qual se encontravam. No entanto, o impacto desses investimentos e da implementação de políticas públicas sobre a pesquisa científica não é de conhecimento público.

Em decorrência dessas mudanças recentes quanto ao investimento e à participação de novos atores, acredita-se que ainda não se conhecem detalhes de como e onde a ciência tem progredido. Para resolver essa questão, este estudo montou e analisou redes sociais de relacionamento entre os agentes do campo da pesquisa científica e tecnológica, desenvolvimento e inovação, levando-se em consideração três diferentes bases científicas, uma coletânea de publicações acadêmicas sobre leishmaniose (LN) e tuberculose (TB), duas das principais doenças negligenciadas no contexto brasileiro. Com os dados coletados foi possível visualizar e analisar a estrutura da rede de colaboração nacional e mundial sobre essas doenças.

Os resultados encontrados demonstram que a ciência, nessas áreas, vem crescendo exponencialmente desde a década de 1980, seja quando medida pelo número de artigos publicados a cada ano, seja quando medida pela taxa de entrada de novos autores e trabalhos de colaboração nacional e internacional. Foi observada também uma mudança recente na rede

devido à entrada e à participação relevante de cientistas provenientes de países em ascensão. Isso modificou, em parte, a organização da ciência em torno de centros de pesquisa previamente estabelecidos.

Ademais, a análise das redes revela a existência de *clusters* ou grupos de cientistas interligados e bastante centralizados que controlam a estrutura das redes em todos os momentos. Esta tese parte do pressuposto que tal centralização limita a capacidade da sociedade de responder às necessidades críticas quanto à adoção de novas ferramentas de controle e tratamento necessárias para tais doenças. Isso porque a expansão de novas técnicas acaba por ser controlada por grupos de pesquisadores relativamente pequenos. A reversão desse quadro depende da implantação de políticas destinadas a alcançar o objetivo de expansão da ciência a serviço da humanidade. Faz-se necessário e oportuno o desenvolvimento de novas ferramentas e métodos de representação e análise de informação que evidenciem, de forma clara, as situações para que a melhor decisão seja adotada pelos implementadores de políticas públicas.

Os investimentos para tratar e erradicar as doenças negligenciadas são insuficientes, o que perpetua a condição de pessoas expostas a essas doenças (MOREL et al., 2005). Para sanar os problemas decorrentes dessa situação é necessário compreender sua complexidade, aqui incluída sua dinâmica, e uma das maneiras de modificar tal situação é pesquisar métodos para analisar o contexto que a envolve. Nesse sentido, este trabalho pretende mostrar a importância do estudo das produções científicas, suas estruturas e dinâmicas, na forma de redes complexas.

O modo como os investimentos na área de doenças negligenciadas têm impactado a pesquisa científica não é de conhecimento público. Acredita-se, portanto, que a análise de redes complexas, por meio da geração de novos tipos de dados e informação científica, pode ajudar aos formuladores de políticas públicas em vários países a avaliar os resultados científicos e econômicos de programas e parcerias em todo o mundo para os países onde essas doenças são endêmicas.

## **1.1 Justificativa**

Segundo Le Coadic (1996), a interação informacional constitui-se em dinâmica característica dos fenômenos do uso da informação, envolvendo a consulta como um indicador da necessidade de informação, o diálogo entre usuários e/ou o sistema de informação e as interações informacionais, que podem ser pessoa-pessoa, pessoa-tecnologia e

pessoa-tecnologia-pessoa. O diálogo em um processo de interação informacional é um ato de comunicação orientado para o objetivo do usuário que impõe regras de cooperação.

O atual cenário mundial das mídias digitais, da era do conhecimento e da informação, do mundo virtual e do ciberespaço oferece novas oportunidades aos indivíduos e às instituições, mas também apresenta grandes desafios. Nesse ambiente informacional e comunicacional, mediado por um sem número de sintomas, as organizações precisam de flexibilidade e agilidade para responder às demandas, oferecendo rapidamente produtos e serviços de qualidade e personalizados.

Choo (2005), Davenport, Prusak (2000), Hansen et al. (1999) afirmam que a gestão do conhecimento está ligada à gestão de pessoas, e que o uso das tecnologias de informação e comunicação e das práticas gerenciais é relevante para a criação de um ambiente adequado ao compartilhamento da informação e do conhecimento.

Nesse contexto, a atuação em redes de colaboração é, por excelência, um meio de ampliar a produção científica e aplicá-la de forma pragmática e intensa. A integração de conhecimentos e competências acelera a geração de inovações na medida em que amplia o horizonte do conhecimento em razão da troca de experiências e da mútua motivação para alcançar novos patamares de saber e realização. As redes são, primordialmente, cooperativas, não competitivas e advêm de objetivos, interpretações e sentidos compartilhados e da disposição de dividir responsabilidades.

Wasserman e Faust (1994) mostram que os atores e suas ações devem ser vistos como interdependentes, e não como unidades independentes. Por sua vez, os relacionamentos devem ser vistos como canais para o fluxo de recursos. Essa interpretação das redes abre novos caminhos para o estudo das demandas por informação e seu fluxo informacional, uma vez que a necessidade do indivíduo pode e deve ser analisada no contexto em que se encontra.

Quanto aos mecanismos de registro e identificação das ações de pesquisa e desenvolvimento tecnológico, estes são geridos por diferentes instâncias e nos diversos contextos mundiais. Essas instâncias podem ser organizações não governamentais, como Drugs for Neglected Diseases Initiative (DNDi) e Bill & Melinda Gates Foundation, estruturas governamentais de Estado, como ministérios da saúde, as indústrias farmacêutica e de biotecnologia, além das entidades de ensino e pesquisa, dentre outros. Os dados e os resultados provenientes de ações dessas diferentes instâncias são organizados por lógicas e objetivos diferentes, o que pode resultar em informações conflitantes, dispersas e de difícil alinhamento com as necessidades de diminuição ou até mesmo de erradicação de algumas dessas doenças.

Conforme as necessidades expressas pela Organização Mundial de Saúde (OMS) no trabalho *Accelerating work to overcome the global impact of neglected tropical diseases – a roadmap for implementation* (WHO, 2012), existem diretrizes para a diminuição do acometimento dessas doenças no mundo. Apoiar essas diretrizes por meio de informações relevantes e que possam apoiar a tomada de decisão é o que se espera com a aplicação de métodos de análise de redes complexas.

É esperado que o progresso na ciência seja uma fonte de crescimento econômico (ADAMS; GRILICHES, 1996). No entanto, os indicadores tradicionais de produtividade científica, tais como o fator de impacto ou o número esperado de publicações em um determinado período de tempo, têm deficiências intrínsecas e não ajudam a ciência a avançar de forma rigorosa.

Redes de colaborações científicas, estudadas por meio do padrão de coautorias de trabalhos, podem ser utilizadas para avaliar as tendências da produção científica. Morel et al. (2009) propuseram recentemente que a análise de redes de coautoria poderia tornar-se uma ferramenta importante para as organizações internacionais ou para os programas de colaboração visando à erradicação de doenças, uma vez que ela fornece informações com base científica relevante para a análise estratégica de planejamento proativo.

Dessa forma, este trabalho estuda as estruturas de interação científicas para as doenças negligenciadas e o impacto produzido na última década sobre essas estruturas e suas dinâmicas. É necessário avaliar não só se a produção científica tem aumentado em países onde as doenças negligenciadas são predominantes, como o Brasil ou a Índia, mas também de que forma a estrutura da rede e seus mecanismos inerentes podem apoiar ou eventualmente impedir o crescimento e a diversificação dessa produção, tão necessários para a inovação científica no tratamento e no controle de tais doenças.

### **1.1.1 Doenças negligenciadas**

As doenças negligenciadas são responsáveis por cerca de 11% da carga global de doenças e são reconhecidas como incapacitantes e deformantes (WHO, 2012). Embora ainda maciçamente subfinanciadas em relação ao seu impacto na saúde global (STEVENS, 2008), essas doenças negligenciadas têm tido um aumento crescente de atenção na última década, com maiores investimentos em pesquisa e produção de novas ferramentas, tais como testes de diagnóstico, medicamentos e vacinas.

A fundação Bill e Melinda Gates tem, por exemplo, um consórcio em colaboração com 13 grandes empresas farmacêuticas, com um fundo de 785 milhões dólares americanos para erradicar até 2020 dez doenças tropicais há muito negligenciadas (BRITO, 2013). Medidas inovadoras de incentivo também foram tomadas, tais como a desvinculação dos custos de pesquisa e desenvolvimento do preço dos produtos finais, a formação de parcerias de pesquisa e o desenvolvimento colaborativo com organizações não governamentais, como a DNDi, bem como a criação de um Observatório Internacional de PD & I em Saúde, sob a égide da OMS, para orientar as prioridades de pesquisa e desenvolvimento nos países em desenvolvimento.

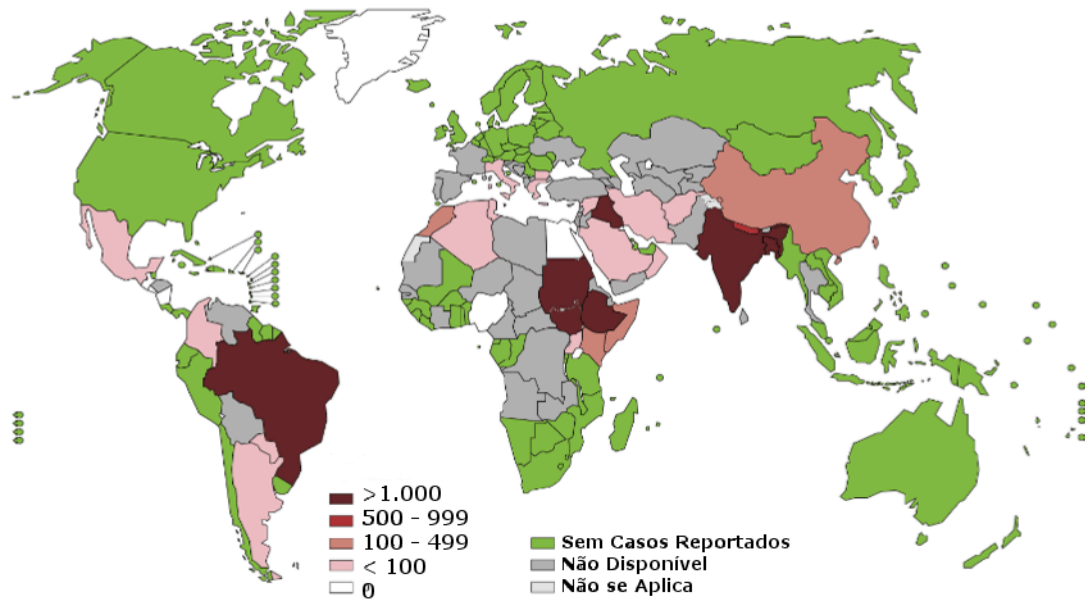
### **1.1.2 Leishmaniose**

A leishmaniose é uma séria ameaça à saúde de cerca de 350 milhões de pessoas. Sua incidência mundial vem aumentando de forma clara e preocupante, e sua distribuição geográfica também se vem ampliando (DESJEUX, 2004; WHO, 2014). Em suas várias formas, a leishmaniose pode ser mutilante, incapacitante, de longa duração ou fatal. Seus tratamentos são antigos, tóxicos e difíceis de gerir (LEMKE, 2012). Além disso, a coinfeção leishmania-HIV parece ser um problema crescente em países como a Etiópia, o Brasil e a Índia, onde ambas as infecções estão se tornando cada vez mais difundidas.

A leishmaniose é uma doença infecciosa causada por protozoários do gênero leishmania, cuja transmissão em humanos se dá pela picada da fêmea de insetos flebotomíneos infectados. A depender da espécie do parasita e da resposta imune do hospedeiro, a doença pode comprometer pele, mucosas e vísceras. Em termos gerais, a leishmaniose é classificada em dois tipos: leishmaniose visceral, também conhecida por Kala-azar, e leishmaniose tegumentar, também conhecida por leishmaniose cutânea.

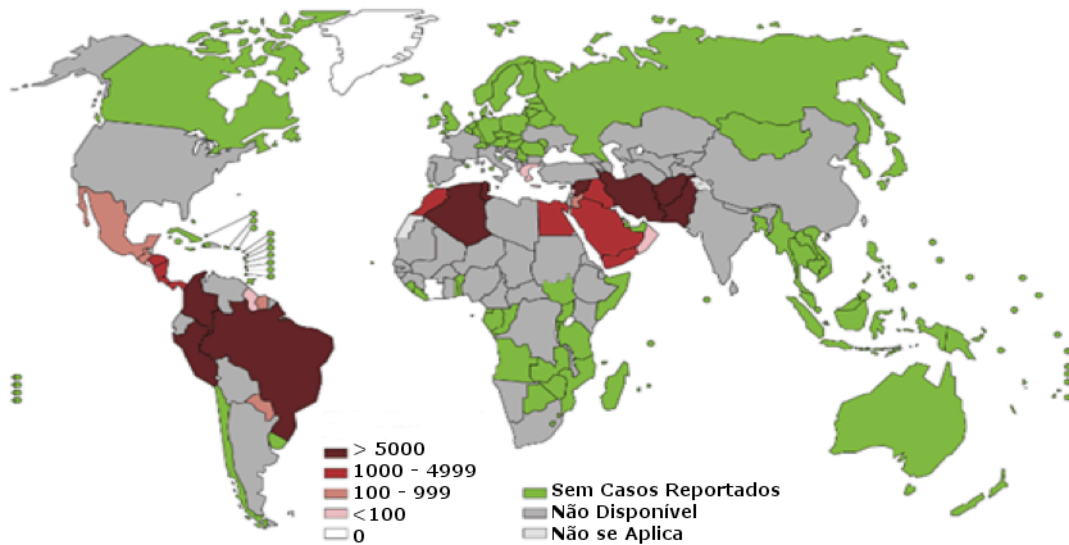
A literatura aponta que os dados de incidência e prevalência da leishmaniose são subestimados em virtude da falta de sistemas de vigilância e da subnotificação dos casos nos países em que elas ocorrem (ALVAR et al., 2012). As Figuras 1 e 2 mostram a incidência da leishmaniose no mundo.

**FIGURA 1. Status da endemicidade da leishmaniose visceral no mundo em 2012**



Fonte: WHO (2014)

**FIGURA 2. Status da endemicidade da leishmaniose tegumentar no mundo em 2012**



Fonte: WHO (2014)

O Brasil, além de ser um dos países com maior incidência e prevalência de leishmaniose, é um dos principais produtores de conhecimento científico sobre essa doença. Com o intuito de participar ativamente do delineamento de estratégias para o desenvolvimento de tecnologias de prevenção ou tratamento, o país tem investido em pesquisas, mas os resultados desse investimento ainda não estão claros quanto a sua efetividade.

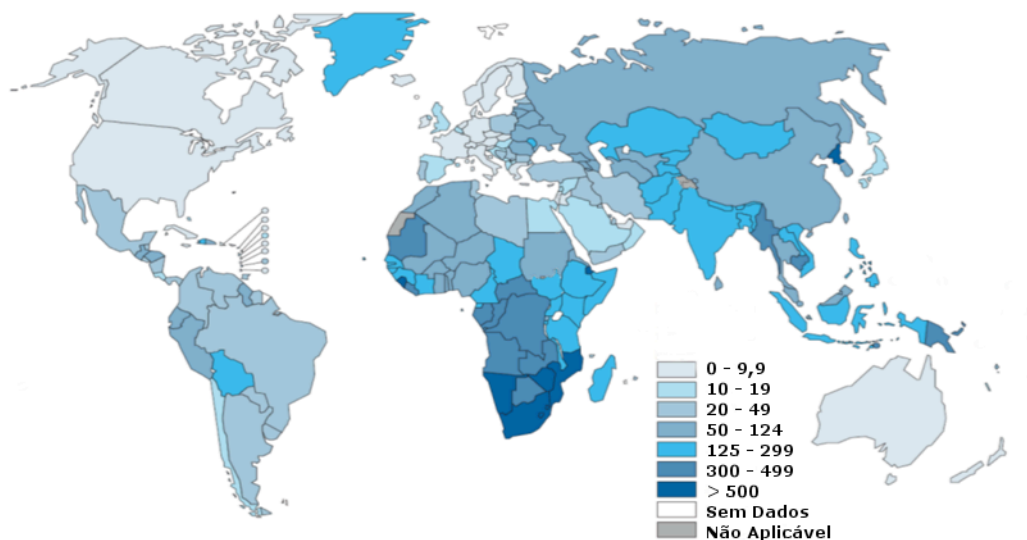
### 1.1.3 Tuberculose

A tuberculose continua sendo um grande problema de saúde global. Ela acomete milhões de pessoas a cada ano e é classificada como a segunda principal causa de morte por doença infecciosa em todo o mundo, antecedida apenas pelo vírus da imunodeficiência humana (HIV) (WHO, 2013). A doença é causada pelo bacilo *Mycobacterium tuberculosis*, que tipicamente afeta os pulmões (tuberculose pulmonar), mas também pode afetar outros órgãos do corpo (tuberculose extrapulmonar). O bacilo é transmitido pelo ar e é capaz de permanecer dormente no corpo humano por anos até que o sistema de imunidade do indivíduo se torna incapaz de afastar a infecção e de destruir o bacilo.

Sem tratamento, as taxas de mortalidade por tuberculose são elevadas. Em estudos sobre a história natural da doença entre os casos de pacientes diagnosticados com TB pulmonar/HIV-negativos, em torno de 70% morreram em dez anos (TIEMERSMA et al., 2011). Em 2012, estima-se que 8,6 milhões de pessoas desenvolveram tuberculose e 1,3 milhão morreu da doença (incluindo 320 mil mortes entre pessoas HIV-positivas) (WHO, 2013).

Um pequeno grupo de países (22) é responsável por 80% de todos os casos da doença no mundo, estando a maioria concentrada em regiões do sudeste da Ásia (29%), na África (27%) e no Pacífico Ocidental (19%). Países como a Índia e a China foram responsáveis por 26% e 12% do total de casos, respectivamente, em 2013. A taxa de incidência de TB varia substancialmente em cada país, com cerca de mil ou mais casos por 100 mil pessoas na África do Sul e na Suazilândia e menos de dez por 100 mil habitantes em partes das Américas, em vários países da Europa Ocidental, no Japão, na Austrália e na Nova Zelândia (Figura 3).

**FIGURA 3. Incidência de novos casos de tuberculose estimados no ano de 2012**



Fonte: WHO (2013)

No Brasil, o número de ocorrências de tuberculose caiu firmemente nas duas últimas décadas, passando de 51,4 casos por 100 mil pessoas-ano em 1990 para 38,2 casos por 100 mil pessoas-ano em 2007 (uma redução de 26%). A mortalidade também caiu nesse período, de 3,6 mortes por 100 mil pessoas-ano para 1,4 mortes por 100 mil pessoas-ano (uma redução de 32%) (BARRETO et al., 2011). Ainda segundo Barreto (2011), “há grandes diferenças regionais na incidência e na mortalidade da tuberculose, com maiores níveis em estados com mais prevalência de infecção por HIV, como o Rio de Janeiro, e em estados com pouco acesso aos serviços de saúde, como aqueles da região amazônica”.

## 1.2 A análise de redes no apoio à avaliação e à gestão do conhecimento

Saracevic (2008) aponta a ocorrência de relações interdisciplinares em processos cujos insumos principais são a informação e o conhecimento, defendendo seu caráter de entrelaçamento, pois certas áreas de conhecimento demandam soluções que servem, exigem ou dependem de outras áreas de conhecimento. Além disso, “novas ferramentas e abordagens são necessárias para facilitar o planejamento de políticas públicas e promover a gestão da inovação nos sistemas de saúde pública dos países” (VASCONCELLOS; MOREL, 2012).

O crescente número de colaborações internacionais encontrados neste trabalho de pesquisa sugere que as parcerias multiorganizacionais são mecanismos promissores para a melhoria da saúde pública. Nesse sentido, a colaboração torna-se um importante mecanismo de promoção de inovações, especialmente para o desenvolvimento de novos produtos ou



terapias voltadas para as necessidades dos sistemas de saúde dos países menos desenvolvidos, que por si só não possuem capacidade tecnológica suficiente para atender à demanda da saúde, haja vista o crescente número de acometimentos de doenças negligenciadas.

Os modelos de Choo (2005) e Nonaka e Takeushi (1995) enfatizam a importância dos laços interpessoais para a criação do conhecimento, mas não destacam o valor das redes sociais existentes no interior das organizações. Para Silva e Ferreira (2007), referência bastante usada no Brasil sobre gestão do conhecimento, há pouco uso da ARS. Ocorre que a ARS é usada para o estudo da troca de recursos entre atores, sendo a informação um dos recursos estudados. Mueller (2007) descreve a ARS como um dos métodos para estudo e pesquisa da Ciência da Informação, e vários trabalhos recentes na área, principalmente na bibliometria, já a vêm utilizando de maneira significativa.

Cross e Parker (2004) descrevem, em uma visão organizacional, algumas aplicações comuns à ARS, dentre elas: análise e suporte de parcerias e alianças entre empresas; avaliação da execução de uma estratégia; melhoria na estratégia de tomada de decisão; integração de redes; promoção da inovação; integração após mudanças significativas e desenvolvimento de comunidades de prática.

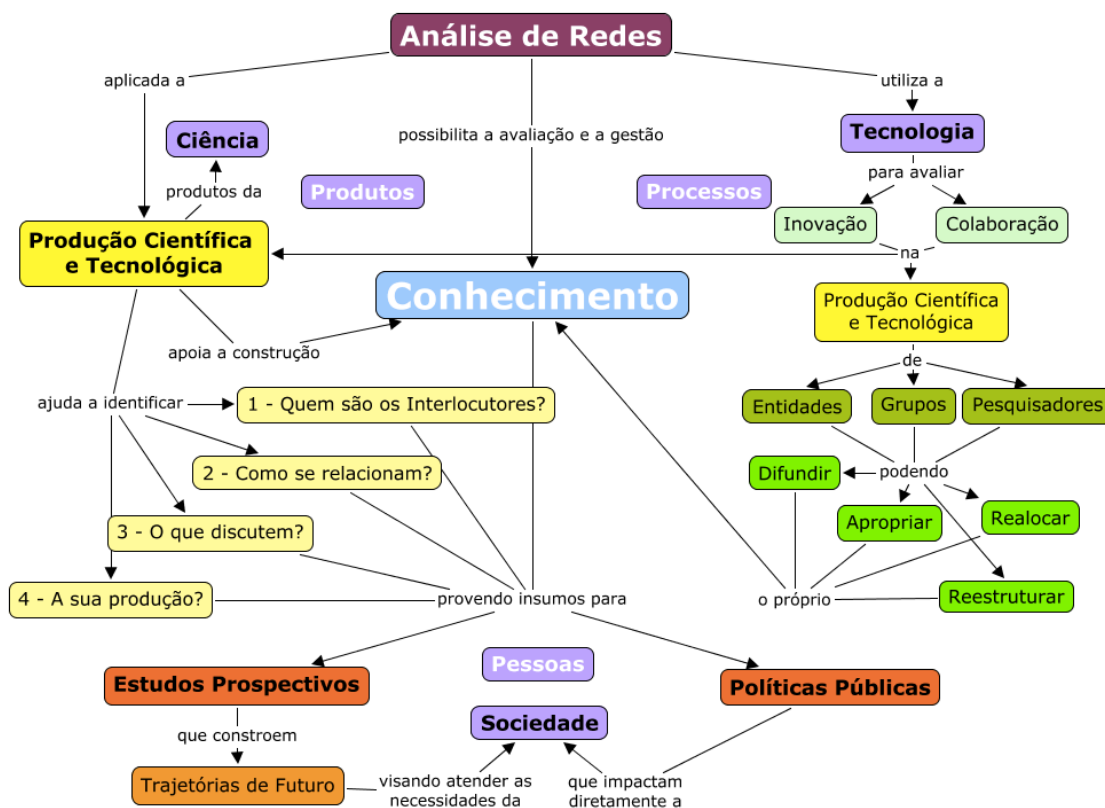
A oportunidade de descobrir novos conhecimentos, a crescente especialização científica, a maior complexidade de equipamentos e novos materiais, bem como a necessidade de combinar diferentes tipos de saberes e competências constituem a base para a colaboração em ciências na área da saúde.

Acredita-se, portanto, que o mapeamento de redes de colaboração e a organização de informações são instrumentos que auxiliam na disponibilização de informações estratégicas de relevância na construção de um futuro alinhado às demandas sociais. A análise aprofundada das redes de cooperação com foco na LN e na TB tem o potencial de identificar seus principais interlocutores, além de mostrar a maneira como se relacionam, o que discutem e o que produzem. A análise de redes permite identificar esses relacionamentos, suas estruturas de poder e as relações centrais e periféricas desempenhadas pelos pesquisadores nas diferentes partes do mundo.

A análise de redes permite ainda acompanhar o caminho da pesquisa, do desenvolvimento científico e tecnológico e o desempenho das competências em saúde existentes, dado que, cada vez mais, a geração de conhecimento e em especial o desenvolvimento tecnológico ocorrem em grandes redes colaborativas, articuladas por toda a cadeia de inovação nos âmbitos nacional e internacional.

A Figura 4 apresenta um mapa conceitual sobre a aplicação dos métodos de análise de redes sobre o conhecimento, tomando como base a tríade produtos, processos e pessoas.

**FIGURA 4. Mapa conceitual da análise de redes sobre a produção científica e tecnológica**



Fonte: elaboração do autor

A tríade produtos, processos e pessoas está relacionada de forma indireta à ciência, à tecnologia e à sociedade, respectivamente. No caso da pesquisa realizada, a ciência tem como foco a produção científica por meio de artigos, a tecnologia envolve as ferramentas utilizadas para avaliar as colaborações e demais informações contidas nesses artigos e a sociedade é afetada diretamente por esses dois outros elementos.

No mapa conceitual apresentado na Figura 4 defende-se que a análise de redes possibilita a avaliação e a gestão do conhecimento. Nesse sentido, o método de análise de redes aplicado à produção científica e tecnológica, que são produtos da ciência, apoia a construção do conhecimento acerca da ciência. Mas de que maneira ocorre esse apoio? Ajudando a identificar *quem* são os interlocutores, *como* se relacionam, *o que* discutem e a *sua* produção científica. Para a realização desse trabalho o analista de redes utiliza a tecnologia para avaliar a inovação e a colaboração na produção científica e tecnológica de entidades, grupos de pesquisa e pesquisadores, podendo com isso difundir, apropriar, realocar

ou reestruturar o próprio conhecimento. Os insumos dessa análise podem ser direcionados a estudos prospectivos que constroem trajetórias de futuro ou a políticas públicas que impactam e visam a atender a sociedade.

### 1.3 Objetivos da pesquisa

A pergunta base da pesquisa aqui relatada é: como ou de que forma os métodos e as técnicas de análise de redes complexas descrevem as estruturas e os processos globais e regionais da pesquisa e sua evolução a partir da investigação das redes de coautoria e outras informações encontradas em publicações científicas no campo das doenças negligenciadas leishmaniose e tuberculose?

#### 1.3.1 Objetivos geral e específicos

Como objetivo geral da pesquisa é proposto: estudar a aplicabilidade de um conjunto de métodos para análise de redes complexas sobre publicações científicas a respeito de doenças negligenciadas. Dessa forma, outros trabalhos que pretendam seguir essa linha de pesquisa podem valer-se do esforço empregado para melhor entender a ciência em suas diferentes áreas de atuação.

Os objetivos específicos têm como foco principal auxiliar na definição e na validação do objetivo geral, que passa pelo entendimento e pela aplicação dos métodos de análise de redes e pelo tratamento de dados. Uma análise dos resultados encontrados nas publicações científicas universo da pesquisa é consequência da validação dos métodos empregados, e em alguns casos pode demandar estudos mais aprofundados para se obter uma visão mais clara do estado da ciência naquele momento. Os objetivos específicos são:

- Definir métodos, processos e ferramentas para a busca e o tratamento dos dados sobre publicações científicas usando a análise de redes complexas.
- Entender os pontos fortes e fracos de diferentes bases científicas na construção de redes complexas.
- Aplicar e avaliar metodologias de análise de redes complexas em diferentes *corpora*.
- Analisar a aplicabilidade prática da análise da estrutura e da dinâmica de redes complexas às pesquisas nas doenças leishmaniose e tuberculose.

- Apresentar a estrutura e a dinâmica das redes de pesquisa das doenças leishmaniose e tuberculose.

### **1.3.2 Pressupostos**

Os pressupostos de pesquisa definidas para esta tese são resultado do estudo prévio de outras áreas de pesquisa e de publicações científicas e do entendimento sobre as possibilidades existentes no estudo da ciência de redes.

A primeira hipótese é que existe um arcabouço de informações relevantes no estudo de publicações científicas que não são exploradas pelos métodos convencionais de análises bibliométricas e cientométricas. Dessa maneira, a análise de redes complexas aplicada a publicações científicas tem muito a contribuir no entendimento da ciência.

A segunda hipótese diz respeito às estruturas existentes nas redes. Apesar de haver uma grande coesão nas redes de publicações científicas, já comprovada em pesquisas anteriores, nas quais a grande maioria dos pesquisadores está conectada entre eles, existe uma estrutura centrada em poucos grupos de pesquisa tradicionais. Essa estrutura tende a dificultar a entrada de novos atores que possam contribuir ou mesmo propor novas formas ou soluções de trabalho.

## 2 Referencial teórico

### 2.1 Ciência de redes

A ciência de redes é uma área acadêmica interdisciplinar e com aplicações práticas em uma diversidade de campos que em alguns casos ainda não foram largamente explorados, como na medicina, por exemplo. A ciência de redes estuda redes complexas, tais como redes de telecomunicações ou tecnológicas, redes biológicas, redes de informação ou conhecimento, além de redes sociais (NEWMAN, 2010). A ciência de redes possui uma série de referências teóricas e metodológicas dos mais diversos campos de pesquisa. Esses campos incluem a sociologia, com o estudo das estruturas sociais, a matemática, com a teoria dos grafos, a computação, com a mineração de dados e a visualização da informação, a Ciência da Informação, com o estudo do fluxo informacional, a física, com a estatística mecânica, e a própria estatística, com a modelagem inferencial. O Conselho Nacional de Pesquisa dos Estados Unidos define a ciência de rede como “o estudo das representações de rede de fenômenos físicos, biológicos e sociais que levam a modelos preditivos desses fenômenos” (NATIONAL RESEARCH COUNCIL, 2005).

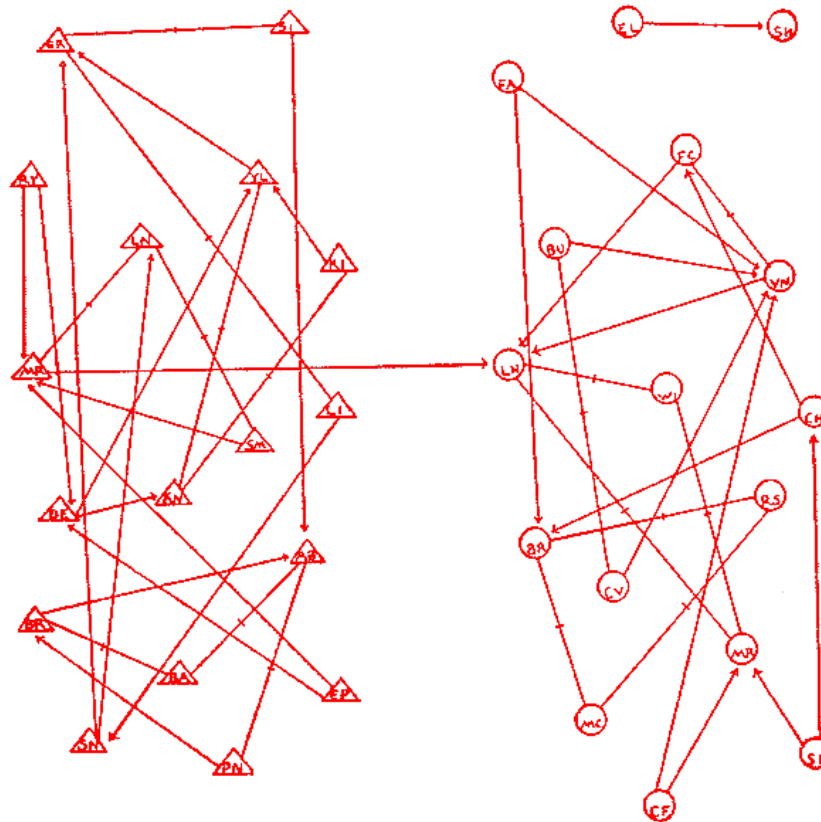
O estudo de redes é aplicado em diversas disciplinas como meio de análise de dados relacionais complexos, diferenciando-se do estudo estatístico clássico, no qual o elemento de estudo e suas características são o principal objeto. A análise de redes, por sua vez, é a aplicação prática dos métodos estudados e descritos na ciência de redes.

O crescimento da área da ciência de redes se deu com o passar dos anos por meio do aprimoramento de diferentes métodos e teorias, que hoje compõem o arcabouço conceitual e metodológico da área. Um dos estudos que deu início à área de pesquisa foi a teoria dos grafos, que teve como seu fundador o matemático Leonhard Euler, que escreveu em 1736 o documento “As sete pontes de Königsberg”. Considerando o desenho das pontes um grafo, Euler demonstrou que era impossível encontrar um caminho para cruzar as sete pontes da cidade sem jamais passar pela mesma ponte duas vezes. Um grafo pode ser descrito como um conjunto de vértices ou nós e um conjunto de linhas entre quaisquer pares dos vértices (NOOY; MRVAR; BATAGELJ, 2011).

Cerca de duzentos anos depois, em 1933, um psicólogo chamado Jacob Moreno apresentou a um grupo de pesquisadores médicos o que ele mais tarde chamaria de sociometria, um método quantitativo para avaliar relações sociais, e com ela o sociograma, que pode ser descrito como a representação gráfica dos relacionamentos sociais que um

indivíduo possui. Os vértices da teoria dos grafos tornam-se com isso indivíduos e as linhas as suas relações interpessoais.

**FIGURA 5. Rede de amizade entre alunos de uma escola de ensino fundamental**



Fonte: MORENO (1953)

Legenda: a figura apresenta círculos e triângulos que denotam meninas e meninos em uma escola primária. Em cada figura há uma legenda com duas letras para designar cada indivíduo, possivelmente relacionados ao nome e ao sobrenome. As linhas de ligação entre as figuras indicam um vínculo de amizade no qual um dos atores indicou o outro ator como seu amigo.

Moreno (1953) afirmou que “antes do advento da sociometria ninguém sabia o que a estrutura interpessoal de um grupo precisamente parecia”. Na primeira representação gráfica do estudo de Moreno, replicado na Figura 5, é apresentada a estrutura social de um grupo de alunos do ensino fundamental em que os meninos se relacionavam com outros meninos e as meninas se relacionavam com outras meninas, com a exceção de um menino, que disse gostar de uma única garota, mas a recíproca não era verdadeira. O sociograma a partir de então encontrou muitas aplicações e tornou-se o campo de análise de redes sociais.

Aproximadamente na mesma época em que Moreno realizava seus estudos sobre a nova ciência, um escritor húngaro de nome Frigyes Karinthy escreveu uma história com o nome “Cadeias” (BARABÁSI, 2003). Na história, o personagem descreve como ele poderia se relacionar com qualquer pessoa no planeta por meio dos amigos dos seus amigos, o que

mais tarde veio a ser conhecido como seis graus de separação. Sem nenhum aparente conhecimento sobre esse conto, um pesquisador de Harvard de nome Stanley Milgram e seus companheiros, em 1967, vieram a desenvolver uma série de estudos que corroboraram a história de Karinthy. Na pesquisa realizada, que tinha como objetivo verificar quantos indivíduos conhecidos são necessários para conectar dois indivíduos selecionados ao acaso, Milgram pede a uma quantidade determinada de indivíduos, normalmente situados em estados do centro dos Estados Unidos, que se prontificaram a participar do experimento, que enviassem uma carta para uma pessoa específica na cidade de Boston, situada no leste dos Estados Unidos. Dentre as diversas variações da pesquisa algo se mantinha como regra: o indivíduo só poderia enviar o pacote para a pessoa-alvo se ele a conhecesse pessoalmente, senão ele deveria enviá-lo apenas para alguém conhecido que ele achasse que teria chance de conhecer a pessoa-alvo. Dentre os diferentes resultados encontrados na pesquisa alguns chamaram a atenção. Poucos foram os pacotes que conseguiram chegar ao seu destino, no entanto alguns dos pacotes que chegaram tinham passado por um número relativamente pequeno de pessoas, que era entre cinco e seis.<sup>1</sup> Além disso, muitos dos pacotes que chegaram ao seu destino acabaram passando por um específico número de pessoas que mais tarde começaram a ser denominados de hubs na análise de redes, pois elas faziam parte de muitos dos caminhos percorridos pelos pacotes.

Voltando um pouco antes do experimento de Stanley Milgram, dois matemáticos, Paul Erdős e Alfred Rényi, escreveram em 1959 um artigo descrevendo um modelo matemático de entrelaçamento de vértices para grafos aleatórios (ERDÖS; RÉNYI, 1959). Esse modelo seguiu como padrão de pesquisa na área de redes até o final dos anos 1990, quando se verificou que a maioria das redes reais não possuía características aleatórias. No entanto, esse modelo é utilizado até hoje na avaliação de medidas de redes na comparação entre grafos complexos e grafos aleatórios, e com isso se descrevem algumas características das redes estudadas. Os dois matemáticos foram reconhecidos em diversas pesquisas, e seus trabalhos são citados até os dias de hoje.

Em 1973, um sociólogo chamado Mark Granovetter escreveu um artigo que veio a tornar-se mais um marco na ciência de redes (GRANOVETTER, 1973) – “A força dos laços fracos”. Nele o autor argumenta que o grau de sobreposição da rede de amizade de dois indivíduos varia de forma direta com a força do relacionamento entre os dois. Isso significa que quanto maior a intensidade de relacionamento entre dois amigos, mais amigos em comum

---

<sup>1</sup> <[http://en.wikipedia.org/wiki/Small-world\\_experiment](http://en.wikipedia.org/wiki/Small-world_experiment)>.

eles terão. Dessa maneira, características sociais como influência, informação, oportunidade de mobilização e organização de comunidades são influenciadas diretamente pelo grau de amizade ou pela força da relação entre os indivíduos.

Nos últimos 15 anos, muitos dos esforços de pesquisa da ciência de redes concentraram-se em descrever matematicamente diferentes topologias de rede. Na pesquisa realizada foram estudados fenômenos ligados principalmente a topologias Livre de Escala, Mundo Pequeno e Centro Periferia.

Uma rede pode ser considerada Livre de Escala se sua distribuição quanto ao número de relações entre os atores seguir uma função matemática chamada de Lei de Potência (ALBERT; BARABÁSI, 2002). O termo Livre de Escala é proveniente da propriedade de funções logarítmicas em que o aumento ou escalonamento no fator da função não influi na sua característica logarítmica, sendo, portanto, Livre de Escala. Esse tipo de rede apresenta um padrão estrutural formado por poucos nós com muitas conexões, chamados de hubs, e muitos nós com poucas conexões.

O modelo Mundo Pequeno (WATTS; STROGATZ, 1998) demonstrou que com a adição de apenas um pequeno número de ligações uma rede simples, em que o diâmetro é proporcional ao tamanho da rede, pode ser transformada em um “mundo pequeno”. Em uma rede Mundo Pequeno, o número médio de linhas entre quaisquer dois vértices é muito baixo e, matematicamente, deve se manter com pequena variação ao ponto que o número de vértices cresça em proporção logarítmica com o tamanho da rede. Uma rede possui características de rede Mundo Pequeno por analogia com o fenômeno do mundo pequeno, popularmente conhecido como seis graus de separação, que começou a ser estudado por Stanley Milgram.

No que diz respeito às redes Centro Periferia, os primeiros estudos datam de 2000 (BORGATTI; EVERETT, 2000). Nele os autores argumentam que as redes com essa topologia possuem um centro ou núcleo (*core*) bem conectado entre si, mas possuindo também fortes conexões com os vértices periféricos. Estudos mais recentes (ROMBACH et al., 2012) apontam que essas redes na verdade possuem vários núcleos que podem conectar-se com grande parte da rede tanto dentro dos seus grupos, ou *cores*, como fora deles.

Dando prosseguimento ao estudo de redes ou ciência de redes, verificou-se a existência das chamadas redes complexas, que possuíam características bem diferentes das redes aleatórias definidas por Paul Erdős e Alfred Rényi e mais próximas das redes definidas por Duncan J. Watts e Steven Strogatz e Albert-László Barabási e Réka Albert. No contexto da teoria de redes, uma rede é dita complexa quando possui características topológicas não triviais, que não ocorrem em redes simples, como redes com entrelaçamentos aleatórios, mas



que muitas vezes ocorrem em redes reais, que podem conter no seu interior muitas outras sub-redes (DOROGOVTSEV; MENDES, 2001; KIM; WILHELM, 2008).

A maioria das redes sociais, biológicas e tecnológicas possui características topológicas não triviais, com padrões de conexão entre seus elementos que não são, na sua grande maioria, distribuições normais nem aleatórias. Tais características incluem uma distribuição de grau com uma cauda dita pesada, ou com muitos vértices, um alto coeficiente de clusterização ou agregação, um entrelaçamento preferencial entre vértices, estruturas de comunidade e estruturas hierárquicas. Essas características não são excludentes, podendo uma rede apresentar uma ou mais dessas características, e estas podem mudar com o passar do tempo (GAY, 2012).

## **2.2 Redes de coautoria**

As redes de coautoria são um poderoso instrumento para a análise de colaborações e parcerias científicas e tecnológicas, proporcionando uma visão dos padrões de cooperação entre indivíduos e organizações (MELIN; PERSSON, 1996; NEWMAN, 2004; GLÄNZEL; LANGE, 2002). A coautoria de um documento representa uma relação oficial do envolvimento de dois ou mais autores ou instituições (GLÄNZEL; SCHUBERT, 2004), e apesar do debate sobre seu significado e interpretação (KATZ; MARTIN, 1997; LAUDEL, 2002) a análise de coautoria ainda é amplamente utilizada para entender e avaliar os padrões de colaboração científica.

A colaboração científica pode, então, ser definida como a interação que ocorre em um contexto social entre dois ou mais cientistas que facilita o compartilhamento de significado e a realização de tarefas com relação a um objetivo mutuamente compartilhado (SONNENWALD, 2007).

Em redes de coautoria, a colaboração é representada de tal maneira que os vértices representam os autores, os departamentos ou as instituições, e dois ou mais autores estão conectados se eles compartilham a autoria de uma publicação (NEWMAN, 2004). Em uma perspectiva organizacional, a cooperação interinstitucional pode ser definida em termos de publicações que tenham autoria compartilhada por indivíduos de duas ou mais instituições (NAGPAUL, 2002).

Estudos empíricos apontam também que tais redes de colaboração tendem a apresentar uma distribuição de ligações extremamente desigual e hierárquica (BARABÁSI et al., 2002;

NEWMAN, 2004; GOYAL, LEIJ, GONZÁLEZ, 2006), com grande concentração da produção científica em alguns poucos autores.

Os estudos sobre redes de coautoria na área de saúde realizados por pesquisadores brasileiros são relativamente recentes e, geralmente, se utilizam de áreas específicas de pesquisa ou de doenças relevantes para programas governamentais ou para a saúde pública (COSTA; MACEDO, 2013; VASCONCELLOS; MOREL, 2012; MOREL et al., 2009; SANTOS, 2012; MOURA; CAREGNATO, 2011).

## **2.3 Análise da colaboração científica por meio de redes de coautoria**

### **2.3.1 Dados bibliométricos**

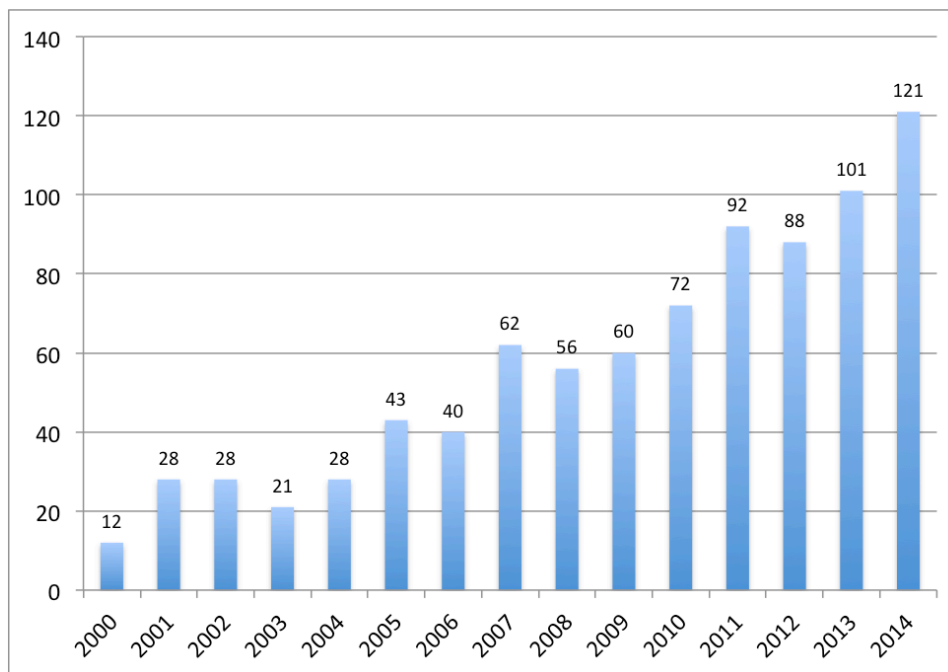
Uma das aplicações práticas da ciência de redes é o apoio na revisão bibliográfica por meio da análise de colaborações científicas. Neste item busca-se fazer a aplicação do método estudado de análise de redes complexas ao universo de pesquisa em redes de coautoria com o intuito de entender se essa área está concentrada em alguns poucos países ou indivíduos e como esses indivíduos se relacionam.

O primeiro passo é a descrição bibliométrica do universo pesquisado, ou a “definição da rede”, como é chamada na metodologia que está sendo utilizada. Para este estudo foi utilizada a base Web of Science, pois além de se poder fazer a análise das colaborações científicas se pode também realizar uma análise de citações e com isso definir os artigos mais relevantes.

A pesquisa foi realizada em dezembro de 2014 e usou como termos de busca em tópicos *co-author\**, *scientific collaboration* e *collaboration network*. O caractere de truncagem “\*” foi utilizado para abarcar todas as formas da palavra coautoria em inglês. Foi utilizado também o operador booleano “OR” entre os termos. A pesquisa resultou em 851 artigos entre o período de 2000 a 2014, com foco nas categorias *Information Science and Library Science*, *Multidisciplinary Science*, *Multidisciplinary Physics*, *Mathematical Physics e Social Sciences Interdisciplinary*. As categorias são uma classificação feita pela Web of Science que está relacionada diretamente às revistas nas quais os artigos são publicados. A escolha por essas categorias foi feita com base em uma análise qualitativa dos artigos, em que os trabalhos relacionados à pesquisa sobre colaboração científica nos moldes feitos para este trabalho de pesquisa se concentravam. Muitos artigos das categorias de ciência da computação foram encontrados, mas estes tinham como foco principal o estudo de métodos e algoritmos para a ciência de redes, e não o estudo de redes de colaboração. A Figura 6 mostra

o crescimento do número de publicações ao longo dos últimos anos na área de pesquisa em colaboração científica.

**FIGURA 6. Número de publicações na área de colaboração científica por ano**

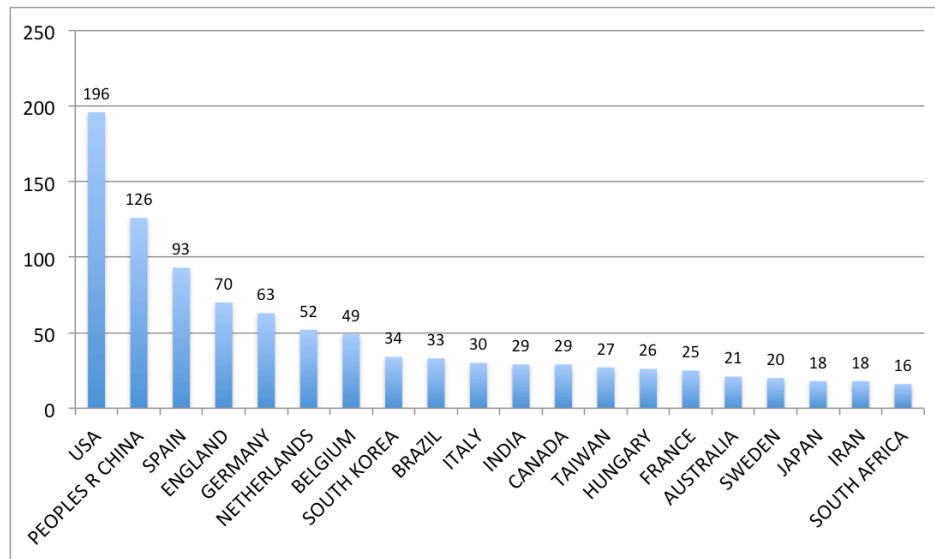


Fonte: elaboração do autor

O que se pode perceber na Figura 6 é um crescimento constante das publicações nos últimos anos. No entanto, esse fator deve ser levado em consideração quanto ao número geral das publicações científicas, que também cresceu em todas as áreas da ciência. No caso das pesquisas em colaboração científica o aumento foi de dez vezes o número de publicações comparando o ano de 2000 com o de 2014. Essa evolução é elevada e relacionada diretamente à ampliação do interesse na área de ciência de redes, que começou seu grande *boom* a partir de trabalhos como os de Watts e Strogatz (1998), Barabási e Albert (1999) e Newman (2001).

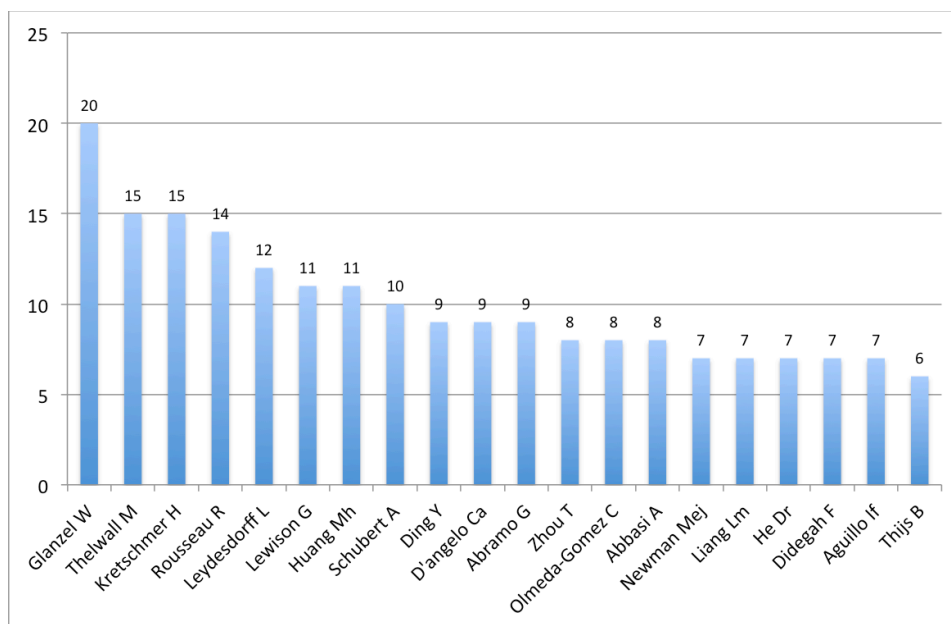
Outro fator que pode ter corroborado esse crescimento é a facilidade de acesso a novas tecnologias que otimizaram o tratamento de grandes bases de dados.

A Figura 7 mostra o número de publicações por país; os países com maior número de publicações são os EUA e a China, seguidos de países europeus, com Coreia do Sul em oitavo e Brasil em nono.

**FIGURA 7. Número de publicações por país**

Fonte: elaboração do autor

A Figura 8 mostra os vinte pesquisadores que mais publicaram nesse período.

**FIGURA 8. Número de publicações por autor**

Fonte: elaboração do autor

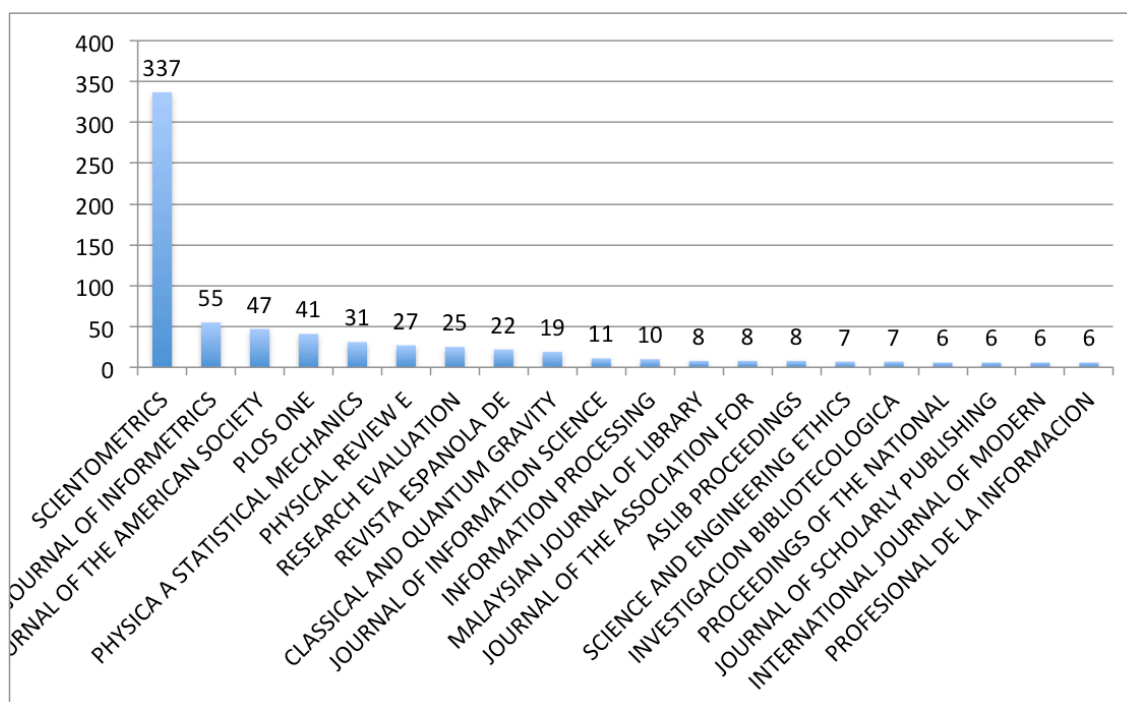
O pesquisador com maior número de publicações na área é Glanzel, com vinte publicações, sendo algumas dessas com grande número de citações. Apesar de a China ser o segundo país com maior número de publicações, seus pesquisadores não aparecem com tanto destaque quanto na figura anterior, tendo quatro pesquisadores na lista dos vinte pesquisadores com maior número de publicações. Da lista de publicações por autor, dois autores destacam-se pela qualidade do trabalho quanto ao número de citações: Mark J.

Newman e Loet Leydesdorff, sendo o primeiro físico e o segundo da área de Ciência da Informação. Esses pesquisadores, além de Glanzel, têm sido uma referência na área de redes de colaboração.

É importante ressaltar que existem pesquisadores renomados na ciência de redes, como, por exemplo, Barabási, que publicou sobre colaboração científica (BARABÁSI et al., 2002), mas o foco de seu trabalho tem sido outras áreas, incluindo a aplicação na saúde (BARABÁSI; OLTVAI, 2004; GOH et al., 2007), e por isso ele não aparece entre os que mais publicaram. Esse mesmo padrão ocorre no caso de outros pesquisadores com nível de citação muito acima da média, mas baixo número de publicações.

Por último é feita uma análise do número de publicações por revista com ênfase nas vinte revistas com maior número (Figura 9).

**FIGURA 9. Número de publicações em coautoria por ano**



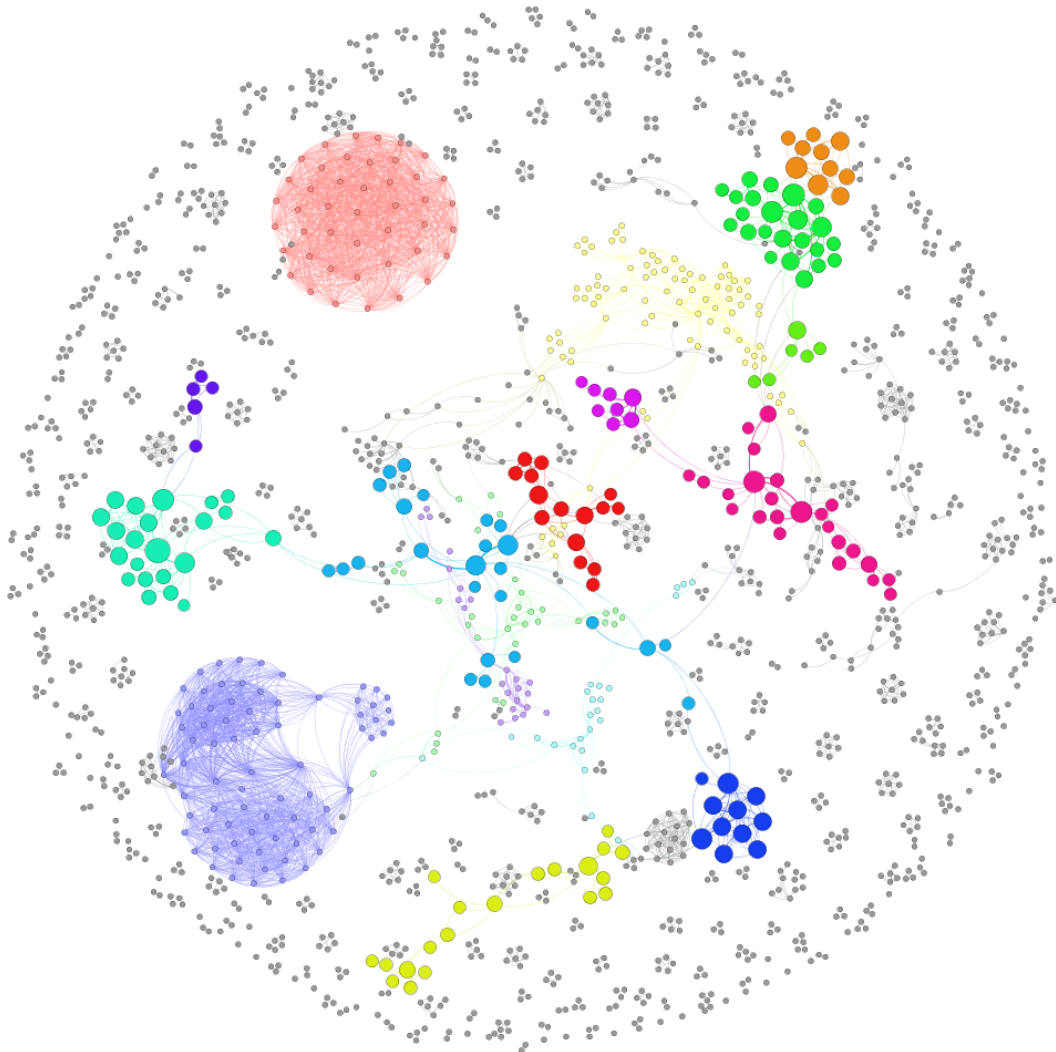
Fonte: elaboração do autor

Das revistas com maior número de publicações sobre o tema temos três da área de Ciência da Informação e da Biblioteconomia, sendo a quarta da área da Saúde e a quinta da Física. Isso demonstra a diversidade de áreas em que o assunto é abordado e trabalhado. No entanto, se formos avaliar o número de publicações, a revista *Scientometrics* possui quase o total da soma de todas as outras 19 revistas, sendo, portanto, uma referência na área.

### 2.3.2 Rede de colaboração

Para a análise da rede de pesquisa em colaboração científica objetivou-se focar na inspeção visual para apoiar a definição das características estruturais das redes e melhor entender os resultados.

**FIGURA 10. Rede de colaboração em pesquisa sobre colaboração científica**



Fonte: elaboração do autor

Legenda: rede de colaboração científica em que os vértices representam os pesquisadores e as linhas, as relações entre elas. Os vértices em cores fazem parte de grupos com número elevado de pesquisadores, enquanto os vértices em cinza fazem parte de grupos menores e não conectados aos demais pesquisadores. Ao centro da figura os vértices com tamanho maior que os demais fazem parte de um mesmo grupo, apesar de diferentes cores nos vértices.

O grafo ou sociograma apresentado na Figura 10 representa a rede de colaboração científica sobre colaborações científicas. No sociograma foram colocados em destaque por meio de cores os grupos nos quais vários pesquisadores estão conectados. O maior grupo de

pesquisadores conectados é chamado de componente gigante e representa nesta rede apenas 9% dos pesquisadores, resultado diferente do que foi encontrado nas redes de pesquisa em leishmaniose e tuberculose, foco de pesquisa desta tese, em que a maior parte dos pesquisadores está conectada em um grande grupo.

Essa estrutura bastante desconectada entre os elementos da rede de pesquisa pode ser explicada pela característica transdisciplinar da área de pesquisa. A análise de redes é na verdade um conjunto de métodos derivados da ciência de redes e utilizados por pesquisadores das mais diversas áreas. Mesmo falando do mesmo assunto, que é neste caso a pesquisa sobre colaboração científica, os participantes desse grupo pertencem a diferentes áreas e têm pouca interação, como pode ser percebido na Figura 10.

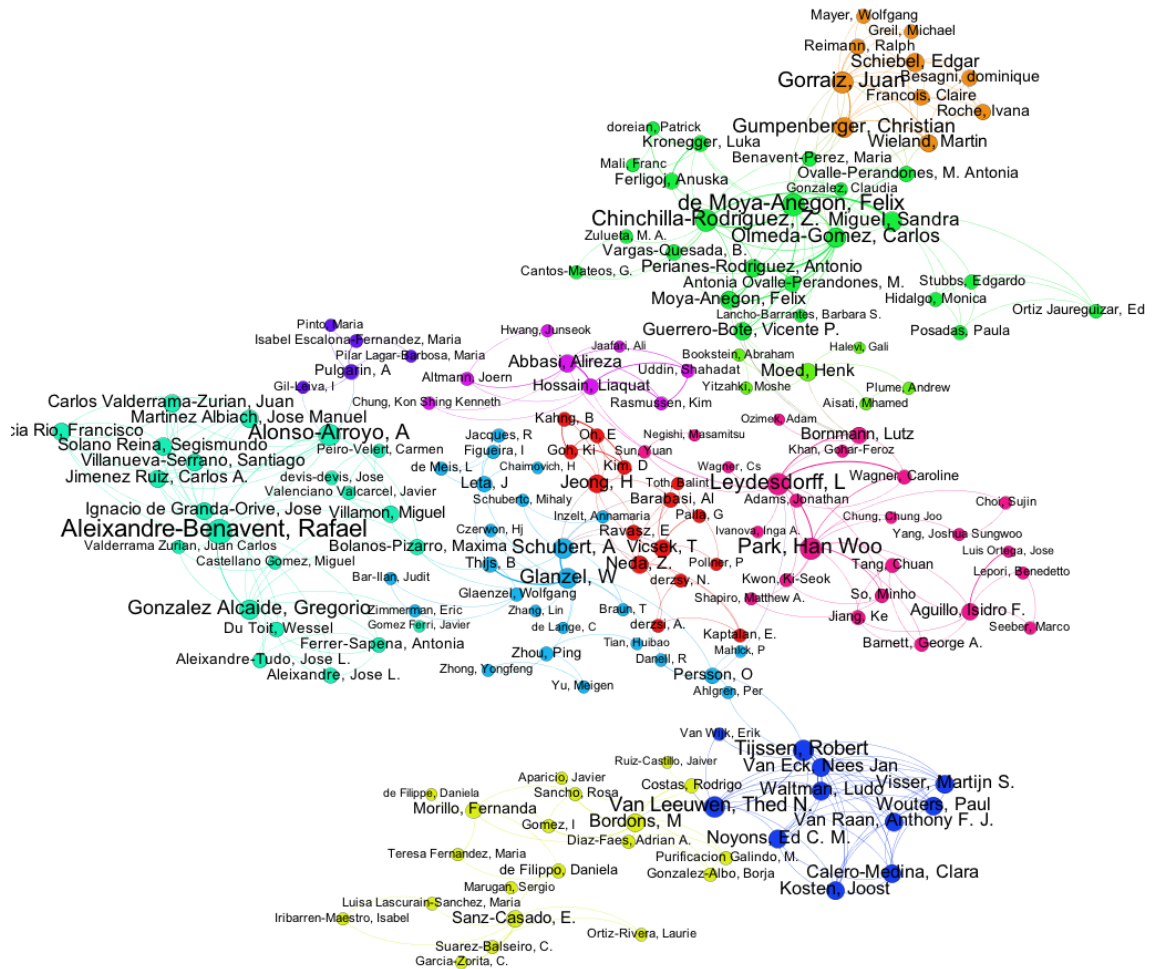
Dois grandes grupos bastante coesos são formados na parte superior e inferior esquerda, respectivamente. Esse tipo de formação se dá por haver artigos com um número muito elevado de pesquisadores; o grupo de cor avermelhada na parte superior tem apenas um artigo no qual todos colaboraram, e o de cor azulada tem alguns poucos artigos.

O grupo de maior interesse ou foco de pesquisa é o que está ao centro, pois apesar de estarem todos conectados ainda formam subgrupos dentro dessa própria rede. Os subgrupos são diferenciados pelas cores e foram definidos como sendo deste grupo por meio de um algoritmo de clusterização descrito no capítulo sobre metodologia.

O que se pode perceber é que a área tem crescido bastante nos últimos anos, mas ainda não se consolidou como uma área específica de pesquisa na qual seus colaboradores trabalham em conjunto de alguma maneira. Áreas de pesquisa já estabelecidas apresentam características de colaboração diferentes do que foi apresentado, lembrando que o recorte utilizado foi bastante específico para o universo que está sendo estudado neste trabalho. Por isso existe um número grande e crescente de pessoas na área, mas que não necessariamente colaboram entre si. Com o passar dos anos a tendência é que aumente essa colaboração por meio de participação em congressos, trabalhos entre diferentes regiões, além de migração de pesquisadores para diferentes regiões, estabelecendo novos relacionamentos.

A Figura 11 apresenta com maiores detalhes os pesquisadores que fazem parte do componente gigante ou do componente com o maior número de pesquisadores que colaboram entre si. Nesta figura procurou-se dar ênfase aos subgrupos, e com isso tentar entender onde estão os principais pesquisadores e como eles influenciam a formação do referido grupo.

FIGURA 11. Número de publicações em coautoria por ano



Fonte: elaboração do autor

Legenda: pesquisadores do componente gigante destacados por cores de acordo com o grupo definido por suas relações. O tamanho de cada nó varia ligeiramente de acordo com o número de colaborações que este possui. As linhas caracterizam as colaborações entre os pesquisadores.

Ao centro da Figura 11 podem-se perceber pesquisadores como Glanzel no grupo azul e Leydesdorf no grupo rosa, mais à direita da figura. Esses dois pesquisadores são apontados como sendo de alta relevância na área devido ao número de artigos bem como artigos com alto número de citações.

A análise da rede de pesquisadores que publicam sobre colaboração científica não foi detalhada neste momento, pois se buscou apenas demonstrar a capacidade de avaliação de resultados bibliométricos associados aos métodos de análise de redes. A interpretação das colaborações entre os diferentes autores e suas redes carece de uma avaliação mais qualitativa, com a busca do entendimento sobre os temas mais específicos além de suas regiões ou grupos de pesquisa.



## 2.4 Colaboração Científica

Com o propósito de se aprofundar em alguns dos pesquisadores apresentados nos grafos de colaboração foi realizado um estudo mais detalhado sobre os autores mais citados na área. A maioria dos artigos analisados foram publicados na revista *Scientometrics*. Esses artigos são apresentados a seguir em duas grandes áreas: Estudos empíricos, nos quais foram avaliados os artigos com aplicação prática quanto à análise de áreas da ciência; e Metodologias, na quais os estudos focavam na avaliação de métodos sobre a área de colaboração científica. Além disso destacamos os trabalhos de pesquisadores que publicaram em revistas brasileiras e autores da área da saúde que utilizaram os métodos de colaboração científica.

### 2.4.1 Estudos empíricos

Como foi demonstrado nos grafos de rede sobre colaboração científica, um dos autores da Ciência da Informação de maior centralidade é Loet Leydesdorf. O autor faz uma série de estudos em diferentes bases científicas como a Web of Science e a Medline, e com isso disponibiliza um conjunto de métodos que são muito utilizados e estudados. Dentre os estudos mais citados e relacionados ao tema colaboração científica temos o estudo sobre colaboração internacional e a criação de poder nas relações científicas (LEYDESDORFF; WAGNER, 2009), a colaboração científica na área das ciências sociais (LEYDESDORFF; PARK; WAGNER 2014) e a avaliação da dinâmica das redes sociais na área de colaboração científica (LEYDESDORFF et al. 2008).

Sobre o tema de colaboração internacional, Glänzel, Leta e Thijs (2006) investigaram o panorama internacional de coautoria entre países e apresentaram uma visão geral do desempenho de investigação científica do Brasil com base em dados bibliométricos. Enquanto na América Latina o Brasil lidera o ranking, a China está aumentando sua colaboração internacional do ponto de vista científico, aproximando-se do desempenho da França e desafiando o Reino Unido e a Alemanha. Durante as últimas décadas, o Brasil tem consolidado um sistema de ciência e tecnologia, bem como uma grande comunidade científica, o que permitiu que se colocasse à frente dos países vizinhos.

Além da colaboração científica internacional, outros estudos de pesquisadores chineses sobre o tema também são comuns na literatura. O estudo de Liang et al. (2001) analisou a colaboração de pesquisadores chineses na computação. Liang realizou estudos para compreender a estrutura da colaboração dos cientistas chineses de ciência da computação por

faixa etária. Concluíram que a maioria dos colaboradores são homens jovens que se relacionam com um grupo menor de colaboradores mais velhos, supondo-se que sejam estudantes e seus orientadores.

Também na china, Wang et al. (2005) analisaram os padrões de coautoria em ciência e tecnologia no seu país, considerando os seguintes tipos de artigo: de autoria única; em coautoria na mesma instituição; em coautoria de autores em diferentes instituições localizadas na mesma região; em coautoria de autores em diferentes regiões da China; e em coautoria de autores de diferentes países. As diferentes regiões têm padrões de colaboração diferentes, correspondentes aos níveis de desenvolvimento econômico, tecnológico e científico, com regiões de maior desenvolvimento tendendo a colaborar mais em um relacionamento direto com o número de publicações.

Outros países como Índia e Itália tiveram suas redes de pesquisadores analisadas para avaliar a colaboração científica (Basu e Kumar 2000; Cainelle et al. 2014). No caso de Cainelle et al. (2014) eles analisam a estrutura de coautoria utilizando a análise de redes sociais para identificar o efeito da colaboração sobre a produção científica dos economistas italianos. Investigaram-se os atributos de vértice como idade, gênero, grau de escolaridade, localização geográfica, entre outros; os atributos relacionais como propensão a cooperar e estabilidade dos padrões de cooperação e as características posicionais como índices de centralidade *betweenness/closeness* e coeficientes de agrupamento em redes de coautoria. Segundo esses autores, o aumento crescente da relevância da autoria múltipla em quase todos os campos da ciência está atrelado à pressão sobre os acadêmicos para publicar artigos. Assim, a chamada política de “publicar ou perecer”, difundida além das fronteiras nacionais e disciplinares, faz um apelo ao aumento de publicações em coautoria entre os pesquisadores.

No caso de Basu e Kumar (2000) eles entendem que a cooperação internacional pode aumentar a visibilidade das publicações científicas de um país. As publicações internacionais em coautoria podem ser consideradas um indicador de cooperação científica entre países. A maior parte da cooperação científica indiana é com as nações desenvolvidas do Ocidente e com o Japão. O estudo apresentado tem como foco a Índia.

Além de estudos relacionados a pesquisadores por região os temas também são um recorte comum para a análise. No caso da Ciência da Informação o estudo de Yan, Ding e Zhu (2009) utiliza essa metodologia e na área de cientometria Hou, Kretschmer e Liu (2007) apresentam dados interessantes nesse recorte.

Yan, Ding e Zhu (2009) estudaram o padrão de colaboração e a estrutura da rede de coautoria de Biblioteconomia e Ciência da Informação (LIS) na China. Os autores analisaram

as características globais de uma rede e também as características locais de elementos individuais em uma rede e discutem a correlação entre métricas de centralidade e contagem de citações. Os autores utilizaram o método de ARS com métricas de centralidade e métricas globais, tais como componentes, distância, clusters, distribuição de grau, tolerância a erros e ataque de rede.

Hou, Kretschmer e Liu (2007) utilizaram a análise de redes sociais para verificar ocorrência, cluster e frequência de palavras na rede de colaboração científica em cientometria. Os autores concluem que a principal área colaborativa em cientometria é a ciência e atividades científicas, enquanto a tecnologia mostrou-se menos atraente. As áreas favoritas foram a física e a química, e os principais métodos são cientometria básica e métodos bibliométricos, principalmente a análise de citações.

#### **2.4.2 Metodologias**

Como já foi apresentado neste capítulo, diversos foram os autores que apresentaram estudos que serviram de base para a ciência de redes. Dentre estes autores alguns deles apresentaram estudos sobre a colaboração científica (Newman 2001; Barabasi et al. 2002). No estudo de Barabási et al. (2002) os pesquisadores fazem a aplicação do modelo por ele definido como redes Livres de Escala a duas áreas da ciência – matemática e neurociência – por um período de oito anos. Os autores comprovam que a rede de coautoria segue a mesma lei de potência vista em outras redes, em que um número pequeno de autores possui o maior número de colaborações da rede, enquanto a grande maioria possui um número pequeno de colaborações. No estudo é demonstrada a evolução da rede de colaboração, com a colaboração preferencial para os autores com maior número de publicações e maior grau de colaboração.

O outro pesquisador, Mark E. J. Newman, é um dos principais pesquisadores de ciência de redes que publica na área de colaboração científica, levando-se em consideração o número de publicações e citações. Em seu estudo (Newman 2001) sobre a estrutura das redes de colaboração científica Newman faz uso da técnica de análise de redes aplicada a três áreas da ciência e se consolida como uma referência para os estudos na área, sendo o artigo mais citado sobre o assunto. O estudo foca nas publicações nas áreas biomédica, da física e da ciência da computação, com isso o autor consegue fazer uma avaliação entre as diferentes áreas e os padrões de colaboração entre os pesquisadores.

A análise da colaboração científica como método tem se solidificado cada vez mais nos últimos anos. Melin e Persson (1996) entendem que a análise bibliométrica de coautoria de artigos científicos é uma abordagem necessária para examinar e acompanhar os indicadores estatísticos de colaboração sensíveis para revelar a estrutura e as mudanças de redes colaborativas, sendo também usada para analisar vários aspectos da colaboração. Kretschmer (2004) analisa a posição dos autores altamente produtivos em uma rede de colaboração. O autor argumenta que o método mais comum para estudar a cooperação é a análise de redes de coautoria. Já Lundberg et al. (2006) entendem que analisar publicações tem se tornado a forma padrão de mensurar colaborações de pesquisa. Os autores alertam para a necessidade do uso devido e criterioso de métodos já reconhecidos para esse tipo de análise, caso contrário existe a possibilidade de erros de interpretação do desenvolvimento de uma colaboração, fornecendo dados incorretos para a tomada de decisões.

Kretschmer (2004) acrescenta que devem ser realizadas investigações sobre mudanças nos clusters, pois o crescimento repentino em pequenos clusters pode estar ligado ao desenvolvimento de um novo campo de pesquisa. A relação de distribuição dos autores altamente produtivos entre os clusters na rede de coautoria está relacionada ao tamanho destes. Autores altamente produtivos têm, em média, baixas distâncias geodésicas e, portanto, caminhos mais curtos em relação a todos os outros autores de uma especialidade se comparados aos autores com baixa produtividade. No entanto, a possibilidade de influenciar cientistas altamente produtivos fica distribuída entre os outros pesquisadores no desenvolvimento de uma especialidade.

Melin e Persson (1996) questionam até que ponto a coautoria reflete dados de colaboração e argumentam que há necessidade de estudos de validação que esclareçam a relação entre análise bibliométrica de coautorias e pesquisa de colaboração. Até se saber exatamente o que as coautorias representam, que tipo de colaboração essas análises não capturam e que proporções os diferentes tipos de colaboração podem atingir, o assunto deve ser tratado com reservas. Não há dúvida de que são necessárias habilidades adequadas e o escrutínio de limites de erros potenciais ao se utilizar coautorias como um indicador de colaboração em pesquisa.

Quanto ao uso dos métodos de análise de redes na colaboração científica internacional Glänzel e De Lange (2002), ao analisarem em que medida o grau de multinacionalidade influencia o impacto de publicações nacionais, concluíram que do ponto de vista do impacto de citação os países, em geral, são beneficiados pela participação em projetos multinacionais, especialmente em ciências da vida. Beaver (2001) acrescenta que a globalização e a

internacionalização caminham para a formação de uma comunidade global na qual a ênfase em cooperação e em vida em grupo se torne cada vez mais comum em detrimento da individualidade. Desse modo, é provável que a crescente globalização aumente a diversidade geográfica de colaboradores, sejam eles indivíduos, laboratórios ou institutos. O autor salienta as vantagens da cooperação científica: eficiência; velocidade; amplitude; sinergia; risco reduzido; flexibilidade; exatidão; feedback; difusão; reconhecimento; e visibilidade. Isso porque, entre outras vantagens, a multiplicidade de pontos de vista energiza e estimula os participantes; podem existir vários projetos em andamento simultaneamente, aumentando as chances de produtividade do grupo; os erros são detectados mais rapidamente quando vários indivíduos diferentes com diferentes perspectivas discutem ou argumentam acerca dos dados e/ou da teoria. As desvantagens são: invisibilidade dos indivíduos; perda de contato com a pesquisa direta; e prejuízo do ethos da pesquisa em razão de sua privatização.

Outros estudos de Glänzel são apresentados defendendo a necessidade de métodos adequados na análise das colaborações internacionais (Glänzel 2001; Glänzel e Schubert 2001). O autor entende que a colaboração científica internacional é um fenômeno complexo e heterogêneo que não pode ser suficientemente caracterizado apenas por indicadores bibliométricos. Isso porque as relações internacionais de coautoria representam uma grande variedade de quadros e motivações, estendendo-se desde programas de cooperação bilaterais ou multinacionais para programas individuais entre cientistas. A complexidade e a heterogeneidade da colaboração científica internacional às vezes contradizem as noções usuais do impacto da colaboração internacional sobre o desempenho nacional. No entanto, os métodos bibliométricos possuem recursos para identificar, de maneira profunda, as características nacionais nas relações internacionais de coautoria. Além disso é apresentado em seus estudos que a coautoria internacional resulta em publicações com maior impacto de citação do que simplesmente artigos nacionais. Embora a coautoria internacional produza um impacto de citação além da média bruta, não existe uma correlação entre a força de uma relação de coautoria e a relativa eminência de citação dos resultados publicados.

Além do uso do método em si e das análises acerca da colaboração internacional os estudos dessa natureza trazem outros aspectos que podem ser analisados na comunicação científica (Persson, Glänzel e Danell 2004). Nesse trabalho os autores analisam diversos aspectos em comunicação científica: se seu número tem aumentado; se os padrões documentados de comunicação científica e colaboração mudaram nas últimas duas décadas; se tais tendências têm características inflacionadas; e o papel de colaboração científica nesse contexto. A análise das tendências em relações de coautoria, de atividade de publicação,

número de referências e frequência de citação em nível de artigos individuais pressupõe que a densidade das redes de co-publicações e citações tem aumentado em todos os campos durante as duas últimas décadas. Os autores argumentam que a normalização adequada das medidas padrão e o uso de indicadores relativos são essenciais para a análise de tendências em estudos em médio ou em longo prazo para garantir a validade das conclusões dos resultados bibliométricos.

Outros métodos também foram sendo desenvolvidos e apresentados com o uso da análise de redes. Mählck e Persson (2000) mapearam as redes de autores de departamentos acadêmicos e introduziram a ideia de mapas sociobibliométricos, que podem ser usados para fazer interpretações sociais de redes bibliométricas. Os autores entendem que o uso combinado de análise estatística de ligações e técnicas de visualização ajuda a aumentar a compreensão de estrutura das redes. No entanto, para se compreender mais profundamente os processos de formação de redes e para saber qual o efeito desses processos nas pesquisas individuais é necessária uma análise qualitativa. As relações de citação entre grupos de pesquisa complementam laços de colaboração, e apenas assim a integração das redes internas aumenta.

No que diz respeito a estrutura da redes científicas a pesquisa de Seglen e Aksnes (2000) analisa a relação entre o tamanho do grupo de pesquisa e a produtividade científica em um ambiente de pesquisa altamente cooperativo, o da ciência biomédica contemporânea em microbiologia norueguesa. Para os autores, a ideia de grupos de pesquisa como entidades dinâmicas, constituídos por um núcleo central de estreitas colaborações entre cientistas cercado por uma mutante e difusa periferia de cooperadores transitórios, parece corresponder muito bem à realidade da ciência. A coautoria é um indicador amplamente utilizado no estudo da colaboração de pesquisa e grupos de pesquisa e se constitui em uma interação científica altamente funcional, devendo fornecer uma base bem fundamentada para a definição de grupos de pesquisa.

### **2.4.3 Redes de coautoria no Brasil**

Como se apresenta a análise de redes em Ciência da Informação no Brasil? Foram selecionadas para esta análise três revistas nacionais da área de Ciência da Informação: *Perspectivas em Ciência da Informação*, editada pela Escola de Ciência da Informação da Universidade Federal de Minas Gerais (ECI/UFMG); *Transinformação*, editada pelo Programa de Pós-Graduação em Ciência da Informação da Pontifícia Universidade Católica

de Campinas (PUC Campinas), e *Informação & Sociedade: Estudos*, editada pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba (UFPB).

O critério para a seleção dessas revistas foi a qualificação A do sistema Qualis da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes). A base utilizada para a busca dos artigos foi a Web of Science (WoS) a partir do portal de periódicos da Capes, considerando-se os descritores ciência de redes (*network science*), rede social (*social network*) e coautoria (*co-authorship*). As produções selecionadas foram publicadas no período de 2009 a 2015.

O *corpus* da pesquisa consistiu de 18 artigos, publicados nas seguintes revistas: *Perspectivas em Ciência da Informação* – 11 artigos, *Transinformação* – três artigos e *Informação & Sociedade: Estudos* – quatro artigos.

Nos artigos avaliados a definição de conceitos sobre análise de redes e colaboração científica se mostrou presente uma vez que o método é relativamente novo se comparado com outros métodos da Ciência da Informação. Ferreira (2011) define rede social de informação como “um conjunto de pessoas com algum padrão de contatos ou interações entre as quais se estabelecem diversos tipos de relações e, por meio delas, circulam diversos fluxos de informação”. Já as autoras Tomaél e Marteleto (2013) consideram a ARS o método de estudo de redes sociais por meio do mapeamento e da descrição das ligações entre os atores envolvidos. Uma outra definição é apresentada por Vilan Filho, Souza e Mueller (2008) que definem autoria múltipla como “o texto científico assinado por mais de um autor”. Eles alertam que a autoria múltipla indica que houve colaboração, sem especificar detalhes a respeito disso, havendo apenas concordância em assumir a responsabilidade conjunta pela produção.

De acordo com Oliveira et al. (2014), a análise de redes de coautoria permite identificar as relações de colaboração científica no ambiente acadêmico e analisar a confiabilidade de redes de coautoria científica considerando a participação dos pesquisadores e a intensidade das relações de coautoria existentes em uma rede. Ferreira (2011) considera que a ARS é um importante instrumento para se estudar relacionamentos que proporcionam o compartilhamento da informação e do conhecimento. Para Vilan Filho, Souza e Mueller (2008) analisar redes permite estudar as interconexões organizacionais e avaliar fenômenos complexos, tais como relações de poder, fluxo de informação e distribuição de recursos.

Os trabalhos dos pesquisadores da Ciência da Informação sobre a colaboração científica também apresentam as possibilidades quando da aplicação dos métodos. Duarte (2011) considera, ao analisar as redes sociais de produção científica com suas conexões, que

estas são redes cooperativas de pesquisadores que contribuem para o desenvolvimento da ciência, que se constitui em uma atividade colaborativa e social. Essa cooperação permite melhorar o resultado, maximizando o potencial do conhecimento e da produção científica. Duarte (2012) entende que as redes sociais de cooperação científica possibilitam o crescimento, o compartilhamento, o desenvolvimento e a inovação. Essas redes sociais cooperativas são formadas por grupos de pesquisa, inclusive pesquisadores geograficamente distantes, com o intuito de obter visibilidade, reconhecimento, experiência, aumento de acesso a fontes competitividade e evitar o isolamento.

Outros pesquisadores defendem o uso da Análise de Redes e seus métodos. Vanz e Stumpf (2010) entendem que a utilização de estudos de redes é importante para o estudo de coautoria, pois estes auxiliam o entendimento do fenômeno. De acordo com Maia e Caregnato (2008), nos estudos a respeito da colaboração científica observa-se um fortalecimento do método de ARS. O método permite identificar relações, vínculos e interações entre atores. Também permite identificar a posição de centralidade, que sugere que atores centrais detêm poder, pois possuem menos restrições e mais oportunidades de acesso à informação em uma estrutura de relações sociais. Pinto, Igami e Bressiani (2010) entendem que comunicar a descoberta científica é imprescindível para a atividade científica e para o progresso da ciência. O principal canal formal de comunicação científica é o periódico científico, e os artigos publicados constituem-se na forma definitiva de publicação dos resultados de uma investigação.

Em um estudo sobre as aplicações dos métodos de ARS com outros métodos da Ciência da Informação, Vanz e Stumpf (2010) apresentam uma série de conceitos importantes. As autoras entendem que observar como os cientistas se comportam, se relacionam, se organizam e como transmitem informações entre si permite compreender a produção e o uso do conhecimento científico. No entanto é feito uma consideração afirmando que embora a literatura relacione a colaboração científica à coautoria ou a trate como sinônimos, a coautoria é uma faceta da colaboração científica porque não mensura a colaboração na sua totalidade e complexidade. Isso porque nem sempre todas as pessoas listadas como autores são as responsáveis pelo trabalho intelectual produzido. De maneira a se aprofundar e validar as colaborações as autoras discorrem que tem sido utilizadas técnicas para a investigação da coautoria, tais como a bibliometria e a cientometria. As autoras consideram as motivações dos pesquisadores para a colaboração científica como produto da sistematização da literatura nacional e internacional (VANZ; STUMPF, 2010, p. 50-51):



1. “desejo de aumentar a popularidade científica, a visibilidade e o reconhecimento pessoal;
2. aumento da produtividade;
3. racionalização do uso da mão de obra científica e do tempo dispensado à pesquisa;
4. redução da possibilidade de erro;
5. obtenção e/ou ampliação de financiamentos, recursos, equipamentos especiais, materiais;
6. aumento da especialização na ciência;
7. possibilidade de ‘ataque’ a grandes problemas de pesquisa;
8. crescente profissionalização da ciência;
9. desejo de aumentar a própria experiência através da experiência de outros cientistas;
10. desejo de realizar pesquisa multidisciplinar;
11. união de forças para evitar a competição;
12. treinamento de pesquisadores e orientandos;
13. necessidade de opiniões externas para confirmar ou avaliar um problema;
14. possibilidade de maior divulgação da pesquisa;
15. como forma de manter a concentração e a disciplina na pesquisa até a entrega dos resultados ao resto da equipe;
16. compartilhamento do entusiasmo por uma pesquisa com alguém;
17. necessidade de trabalhar fisicamente próximo a outros pesquisadores por amizade e desejo de estar com quem se gosta”.

Dentre os artigos analisados foi percebido a utilização prática dos métodos em diferentes áreas sendo a principal delas a própria Ciência da Informação e os seus autores. Silva, Pinheiro e Reinheimer (2013) mapearam o desenvolvimento da pesquisa na área da comunicação científica com base nos artigos publicados nas principais revistas de CI no Brasil. Os autores relatam uma rede altamente conectada e a aglutinação de alguns trabalhos e autores em torno dos autores mais influentes, resultados coerentes com a de outros pesquisadores.

Ainda na Ciência da Informação Graeml (2010) analisa as estruturas social e intelectual (relacionamentos sociais) de pesquisadores com base em redes de coautoria e em redes de citação, utilizando medidas de coesão da análise de redes em redes de coautoria, a saber: densidade, distância e coeficiente de agrupamento. Identificam-se rede de coautoria fragmentada e formação de grupos menores de colaboração entre pesquisadores, resultando

em descentralização da produção científica na área pesquisada e na construção de redes de autoria mais sólidas.

Os resultados de Silva, Barbosa e Duarte (2012) mostram outro recorte dentro da CI. Os autores analisam as redes sociais de coautoria dos pesquisadores mais produtivos na área de CI no Brasil utilizando a análise de redes sociais. A parte mais significativa dessa rede constituiu-se de pesquisadores mais produtivos, e a maioria das produções foi elaborada em coautoria.

A pesquisa de Oliveira, Souza e Castro (2014) já utilizam a análise sociométrica para verificar padrões de interações entre as bibliotecas do Consórcio das Universidades Federais do Sul-Sudeste de Minas Gerais. O estudo buscou identificar as bibliotecas que apresentam posição favorável para a busca e trocas de informações e, conseqüentemente, possuem melhores condições para gerar inovação na rede.

Outros trabalhos na CI discorrem sobre o uso das métricas como densidade, diâmetro e distância, centralidade e componentização ou clusterização de grupos. Bordin, Gonçalves e Todesco (2014) analisaram a relação entre pesquisadores em redes em coautoria para diagnosticar a produção e a colaboração científica utilizando métricas de análise de rede social.

Acerca dos problemas com o uso do método Barbastefano (2013) alerta para as ambigüidades nos nomes dos autores em estudos de redes sociais baseadas em coautoria, tais como homônimos, grafias diferentes, mudanças de nomes, nomes incompletos e abreviações.

#### **2.4.4 Redes de coautoria na saúde**

Na saúde, os estudos sobre redes podem ser descritos por alguns autores que utilizaram esses conceitos aplicando-os prioritariamente sobre as redes de pesquisa na saúde. Uma das primeiras iniciativas nacionais que sugeriram que a análise de coautoria poderia fornecer informações relevantes para análise e planejamento estratégico de organizações que visam à erradicação de doenças foi o estudo de Morel et al. (2009). Esse estudo analisa as redes de coautoria de pesquisadores brasileiros para avaliar a colaboração científica na pesquisa em seis doenças tropicais negligenciadas (dengue, doença de chagas, leishmaniose, hanseníase, malária e tuberculose), alvos de um programa lançado pelo governo brasileiro para financiar a pesquisa, o desenvolvimento e a capacitação na área. Por meio da ARS os autores foram capazes de identificar as instituições-chave que poderiam atuar como pontes entre a comunidade de pesquisa, as comunidades de pesquisa mais ativas e sua localização, a

abordagem dada às doenças que têm causado sérios problemas para a saúde pública e outras informações importantes para a gestão do programa (MOREL et al., 2009).

O trabalho de Moura e Caregnato (2011) avaliou redes de coautoria em artigos e coinvenção em patentes de pesquisadores brasileiros que possuíam tanto pedidos de patentes como artigos publicados na área de biotecnologia. Assumindo a premissa de que, embora apresentem diferenças significativas, artigos e patentes são expressões da pesquisa, o objetivo do trabalho foi verificar a relação entre a produção científica e a produção tecnológica desses pesquisadores. Como resultado, as autoras não só observaram um percentual de 70,7% de interação entre a produção científica e tecnológica, mostrando que o depósito de patentes e a publicação de artigos não são atividades excludentes, mas também que os autores mais envolvidos em patentes são também os que mais publicam artigos.

O estudo de Vasconcellos e Morel (2012) analisou as redes de publicações e patentes na área da tuberculose no Brasil. Entre outros resultados, o trabalho demonstra o papel ativo das universidades em iniciativas inovadoras nesse campo, mas também que a participação da indústria nos depósitos de patentes é fraca, e que há uma falta de cooperação entre universidades e empresas, o que demonstra a ausência de comunicação entre a academia e a indústria na área. Os autores ressaltam a importância das análises conjuntas entre publicações e patentes em países como o Brasil, que consolidaram sua base científica e no momento contam com os mesmos atores para conduzir os desafios da inovação tecnológica.

A tese de doutorado de Santos (2012) abordou as redes de coautoria (com foco em pesquisadores brasileiros) e de coinvenção (com foco em patentes depositadas no Brasil) relacionadas a vacinas contra o papilomavírus humano (HPV) e a dengue. A autora encontra uma grande diferença quando as redes de publicações e de patentes são comparadas: as redes de coautoria são muito densas e interligadas, enquanto as redes de coinvenção são escassas, possuindo poucas conexões e vários componentes. Santos também verificou uma baixa inserção dos pesquisadores brasileiros nas redes globais de P & D relativas a esses temas. Ela sugere o aumento dos esforços de coordenação para promover a aprendizagem tecnológica desses atores.

O trabalho de Costa, Pedro e Macedo (2013) investigou a colaboração científica em biotecnologia na Região Nordeste do Brasil. Nos últimos anos, a pesquisa científica e tecnológica em biotecnologia foi colocada como um dos pilares para conduzir os indicadores de ciência, tecnologia e inovação da região, como forma de descentralizar a produção de conhecimento no país (BRASIL, MINISTÉRIO DA SAÚDE, 2012). O estudo mostrou que a colaboração nesta área ocorre principalmente no nível intrainstitucional e que a dinâmica da

cooperação na região apresenta baixa diversidade institucional e uma representação mínima de empresas na produção científica. O autor sugere que a colaboração científica entre as instituições deve ser reforçada a fim de melhorar o fluxo intrarregional, inter-regional e internacional de informação.

### **3 Metodologia**

Avalia-se como de grande importância a realização de uma pesquisa com o objetivo de aprofundar o conhecimento sobre as redes de colaboração científica e sobre as mudanças no padrão de investimentos recentes. Este projeto de pesquisa é essencial para se entender os efeitos dessas ações relacionadas à produção de artigos científicos e avaliar o impacto da mudança na estrutura e na dinâmica das redes.

A colaboração científica é uma característica marcante da pesquisa contemporânea. Ao colaborarem, os pesquisadores podem estabelecer redes de comunicação acadêmica, compartilhar ideias, recursos e informações, gerar e disponibilizar novos conhecimentos e, em última instância, criar inovações, reduzindo o custo e aumentando a produtividade da pesquisa (ROYAL SOCIETY, 2011).

O avanço da tecnologia tem permitido uma análise mais ampla dessas colaborações, especialmente em grandes bases de dados. A análise de redes com dezenas ou centenas de milhares de nós e linhas era incomum e restrita a pequenos grupos de pesquisadores. Atualmente, com as atividades de pesquisa e os métodos apresentados neste estudo, várias oportunidades e mecanismos de gestão podem ser desenvolvidos para subsidiar o processo decisório, inclusive a governança de ações públicas.

A pesquisa realizada demonstrou que existe uma mudança substancial na forma de realização da pesquisa coletiva e um crescente domínio de grupos de pesquisa sobre os demais cientistas na produção de conhecimento (WUCHTY; JONES; UZZI, 2007). Como consequência desse domínio, o número de trabalhos em coautoria vem aumentando drasticamente nos últimos anos, influenciando interações cada vez mais coesas e uma possível remodelagem da organização da prática científica.

#### **3.1 Modelo exploratório de análise de redes**

Na realização das atividades de pesquisa deste estudo foi utilizado como referência o modelo exploratório de análise de redes proposto por Nooy et al. (2011), no qual são definidas quatro etapas cíclicas na pesquisa de análise de redes sociais, a saber: a definição da rede; a manipulação da rede; a determinação das características estruturais; e a inspeção visual. Em seu trabalho, Nooy et al. (2011) consideram a detecção e a interpretação de padrões sociais nas relações entre os atores da rede o principal objetivo da análise de redes sociais. É importante ressaltar que existem diferentes tipos de redes, tais como as biológicas, as de informação ou as tecnológicas, além das redes sociais (NEWMAN, 2010, p. 14). O foco

do modelo e do estudo realizado é nas redes sociais, nas quais os pesquisadores identificados nas publicações científicas são os atores ou os nós das redes. O tipo de pesquisa portanto é exploratória onde o objeto são as “Relações profissionais de colaboração científica”, em que se considera a existência de uma relação entre dois pesquisadores caso estes tenham colaborado em um artigo ou publicação científica. Em anos recentes, tem sido percebida em alguns casos uma forma de colaboração com menor interação social, em que em um mesmo artigo dezenas ou centenas de pesquisadores acabam por apoiar o trabalho realizado, mas com pouca ou nenhuma interação entre eles. Esses casos são identificados e apontados quando necessário.

O tipo da pesquisa é dito exploratório por não haver nenhuma hipótese específica acerca das características da estrutura da rede. Apesar de já existirem diversos estudos sobre redes de coautoria ou colaboração científica, os resultados específicos são dependentes do universo estudado, seja ele temporal, seja por área ou por qualquer outro fator. Uma forma de explicar tal fato nas redes estudadas seria, por exemplo, quanto à formação de grupos de pesquisa, muito explorado neste trabalho. Uma hipótese poderia ser definida quanto ao percentual de colaborações dentro e fora de um espaço geográfico ou instituição de pesquisa por seus pesquisadores. No entanto, como para cada área de pesquisa ou recorte existem diferentes características, essa hipótese não teria nenhum embasamento estatístico, seria uma hipótese apenas empírica.

No estudo de redes, outro fator importante está relacionado às estimativas ou à falta delas, por assim dizer. Não se podem obter estimativas sobre características ou comportamentos da rede com base em amostras de uma rede, como é comumente feito em análises estatísticas. Isso quer dizer que se faz obrigatório o estudo de toda a rede existente? Obviamente não! No caso da pesquisa científica seria o mesmo que considerar obrigatório o estudo de todas as publicações científicas existentes em todas as áreas, procedimento inviável no espaço de tempo disponível.

Uma vez definido um universo de pesquisa, ele é o objeto de trabalho, não sendo este a representação de algo maior ou similar. Um exemplo da impossibilidade de se realizar estimativas está relacionado ao número de colaborações de determinados cientistas na rede de pesquisadores. Como será exposto no Parte II de resultados, a distribuição do número de colaborações por autor segue a lei de potência, em que poucos autores possuem muitas colaborações, enquanto muitos autores possuem poucas colaborações. Dessa maneira, se a amostra fosse feita se valendo dos autores com muitas colaborações, esta certamente não seria uma representação da característica da rede.

### 3.1.1 Definição da rede

Um grafo pode ser direcionado ou não quando a linha que une dois vértices carrega uma informação unilateral ou bidirecional. Um exemplo desses diferentes tipos de direcionalidade no grafo na pesquisa científica seria uma rede de citações e uma rede de coautoria. No primeiro caso o grafo é direcionado, pois quando um pesquisador cita outro autor a recíproca não é necessariamente verdadeira. No caso de uma grafo de coautoria, este seria não direcionado, pois se pressupõe a colaboração de ambas as partes.

Outro elemento importante é a permissão ou não de linhas múltiplas. No caso da rede de coautoria, são permitidas as linhas múltiplas, uma vez que um pesquisador pode colaborar diversas vezes com um mesmo companheiro, sendo levado em consideração o número de colaborações. Outro elemento existente nos grafos, mas que não é considerado nas nossas análises, são os *loops*, ou seja, quando um vértice se relaciona consigo mesmo. No caso da colaboração isso não é levado em consideração, mas em uma rede de citação esse elemento pode ser analisado.

Uma vez que temos o grafo e seus elementos, o próximo passo é a definição da rede, que nada mais é do que o grafo com informações adicionais sobre os vértices, os nós e as linhas. Dessa maneira, quando se definem os nomes, os países, os grupos ou o peso das linhas na rede de pesquisa científica estão sendo definidos o contexto e a estrutura informacional da rede.

As redes analisadas, portanto, são compostas de nós, que representam os pesquisadores em leishmania e tuberculose que tiveram publicações científicas indexadas nas bases PubMed, Web of Science e SciELO. Nesta última base as revistas também foram consideradas nós em uma rede de dois modos, situação a ser explicada logo adiante. As relações são as colaborações nos artigos, sendo que no caso do SciELO as relações foram também os próprios pesquisadores. O universo temporal e geográfico analisado variou conforme os resultados apresentados, descritos em cada capítulo.

### 3.1.2 Manipulação da rede

Quando se trabalha com análise de redes muitas vezes são necessários ajustes no universo pesquisado por conta do tamanho da rede ou de particularidades que podem impactar no entendimento da estrutura apresentada. Essas manipulações devem ser declaradas para evitar análises tendenciosas. O modelo exploratório proposto funciona muito bem com redes de pequeno ou médio porte com até algumas centenas de nós. Redes com dezenas de milhares

de nós acabam por gerar certa dificuldade, e por isso foram realizados alguns recortes no estudo.

Dentre as bases de dados trabalhadas, usou-se a manipulação da rede quando vários pesquisadores possuíam apenas uma ou duas publicações na área. Em alguns desses casos decidiu-se eliminar esses pesquisadores, uma vez que o foco era a colaboração científica. Nos casos em que o número de pesquisadores foi reavaliado com base no número de publicações, essa informação foi registrada na definição do universo da pesquisa. Os pesquisadores eliminados poderiam ter um número elevado de publicações em outras áreas de pesquisa, mas como o recorte realizado teve como foco a LN e a TB, os pesquisadores com poucas publicações nessas áreas acabaram por apresentar pouca relevância ao estudo.

Outros tipos de relações podem ser levados em consideração na manipulação da rede como região ou instituição em que trabalha e revistas ou áreas de publicação. No caso da pesquisa científica pode-se também trabalhar com a análise de citações, conforme mencionado. Nesse caso, pode-se optar primeiro por um tipo de relação, como a colaboração, para que se tenha um melhor entendimento da rede para depois contemplar o segundo tipo de relação, como a citação, de forma que a pesquisa seja vista sob outra perspectiva. No caso do estudo apresentado não foram levadas em consideração outras relações nas bases PubMed e Web of Science, mas foi analisado outro tipo de relação na base SciELO, que foi a rede de dois modos, levando-se em consideração as revistas nas quais foram publicados os artigos analisados.

### **3.1.3 Determinação das características estruturais**

Nas pesquisas sobre análise de redes sociais ou ciência de redes, muitas características estruturais foram quantificadas por meio de medidas, como, por exemplo, a centralidade ou a densidade. As medidas utilizadas no percurso deste trabalho estão descritas nos subitens que virão a seguir. Algumas medidas dizem respeito a toda a rede, e outras medidas têm como foco os nós ou as linhas. O estudo das características estruturais é muito mais preciso e conciso do que as análises visuais e serve como apoio para a análise visual.

No caso desta pesquisa, a análise da estrutura foi fundamental para o entendimento das diferentes redes de colaboração e seus subgrupos. Foi visto de forma clara que existem estruturas de poder muito bem estabelecidas com base no número de relações entre os pesquisadores. Houve um grupo da pesquisa na base PubMed, chamado de hubs, que se mostrou muito bem relacionado, atravessando as fronteiras geográficas e criando o que foi



chamado de “clube dos ricos” na escala mundial. No contexto brasileiro, o recorte mostrou a penetração efetiva de pesquisadores de uma mesma instituição no contexto da pesquisa em doenças negligenciadas.

A análise da estrutura por meio das medidas permitiu ao estudo identificar características e mudanças de uma estrutura Livre de Escala para uma estrutura Mundo Pequeno, o que caracteriza o embasamento da rede de pesquisa em grupos de pesquisa mais do que em pesquisadores individuais, como acontecia há dez ou vinte anos. Isso, no entanto, não diminui o poder dos hubs, que conseguem relacionar-se bem tanto dentro dos grupos quanto com os elementos fora dos seus grupos. Além disso, pôde-se avaliar a existência de uma estrutura centro-periferia na qual se pressupõe que todos os nós estão de certa forma conectados, alguns desses em grupos coesos ao centro do “poder”, e outros pesquisadores na periferia, mas mesmo assim conectados ao resto da rede.

Uma característica que acaba por se valer também da manipulação da rede é o chamado “componente gigante”, que vem a ser o maior elemento da rede, onde os nós daquele grupo se encontram, todos conectados entre si. Quanto à estrutura, o estudo do componente gigante ajuda a identificar se a rede é bastante coesa ou esparsa e dividida em pequenos grupos. Estudos sobre redes de colaboração científica revelam a tendência de colaboração intensa entre os diferentes atores, o que torna o componente gigante o elemento de maior relevância na rede, uma vez que a grande maioria dos pesquisadores se encontra conectada entre si. A maioria das medidas utilizadas na análise de redes complexas prevê a utilização apenas desse componente. No que diz respeito à manipulação da rede, os nós que não fazem parte desse grupo coeso acabam por ser eliminados da análise final.

#### **3.1.4 Inspeção visual**

A visualização das redes ou inspeção visual é um elemento chave na análise exploratória, possibilitando o mapeamento e a apresentação de padrões de relacionamentos. As redes com milhares de nós prejudicam ou dificultam essa etapa do processo. Isso ocorre porque o olho humano pode ser facilmente manipulado, sendo levado a acreditar em informações que não necessariamente estão presentes.

Para evitar esse tipo de resultado faz-se necessário uma série de precauções metodológicas visando a garantir a melhor apresentação de resultados possíveis para que se consiga obter a mais fidedigna interpretação. A distribuição dos nós na rede é um elemento essencial para essa análise. Algoritmos de distribuição são largamente reconhecidos, por isso

apresentam-se suas características para que os resultados possam ser mais bem apreendidos. Dois dos algoritmos mais utilizados neste trabalho foram o Fruchterman Reingold<sup>2</sup> e o Force Atlas.<sup>3</sup> O primeiro faz uma distribuição circular dos nós, colocando os elementos com maior centralidade no centro da esfera, e o segundo foca nos relacionamentos mais fortes e evidentes, demonstrando a coesão dos grupos. Além disso, foi largamente utilizado um algoritmo de clusterização quanto à definição de subgrupos dentro das redes. Estes subgrupos foram identificados com cores diferentes para caracterizar a participação em um determinado *cluster*.

O tamanho dos nós e a espessura das linhas de ligação entre os pesquisadores também foram utilizados para assinalar o número de elementos e ligações ou o grau de centralidade em alguns casos. Todas as características apresentadas nas figuras foram explicadas na sua interpretação.

### 3.2 Coleta e tratamento dos dados

Como consequência do avanço tecnológico, existe hoje um volume substancial de trabalho em redes de colaboração científica. Muitas redes de coautoria demonstram uma distribuição desigual no padrão de colaboração e no envolvimento dos pesquisadores, havendo alguns cientistas com um número muito grande de colaboradores enquanto muitos outros possuem poucas colaborações (NEWMAN, 2004; BARABÁSI et al., 2002).

A busca dos dados levou em consideração os padrões de trabalho já realizados anteriormente com o auxílio das ferramentas disponíveis para tratamento destes. O resultado foi a capacidade de tratar um elevado número de publicações ainda se valendo do modelo exploratório de análise proposto.

#### 3.2.1 PubMed

Na primeira etapa da pesquisa foram utilizados os dados provenientes da base de dados PubMed,<sup>4</sup> um serviço do National Center for Biotechnology Information (NCBI), que indexa milhões de publicações da Medical Literature Analysis and Retrieval System Online (Medline), e outras revistas de ciências para artigos biomédicos.

---

<sup>2</sup> <<https://wiki.gephi.org/index.php/Fruchterman-Reingold>>.

<sup>3</sup> <<https://wiki.gephi.org/index.php/ForceAtlas>>.

<sup>4</sup> Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed> accessed on october 2013>.

Os termos/descriptores de busca utilizados para a primeira etapa foram “leishmania” e “leishmaniasis” nos títulos e nos resumos de trabalhos, com um retorno de 19.437 artigos. Após serem considerados o crescimento gradual da frequência anual dos artigos e a regularidade quanto ao tempo médio dos pesquisadores envolvidos nos artigos recuperados, decidiu-se utilizar os dados de 1981 a 2012, que continham 16.490 artigos.

Na pesquisa sobre a tuberculose foi utilizado o termo/descriptor “tuberculosis”, com 103.162 artigos no mesmo período de estudo da LM. Como o número de artigos era muito elevado, decidiu-se restringir o período analisado da TB para 16 anos, em vez de 32 anos. O número de artigos então ficou reduzido a 68.958. O número total de pesquisadores avaliados foi de 9.577, levando-se em consideração apenas os autores com mais de uma publicação na área.

A seguir, os artigos foram separados por períodos intervalares de quatro anos contíguos, resultando em oito períodos. Essa divisão foi essencial para que pudesse ser feita uma análise temporal ou dinâmica da transformação dos grupos e da utilização das palavras-chave. O período de quatro anos foi definido com base no tempo médio de publicação dos pesquisadores.

A definição sobre o recorte temporal e a análise em períodos de quatro anos foi avaliada com base nos dados disponíveis. O elevado número de autores foi um elemento decisivo para o recorte acerca dos pesquisadores com apenas uma publicação e que representavam quase 50% da rede.

### **3.2.2 Web of Science**

A segunda base utilizada foi a Web of Science (WoS), chamada anteriormente de Web of Knowledge e mantida pela Thomson Reuters. A base é conhecida pela indexação de citações em publicações científicas e acessada por meio de uma assinatura *online*. O acesso a esta base foi possível por meio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), que disponibiliza o acesso à base a estudantes de nível de pós-graduação *stricto sensu* e a entidades de ensino e pesquisa.

A WoS foi a base de escolha por abranger um grande número de periódicos acadêmicos e por fornecer informações exclusivas sobre afiliação e/ou endereços dos autores, permitindo o acompanhamento do comportamento colaborativo dos membros da população-alvo, que geralmente publicam em revistas de alto impacto.

A busca sobre leishmaniose foi realizada em setembro de 2014 e abrangeu o período de 1984 a 2013. Utilizaram-se como parâmetros de delimitação da busca os seguintes campos e descritores: OG = (Fundação Oswaldo Cruz), OR CU = (Brasil OR Brazil), AND TS = (leish\*) OR Kala Azar.

A coleta de dados da tuberculose abrangeu o período de 2003 a 2014. A consulta foi feita utilizando-se os seguintes termos: TS = (Tuberculos\*) AND OG = (Fundação Oswaldo Cruz) OR CU = (Brasil OR Brazil). Nos tipos de publicação foram incluídos apenas artigos publicados em periódicos (*articles/articles in press*).

O universo da pesquisa foi restrito aos artigos que continham pesquisadores filiados a instituições brasileiras. O motivo do recorte foi o grande número de publicações na área e a capacidade de análise ser delimitada pelo espaço de tempo e pela envergadura tecnológica. Foram utilizados na pesquisa dados relacionados à categorização de grandes áreas de conhecimento dos artigos com o intuito de avaliar as estruturas de colaboração sob outro tipo de caracterização.

O recorte temporal seguiu uma definição diferente da realizada durante a pesquisa da PubMed. No caso da TB o foco foi a avaliação das mudanças ocorridas durante a última década, que concentrou uma série de mudanças nas políticas públicas aplicadas à pesquisa, ao tratamento e ao controle da doença no Brasil. Esse recorte foi indicado por um especialista na área cujo interesse era investigar esses acontecimentos. Quanto ao recorte em LN, a base WoS apresentou um crescimento drástico comparado aos anos anteriores, o que reforçou a necessidade de se iniciar no período descrito.

### 3.2.3 SciELO

A terceira base utilizada foi a Scientific Electronic Library Online (SciELO), uma base originalmente brasileira que suporta dados bibliográficos com revistas de acesso aberto. O SciELO indexa hoje revistas primariamente brasileiras e de uma série de outros países latino-americanos por vezes não indexadas pelas grandes bases americanas.

O foco no SciELO foram os artigos de LN e TB e as revistas utilizadas para a publicação desses artigos. Neste caso foi utilizada uma rede de dois modos na qual os autores se relacionavam com as revistas em que já haviam publicado seus artigos. Na transformação da rede de dois modos para uma rede de um modo as revistas agora se relacionam tomando como base os autores. Por exemplo, se um autor publicou em duas ou mais revistas diferentes, estas revistas possuem um “relacionamento”, que neste caso seria o autor, e com isso

compartilham algo em comum. Com isso podem ser observadas quais revistas possuem assuntos semelhantes e quais possuem seus próprios públicos específicos.

A busca na LN utilizou o termo (leish\*), com retorno de 1.559 artigos de 2000 a 2014 e um total de 1.038 pesquisadores com duas ou mais publicações na área. Por sua vez, a busca em TB utilizou o termo (tuberculos\*), com 3.578 artigos de 2000 a 2014 e 2.045 pesquisadores com duas ou mais publicações. Das três bases analisadas esta foi a única que utilizou o mesmo recorte de tempo que as duas outras bases.

### **3.3 Elementos e medidas de redes**

#### **3.3.1 Elementos**

Um grafo é constituído por dois tipos de elementos, a saber: vértices e arestas, também chamados de nós e linhas. Cada linha tem dois pontos finais no conjunto de vértices, e é dito que esta conecta ou junta os dois nós. Uma aresta pode ser definida como um conjunto de dois vértices (ou um par ordenado, no caso de um grafo direcionado). As duas extremidades de uma aresta também são referidas como adjacentes umas às outras.

Um vértice é desenhado simplesmente como um nó ou um ponto. Uma aresta (um conjunto de dois elementos) é desenhada como uma linha que conecta dois vértices, chamados de pontos finais ou vértices finais. O tamanho de um grafo equivale ao número dos seus nós.

##### ***3.3.1.1 Número de vértices***

O número de vértices de uma rede reflete o número de indivíduos ou de instituições, dependendo da unidade de análise, envolvido na rede em cada período analisado.

##### ***3.3.1.2 Número de links (linhas)***

O *link* representa as ligações (número de conexões) entre os vértices, é o atributo relacional que os conecta. Um nó pode estar ligado a outro porque ambos colaboraram em um mesmo artigo científico, participaram da mesma banca de tese, trabalharam na mesma área, etc.

### **3.3.2 Medidas de rede**

Muitas vezes as redes têm certos atributos que podem ser calculados para analisar suas propriedades e suas características. Essas propriedades de rede frequentemente definem os modelos de rede e podem ser usadas para analisar como certos modelos contrastam com outros.

#### ***3.3.2.1 Densidade***

Esta métrica tem por objetivo medir a conectividade dentro da rede com base no número de ligações totais possíveis. A densidade equivale ao percentual entre o número de ligações e o número máximo de ligações possíveis em uma dada rede (WASSERMAN; FAUST, 1994). Desse modo, redes densas são aquelas onde há grande quantidade de conexões, e redes esparsas são aquelas com pequena quantidade de ligações. Quanto maior a densidade, maior a coesão entre o grupo.

#### ***3.3.2.2 Caminho médio da rede***

O comprimento do caminho médio da rede é calculado para se encontrar o caminho mais curto entre todos os pares de nós. A medida é calculada adicionando-se a distância entre todos os nós da rede e em seguida dividindo-se pelo número total de pares. Isso apresenta, em média, o número de passos necessários para se chegar de um membro da rede a outro, em média. Essa medida aplicada às relações das pessoas no mundo é conhecida como seis graus de separação.

#### ***3.3.2.3 Diâmetro da rede***

O diâmetro é outra forma de medir grafos de rede. Define-se o diâmetro de uma rede como o caminho mais longo de todos os caminhos mais curtos calculados em uma rede. Em outras palavras, uma vez que o comprimento do percurso mais curto a partir de cada nó para todos os outros nós é calculado, o diâmetro é o mais longo de todos os comprimentos dos percursos calculados. O diâmetro é a representação da dimensão linear de uma rede (WASSERMAN; FAUST, 1994).

#### **3.3.2.4 Coeficiente de agrupamento ou clusterização**

O coeficiente de agrupamento é uma medida sobre “o quanto todos os meus amigos conhecem um ao outro”. Isso às vezes é descrito como o quanto os amigos dos meus amigos são meus amigos. Mais precisamente, o coeficiente de agrupamento de um nó é a proporção de ligações de conexões existentes entre os vizinhos de um nó para o outro vizinho até o número máximo possível de tais ligações (NEWMAN, 2010). O coeficiente de agrupamento para a totalidade da rede é a média dos coeficientes de agrupamento de todos os nós. Um alto coeficiente de agrupamento para uma rede é outra indicação de um mundo pequeno.

#### **3.3.2.5 Coesão**

A maneira como a rede está conectada tem um papel bastante relevante na forma como as redes são analisadas e interpretadas. Dentre as medidas de coesão estão os cliques ou grafos completos: uma rede completamente conectada, na qual todos os nós estão ligados a todos os outros nós (NOOY; MRVAR; BATAGELJ, 2011). Outra medida é o componente gigante – o percentual do maior componente de rede em que um único componente e suas conexões contêm a maioria dos nós da rede.

#### **3.3.2.6 Grau médio**

O grau  $k$  de um nó corresponde ao número de linhas ligadas a ele. O grau médio está estreitamente relacionado à densidade de uma rede. No modelo gráfico aleatório Erdős Rény pode-se calcular  $\langle k \rangle = p(N-1)$ , em que  $p$  é a probabilidade de dois nós se conectarem (NEWMAN, 2010).

#### **3.3.2.7 Centralidade**

Na ARS, atores importantes são tais que estão frequentemente envolvidos na relação com outros atores. Esses envolvimento os tornam mais visíveis, possuindo a maioria do acesso ou controle, sendo considerados mais centrais na rede. É nesse sentido que as medidas de centralidade tentam descrever a localização de um ator na rede. Essas medidas levam em consideração as diferentes maneiras com as quais um ator interage e se comunica com o restante da rede, sendo mais importantes, ou centrais, aqueles vértices localizados em posições mais estratégicas na rede.

A noção de centralidade em redes sociais foi introduzida por Bavelas em 1948 (BAVELAS, 1948 apud FREEMAN, 1979). Ele afirma que, num grupo de pessoas, um indivíduo em particular que se encontra estrategicamente localizado num caminho mais curto de comunicação entre pares de indivíduos está numa posição mais central na rede. Ele será responsável por transmitir, modificar ou reter a informação entre membros do grupo e será tão mais influente quanto mais central estiver posicionado.

Ao longo de décadas, pesquisadores de redes têm introduzido um grande número de medidas de centralidade com o objetivo de medir a variação da importância dos vértices de acordo com critérios preestabelecidos. Essas diferentes medidas e algoritmos foram resumidos, expandidos e desenvolvidos em um artigo escrito por Freeman (1979), no qual ele introduz a tipologia moderna de medidas de redes ao especificar três definições clássicas de centralidade: centralidade de grau (*degree centrality*), centralidade de proximidade (*closeness centrality*) e centralidade de intermediação (*betweenness centrality*). O grau é uma medida da influência direta que um vértice tem em relação a seus contatos, a proximidade está relacionada com o tempo que uma informação leva para ser compartilhada por todos os vértices na rede, e a intermediação de um vértice pode ser considerada o controle da comunicação entre todos os demais pares de vértices da rede. As medidas de centralidade de proximidade e centralidade de intermediação são baseadas na suposição de que a informação (ou qualquer conteúdo da ligação) é transmitida somente ao longo dos possíveis caminhos mais curtos.

Bonacich (1987) propôs uma medida denominada centralidade de autovetor (*eigenvector centrality*), cujo objetivo é medir a importância de um nó em função da importância de seus vizinhos. Isso quer dizer que, mesmo se um nó estiver conectado somente a alguns outros vértices da rede (tendo assim uma baixa centralidade de grau), esses vizinhos podem ser importantes e, conseqüentemente, o nó será também importante, obtendo uma centralidade de autovetor elevada.

O conceito de centralidade pôde ser então formalizado e aplicado a qualquer rede. Tomando por base o resultado de determinada medida de centralidade de cada nó da rede é possível ordenar os vértices dessa rede em razão de sua importância relativa. Essa importância está associada ao significado de cada medida, ou seja, como cada uma delas considera um nó mais importante que outros.

As medidas de centralidade também podem ser utilizadas para descrever o quão uma rede é centralizada. A centralização é uma medida do quanto os *links* de uma determinada rede estão concentrados em um ou em poucos vértices dessa rede. Uma rede centralizada,



muitas vezes denominada de rede hierarquizada, concentra seus *links* em uma ou poucas pessoas, enquanto uma rede descentralizada possui *links* igualmente distribuídos entre seus vértices (VALENTE, 2010).

Uma das maneiras de medir a centralização de uma rede é calcular o desvio-padrão das medidas de centralidade dos vértices individuais. Uma rede centralizada apresentará um alto desvio-padrão das medidas de centralidade dos vértices porque alguns indivíduos possuirão alta centralidade, enquanto muitos outros possuirão baixa centralidade (WASSERMAN; FAUST, 1994).

### **3.4 Tipos e estrutura das redes**

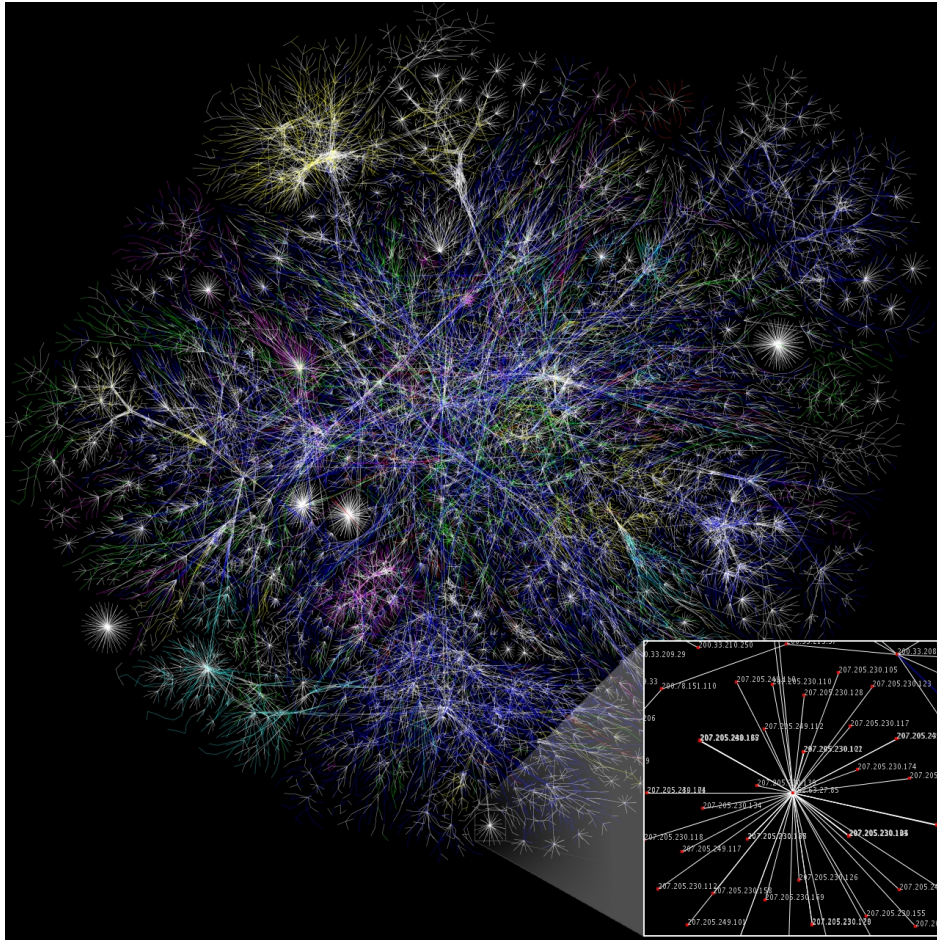
#### **3.4.1 Tipos de redes**

Os tipos de redes comumente estudados podem ser divididos em quatro grandes grupos segundo Newman (2010) no estudo empírico de redes: redes tecnológicas, redes sociais, redes de informação ou conhecimento e redes biológicas. Em algumas situações essas redes podem inclusive se sobrepor, como no caso de redes de informação e redes sociais, por exemplo.

##### **3.4.1.1 Redes tecnológicas**

As primeiras redes a serem estudadas em larga escala foram as redes tecnológicas. O aumento da capacidade de processamento de dados e o surgimento de novas tecnologias possibilitaram o estudo de redes constituídas por centenas de milhares de nós e linhas. Dentre as redes tecnológicas, um dos exemplos mais comuns e estudados por vários pesquisadores no início do século XXI é a internet. O acesso a informações relacionais em grande escala foi facilitado, e o mapeamento da internet começou a ser avaliado com a utilização de *crawlers* ou rastreadores. A Figura 12 apresenta o mapa da internet desenhado com base na utilização de um desses rastreadores, com a avaliação das distâncias em tempo entre diferentes roteadores e endereços IP com base no tempo de resposta quando do envio de um pacote de dados. A rede é dividida em cores, que representam os diferentes endereços de internet, como .com, .br e outros.

**FIGURA 12. Mapa da internet**



Fonte: WIKIMIDIA COMMONS

Legenda: mapa parcial da internet com base nos dados de 15 de janeiro de 2005 encontrado na opte.org. Cada linha é desenhada entre dois nós, que representam dois endereços IP. O comprimento das linhas é indicativo do *delay* no envio de pacotes de dados entre esses dois nós. Este gráfico representa menos de 30% das redes de classe C alcançáveis pelo programa de coleta de dados no início de 2005. As linhas são coloridas de acordo com seu código correspondente no RFC 1918 e descrito a seguir: azul escuro: net, ca; verde: com, org; vermelho: mil, gov, edu; amarelo: jp, cn, tw, au; magenta: uk, it, pl, fr; gold: br, kr, nl; branco: desconhecido.

Outros tipos de redes tecnológicas cujo estudo tem trazido resultados práticos para a sociedade são as redes de telefonia e as redes de energia elétrica. Pesquisas nessa área têm colaborado com a diminuição de problemas de sobrecarga e fluxo constante da demanda. Outras redes que podem ser agrupadas nesse modelo são as de transporte ferroviário, aéreo, fluvial ou por rodovias.

### **3.4.1.2 Redes sociais**

Um segundo tipo de rede de acordo com Newman são as redes sociais. Estas redes vêm sendo estudadas há vários anos por pesquisadores como Jacob Moreno, Mark Granovetter, dentre outros. A diferença apresentada na última década é a facilidade de acesso a dados relacionais entre indivíduos, principalmente com o surgimento de diversas mídias

sociais, tais como Facebook, Twitter, etc. Dentre os mais citados pesquisadores que conseguiram melhor definir os métodos para o estudo das redes sociais estão Stanley Wasserman e Katherine Faust, cujo trabalho é focado nesse tipo de rede.

Uma rede social consiste em um ou mais conjuntos finitos de atores (e eventos) e todas as relações definidas entre eles (WASSERMAN; FAUST, 1994). A análise de redes sociais compreende um conjunto de técnicas interdisciplinares que propiciam uma leitura dinâmica das interações sociais. É uma alternativa à interpretação “estática” do papel social do indivíduo ou grupo dentro de um contexto. Pressupõe uma análise de tendências em um determinado contexto, e não afirmações absolutas, isto é, a ARS é utilizada para oferecer uma perspectiva (GUIMARÃES; MELO, 2005).

Segundo Wasserman e Faust (1994), os tipos de redes podem ser definidos conforme a natureza dos atores ou dos conjuntos de atores, se houver um tipo de ator ou vários tipos, bem como em função das propriedades das ligações e das relações entre eles. A maioria das redes em análise são aquelas de um modo (também chamadas de redes monomodais), mas as redes de dois ou mais modos (duomodais ou multimodais) têm sido observadas em diversos casos reais, dessa maneira é importante criar modelos para sua análise.

Nas redes de um modo, nas quais o modo determina o conjunto único de atores que se relacionam entre si, cada ator pode ter, ou não, ligações com quaisquer atores pertencentes à sua rede. Todos os atores participam da rede com ligações entre si, ou seja, um ator pode alcançar qualquer outro ator. Um exemplo, já citado, é a rede de colaboração entre pesquisadores cujo trabalho conjunto na pesquisa e na confecção de um artigo define o tipo de relação.

As redes de dois modos são constituídas de dois conjuntos diferentes de atores, de um conjunto de atores e um conjunto de eventos e das relações entre os atores de um conjunto e os atores (ou eventos) do outro.

Redes de afiliação são típicas de dois modos. Nelas atores estão relacionados com eventos. A estrutura dessas redes pode ser analisada com técnicas de rede padrão, mas vários conceitos estruturais têm de ser redefinidos ou devem ser interpretados de forma diferente quando aplicados à rede de dois modos (NOOY; MRVAR; BATAGELJ, 2011).

Uma rede por afiliação (*affiliation networks*), também denominada *membership network*, é um tipo especial de rede de modo duplo na qual existe um conjunto de atores e um conjunto de eventos ou atividades (MATHEUS; SILVA, 2006).

### 3.4.1.3 *Redes de informação ou conhecimento*

Nas redes de informação ou conhecimento duas são as mais comumente estudadas. A primeira é a rede mundial de computadores, conhecida como *World Wide Web* (WWW). Diferente dos estudos sobre roteadores e endereços IPs realizados nas redes tecnológicas, o estudo da WWW foca nos *hyperlinks* e nas informações contidas nas páginas web disponíveis. Quando se faz uma busca em uma ferramenta de busca como o Google ou o Yahoo!, por exemplo, as informações disponibilizadas e encontradas podem ser definidas ou organizadas em rede. Os *hiperlinks* encontrados nas páginas web também podem servir como uma forma de relacionamento entre as diferentes páginas e informações.

A segunda rede muito utilizada por estudiosos na Ciência da Informação são as redes de citações, que utilizam artigos científicos e suas respectivas bibliografias para determinar a fonte de onde o conhecimento apresentado em um determinado artigo surgiu. Nos últimos anos, tem-se percebido um crescente interesse no estudo de redes de pesquisa acadêmica, em que um nó geralmente denota uma entidade acadêmica, como um artigo, um jornal ou um autor, e uma linha ou arco geralmente representam relacionamentos entre os nós com base na citação, na cocitação, na coautoria ou no acoplamento bibliográfico (YAN; DING, 2012), dentre outros tipos de relações. No caso das redes de citação, cocitação ou acoplamento bibliográfico não existe, na maior parte das vezes, uma relação social entre os nós, uma vez que o relacionamento se baseia no referenciamento à informação transmitida ou utilizada e não no pesquisador propriamente dito. No caso da coautoria, assume-se uma relação social entre os autores, uma vez que se pressupõe o trabalho conjunto. Outro espaço no qual as redes de citações têm sido utilizadas com bons resultados é no estudo de patentes.

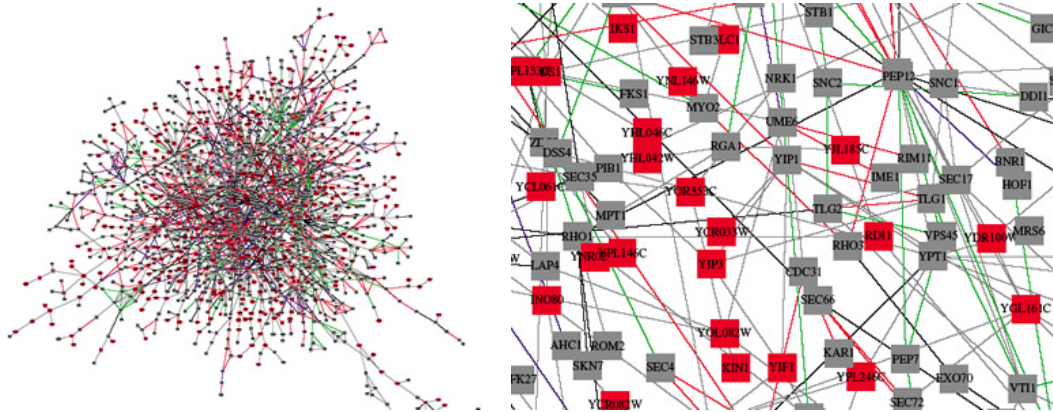
### 3.4.1.4 *Redes biológicas*

A área de estudos de redes biológicas é extremamente vasta, e segundo especialistas na área de ciência de redes este é provavelmente um dos maiores campos a serem explorados nos próximos anos. Segundo Newman (2010), as redes biológicas podem se subdividir em redes bioquímicas, redes neurais e redes ecológicas.

Dentre as redes bioquímicas destacam-se as redes metabólicas, que podem ser definidas como um conjunto completo de processos metabólicos e físicos que determinam as propriedades fisiológicas e bioquímicas de uma célula. Como tal, essas redes compreendem as reações químicas do metabolismo, as vias metabólicas, assim como as interações reguladoras que orientam essas reações. O sequenciamento de genomas completos tornou possível a

reconstrução da rede de reações bioquímicas em muitos organismos, desde bactérias a humanos. Além das redes metabólicas, existem as redes de interação entre proteínas, também conhecidas como rede PPI (*protein-protein interaction*). A Figura 13 exibe uma rede PPI.

**FIGURA 13. Interação proteína-proteína**



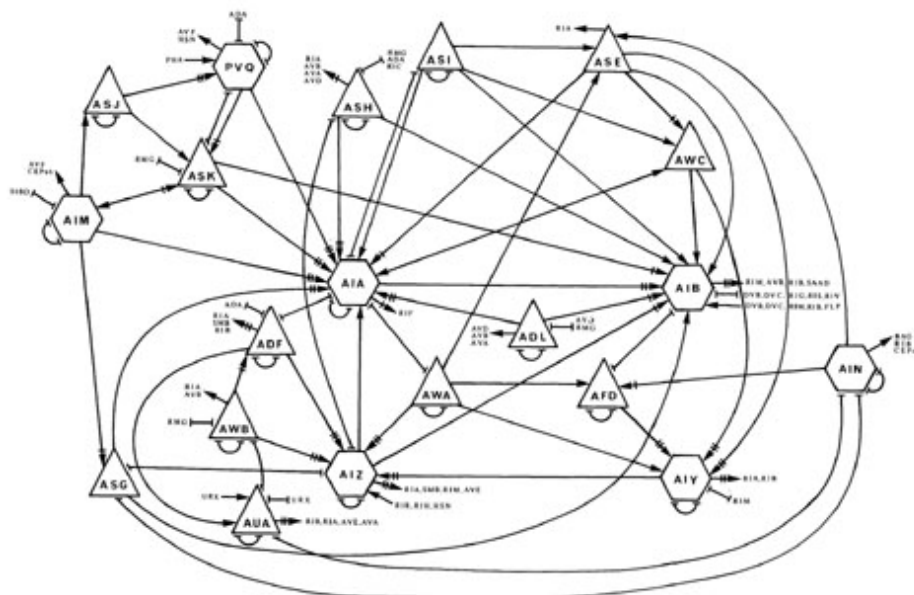
Fonte: SCHWIKOWSKI; UETZ; FIELDS (2000). Uma rede de proteínas que interagem em levedura. *Nat. Biotechnol*, 18, p. 1257-1261

Legenda: um mapa de interações da proteína de levedura montado com base em interações. O mapa contém 1.548 proteínas (caixas) e 2.358 interações (linhas de ligação). As proteínas são coloridas de acordo com seu papel funcional: proteínas envolvidas na fusão da membrana – azul; estrutura da cromatina – cinzento; estrutura da célula – verde; metabolismo dos lípidos – amarelo; e citocinese – vermelho.

As redes regulatórias genéticas, como o DNA, são também uma rica fonte de dados para o estudo das redes bioquímicas.

Quanto às redes neurais – estudos do cérebro –, estas têm crescido enormemente no campo da medicina, e a aplicação de métodos da ciência de redes tem propiciado avanços bastante consideráveis na área (Figura 14).

**FIGURA 14. Circuito neural de uma minhoca**



Fonte: WHITE et al. (1986)

Legenda: diagrama de uma parte do conjunto de circuitos do cérebro de uma minhoca. Parte do circuito neural do nemátodo *Caenorhabditis elegans*, reconstruído à mão com base em micrografias eletrônicas de fatias do cérebro do verme.

### 3.4.2 Estrutura das redes

Uma das principais análises de redes aplicadas ao estudo diz respeito à estrutura das redes e à sua evolução durante os diferentes períodos. Após alguns anos de pesquisa na área de redes complexas descobriu-se que as redes ditas “reais” possuem características estruturais e dinâmicas diferentes das preconizadas pela teoria clássica de redes aleatórias proposta por Erdős e Rényi (1959). As estruturas mais comuns dessas redes reais são: (i) a propriedade do Mundo Pequeno – *Small World* (há caminhos relativamente curtos entre quaisquer dois nós e um grande coeficiente de agrupamento) e (ii) a propriedade de redes Livres de Escala (há uma distribuição de probabilidades do número de ligações em um nó seguindo a lei de potência).

Em razão da descoberta dessas e de outras propriedades foram desenvolvidos novos modelos teóricos para reproduzir as propriedades estruturais observadas em topologias de redes reais. Redes com alto coeficiente de agrupamento e pequeno caminho médio podem ser geradas com uma evolução do modelo do Mundo Pequeno de Watts e Strogatz (1998), enquanto redes com distribuição de grau que seguem a lei da potência podem ser geradas com uma evolução do modelo Livre de Escala de Barabási e Albert (1999).

Quanto à formação de grupos ou estruturas coesas detectadas nas redes, foram utilizados os algoritmos de clusterização e modularidade definidos por Newman (2006). Os algoritmos foram escolhidos em razão da possibilidade de análise de redes com grandes volumes de dados e alto coeficiente de modularidade. A consequência desse resultado é o agrupamento de atores com base no autovetor das linhas de cada nó.

#### 3.4.2.1 Redes aleatórias ou randômicas

Em matemática, um grafo aleatório é o termo geral para se referir às distribuições de probabilidades sobre as redes. Redes aleatórias podem ser descritas simplesmente por uma distribuição de probabilidades ou por um processo aleatório que as gera (BOLLOBÁS, 2001). Do ponto de vista da ciência de redes, grafos aleatórios são usados para responder a perguntas sobre as propriedades de redes típicas. Suas aplicações práticas são encontradas em todas as áreas nas quais as redes complexas precisam ser modeladas. Um grande número de modelos de redes aleatórias é conhecido, o que reflete os diversos tipos de redes complexas

encontradas em áreas diferentes. Em um contexto matemático, o gráfico aleatório refere-se quase exclusivamente ao modelo gráfico aleatório Erdős; Rényi (1959).

### **3.4.2.2 *Redes Mundo Pequeno ou Small World***

Uma rede Mundo Pequeno é um tipo de gráfico matemático em que os nós, em sua maioria, não são vizinhos uns dos outros, mas a maioria dos nós pode ser alcançada pelos demais nós por um pequeno número de saltos ou passos. Especificamente, uma rede Mundo Pequeno é uma rede na qual a distância  $L$  (o número de etapas necessárias) entre os dois nós escolhidos aleatoriamente cresce proporcionalmente em escala com o crescimento logarítmico do número de nós da rede (WATTS; STROGATZ, 1998).

No contexto de uma rede social, isso resulta no fenômeno mundo pequeno, em que pessoas diversas estão ligadas por meio dos amigos dos seus amigos. Muitos gráficos empíricos são modelados de acordo com a característica de mundo pequeno. As redes sociais, a conectividade da internet, as wikis, como a Wikipedia, e as redes de genes, todas exibem características da rede Mundo Pequeno.

Certa categoria de redes Mundo Pequeno foi identificada como uma classe de grafos aleatórios por Duncan Watts e Steven Strogatz em 1998. Eles observaram que os gráficos poderiam ser classificados de acordo com duas características estruturais independentes: o coeficiente de agrupamento e o nó distância média de  $n$ -para-nó (também conhecido como comprimento do caminho médio mais curto).

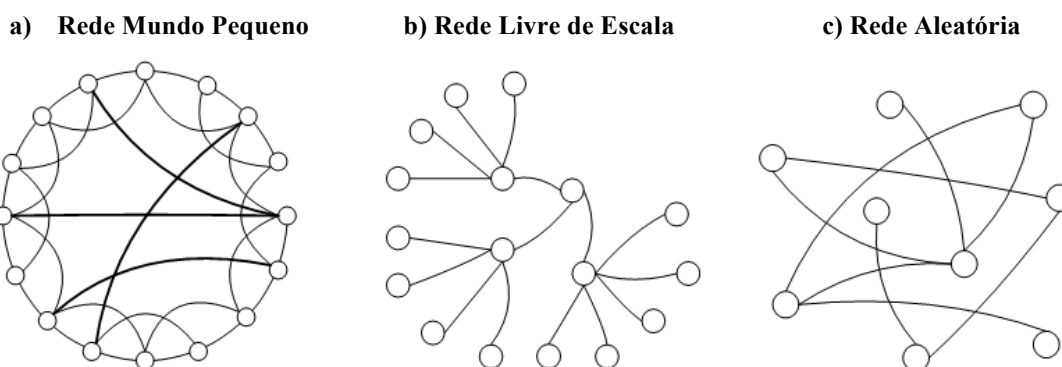
### **3.4.2.3 *Redes Livre de Escala ou Scale Free***

A característica mais notável em uma rede Livre de Escala é o aparecimento comum de nós ou vértices com um grau que excede em muito a média dos demais nós da rede. Os nós de mais alto grau são frequentemente chamados de hubs, e normalmente têm uma função e propósitos muito específicos em suas redes, mas isso pode variar bastante, dependendo do domínio.

A propriedade Livre de Escala se correlaciona fortemente com a robustez da rede quanto à resistência a ataques. Acontece que os principais hubs são seguidos de perto pelos menores. Esses centros menores, por sua vez, são seguidos por outros nós com um grau ainda menor, e assim por diante. Essa hierarquia permite um comportamento tolerante a falhas. Se as falhas ocorrem aleatoriamente e a vasta maioria dos nós são aqueles com pequeno grau, a probabilidade de que um hub seja afetado é quase insignificante. Mesmo que ocorra uma

falha em um dos hubs, a rede geralmente não vai perder sua conexão devido aos hubs restantes. No entanto, se escolhermos alguns grandes centros e os tirarmos da rede, a rede é transformada em um conjunto de grafos bastante isolados. Assim, as propriedades dos hubs são tanto uma força como uma fraqueza de redes Livre de Escala. A Figura 15 apresenta o exemplo de três redes segundo suas características estruturais.

**FIGURA 15. Estrutura de redes Mundo Pequeno, Livre de Escala e Aleatória**



Fonte: elaboração do autor com base em Chung-Yuan Huang (2005)

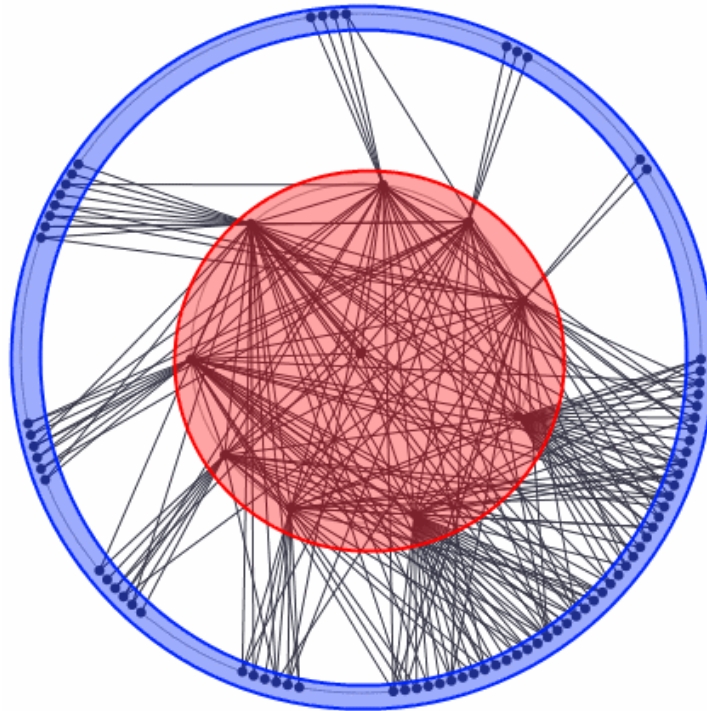
Legenda: a) rede Mundo Pequeno, em que qualquer nó pode ser alcançado pelos demais nós por meio de um número pequeno de passos; b) rede Livre de Escala, em que alguns poucos nós possuem um número muito mais elevado de relações que a maioria dos demais nós da rede; c) rede Aleatória, na qual as relações entre quaisquer dois nós seguem um padrão aleatório.

#### 3.4.2.4 *Redes centro-periferia*

Outro conceito comum, mas informal, em análise de redes sociais ou complexas é o de uma estrutura de núcleo e periferia (Figura 16). A concepção implica uma rede com um núcleo denso e coeso e uma periferia esparsa e desconectada. De acordo com Borgatti e Everett (2000), “a noção de uma rede Núcleo-Periferia pode ser vista como um centro com vários núcleos e uma comunidade periférica ligada aos diferentes núcleos”. Esse conceito pode ser avaliado de outra forma, buscando detectar o coeficiente do Clube dos Ricos (COLIZZA et al., 2006), que avalia a proporção de relacionamentos dentro e fora de grupos específicos. A ideia é que grupos com maior número de relacionamentos, ou mais “ricos”, se relacionam mais entre eles do que com atores fora desses grupos. Sob esse aspecto foi estudado como as redes de pesquisa se relacionam, e descobriu-se que há uma ligação direta dos grupos mais centrais com a grande maioria dos demais pesquisadores de uma área específica, podendo denotar controle de alguns grupos ou pesquisadores específicos, mas permitindo o acesso dos pesquisadores menos influentes a redes mais bem estruturadas.



**FIGURA 16. Rede centro-periferia**



Fonte: elaboração do autor com base em Borgatti e Everett (2000)

Legenda: rede núcleo-periferia gerada usando o modelo Erdos-Renyi modificado. Os nós na área vermelha pertencem ao núcleo, enquanto os nós na área azul estão em periferia.

#### **3.4.2.5 Formação de comunidades**

A estrutura de comunidades é uma característica comumente observada em redes em diversos contextos. Por meio do estudo da estrutura das comunidades de uma rede pode-se analisar o comportamento de suas unidades funcionais, possibilitando assim a compreensão do papel de cada elemento para o funcionamento da rede como um todo.

A estrutura de comunidades é a divisão de uma rede em grupos cujas conexões internas são densas e as conexões externas são mais esparsas. Entre as medidas de avaliação de estruturas de comunidades em redes complexas a mais utilizada atualmente é a modularidade (NEWMAN; GIRVAN, 2004). A ideia principal da modularidade é que um determinado subconjunto de vértices pode ser considerado uma comunidade se o número de conexões internas entre eles for maior que o número de conexões esperadas entre esses vértices no caso de uma formação aleatória na estrutura. Assim, se confirmada essa suposição, entende-se que há uma estrutura organizada que pode ser considerada uma comunidade.

### 3.4.2.6 *Sub-redes K-core*

O *k-core* é uma sub-rede em que cada nó se relaciona a outros nós, com pelo menos um número mínimo  $k$  de relacionamentos com outros nós nessa mesma sub-rede (NEWMAN, 2010). O *k-core* divide os grupos com base nos seus relacionamentos.

Os *softwares* Microsoft Excel,<sup>5</sup> R Studio<sup>6</sup> e Gephi<sup>7</sup> foram utilizados como ferramentas para a elaboração dos grafos para análise e cálculo das métricas de redes. As interpretações dos resultados deram-se com base na avaliação dos grafos e das medidas consideradas relevantes para essa análise de rede. Desse modo, os dados coletados foram analisados por meio de algoritmos e rotinas de análise de redes que permitiram verificar a possível propagação de informações da rede.

## 3.5 Organização dos capítulos de resultados e discussão

No *corpus* utilizado neste estudo são descritos os padrões de colaborações científicas das redes de pesquisadores em leishmaniose e tuberculose em três bases científicas distintas e ao longo de diferentes períodos de tempo. Cada capítulo da discussão teve uma base de dados como referência, sendo analisados os dois universos de pesquisa, LN e TB, nos três capítulos. A seguir o esboço acerca dos capítulos de discussão.

No capítulo sobre a base PubMed foi avaliado como ponto principal que a organização hierárquica nas redes de colaboração depende de alguns cientistas-chave com alto número de colaboradores, ou hubs, que orientam a arquitetura e o comportamento das redes. O excesso de ligações normalmente é resultante de pesquisadores mais antigos em termos de publicação e vida útil. Foi demonstrado também que as redes possuem uma arquitetura modular dividida em *clusters* e que essas comunidades científicas podem coexistir com um “clube dos ricos”, que seriam os hubs interligados. Na evolução temporal da estrutura das redes foi analisada a dinâmica de conectividade dentro e entre as comunidades, além da avaliação geográfica dos elementos da rede.

---

<sup>5</sup> Microsoft Excel é uma aplicação utilizada para criar, editar e processar planilhas, entre outras funções. <<http://office.microsoft.com/pt-br/excel/>>.

<sup>6</sup> RStudio é um ambiente de desenvolvimento integrado (IDE) para R. Ele inclui um console, um editor-realce de sintaxe que suporta a execução de código direta, bem como ferramentas para traçar, historiar, depurar e gerenciar o espaço de trabalho. Fonte: <<http://www.rstudio.com/>>.

<sup>7</sup> Gephi é uma ferramenta interativa de visualização e exploração para todos os tipos de redes complexas e gráficos dinâmicos, além de ser um *software* livre com muitas aplicações para análise e geração de grafos. Fonte: <<https://gephi.github.io/>>.

No capítulo sobre a Web of Science foi realizado outro recorte dentro do universo de pesquisa nas duas áreas, levando-se em consideração apenas pesquisadores de instituições brasileiras e suas relações com pesquisadores de instituições nacionais e internacionais. As relações entre os demais pesquisadores internacionais foram descartadas, a menos que houvesse algum pesquisador de instituição brasileira colaborando no artigo. Além disso, foi realizada uma análise mais detalhada sobre a Fundação Oswaldo Cruz (Fiocruz) e seus colaboradores. Como a WoS possui uma qualificação dos dados dos pesquisadores mais detalhada que a PubMed, esse recorte foi possível e nos permitiu identificar subgrupos de pesquisas mais refinados dos que os apresentados no estudo anterior. Os resultados foram apresentados levando-se em consideração o crescimento temporal ou a dinâmica dos dados em uma análise bibliométrica, utilizando a seguir os indicadores de análise de redes. Foram avaliadas as redes brasileiras e da Fiocruz, bem como as cooperações internacionais dessas duas redes. Por último foi feita uma avaliação dos subgrupos formados levando-se em consideração suas localizações geográficas, suas afiliações e as categorias definidas pela WoS para classificar os artigos publicados.

No capítulo sobre o SciELO foi feito um estudo levando-se em consideração as revistas utilizadas na publicação dos artigos. Como o SciELO é uma base brasileira que indexa prioritariamente revistas latino-americanas em geral, esse recorte ajudou-nos a entender melhor a situação dessa região, do Brasil e dos países vizinhos.

Primeiro foram avaliadas quais as revistas de maior relevância quanto ao número de publicações entre os pesquisadores nas duas áreas de pesquisa e a diferença de concentração dessas revistas científicas entre LN e TB. O segundo passo foi avaliar, com base em uma rede de dois modos, como essas revistas se relacionavam entre si quando a relação estabelecida eram os autores. Caso um mesmo pesquisador houvesse publicado em duas revistas, estas teriam um *link* entre si. O que se percebeu é que existe uma separação muito clara na escolha das revistas relacionada a sua posição geográfica e pouca interação entre pesquisadores brasileiros e revistas de países vizinhos.

## PARTE II - RESULTADOS

### 4 Resultados da pesquisa na base PubMed

Essa etapa da pesquisa teve como ponto de partida as publicações científicas coletadas da base de dados PubMed,<sup>8</sup> uma das mais importantes no estudo realizado em virtude de sua abrangência e seu elevado número de publicações nas áreas biomédica e de ciências da saúde. A base de dados PubMed, um serviço do National Center for Biotechnology Information (NCBI), indexa milhões de publicações da Medical Literature Analysis and Retrieval System Online (Medline) e outras revistas de ciências para artigos biomédicos.

Dentre os pontos fortes da PubMed destaca-se sua abrangência nos tópicos da saúde. Outras bases científicas possuem dados sobre pesquisas na área da saúde, mas nenhuma com os números e a diversidade apresentada na PubMed. Outro elemento de destaque é a utilização dos Medical Subject Headings (MeSH), um thesaurus e vocabulário controlado que tem como objetivo indexar os artigos da PubMed. O MeSH é mantido pela National Library of Medicine (NLM) dos Estados Unidos, sendo amplamente aceito como referência pela comunidade científica.

O ponto fraco da PubMed é a qualificação dos dados, pois em comparação com outras bases científicas, como a Web of Science, por exemplo, possui um menor controle sobre os autores e suas afiliações. Essa menor qualidade nos dados referente às filiações dos autores pode resultar em distorções quanto à avaliação de colaborações internacionais bem como à participação de países em publicações científicas. Segundo informações recentes, a PubMed tem como objetivo a melhor qualificação dos dados e dos artigos indexados a partir do ano de 2014.

Uma vez definida a base a ser utilizada na pesquisa foram iniciadas as buscas dos dados nas duas doenças foco do trabalho – leishmaniose (LN) e tuberculose (TB). O número de publicações sobre as duas doenças divergiu bastante, tendo a TB um número aproximadamente seis vezes maior de publicações que a LN. Essa diferença influenciou diretamente o período de avaliação aplicado às duas áreas, uma vez que a capacidade de processamento e tratamento dos dados passou a ser um elemento limitador da pesquisa. No caso da LN foram analisados os dados de publicações científicas de 1981 a 2012, com um total de 32 anos, enquanto para a TB foram analisadas as publicações científicas de 1997 a 2012, com um total de 16 anos. Como o foco do trabalho de pesquisa é a utilização dos

---

<sup>8</sup> <<http://www.ncbi.nlm.nih.gov/pubmed> accessed on october 2013>.

métodos de análise de redes e não a comparação direta das publicações científicas entre as duas áreas, acredita-se que essa diferença não tenha tido impacto relevante nos resultados apresentados.

Após a coleta dos dados, estes foram tratados de forma que as informações imprescindíveis à nossa análise pudessem ser reunidas. O objetivo foi transformar a lista de artigos em redes de colaboração, com os respectivos atributos que se fizeram necessários e possíveis em virtude da qualidade dos dados. A colaboração entre dois atores denota uma relação social entre eles, podendo com isso ser criada a rede de coautoria. Um dos fatores de impacto na criação das redes está relacionado à desambiguação dos nomes dos autores. A homonímia e a sinonímia possuem um grande impacto na análise das publicações científicas. No caso da homonímia, autores diferentes possuem uma mesma nomenclatura, o que pode ser acentuado em virtude das abreviações. No caso da sinonímia, um mesmo autor pode possuir diferentes estruturas no seu próprio nome e em suas diferentes publicações, sendo necessária uma revisão detalhada para se perceber quando dois nomes diferentes dizem respeito a uma mesma pessoa. O trabalho de desambiguação na base foi bastante extenso em virtude do grande número de publicações analisadas. No entanto, acredita-se que ainda existam problemas a esse respeito que não puderam ser tratados durante o percurso desta pesquisa.

No que diz respeito à análise geográfica, optou-se pelo aprofundamento somente nos casos dos pesquisadores mais importantes da rede. A base PubMed tem a tendência de publicar somente o país de afiliação do primeiro autor do artigo. Quando da análise dos pesquisadores mais influentes foi levado em consideração o país pelo qual o autor mais publicou. Nesse sentido, um autor brasileiro que trabalha em uma instituição de ensino e pesquisa dos Estados Unidos, por exemplo, terá seu trabalho vinculado ao país da instituição à qual está afiliado, neste caso os Estados Unidos. Não existem dados sobre a nacionalidade dos autores, e sim sobre a afiliação quando da publicação dos seus trabalhos.

Na análise das revistas científicas utilizadas para a publicação dos artigos foi apresentado apenas um resumo do número de publicações nas revistas com maior frequência e seu fator de impacto. Esses dados podem auxiliar quando da escolha de um universo de pesquisa mais focado. Um recorte de pesquisa mais objetivo sobre quais revistas são mais relevantes para cada área ajudaria a diminuir o elevado número de artigos com menor relevância. Outro dado importante nas pesquisas bibliográficas diz respeito ao fator de impacto e à notoriedade das referidas revistas. Esse fator pode influenciar bastante o entendimento sobre os assuntos em voga, levando-se em consideração que os pesquisadores mais renomados e possivelmente com maior conhecimento na área publicam nessas revistas.

Um dos estudos que nos pareceu bastante promissor, mas que não foi aprofundado, diz respeito ao uso das MeSH e sua evolução temporal pelos *clusters* ou pela comunidade de autores. Um dos pesquisadores que publica na área da Ciência da Informação e que produziu vários artigos sobre esse tópico foi Loet Leydesdorff. Acredita-se que essa área de pesquisa ainda pode ser muito explorada com a possibilidade de se melhor entender os rumos da pesquisa na área biomédica e como os *clusters* ou grupos de pesquisa podem ser representados quanto à utilização desses elementos.

Visando a um melhor entendimento este capítulo de pesquisa foi subdividido em etapas: na primeira parte constam os dados bibliométricos acerca das publicações científicas, como número de publicações, revistas utilizadas nessas publicações, número de autores no geral e por publicação, número de países com publicações registradas e palavras-chaves ou MeSH. A maior parte dessas análises é feita levando-se em consideração a evolução temporal e as mudanças percebidas durante o período analisado.

Na segunda parte do capítulo são apresentados os resultados das análises de rede nas duas áreas de pesquisa e suas similaridades no que diz respeito às redes de coautoria. São apresentados os indicadores de redes e a forma como a rede científica é coesa e bastante conectada. O resultado deste capítulo é um entendimento sobre a organização hierárquica da rede na qual alguns poucos pesquisadores possuem um número muito elevado de relações, podendo com isso exercer um controle direto sobre os rumos da pesquisa na área.

Na terceira parte são discutidos a formação de grupos e o controle que pode ser empregado na rede por meio dos seus atores principais ou hubs, como são denominados. É feita uma análise sobre as possibilidades de interação na estrutura centro-periferia, que foi encontrada na pesquisa na qual um grupo de pesquisadores situados de forma centralizada, segundo os conceitos de análise de redes, pode exercer influência direta nos demais integrantes dessa estrutura, que se encontram na periferia da rede, mas com uma relativa proximidade do centro. Relacionado a este conceito é analisado também o fenômeno “clube dos ricos”, do inglês “*rich club*”, no qual pessoas influentes tendem a se relacionar com mais frequência com outras pessoas também influentes.

Na quarta parte são debatidas a evolução das redes e a mudança nas suas estruturas. Uma visão dinâmica da rede é necessária para melhor entender as mudanças ocorridas no período analisado, mas principalmente para aumentar a capacidade de previsão das tendências futuras. A rede não é um elemento estático, e não utilizar uma avaliação temporal pode gerar uma visão distorcida quando da avaliação dos resultados.

No quinto subitem discute-se o fluxo de interação/comunicação entre os diferentes grupos ou *clusters* dentro da rede. Nessa análise é levada em consideração a localização geográfica da maioria dos atores que representam esses *clusters*. O entendimento sobre essa rede de colaboração entre os diferentes *clusters* é essencial na definição de padrões de colaboração. Estudos indicam (LEYDESDORFF; WAGNER, 2009) que o aumento na colaboração internacional ou entre grupos pode influenciar diretamente o exercício do poder no que diz respeito ao controle e à transmissão da informação no caso de uma rede científica de colaboração.

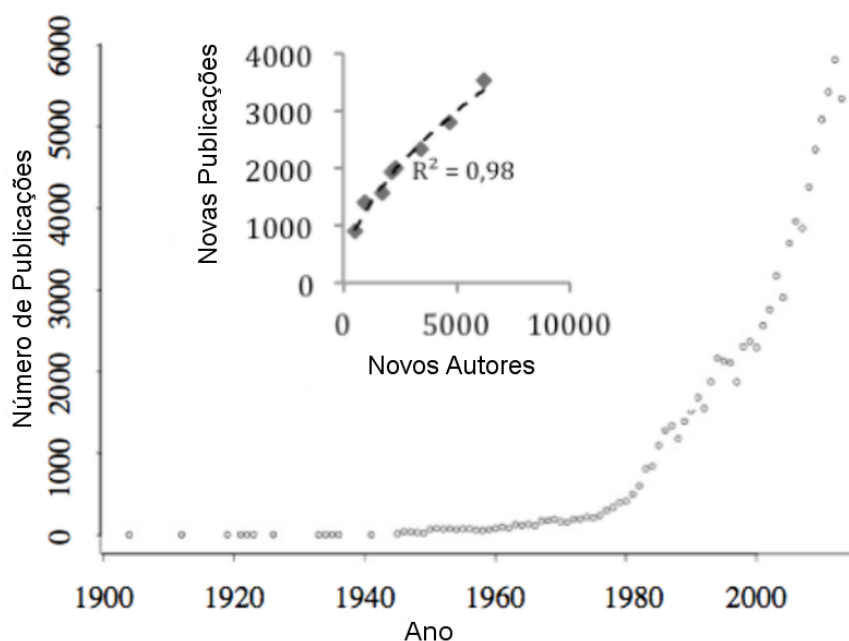
## 4.1 Resultados bibliométricos

### 4.1.1 Dados bibliométricos – leishmaniose

Após a coleta e o tratamento dos dados, foram analisados os resultados preliminares quanto às revistas científicas, os países e às palavras-chave utilizadas, apresentados a seguir. O total de artigos coletados durante o período entre 1981 e 2012 foi de 16.482, utilizando-se o termo de busca “leishmania\*” no título, no resumo e nas palavras-chaves. Além disso, foi feito um filtro com as palavras *clinical* e *preclinical*. Os artigos nos quais essas palavras apareciam no título ou no resumo foram retirados nesse momento por não serem relevantes ao estudo, que tinha como objetivo entender como a pesquisa e a inovação ocorrem em uma área específica. Além disso, em oito artigos o número de autores era muito elevado, chegando, em alguns casos, a mais de cem. Esses artigos foram avaliados e descartados para evitar distorção na análise da coautoria e por não apresentarem pesquisas com caráter científico, mas sim testes clínicos.

A escolha de um intervalo de 32 anos para o estudo da rede LN foi determinada pelo súbito crescimento de publicações a partir do ano de 1981 e compreende a recente década de aumento nos investimentos em pesquisa de acordo com recentes políticas públicas para pesquisa (Figura 17). De acordo com estudos sobre redes de coautoria (CLAUSET; SHALIZI; NEWMAN, 2009), observou-se também que os números anuais de novos artigos seguem uma escala da lei da potência com o número correspondente de novos autores, com  $\Delta\text{Publicações} = 30,7 (\Delta\text{Autores})^{\alpha = 0,54}$  (Figura 17). O expoente  $\alpha < 1$  significa retornos decrescentes de escala, que destaca que os novos artigos exigem cada vez maiores esforços em termos de número de cientistas no campo.

**FIGURA 17. Número de publicações científicas por ano (gráfico externo).  
Número de publicações científicas por autor (gráfico interno)**



Fonte: elaboração do autor

Legenda: número de artigos publicados por ano sobre leishmaniose no período de 1904 a 2012. O gráfico interno exibe a curva de produtividade em pesquisa na área de leishmaniose no período de 1981 a 2012 em uma escala de potência.

O número de revistas científicas que tiveram esses artigos publicados foi de 1.498, e apenas 195 dessas revistas apresentavam dez ou mais artigos. Uma análise sobre os tipos de revista foi bastante relevante, uma vez que existem revistas presentes durante todo o período pesquisado, e outras apenas em períodos específicos. Acredita-se que uma análise nesses moldes deve ser realizada, levando-se em consideração a temporalidade e a utilização das palavras-chave, que podem demonstrar, ou não, um caráter inovador. No entanto, esse estudo não foi efetivado neste trabalho.

As dez primeiras revistas mais representativas no que diz respeito ao número de publicações são: *Molecular and Biochemical Parasitology*, com 612 publicações e avaliação WebQualis variando entre B1 e A2 para as áreas de Ciências Biológicas e Saúde Coletiva respectivamente; *Transactions of The Royal Society of Tropical Medicine and Hygiene*, com 543 publicações e avaliação WebQualis variando entre B2, B1 e A2 para as áreas de Ciências Biológicas, Medicina e Saúde Coletiva respectivamente; *Infection and Immunity*, com 383 publicações e avaliação WebQualis entre A2 e A1 para Ciências Biológicas e Medicina; *Journal of Immunology*, com 380 publicações e WebQualis variando de C1 para Ciências Biológicas e A1 para Medicina, com o fator relevante de que existem diversas revistas com o título similar mas por país; *The American Journal of Tropical Medicine and Hygiene*, com



379 publicações e extrato WebQualis B1 para Química e A2 para Engenharias; *Memórias do Instituto Oswaldo Cruz*, com 363 publicações e extrato WebQualis variando entre B2 e A1 em diversas áreas da saúde. Essa revista tem um fator importante na pesquisa por ser brasileira, revelando a relevância desse tópico para o país, que é um dos principais atores nessa área; *The Journal of Biological Chemistry*, com 298 artigos e extrato WebQualis variando entre A2 e A1 em diversas áreas biomédicas; *Experimental Parasitology*, com 294 publicações e extrato WebQualis B1; *Annals of Tropical Medicine and Parasitology*, com 249 publicações e extrato WebQualis B2; e por último, na lista das dez revistas com maior número de publicações, *Acta Tropical*, com 184 publicações sobre leishmaniose e extrato WebQualis entre B2 e A1 para Ciências Biológicas e Medicina respectivamente. Essas revistas mantiveram um alto número de publicações durante todo o período estudado.

A Tabela 1 apresenta uma lista com as 120 revistas com maior número de publicações na área; ressalte-se que todas tiveram mais de trinta publicações durante os 32 anos analisados.

**TABELA 1. Revistas científicas com maior número de publicações sobre leishmaniose**

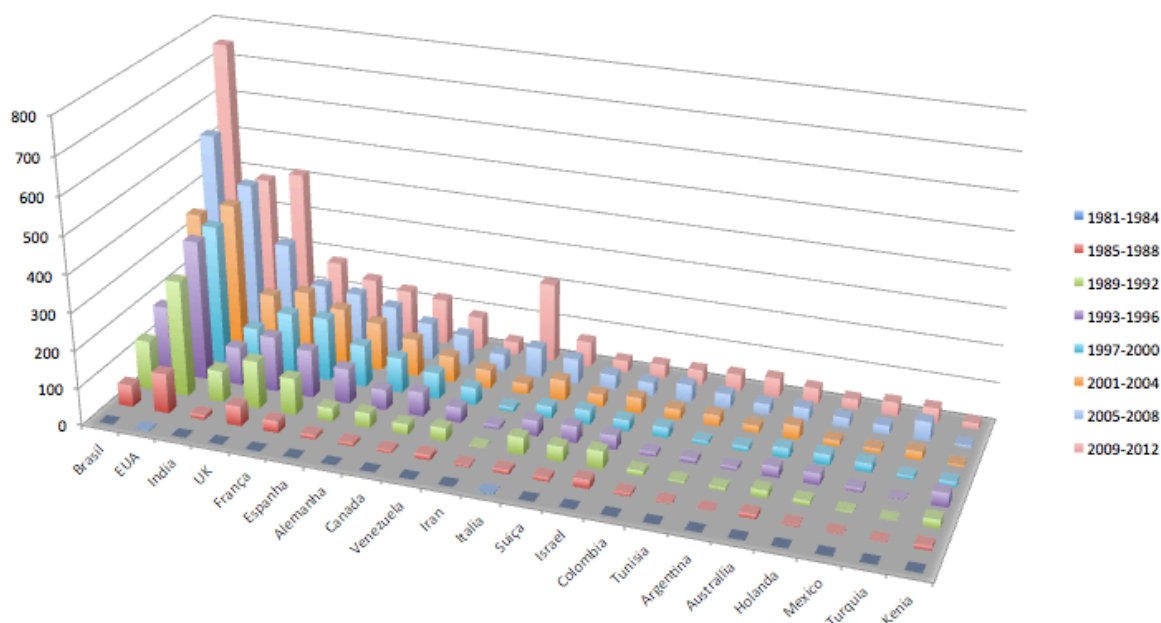
1 - <i>Molecular and Biochemical Parasitology</i>	61 - <i>Bulletin de la Societe de Pathologie Exotique (1990)</i>
2 - <i>Transactions of The Royal Society of Tr Md Hy</i>	62 - <i>Brazilian Journal of Medical and Biological Research</i>
3 - <i>The American Journal of Tropical Medicine and Hygiene</i>	63 - <i>Planta Médica</i>
4 - <i>Experimental Parasitology</i>	64 - <i>Emerging Infectious Diseases</i>
5 - <i>Memórias do Instituto Oswaldo Cruz</i>	65 - <i>Medecine Tropicale: Revue du Corps dd Sante Colonial</i>
6 - <i>Journal of Immunology</i>	66 - <i>Revista De Saúde Pública</i>
7 - <i>Infection and Immunity</i>	67 - <i>Gene</i>
8 - <i>Annals of Tropical Medicine and Parasitology</i>	68 - <i>Journal of Natural Products</i>
9 - <i>The Journal of Biological Chemistry</i>	69 - <i>The British Journal of Dermatology</i>
10 - <i>Acta Tropical</i>	70 - <i>Parasitology International</i>
11 - <i>Meditinskaiia Parazitologiia I Parazitarnye Bolezni</i>	71 - <i>Journal of Ethnopharmacology</i>
12 - <i>Parasitology Research</i>	72 - <i>Journal of Leukocyte Biology</i>
13 - <i>Revista da Sociedade Brasileira de Medicina Tropical</i>	73 - <i>Bulletin of the World Health Organization</i>
14 - <i>Parasitology</i>	74 - <i>Hospital (Rio de Janeiro, Brazil)</i>
15 - <i>Antimicrobial Agents and Chemotherapy</i>	75 - <i>Scandinavian Journal of Immunology</i>
16 - <i>The Journal of Parasitology</i>	76 - <i>European Journal of Biochemistry/Febs</i>
17 - <i>Plos Neglected Tropical Diseases</i>	77 - <i>Biomedica: Revista del Instituto Nacional de Salud</i>
18 - <i>International Journal for Parasitology</i>	78 - <i>The Journal of Eukaryotic Microbiology</i>
19 - <i>European Journal of Immunology</i>	79 - <i>Plos Pathogens</i>
20 - <i>Veterinary Parasitology</i>	80 - <i>Molecular and Cellular Biology</i>
21 - <i>Journal of The Egyptian Society Of Parasitology</i>	81 - <i>Immunology</i>
22 - <i>Revista do Instituto de Medicina Tropical de São Paulo</i>	82 - <i>Eukaryotic Cell</i>
23 - <i>Parasite Immunology</i>	83 - <i>Archives of Dermatology</i>
24 - <i>The Journal of Infectious Diseases</i>	84 - <i>The Journal of Tropical Medicine and Hygiene</i>
25 - <i>Pnas of Usa</i>	85 - <i>Nature</i>
26 - <i>Plos One</i>	86 - <i>Cellular Microbiology</i>
27 - <i>Journal of Medical Entomology</i>	87 - <i>Advances in Experimental Medicine and Biology</i>
28 - <i>Vaccine</i>	88 - <i>The Journal of Communicable Diseases</i>
29 - <i>The Journal of Experimental Medicine</i>	89 - <i>Journal of Cell Science</i>
30 - <i>International Journal of Dermatology</i>	90 - <i>Immunology Letters</i>
31 - <i>Clinical and Experimental Immunology</i>	91 - <i>Journal of Vector Ecology</i>
32 - <i>The Journal of Protozoology</i>	92 - <i>Fems Microbiology Letters</i>

33 - <i>Nucleic Acids Research</i>	93 - <i>Biochemistry</i>
34 - <i>Parasitology Today</i>	94 - <i>The Veterinary Record</i>
35 - <i>Bulletin de la Societe de Pathologie Exotique et de Ses Filiales</i>	95 - <i>The Embo Journal</i>
36 - <i>Journal of Clinical Microbiology</i>	96 - <i>Biochemical Pharmacology</i>
37 - <i>Biochimica et Biophysica Acta</i>	97 - <i>Parazitologiia</i>
38 - <i>The Biochemical Journal</i>	98 - <i>Molecular and Cellular Biochemistry</i>
39 - <i>Trends in Parasitology</i>	99 - <i>Journal of the American Academy of Dermatology</i>
40 - <i>Parasites &amp; Vectors</i>	100 - <i>Journal of Molecular Epidemiology and Evol Genetic Diseases</i>
41 - <i>Medical and Veterinary Entomology</i>	101 - <i>Archives de L'institut Pasteur de Tunis</i>
42 - <i>Clinical Infectious Diseases</i>	102 - <i>Enfermedades Infecciosas Y Microbiologia Clinica</i>
43 - <i>Bioorganic &amp; Medicinal Chemistry</i>	103 - <i>BMC Genomics</i>
44 - <i>Biochemical and Biophysical Research Communications</i>	104 - <i>Vector Borne and Zoonotic Diseases</i>
45 - <i>Journal of Medicinal Chemistry</i>	105 - <i>The Journal of Infection</i>
46 - <i>The Journal of Antimicrobial Chemotherapy</i>	106 - <i>Medical Microbiology and Immunology</i>
47 - <i>Tropical Medicine &amp; International Health: Tm &amp; Ih</i>	107 - <i>Journal of Vector Borne Diseases</i>
48 - <i>Microbes and Infection/Institut Pasteur</i>	108 - <i>Febs Letters</i>
49 - <i>European Journal of Medicinal Chemistry</i>	109 - <i>Aids (London, England)</i>
50 - <i>Molecular Microbiology</i>	110 - <i>Veterinary Immunology and Immunopathology</i>
51 - <i>Annales de Parasitologie Humaine et Comparee</i>	111 - <i>Cellular Immunology</i>
52 - <i>Cadernos de Saúde Pública</i>	112 - <i>Vestnik Dermatologii I Venerologii</i>
53 - <i>Parassitologia</i>	113 - <i>Presse Medicale (Paris, France, 1983)</i>
54 - <i>Parasite (Paris, France)</i>	114 - <i>Phytotherapy Research: Ptr</i>
55 - <i>Chinese Journal of Parasitology &amp; Parasitic Diseases</i>	115 - <i>Phytochemistry</i>
56 - <i>The Indian Journal of Medical Research</i>	116 - <i>Zeitschrift Fur Parasitenkunde (Berlin, Germany)</i>
57 - <i>Bioorganic &amp; Medicinal Chemistry Letters</i>	117 - <i>Medicina Clinica</i>
58 - <i>Lancet</i>	118 - <i>Eastern Mediterranean Health Journal</i>
59 - <i>East African Medical Journal</i>	119 - <i>Clinical and Diagnostic Laboratory Immunology</i>
60 - <i>Science (New York, NY)</i>	120 - <i>Turkiye Parazitolojii Dergisi</i>

Fonte: elaboração do autor

Outra análise realizada com base nos resultados bibliométricos diz respeito ao número de países representados pelos pesquisadores, com um total de 118 países. Na Figura 18 é mostrada a evolução por períodos de 21 países, tendo Brasil, Índia e Irã um crescimento acima da média no último período – 2009 a 2012. Esse é um fator fundamental para entender a mudança no processo produtivo.

**FIGURA 18. Número de publicações científicas sobre leishmaniose por país, por período**



Fonte: elaboração do autor

Legenda: número de artigos publicados nos últimos 32 anos por período de quatro anos para os 21 países com maior número de publicações na área de leishmaniose.

Alguns países europeus e os EUA, apesar de grande participação na pesquisa científica na área da leishmaniose, mantiveram constante o número de publicações sobre esse assunto nos últimos três ou quatro períodos (12 a 16 anos). Ao levar em consideração os locais de ocorrência da doença e suas variações, pode-se entender essa mudança como um resultado positivo, na medida em que países mais afetados pelo problema estão investindo mais em pesquisa.

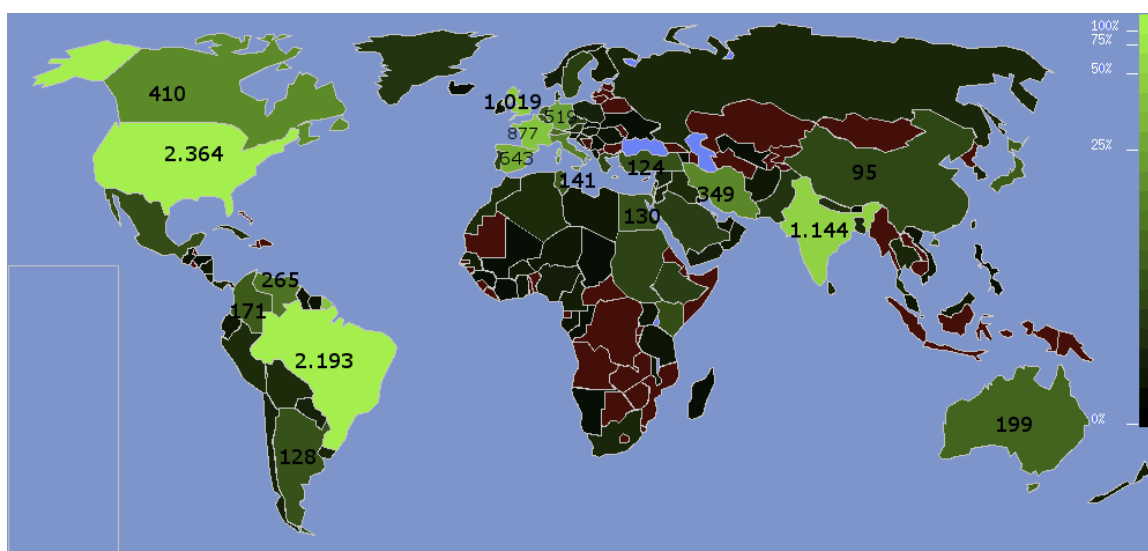
Para melhor entender os motivos dessa mudança de padrão de pesquisa nos últimos anos faz-se necessário um maior aprofundamento da pesquisa, pois aqui foram levantadas apenas algumas possibilidades. O primeiro fator está relacionado ao crescimento de países como Brasil e Índia, que com o passar dos anos tiveram um maior incentivo à pesquisa, aumentando com isso a contribuição internacional. Como esses países sofrem diretamente as consequências da leishmaniose, investimentos nessa área foram necessários, o que repercutiu na produção científica.

Quanto à manutenção dos níveis de publicação nos países desenvolvidos, vários fatores podem ser levados em consideração, como, por exemplo, a recente crise econômica, que pode ter afetado a disponibilização de capital para pesquisa, ou mesmo o interesse em outras áreas de pesquisa, uma vez que as doenças negligenciadas não fazem parte dos

problemas existentes em seu território, ou até mesmo o fato de os pesquisadores que antes trabalhavam somente nos centros de pesquisa desses países terem encontrado espaço em outras regiões.

Na Figura 19 pode-se observar um mapa com todos os países e o número total de suas publicações durante os trinta anos analisados. Os países coloridos com o tom de verde mais claro possuem maior produção. Alguns países tiveram o número de produções científicas descrito no mapa para melhor compreensão dos desvios de tonalidade.

**FIGURA 19. Mapa-múndi da quantidade de publicações científicas sobre leishmaniose por país**



Fonte: elaboração do autor

Legenda: mapa-múndi com diferenças de tonalidade do verde mais claro para o mais escuro relacionadas ao número de publicações por país. Os países na cor vermelha não tiveram publicações registradas no período analisado. Quanto mais claro o verde, maior o número absoluto de publicações para o país.

Foram analisadas também as palavras-chave, ou MeSH, que chegaram a um total de 7.356. Destas, 2.425 foram utilizadas apenas uma vez, e 1.072, apenas duas vezes. No outro espectro verificou-se que quatro MeSH tiveram mais de 10 mil aparições, sendo elas *immunology*, *metabolism*, *animals* e *genetics*. Um dos objetivos da pesquisa foi entender como são produzidos novos conhecimentos científicos em uma área específica. Partiu-se do pressuposto de que os MeSH denotam o conhecimento utilizado nos respectivos trabalhos. Na Tabela 2 é apresentado o número de palavras-chave por período, o número de novas palavras naquele mesmo período comparado ao período anterior e dessas, quantas apareceram apenas uma vez naquele período.

**TABELA 2. Número de MeSH por período em pesquisas sobre leishmaniose**

Período	Número de MeSH	Número de novas MeSH	Novo MeSH uma única vez
1981-1984	1.272	---	---
1985-1988	1.728	864	126
1989-1992	2.086	797	158
1993-1996	2.483	858	206
1997-2000	2.860	801	248
2001-2004	3.293	845	340
2005-2008	3.895	1.000	565
2009-2012	4.145	913	693

Fonte: elaboração do autor

Pode-se perceber que, apesar do crescimento do número de publicações, conforme apresentado na Figura 20, e conseqüentemente do número total de MeSH por período apresentado na Tabela 2, o número de novos MeSH não tem acompanhado a mesma proporção. Ademais, existe o agravante quanto ao aumento de palavras-chave que apareceram apenas uma única vez durante um período de quatro anos, o que pode demonstrar sua baixa relevância.

Um melhor entendimento sobre a utilização dos MeSH só pode ocorrer com o acompanhamento de um especialista na área, o qual opinará sobre a relevância ou não de certas palavras, tanto as muito utilizadas, como *imunology*, como aquelas que aparecem apenas uma vez durante certo período.

Outro elemento importante nessa análise é a forma como os MeSH são organizadas de forma hierárquica. O padrão de organização definido pela NLM faz com que palavras (conceitos) mais abrangentes fiquem no topo da lista, e assuntos mais específicos sejam colocados em um nível hierárquico inferior. Uma vez que uma palavra de um nível menor é utilizada, seus antecessores também são utilizados para representar aquele artigo. Isso explica o motivo pelo qual palavras como *immunology* chegam a aparecer mais de 10 mil vezes em um universo de 16.482 artigos.

#### 4.1.2 Dados bibliométricos – tuberculose

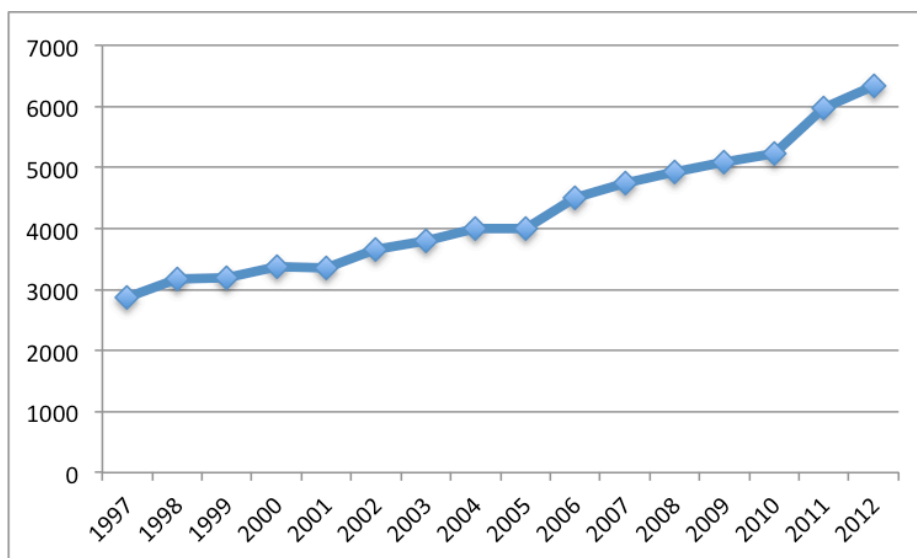
O número de publicações sobre tuberculose é mais de seis vezes maior que o número de publicações sobre leishmaniose. Além disso, a TB não chega a ser considerada uma doença negligenciada segundo a lista da Organização Mundial da Saúde,<sup>9</sup> com acometimentos da doença nas mais diversas áreas do planeta. Dessa maneira, outros países no mundo também

<sup>9</sup> Doenças consideradas negligenciadas pela Organização Mundial da Saúde. <[http://www.who.int/neglected\\_diseases/diseases/en/](http://www.who.int/neglected_diseases/diseases/en/)>.

possuem grande interesse na pesquisa desta doença. Outro elemento importante quanto ao número de publicações está no fato de que países do bloco dos Brics (Brasil, Rússia, Índia, China e África do Sul) também participam de forma efetiva quanto à pesquisa nessa área, aumentando o número de países interessados e com capital investido na pesquisa, não se restringindo apenas aos EUA e aos países da Europa ocidental. No Brasil, a TB é considerada uma das doenças negligenciadas e de grande interesse para os órgãos da saúde quanto ao seu tratamento e controle.

O acesso à base PubMed para coleta dos dados da TB ocorreu em outubro de 2014, e se decidiu por manter a data final em 2012 para se ter um paralelo com os dados da LN. Como elemento de comparação, o número de artigos encontrados com a pesquisa ou unitermo “tuberculosis” foi de 103.162 no mesmo período de estudo da LN, que teve 16.482. Assim, resolveu-se restringir o período analisado da TB para 16 anos em vez de 32 anos. O número de artigos então ficou reduzido a 68.958.

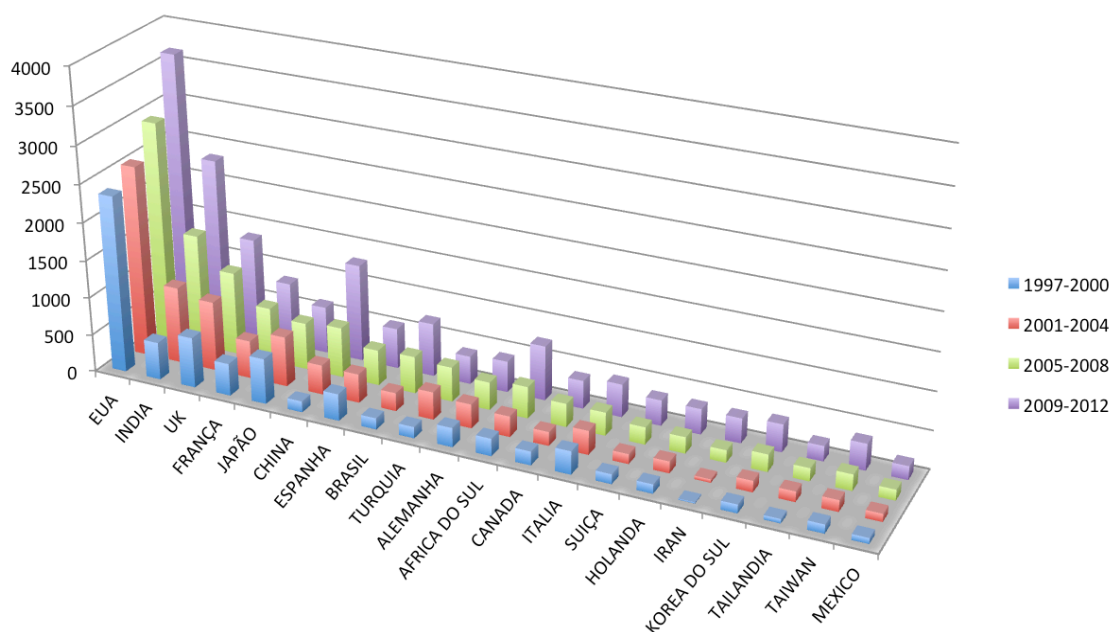
A Figura 21 demonstra o crescimento no número de publicações por ano durante o período analisado. Pode-se perceber que esse número mais que dobrou: de 2.868 publicações em 1997 para 6.328 publicações no ano de 2012. Deve-se levar em consideração nesse aumento que o número de publicações no mundo também cresceu. Segundo estudos recentes que analisaram o crescimento das publicações científicas no período de 1980 a 2012 (BORNMANN; MUTZ, 2014), as publicações em geral pouco mais que dobraram nesse espaço de tempo. No entanto, nos dados apresentados sobre a TB essa duplicação quanto ao número de publicações ocorreu em praticamente metade do tempo, fato apresentado no estudo de Bormann e Mutz (2014).

**FIGURA 20. Número de publicações científicas sobre tuberculose por ano**

Fonte: elaboração do autor

Uma segunda análise quanto ao número de publicações foca nos países onde os pesquisadores estão afiliados por meio das suas instituições, lembrando sempre do limitador do item afiliação na base PubMed. Foi feita uma análise em períodos de quatro anos na qual os países que aparecem com o maior número de publicações são os EUA, a Índia e o Reino Unido (UK), situação exposta na Figura 21. O Brasil, nesta análise, encontra-se em oitavo lugar quanto ao número de publicações, diferente do que ocorreu em LN – considerando esta doença ele está ou em primeiro ou em segundo lugar, dependendo do recorte temporal. Ao focar no crescimento percentual entre os dois últimos períodos (2005 a 2008 e 2009 a 2012), pode-se perceber que alguns países tiveram um crescimento mais acentuado que os demais, sendo estes Índia, China, Brasil e África do Sul, justamente os países pertencentes ao bloco dos Brics. Isso quer dizer que se for mantida a taxa de crescimento do Brasil, nos próximos anos ele pode se tornar mais atuante quanto ao número de publicações científicas em comparação com outros países como França, Japão e Espanha, que apresentaram um crescimento no último período não tão acentuado. Apesar de a Rússia não aparecer entre os vinte países que mais publicam, ela possui revistas que estão entre as que mais possuem artigos na área. Esse fato pode indicar um problema quanto à afiliação dos pesquisadores russos em decorrência da semântica.

**FIGURA 21. Número de publicações científicas sobre tuberculose por país, por período**



Fonte: elaboração do autor

Legenda: número de artigos publicados nos últimos 16 anos por período de quatro anos para os vinte países com maior número de publicações na área de tuberculose.

Quanto ao estudo sobre as revistas científicas com maior número de publicações, foram encontradas algumas diferenças interessantes a respeito das revistas com artigos em LN. Um dos primeiros fatores que chamaram a atenção está relacionado ao número de revistas específicas em TB e de países diferentes dos países da LN, como Japão, Rússia e China, entre as vinte revistas com maior número de publicações na área. Outro dado importante encontrado diz respeito ao surgimento e ao desaparecimento de algumas revistas durante o período de análise, como a *Plos One*, que até 2005 não contabilizava nenhuma publicação na área de TB, mas que em 2012 teve o maior número de publicações, com 381 somente naquele ano. Em contrapartida, a revista *Problemy Tuberkuleza*, que era bastante utilizada durante os anos de 1997 a 2003, deixou de ter publicações a partir de então, provavelmente por ter mudado de nome para *Problemy Tuberkuleza i Boleznei Legkikh*, mas que também deixou de publicar na área de TB em 2009. A Tabela 3 exhibe a evolução do número de publicações nas vinte revistas que tiveram o maior número de publicações em TB durante o período analisado.



TABELA 3. Número de publicações por ano das vinte revistas mais bem cotadas

	012	011	010	009	008	007	006	004	005	003	002	000	001	999	998	997
<i>The International Journal of Tuberculosis and Lung Disease</i>	97	87	49	47	39	98	24	20	34	37	82	32	87	12	02	14
<i>Journal of Clinical Microbiology</i>	7	8	8	4	0	8	6	0	1	9	9	2	1	5	2	5
<i>Kekkaku: [Tuberculosis]</i>	6	9	6	7	9	9	2	8	0	4	0	7	8	5	0	1
<i>Problemy Tuberkuleza I Boleznei Legkikh</i>				33	36	50	38	55	48	22						
<i>Plos One</i>	81	39	26	00	3	6										
<i>Problemy Tuberkuleza</i>										0	96	03	81	06	28	56
<i>Tuberculosis (Edinburgh, Scotland)</i>	4	10	4	8	05	9	3	4	8	2	9		7			
<i>Infection And Immunity</i>	9	0	1	7	8	3	6	1	9	8	0	3	0	0	5	8
<i>Clinical Infectious Diseases</i>	6	6	7	4	0	7	6	3	9	8	3	3	2	8	1	4
<i>Lancet</i>	4	9	6	9	4	4	9	9	2	0	0	1	9	8	9	5
<i>American J of Respiratory and Critical Care Medicine</i>	4	2	4	6	7	5	6	9	7	3	5	5	7	6	4	3
<i>Journal of Immunology</i>	8	1	2	2	1	2	2	3	2	9	6	3	5	9	4	6
<i>Antimicrobial Agents And Chemotherapy</i>	5	5	3	7	5	4	7	8	3	1	8	9	0	0	8	8
<i>Chinese Journal of Tubercul and Respiratory Diseases</i>	4	2	8	0	4	6	2	6	8	7	6	1	6	2	1	7
<i>The Journal of Infectious Diseases</i>	5	7	8	5	0	4	1	3	9	5	1	8	8	0	0	3
<i>The Journal of Biological Chemistry</i>	8	0	6	6	3	3	2	7	8	5	3	8	2			
<i>The European Respiratory Journal</i>	4	8	0	5	4	9	2	3	7	3		2		0	0	9
<i>Chest</i>	1		3	4	0	7	6	3	7	9	7	6	6	1	2	4
<i>Journal of Bacteriology</i>	7	6	6	1	9	3	0	3	8	2	5	3	6	3	4	2
<i>Vaccine</i>	0	1	3	0	1	9	7	5	8	8		1	1			

Além da análise quanto ao número de publicações, aos países que mais publicaram nos últimos anos e às revistas mais utilizadas, foi feita também uma análise da MeSH por período. As semelhanças com a LN são grandes, uma vez que na escala hierárquica definida no vocabulário controlado da NLM as duas doenças possuem diversos fatores em comum. Existem palavras como imunologia, metabolismo e genética, por exemplo, que aparecem tanto em TB como em LN entre as mais frequentes. A Tabela 4 apresenta os cem MeSH mais frequentes em ordem decrescente geral. Dessa maneira, pode-se notar que algumas palavras começaram a ser mais utilizadas nos últimos anos em detrimento de outras, mesmo que no geral elas apareçam em uma sequência inferior. Isso pode ser notado com bastante frequência nos MeSH *young adult* ou *drug resistant*, que ocorrem mais vezes nos últimos anos que palavras com similar número de ocorrências no geral.





	11	49	12	45	85	65	71	49	33	48	05	08	20	8	28	04
<i>Biosynthesis</i>	42	81	71	71	87	02	11	51	71	69	87	64	39	40	02	12
<i>Bcg vaccine</i>	44	67	76	70	69	62	88	83	51	37	49	58	88	30	19	19
<i>DNA, bacterial</i>	02	76	77	21	45	53	67	67	44	50	48	13	23	28	41	34
<i>Pathogenicity</i>	64	63	82	75	70	53	58	65	70	68	31	04	01	3	6	5
<i>Molecular sequence data</i>	56	70	54	66	67	31	48	75	47	42	46	27	25	29	17	8
<i>Drug therapy, combination</i>	48	67	31	26	35	47	22	22	53	39	28	17	29	36	66	14
<i>Lung</i>	60	87	47	80	63	34	44	27	47	42	15	9	15	07	03	9
<i>Mortality</i>	92	89	51	57	60	42	43	27	35	14	25	08	06	13	5	03
<i>Follow-up studies</i>	72	21	48	50	32	43	29	28	21	23	19	01	5	23	44	32
<i>Growth &amp; development</i>	90	82	36	37	29	36	38	25	18	09	36	13	17	04	5	0
<i>Prospective studies</i>	00	79	69	57	42	23	24	5	7	07	2	7	18	0	02	9
<i>Anti-bacterial agents</i>	88	72	60	54	38	40	20	26	23	02	00	8	9	4	3	9
<i>Standards</i>	54	81	29	26	30	34	36	08	05	09	19	08	5	2	1	3
<i>Physiopathology</i>	12	65	36	21	48	66	46	16	24	07	01	9	9	6	5	7
<i>Drug resistance, bacterial</i>	16	98	51	64	28	48	39	41	45	11	02	9				
<i>Mycobacterium</i>	23	36	53	31	20	08	5	3	8	7	6	2	02	7	04	11
<i>Economics</i>	79	59	61	27	03	07	14	6	9	7	7	7	0	7	4	3
<i>Cattle</i>	51	58	12	29	13	13	68	07	9	2	5	9	2	2	7	6
<i>Macrophages</i>	60	49	26	18	42	27	15	8	0	24	0	7	7	6	6	2
<i>India</i>	85	80	36	45	32	29	3	5	1	04	5	8	6	1	2	2
<i>Amino acid sequence</i>	10	18	11	17	16	00	06	15	9	09	4	5	9	9	9	4
<i>Antagonists &amp; inhibitors</i>	03	81	58	37	29	49	09	4	8	2	1	2	6	7	5	5
<i>Magnetic resonance imaging</i>	25	20	22	2	8	10	14	1	7	4	9	7	0	1	3	2
<i>Organization &amp; administration</i>	36	24	28	09	20	12	7	9	10	02	1	0	7	1	7	6
<i>Genotype</i>	13	80	61	44	21	05	11	4	6	6	6	2	8	2	2	
<i>Veterinary</i>	55	53	17	21	4	1	24	7	2	3	1	7	9	2	3	8
<i>Mutation</i>	32	28	9	15	08	00	13	00	7	8	8	2	8	6	6	9
<i>Mice, inbred C57Bl</i>	12	11	9	08	12	0	06	3	5	1	6	6	9	9	9	2
<i>Cross-sectional studies</i>	96	69	32	26	07	7	7	3	6	0	1	1	1	6	0	2
<i>History</i>	8	04	2	9	6	2	6	02	5	8	00	4	6	4	1	6
<i>Enzyme-linked immunosorbent assay</i>	3	00	18	42	28	25	0	7	8	8	1	5	4	2	3	1

## 4.2 Organização hierárquica da rede

As redes estudadas apresentam uma estrutura hierárquica baseada nos tipos de relações que existem ou são construídas pelos pesquisadores ao longo de sua carreira científica. Quanto maior o número de publicações maior o número de colaborações de um autor, levando-se em consideração o número médio de colaboradores por artigo, que tem aumentado nos últimos anos. Com o crescimento gradativo do número de publicações e colaborações um pesquisador passa a exercer maior controle sobre sua rede de colaboração direta e indireta, uma vez que este se encontra exposto a uma diversidade maior de conhecimento e de parceiros de pesquisa.

Para entender as estruturas hierárquicas de pesquisadores influentes, faz-se necessário compreender as medidas de rede que podem contribuir com essa análise. Duas medidas importantes no sentido da formação de grupos são o coeficiente de clusterização e a assortatividade. O coeficiente de clusterização demonstra qual é a probabilidade média de que dois coautores também sejam coautores de um terceiro pesquisador, e a assortatividade, que é o coeficiente de correlação de Pearson, analisa a possibilidade de que dois pesquisadores com relacionamentos similares tenham relações entre si. Outras estatísticas dizem respeito ao número de colaboradores ou ao número de colaborações, a chamada centralidade de grau, onde no segundo caso são levadas em consideração as múltiplas colaborações com um mesmo autor.

Essas medidas demonstram que quanto mais colaborações um autor possui maior a chance de que novas colaborações se estabeleçam. Nesta análise não foi levado em consideração o ciclo de vida de um pesquisador, que no final da carreira tende a diminuir o número de publicações devido a fatores como aposentadoria ou mesmo ao aumento de atividades gerenciais em alguns casos.

Um dos elementos encontrados nas pesquisas e que corrobora estudos de coautoria (NEWMAN, 2004; BARABÁSI, 2002) é que a rede de pesquisadores em ciência tende a ser muito coesa e com muitos relacionamentos em todas as esferas. O fato de que a maioria dos cientistas não funciona de forma isolada da corrente principal da sua área de pesquisa é reconfortante. Contudo, é imprescindível uma análise mais aprofundada para perceber se esse forte relacionamento tende a ajudar ou não a evolução do conhecimento de áreas específicas.

Dando continuidade ao trabalho de análise da colaboração, foca-se na distribuição da centralidade de grau, que é uma propriedade quantitativa básica de redes que permite revelar se estas possuem características de interação em que um pequeno número de nós ou hubs

possui uma alta centralidade de grau enquanto a grande maioria da rede possui poucas colaborações. Esta propriedade é importante porque muitos comportamentos de sistemas em rede são dominados por seus hubs. Esses resultados definem uma hierarquia topológica em que os autores, ou hubs, com muitas relações com outros autores estão no topo da cadeia de conhecimento, influência e poder.

#### 4.2.1 Medidas da rede leishmaniose

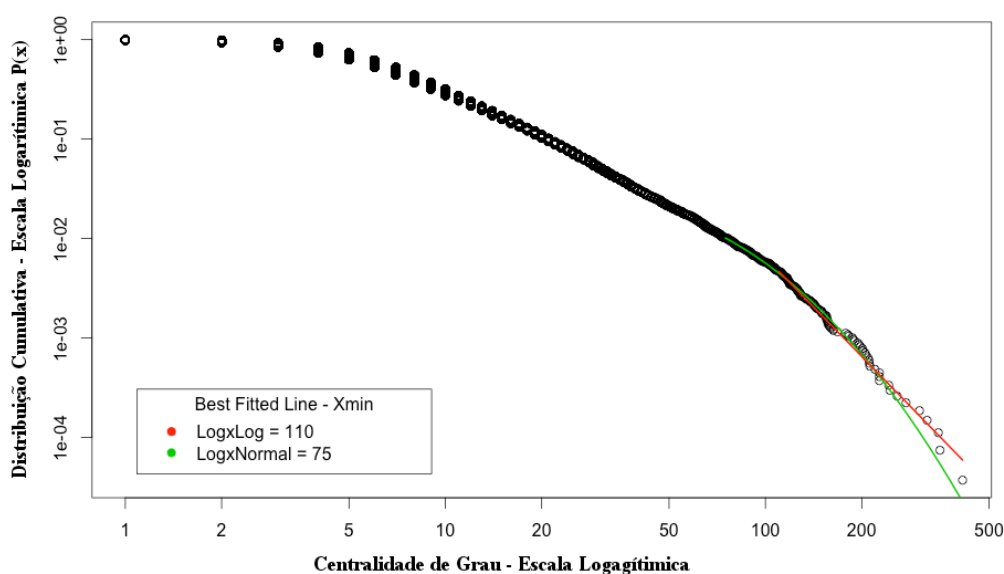
As estatísticas básicas e agregadas da rede LN são apresentadas na Tabela 5. Segundo os dados levantados, os pesquisadores conduzem atividades conjuntas sobre leishmaniose com uma média de 4,85 coautores por artigo. O maior componente, ou o conjunto máximo de cientistas ligados uns aos outros, preenche a maior parte da rede de coautores (86%), que é chamado de componente gigante. O coeficiente de clusterização, a medida da extensão em que os cientistas na rede tendem a agrupar-se em conjunto, é também elevado – cerca de 5.320 vezes superior ao de uma rede aleatória equivalente que normalmente é utilizada como referência. Além disso, 85% dos cientistas de LN já escreviam seus trabalhos em equipe durante o período de 1981 a 1984, com um número médio de 2,4 autores por artigo. No período de 2009 a 2012, apenas 3% dos pesquisadores de LN escreveram seus artigos individualmente, enquanto o tamanho da equipe média havia aumentado para 4,6, o que representa o dobro do tamanho médio da equipe ao longo do período de 32 anos. Outro elemento de análise é a média do caminho mais curto, muito lembrado quando se discutem os seis graus de separação. Essa medida analisa quantas conexões são necessárias para que um pesquisador da rede se conecte a qualquer outro pesquisador da mesma rede.

**TABELA 5. Estatísticas agregadas básicas para a rede de colaboração em leishmaniose**

	Leishmaniose
Número total de autores	31.284
Número total de ligações entre autores	151.567
Número total de artigos	16.482
Artigos com um único autor	6%
Número máximo de artigos por autor	172
Número médio de colaboradores por autor	4,85
Número máximo de colaboradores por autor	413
Número médio de colaborações por autor	6,55
Número máximo de colaborações por autor	718
Tamanho do componente gigante (%)	86
Média do caminho mais curto	4,8
Coefficiente de clusterização	0,824
Assortatividade	0,025

A Figura 22 mostra uma representação gráfica em escala logarítmica da distribuição de graus, ou número de coautores, que os pesquisadores possuem na rede LN. Embora a rede não siga um padrão da lei de potência para todos os valores de grau  $k$ , observa-se na parte inferior do grafo, ou na sua cauda, que esta adota a distribuição de uma lei de potência com  $p(k) = k^{-\alpha}$  para valores acima de  $k_{min} = 110$  e com valor de  $p = 0,72$  no teste de hipótese (CLAUSET; SHALIZI; NEWMAN, 2009).

**FIGURA 22. Curva de distribuição de grau dos autores na rede sobre leishmaniose**



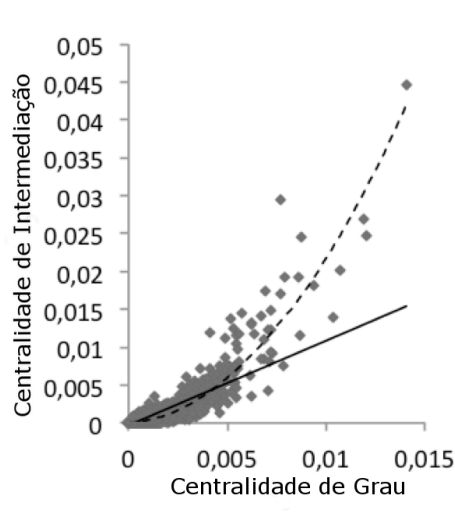
Fonte: elaboração do autor

Legenda: o gráfico é a distribuição de grau dos autores em leishmaniose em escala logarítmica. O eixo esquerdo apresenta uma escala cumulativa onde  $1e + 00$  representa 100% dos autores, indicando que toda rede possui pelo menos uma colaboração, o que pode ser visto no eixo inferior. No outro extremo, apenas  $1e - 04$  ou 0,0001 dos pesquisadores possui mais de quatrocentas colaborações ou centralidade de grau. A linha vermelha no gráfico indica uma curva logarítmica verdadeira, também chamada de LogxLog, onde os pesquisadores com grau de centralidade acima de 110 estão alinhados com essa distribuição. A linha verde apresenta uma curva logarítmica normal, onde os pesquisadores com grau acima de 75 estão alinhados com essa distribuição. Esse resultado indica que a curva de distribuição de grau dos autores não segue na sua totalidade a distribuição de potência.

Foram encontrados 125 hubs que possuem a centralidade de grau acima do  $k_{min}$ . A primeira indicação da sua importância é que esses hubs, que representam apenas 0,47% de todos os pesquisadores do componente gigante, interagem com 39% dos demais pesquisadores da rede.

Para confirmar que os hubs desempenham um papel-chave na rede LN, investigou-se a correlação entre a centralidade de grau e a de intermediação (Figura 23). A centralidade de intermediação é considerada uma medida do quanto um agente pode influenciar ou controlar os outros. Cientistas com maior centralidade de grau possuem também uma alta centralidade de intermediação, o que implica que eles têm poder sobre a rede.

**FIGURA 23. Correlação entre centralidade de grau e centralidade de intermediação**



Fonte: elaboração do autor

Legenda: centralidade de intermediação do nó *versus* centralidade de grau para o período de 1981 a 2012. Nós com alta centralidade de grau também possuem alta centralidade de intermediação, tanto em sua comunidade quanto em outras comunidades. A linha reta representa uma avaliação linear dos dados, enquanto a linha pontilhada segue a tendência. Os valores de grau e de intermediação estão normalizados.

#### 4.2.2 Medidas da rede tuberculose

Os dados de rede apresentados sobre a tuberculose possuem uma diferença perceptível quanto aos dados da rede leishmaniose. Isso se deve ao fato de o recorte feito na rede TB ter sido diferente, de muitas maneiras, do recorte realizado na rede LN. O primeiro fator diz respeito ao recorte temporal, que foi metade (16 anos) para TB em relação à LN (32 anos) devido ao grande número de artigos em TB. O segundo e mais importante fator para análise de redes foi o recorte quanto ao número de pesquisadores. O número total de autores para os 68.958 artigos em TB foi de 141.212. Devido a restrições na capacidade de processamento de redes com esse número de nós, resolveu-se analisar somente os pesquisadores com três ou mais publicações na área. Por essa razão, autores que publicaram somente uma ou duas vezes durante os últimos 16 anos foram retirados da análise. Esse recorte acabou por deixar a rede muito mais conectada e com menos autores na periferia.

As estatísticas básicas e agregadas da rede TB são apresentadas na Tabela 6. Um dos elementos que apresentaram uma diferença representativa em comparação à rede LN foi o componente gigante. Como agora todos os autores possuem três ou mais publicações, o percentual de autores não conectados na rede foi de apenas 2%, deixando o componente gigante com 98% da rede.



Outro elemento de destaque é a razão entre o número de autores e o número de linhas. Enquanto a LN tinha em média cinco vezes mais linhas do que autores, a TB possui cerca de 12 vezes mais linhas que autores, com 124.915 linhas de ligações e 9.577 pesquisadores na rede estudada. O número máximo de artigos por autor foi também maior que o da LN, com 250 publicações, mesmo com um menor recorte temporal. O número médio de colaboradores e colaborações ficou em 26 e 57 respectivamente. Esse número médio deveria ser bem menor caso os autores com uma e duas publicações tivessem sido incluídos na análise. O número máximo de colaboradores e colaborações não deve ter tido um impacto muito grande e ficou avaliado em 536 colaboradores e 1.357 colaborações. O número de colaboradores e colaborações está ligado diretamente à centralidade de grau. No entanto, na pesquisa de publicações científicas deve-se levar em consideração que alguns artigos possuem um número muito mais elevado de colaboradores que a média dos demais artigos. Se a análise de uma rede social se baseia no fato de os pesquisadores, quando colaboram em um mesmo artigo, criarem uma relação de trabalho, esse discurso deve ser avaliado quando isso não acontece em casos de artigos com centenas ou mais colaboradores.

**TABELA 6. Estatísticas agregadas básicas para a rede de colaboração em tuberculose**

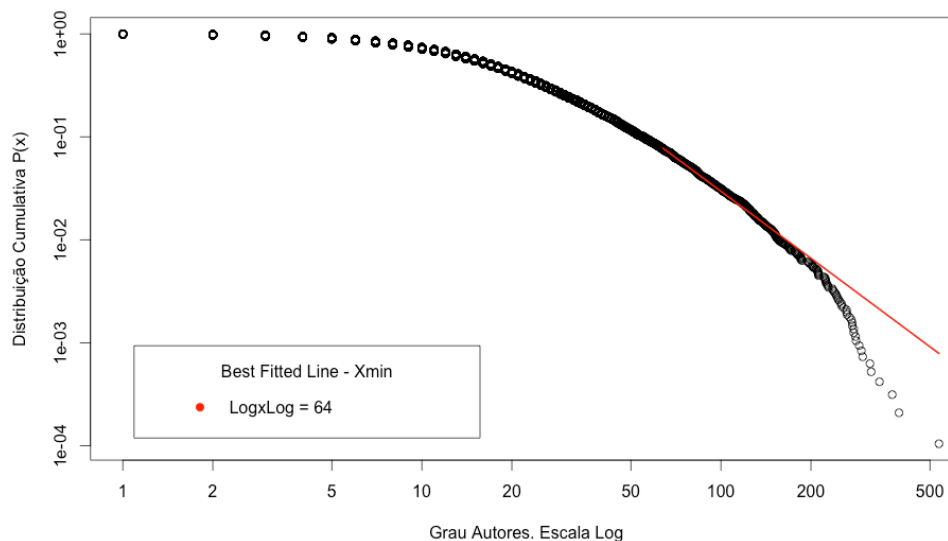
	Tuberculose
Número total de autores	9.577
Número total de ligações entre autores	124.915
Número total de artigos	68.958
Número máximo de artigos por autor	250
Número médio de colaboradores por autor	26
Número máximo de colaboradores por autor	536
Número médio de colaborações por autor	57
Número máximo de colaborações por autor	1.357
Tamanho do componente gigante (%)	98
Média do caminho mais curto	3,6
Coefficiente de clusterização	0,9182
Assortatividade	0,0835

Fonte: elaboração do autor

A Figura 24 exibe a representação gráfica em escala logarítmica da distribuição de graus, ou número de coautores, que os pesquisadores possuem na rede TB, assim como foi feito na rede LN. A alta coesão dos autores na rede descortina também uma diferença no limite mínimo para uma distribuição com características da lei de potência. No caso da rede TB, o recorte ficou com os autores com 64 colaboradores (grau) ou mais, que seguem uma escala logarítmica quanto a sua distribuição. Apesar desse recorte,  $k_{min}$ , na parte de menor grau no gráfico pode-se perceber, na outra ponta, uma divergência quanto aos autores com maior grau da linha de potência real demonstrada pela linha vermelha. A rede, portanto, segundo o algoritmo de Clauset, Shalizi e Newman (2009), atende em parte uma distribuição

logarítmica alinhada com outras redes de pesquisa já analisadas com  $p(k) = k^{-\alpha}$  para valores acima de  $k_{min} = 64$  e com valor de  $p = 0,90$  no teste de hipótese.

**FIGURA 24. Curva de distribuição de grau dos autores na rede sobre tuberculose**



Fonte: elaboração do autor

Legenda: o gráfico é a distribuição de grau dos autores em tuberculose em escala logarítmica. O eixo esquerdo apresenta uma escala cumulativa onde  $1e + 00$  representa 100% dos autores, indicando que toda rede possui pelo menos uma colaboração, o que pode ser visto no eixo inferior. No outro extremo, apenas  $1e - 04$  ou 0,0001 dos pesquisadores possui mais de quinhentas colaborações ou centralidade de grau. A linha vermelha no gráfico indica uma curva logarítmica verdadeira, também chamada de LogxLog, onde os pesquisadores com grau de centralidade acima de 64 estão alinhados com essa distribuição. Esse resultado indica que a curva de distribuição de grau dos autores segue em parte a distribuição de potência.

### 4.3 Formação de grupos e suas relações

No que diz respeito à colaboração, existe uma tendência de os pesquisadores trabalharem com mais frequência com pessoas próximas ao seu círculo de pesquisa, começando assim a formação de grupos ou de comunidades de rede com base nos grupos formais ou não formais de ensino e pesquisa já existentes. Esses grupos podem ser vistos em uma escala mais macro quanto as suas fronteiras geográficas, com países se tornando um tipo de barreira ou clube. Em uma análise mais aprofundada, podem-se perceber agrupamentos mais específicos relacionados a instituições de ensino e pesquisa dentro dessas barreiras geográficas, ou o que se pode chamar de subgrupos ou sub-redes dentro de um grupo maior. Uma das características das redes de pesquisa científicas é a colaboração mais intensa entre os pares mais próximos e menos intensa com pesquisadores que participam de outros grupos de pesquisa. A diferença, no entanto, encontra-se nos pesquisadores mais renomados ou com o maior número de publicações e colaborações, que acabam por fazer a ponte entre os grupos existentes, colaborando com intensidade similar dentro e fora dos seus grupos de origem.

Todos os resultados de análise de redes de coautoria até o momento culminam para um estudo exploratório acerca das relações e do fluxo informacional entre os diferentes grupos de pesquisa. A hipótese trabalhada nesta tese é que os grupos possuem pessoas-chave, definidas pelas estruturas de hubs ou *k-core*, que são responsáveis pelo direcionamento das pesquisas, uma vez que sua influência sobre os demais representantes da rede é muito grande. Além disso, acredita-se que cada comunidade possui um universo mais ou menos definido de pesquisa e que sua relação com outros grupos compõe um elemento de definição da sua área.

#### 4.3.1 Rede leishmaniose

Durante a análise da rede LN foram propostos recortes devido ao alto número de pesquisadores. Dois recortes foram utilizados, sendo o primeiro com base no *k-core*, representando cerca de 10% do total de pesquisadores, e o segundo recorte com os hubs e baseado na distribuição de grau, representando cerca de 0,47% dos pesquisadores. O *k-core* é uma sub-rede em que cada nó se relaciona a, pelo menos, um número mínimo  $k$  de outros nós nessa mesma sub-rede (NEWMAN, 2010). O *k-core* divide os grupos com base nos seus relacionamentos.

Para analisar as estruturas hierárquicas da rede de coautoria sobre pesquisa em leishmaniose foram feitos dois recortes na rede utilizando-se o *k-core* e os hubs. A distribuição ficou com 2.592 nós para o *k-core* e com 125 nós para os hubs. O segundo passo foi analisar o nível de controle e de relacionamentos desses nós com o resto da rede. Os hubs demonstraram ter ligação direta com 10.333 nós. Destes, 7.741 estão fora dos grupos de hubs e de *k-core*. Quanto à segunda estrutura, *k-core*, esta possui relacionamentos diretos com 19.509 nós da rede, estando fora dos grupos de *k-core* ou de hubs 16.917 pesquisadores. Isso quer dizer que as sub-redes formadas pelos hubs e pelos autores do *k-core* se relacionam tanto com membros do seu próprio grupo ou centro (*core*) quanto com membros fora desses grupos ou da periferia.

Pode-se concluir, portanto, que a estrutura de hubs, apesar de ser vinte vezes menor que a do *k-core* (125 e 2.592 respectivamente), consegue uma abrangência acima da metade da do *k-core* (10.333 e 19.509 respectivamente). O segundo ponto é que em ambas as estruturas hierárquicas existe um forte entrelaçamento entre os nós, mas também com os elementos da periferia, corroborando a estrutura centro-periferia definida por Borgatti E Everet (1999).

No que diz respeito à formação de comunidades, existem muitos algoritmos que buscam a divisão da rede de forma automatizada. Com base nos relacionamentos decidiu-se usar o algoritmo *leading eigenvector*, de Newman (2006). A decisão foi tomada em razão da possibilidade de aplicação em redes de grande porte, o que não é possível para outros algoritmos, mantendo-se um nível similar de modularidade, que é uma medida da qualidade da divisão da rede.

A divisão em comunidades é um método importante na análise de redes, pois ajuda a identificar grupos mais coesos e relações entre diferentes grupos. As medidas de modularidade ajudam a identificar o quanto a divisão proposta está mais ou menos aderente ao conceito de maior número de relações dentro de um grupo e menor número de relações entre grupos. Diferentes algoritmos podem gerar diferentes resultados, além do que o tamanho dos grupos, ou o número de autores em cada grupo, pode divergir bastante.

**TABELA 7. Comunidades e distribuição dos hubs e dos nós do *k-core***

<b>Grupo</b>	<b>Nós</b>	<b><i>K-core</i></b>	<b>Hubs</b>	<b>Grupo</b>	<b>Nós</b>	<b><i>K-core</i></b>	<b>Hubs</b>	<b>Grupo</b>	<b>Nós</b>	<b><i>K-core</i></b>	<b>Hubs</b>
<b>1</b>	2057	451	32	<b>16</b>	271	37	1	<b>31</b>	90	10	0
<b>2</b>	394	56	5	<b>17</b>	1077	84	5	<b>32</b>	52	5	0
<b>3</b>	2128	448	21	<b>18</b>	28	1	0	<b>33</b>	347	30	1
<b>4</b>	1704	201	13	<b>19</b>	236	38	0	<b>34</b>	11	1	0
<b>5</b>	1139	122	7	<b>20</b>	651	42	2	<b>35</b>	56	8	0
<b>6</b>	734	112	5	<b>21</b>	21	8	0	<b>36</b>	17	0	0
<b>7</b>	2164	199	11	<b>22</b>	1	0	0	<b>37</b>	220	16	1
<b>8</b>	229	35	1	<b>23</b>	146	7	0	<b>38</b>	125	4	0
<b>9</b>	29	10	0	<b>24</b>	363	22	0	<b>39</b>	41	0	0
<b>10</b>	1733	214	11	<b>25</b>	144	13	0	<b>40</b>	100	10	0
<b>11</b>	6006	50	1	<b>26</b>	57	2	0	<b>41</b>	174	19	0
<b>12</b>	60	10	0	<b>27</b>	67	7	0	<b>42</b>	170	18	0
<b>13</b>	1863	162	5	<b>28</b>	314	9	0	<b>43</b>	120	3	0
<b>14</b>	333	40	0	<b>29</b>	91	8	0	<b>44</b>	204	6	0
<b>15</b>	901	55	3	<b>30</b>	3	0	0	<b>45</b>	181	19	0

Fonte: elaboração do autor

Aplicando o algoritmo na rede foram obtidas 45 comunidades. Cada uma das comunidades está descrita em números crescentes e em negrito na Tabela 7. O número de nós de cada grupo é apresentado na coluna seguinte, na qual se pode perceber a diferença de tamanho de cada grupo. Em cada comunidade foi analisada a distribuição dos nós das estruturas de hubs e de *k-core*. As colunas 3 e 4, com os nomes *k-core* e hubs, apresentam o número de atores dessas duas estruturas naquele grupo específico. Podem-se perceber grandes diferenças no tamanho das comunidades. Todavia, a média de nós das duas estruturas, *k-core* e hubs, em cada grupo obedece a um padrão, seguindo um percentual similar para os grupos. Esse padrão possui algumas exceções, como o grupo 11, por exemplo, que tem muitos nós, mas poucos hubs ou *k-core*. Ou seja, as estruturas de hubs e *k-core* estão representadas em quase todas as comunidades, respeitando-se o tamanho destas.

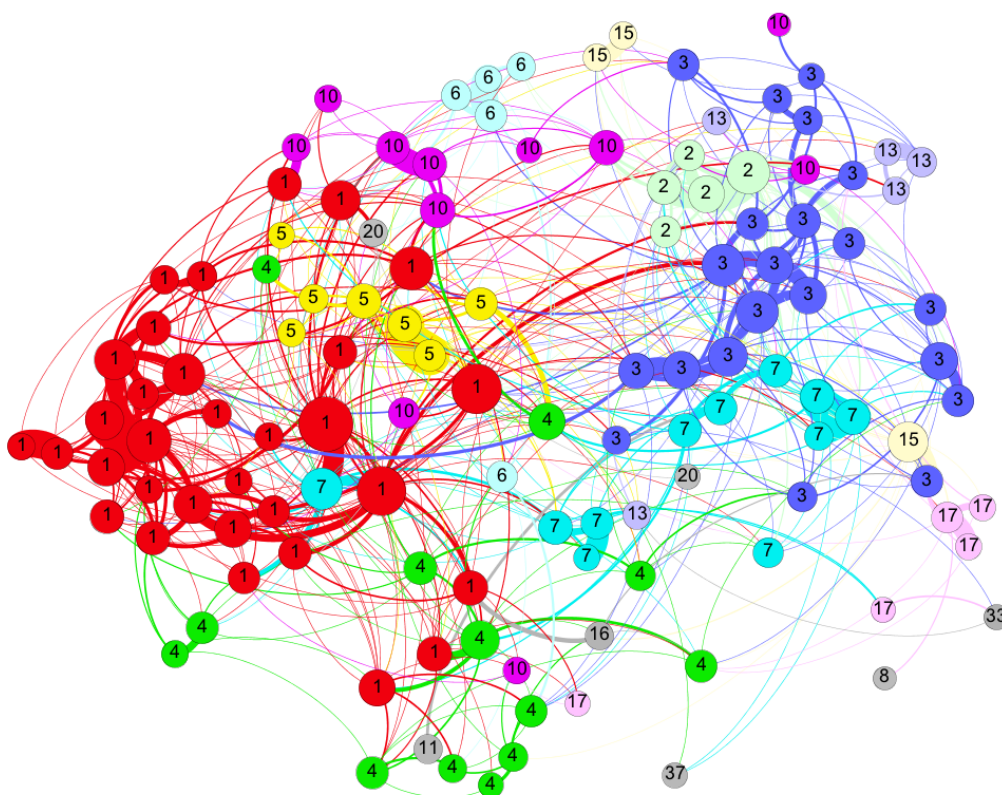
Outro resultado que chama a atenção é o número de nós nas primeiras 17 comunidades, que representam cerca de 85% da rede. Essa característica está relacionada ao algoritmo que busca os nós com maior centralidade de autovetor e seus pares e os coloca em primeiro lugar. Uma grande diferença no tamanho e na representatividade das duas estruturas pode ser percebida novamente na comunidade 11. Um estudo mais aprofundado dessa comunidade demonstrou que seus elementos fazem parte de várias minicomunidades periféricas espalhadas. Por possuírem características similares, o algoritmo acabou por concentrá-las em um só grupo.

Uma vez definidos os hubs e sua capacidade de relação e intermediação sobre a rede foi estudada então sua interação. A ideia é que se os hubs controlam parte da rede com a qual mantêm contato e se entre si mantêm algum tipo de relação, conseqüentemente podem controlar toda a rede, formando o que é chamado de fenômeno “clube dos ricos”. Tal fenômeno é uma propriedade fundamental, pois representa a formação de grupos dominantes nas ciências sociais (NEWMAN, 2004; WASSERMAN; FAUST, 1994).

Foi descoberto que os hubs possuem uma forte relação entre si, mas também possuem muitas colaborações com seus respectivos grupos de pesquisa. Dessa forma, entende-se que os hubs na rede LN têm a capacidade de presidir ou controlar comunidades distintas e ainda sim se relacionar com outros hubs de diferentes comunidades.

Se os cientistas-chave efetivamente controlam diferentes grupos ao mesmo tempo em que coordenam entre si, o próprio clube dos ricos deve mostrar uma organização modular, e esta organização deve ser coerente com a de toda a rede. A Figura 25 exhibe uma parte da rede LN contendo os hubs e as conexões entre eles. A representação de rede aponta que hubs são altamente interligados, formando um subgrafo coeso. Além dessa interconexão geral, os cientistas-chave também formam claramente subgrupos dentro desse núcleo da rede, no qual muitas de suas conexões que integram esses subgrupos estão de acordo com a estrutura da comunidade da rede global.

FIGURA 25. Rede de relações entre hubs



Fonte: elaboração do autor

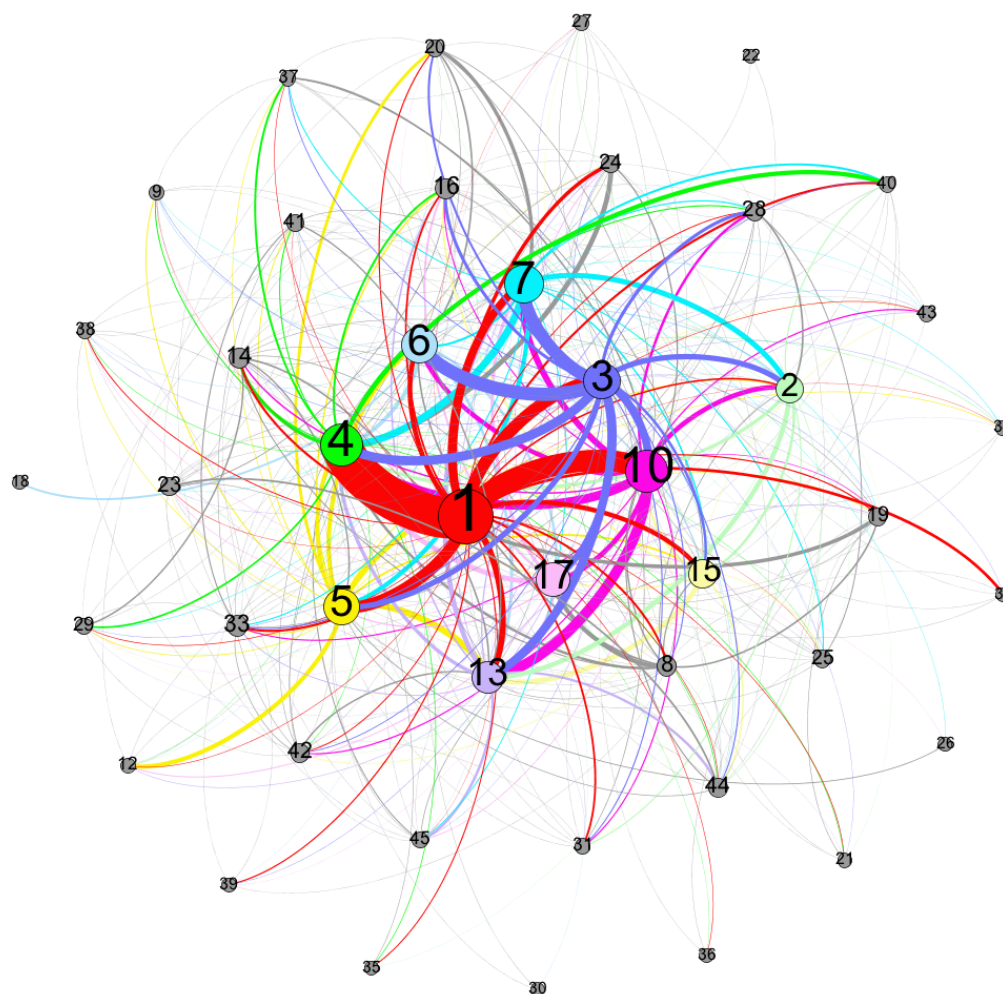
Legenda: organização estrutural dos hubs. Cada nó e cada linha são coloridos de acordo com a comunidade da qual participam na rede, o que também é demonstrado pelo número dentro do nó. A distribuição de comunidades foi determinada utilizando-se o algoritmo *Leading Eigenvector*. O tamanho dos nós representa o grau na rede. A espessura das linhas representa o número de colaborações entre dois nós, com a cor determinada para a linha seguindo a de um dos nós ao qual está ligada. A distribuição da rede segue o algoritmo *Force Atlas 2*, que demonstra aproximações estruturais de forma visual.

O grafo da Figura 25 demonstra as relações entre os hubs, com o número e a cor dos nós denotando os diferentes *clusters* ou grupos encontrados. Apenas os grupos que possuem hubs são representados na figura. O tamanho de cada nó da rede está relacionado ao número de colaboradores ou ao grau que possui nessa rede, enquanto a espessura das linhas demonstra o número de colaborações. Quanto mais espessa a linha maior o número de colaborações.

Na Figura 26 foi avaliada a arquitetura modular do clube dos ricos com a rede global LN. Para maior clareza, sua representação foi reduzida, ficando o foco apenas nas comunidades e nas ligações entre estas. O grafo indica que a rede LN está organizada em torno das comunidades que contêm hubs, que formam seu núcleo, enquanto todos os outros grupos ou *clusters* são periféricos. Em particular, a comunidade 1 destaca-se no sentido de que sua posição é central (elevado número de *links* para muitas comunidades) no âmbito do “centro de poder”. Esses resultados sugerem que a arquitetura geral da rede corresponde a

unidades funcionais conduzidas por centros específicos ou por cientistas de renome. A comunidade 1 tem também a maior atividade interna em termos de ligações. O fato de os hubs desempenharem um papel dominante no seio das comunidades é ainda confirmado pela verificação de que uma correlação entre centralidade de grau e intermediação existe nas comunidades de base.

**FIGURA 26. Rede de relacionamentos entre grupos sobre leishmaniose**



Fonte: elaboração do autor

Legenda: relacionamentos entre grupos da rede. O grafo apresenta cada nó como o grupo encontrado pelo algoritmo de clusterização, e as linhas, as relações entre os grupos. Os grupos ao centro estão representados por uma cor e são os que têm o maior número de nós e, conseqüentemente, o maior número de relações. Os números em cada nó representam seu grupo. As linhas entre dois nós representam a ligação entre os dois grupos, e a espessura, o número de relações existentes. O tamanho de cada nó representa o número de relações com outros grupos.

A Figura 26 apresenta as relações entre os diferentes grupos ou *clusters*. O tamanho de cada nó representa o número de autores presente no grupo em uma escala normalizada, e a espessura das linhas entre dois nós, o número de relações ou coautorias entre os membros dos

dois grupos. Por exemplo, os grupos 1 e 4 possuem uma forte relação entre si, pois nos dois grupos existem pesquisadores brasileiros.

### 4.3.2 Rede tuberculose

Para a análise dos grupos e de suas relações, os mesmos procedimentos aplicados à rede LN foram aplicados à rede TB. Primeiro foram definidos os hubs com base na curva de distribuição de grau, e depois foi definido o grupo *k-core* com base no número de atores da rede. Na divisão por grupos encontraram-se duas peculiaridades em decorrência da distribuição e do algoritmo utilizado. A primeira diz respeito ao grupo de número 2, que apesar de relativamente grande não possuía um número significativo de autores que eram hubs ou *k-core*. Em uma análise mais aprofundada, descobriu-se que a formação desse grupo se deve ao fato de haver artigos com um número elevado de autores, o que acaba por criar uma relação muito coesa, pois todos se relacionam com todos.

Mesmo fazendo o recorte do universo de pesquisa apenas com pesquisadores com três ou mais publicações, os artigos com centenas de autores não foram retirados, conforme ocorreu na rede LN, daí essa distorção. Outro grupo fora da curva é o grupo 25, que, conforme consoante encontrado na rede LN (grupo 11), acaba por ser uma junção de vários pesquisadores espalhados e colocados todos em um mesmo grupo por conta da forma como o algoritmo aplica sua avaliação, ou seja, levando em consideração o valor do autovetor. Nos demais grupos acredita-se que se manteve um padrão de tamanho do grupo com o número de pesquisadores relevantes, fossem eles hubs ou *k-core*.

**TABELA 8. Comunidades e distribuição dos hubs e dos nós do *k-core***

Grupo	Nós	<i>K-core</i>	Hubs	Grupo	Nós	<i>K-core</i>	Hubs	Grupo	Nós	<i>K-core</i>	Hubs
<b>1</b>	393	119	17	<b>11</b>	798	47	5	<b>21</b>	92	0	0
<b>2</b>	556	1	1	<b>12</b>	110	2	4	<b>22</b>	21	0	0
<b>3</b>	645	91	17	<b>13</b>	49	3	2	<b>23</b>	3	0	0
<b>4</b>	520	76	14	<b>14</b>	35	0	0	<b>24</b>	1	0	0
<b>5</b>	1194	191	22	<b>15</b>	15	0	0	<b>25</b>	1032	0	0
<b>6</b>	594	19	7	<b>16</b>	356	36	0	<b>26</b>	25	0	0
<b>7</b>	2054	177	25	<b>17</b>	129	3	1	<b>27</b>	4	0	0
<b>8</b>	89	0	0	<b>18</b>	1	1	0	<b>28</b>	8	0	0
<b>9</b>	183	8	0	<b>19</b>	171	1	0	<b>29</b>	11	0	0
<b>10</b>	393	17	17	<b>20</b>	89	1	0	<b>30</b>	6	0	0

Fonte: elaboração do autor

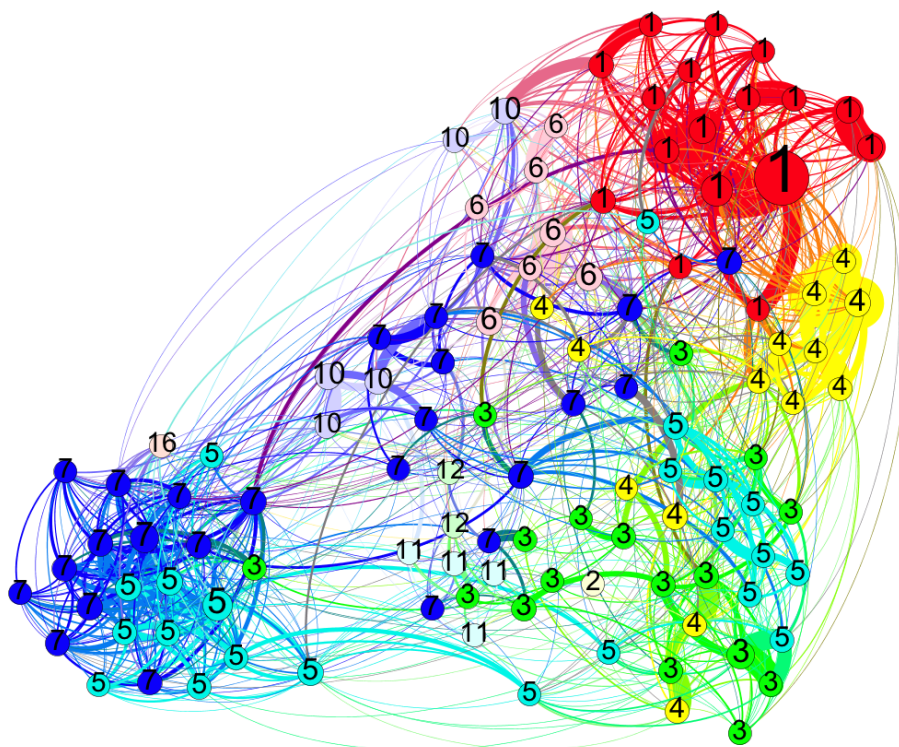
A Tabela 8 apresenta os resultados e a distribuição dos grupos. Foram encontrados 115 pesquisadores com 150 ou mais colaborações. Esses pesquisadores são os hubs da rede TB. Esse recorte foi feito de acordo com a curva de distribuição de grau, com o posicionamento dos autores na curva começando a divergir da curva logarítmica real, em vermelho, na Figura 24. Esses pesquisadores representam 1,2% da rede, no entanto eles



possuem relacionamento direto com 5.857 outros pesquisadores, que representam 61% da rede. No que diz respeito ao *k-core*, foi utilizado o recorte 25-core, ou seja, somente os pesquisadores que tinham relacionamentos com 25 outros pesquisadores do mesmo nível de colaboração estavam nesse recorte. O total de pesquisadores nesse grupo foi de novecentos e com relacionamento direto com 74% da rede. Novamente se segue um padrão similar ao da rede LN, e apesar de se ter um número muito menor de pessoas, os hubs possuem um número aproximado de relações com os demais membros da rede em comparação ao *k-core*.

A Figura 27 demonstra os relacionamentos entre os hubs, ou o chamado clube dos ricos. Pode-se perceber o grande número de interações entre esses pesquisadores, mesmo fazendo parte de grupos distintos. Observa-se uma forte relação entre os grupos 5 e 7, e do grupo 1 com os grupos 4 e 6.

**FIGURA 27. Rede de relações entre hubs**



Fonte: elaboração do autor

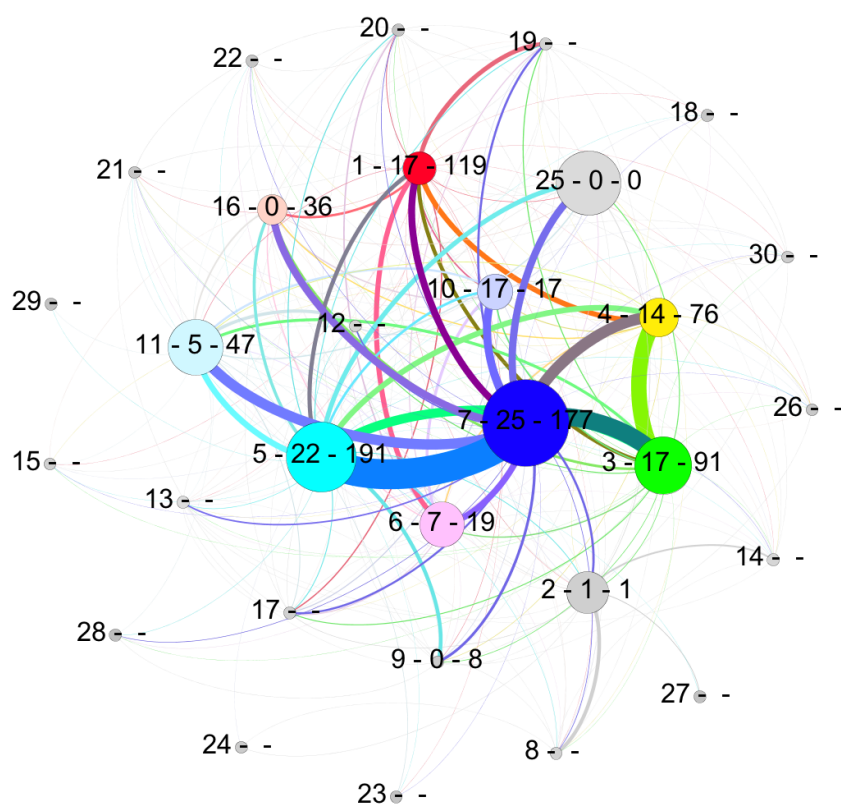
Legenda: organização estrutural dos hubs. Cada nó e cada linha são coloridos de acordo com a comunidade da qual participam na rede, o que também é demonstrado pelo número dentro do nó. A distribuição de comunidades foi determinada utilizando-se o algoritmo *Leading Eigenvector*. O tamanho dos nós representa o grau na rede. A espessura das linhas representa o número de colaborações entre dois nós, onde a cor determinada para a linha segue a de um dos nós ao qual está ligada. A distribuição da rede segue o algoritmo *Force Atlas 2*, que demonstra aproximações estruturais de forma visual.

Na Figura 27 estão representados todos os 150 hubs da rede TB. As cores e a numeração de cada nó representam seu grupo ou comunidade, e as linhas, as relações entre os

pesquisadores. Quanto mais espessa a linha maior o número de relações. A primeira análise é que os hubs de um mesmo grupo se relacionam de forma bastante coesa, mas também possuem relacionamentos com os membros de outras comunidades, o que possibilita a troca de conhecimento e informações entre os grupos. Como esses mesmos hubs possuem relacionamentos com 61% da rede, todas as informações coordenadas por esse grupo acabam por refletir de forma direta nos demais membros da rede.

A segunda análise de relacionamentos entre os grupos foi realizada levando-se em consideração as relações totais entre os grupos. A Figura 28 apresenta todos os grupos e a intensidade de colaboração. Pode ser notado, assim como na rede LN, que os grupos com maior número de hubs e atores do *k-core* se encontram no centro da rede e com o maior número de colaborações.

**FIGURA 28. Rede de relacionamentos entre grupos de pesquisa sobre tuberculose**



Fonte: elaboração do autor

Legenda: relacionamentos entre grupos da rede. O grafo apresenta cada nó como o grupo encontrado pelo algoritmo de clusterização, e as linhas, as relações entre os grupos. Os grupos ao centro estão representados por uma cor e são os grupos com maior número de nós e, conseqüentemente, maior número de relações. As linhas entre dois nós representam a ligação entre os dois grupos, e a espessura, o número de relações existentes. O tamanho de cada nó representa o número de relações com outros grupos. Os números em cada nó representam em primeiro lugar seu grupo, em segundo o número de hubs e em terceiro o número de elementos *k-core*. Os nós que não possuem os segundos e/ou terceiros números não possuem hubs ou elementos *k-core*.

A Figura 28 apresenta os grupos e suas relações. A cor de cada nó está relacionada ao seu grupo, seguindo o mesmo padrão da Figura 27. A espessura das linhas diz respeito ao número de relações entre dois grupos. Assim como foi visto na Figura 27, verifica-se uma forte relação entre os grupos 6 e 7. Esse mesmo padrão pode ser visto novamente na Figura 28. Os números à frente dos nós referem-se ao número da comunidade da qual fazem parte, o segundo número é a quantidade de hubs naquele grupo, e o terceiro número é a quantidade de atores do *k-core* que pertencem àquele grupo. O tamanho do nó está relacionado ao número de atores daquele grupo. Pode-se perceber, por exemplo, que apesar de o grupo 1 não ter muitos pesquisadores, o tamanho do nó indica que ele possui muitos hubs e atores do *k-core*, o que o torna um grupo central.

#### 4.4 Evolução temporal das redes e a mudança em suas estruturas

Para que se possa decodificar a estrutura das redes é importante que se faça uma análise temporal da sua evolução, e não apenas uma visualização estática das suas relações. Na Tabela 9 consta a evolução das redes em cada período de quatro anos. Este estudo foi aplicado somente à rede LN.

**TABELA 9. Autores, publicações e relacionamentos de pesquisa sobre leishmaniose**

Período	Número de autores	Número de publicações	Número médio de autores por publicação	Publicações com um autor (%)	Autores do component e gigante	Linhas do component e gigante	Autores do component e gigante (%)
1981-1984	1.526	878	3,13	15	555	1.504	36
1985-1988	2.490	1.377	3,55	10	1.192	4.090	48
1989-1992	3.459	1.561	3,93	9	2.127	7.410	61
1993-1996	4.481	1.929	4,29	8	2.924	11.256	65
1997-2000	5.201	1.981	4,47	7	3.382	13.409	65
2001-2004	6.672	2.330	4,96	5	4.943	22.139	74
2005-2008	9.166	2.922	5,4	4	6892	33.967	75
2009-2012	11.891	3.646	5,89	3	9.489	51087	80
<b>1981-2012</b>	<b>30.844</b>	<b>16.482</b>	<b>4,76</b>	<b>6</b>	<b>26.852</b>	<b>144.579</b>	<b>86</b>

Fonte: elaboração do autor

Na primeira coluna da Tabela 9 aparece o número de autores por período e seu crescimento gradual por período em cada linha, e na última linha da tabela é apresentado o total para todo o período pesquisado. Ao se comparar o número de autores com o número de publicações por período, pode-se perceber um percentual de crescimento maior no número de autores. Isso pode ser explicado pelo crescimento do número médio de autores por publicação e pela diminuição no percentual de publicações com apenas um autor. Com o passar dos anos,

torna-se evidente o aumento do número de colaborações nos diversos níveis, seja entre instituições, seja entre países ou até mesmo entre áreas, devido ao crescimento da interdisciplinaridade.

Na análise de redes trabalha-se com o conceito de componente gigante, que representa o maior componente com nós conectados entre si. Esta análise foca apenas nesse tipo de componente como referência para o estudo de redes complexas e suas estruturas. Os números de nós (autores) e de linhas da rede, no que diz respeito ao componente gigante, são apresentados nas colunas 5 e 6 da Tabela 10. O percentual de nós do componente gigante com relação aos nós da rede completa aparece na coluna 7. O crescimento percentual do componente gigante na rede de coautoria, disposto na última coluna da tabela, mostra que, com o passar dos anos, essa rede vem se tornando cada vez mais conectada, com menos de 20% dos nós sem relacionamento direto com os demais pesquisadores no último período (2009 a 2012). Isso demonstra uma característica das redes de pesquisa – a transferência de conhecimento e a colaboração.

Na Tabela 10 são apresentadas outras medidas de rede, como os graus médio e máximo dos nós, além dos coeficientes de agregação e o caminho médio da rede. Nas duas últimas colunas aparece a razão entre esses fatores comparados com os mesmos fatores calculados para uma rede aleatória. O coeficiente de clusterização calcula a restrição agregada da rede ou a possibilidade de que dois nós tenham uma relação caso tenham algum conhecido em comum. A comparação com uma rede aleatória demonstra que a clusterização tem um crescimento constante a cada período, com as redes tornando-se cada vez mais coesas, e os pesquisadores, mais conectados entre si.

**TABELA 10. Medidas de rede: graus médio e máximo, coeficiente de agregação e distância média da rede sobre leishmaniose e de uma rede randômica de tamanho similar (nós e linhas)**

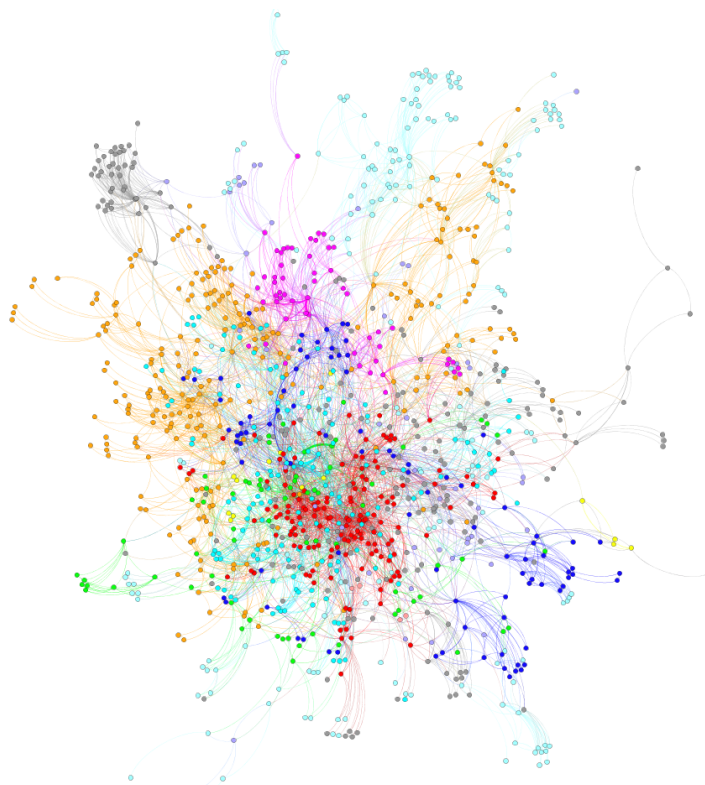
Período	Grau médio	Grau máximo	Coeficiente de agregação	Distância média da rede	Coef. Clu./ Coef. Agre. de uma rede randômica	Dist. Med./ Dist. Med. de uma rede randômica
1981-1984	5,42	37	0,43	7,32	40,46	1,88
1985-1988	6,86	56	0,44	6,43	81,41	1,65
1989-1992	6,97	63	0,45	6,61	154,98	1,59
1993-1996	7,70	71	0,46	6,15	170,02	1,48
1997-2000	7,93	1,00	0,50	6,63	277,11	1,60
2001-2004	8,96	1,14	0,47	5,74	267,16	1,39
2005-2008	9,86	1,06	0,50	6,04	378,18	1,47
2009-2012	10,77	1,87	0,43	5,58	347,97	1,36
1981-2012	10,77	4,13	0,24	4,81	690,25	1,06

Fonte: elaboração do autor

Pesquisas recentes têm apontado diversos tipos de redes complexas, como redes Livre de Escala. No entanto, há suspeitas de que muitas dessas redes não são de fato Livres de Escala e carecem de um estudo mais apurado para validar seus argumentos (CLAUSET; SHALIZI; NEWMAN, 2009). Para verificar a estrutura das redes de LN e TB foi aplicado o algoritmo sugerido no trabalho de Clauset, indicando que as redes de coautoria LN nos períodos de 1981 a 1996 tinham características Livres de Escala, mas as perderam após 1997. Essa mudança na estrutura é um fator de grande importância para o estudo de grupos de pesquisa, pois indica uma alteração na forma de controle e repasse de informações dentro da rede.

Quanto à avaliação de uma rede Mundo Pequeno, os resultados da rede LN obtidos e apresentados na Tabela 10 demonstram a manutenção do coeficiente de agregação durante os diferentes períodos, bem como a diminuição da distância média, mesmo com o crescimento vertiginoso da rede. Com base nessas características, fica claro que a rede de coautoria possui uma topologia de acordo com o modelo Mundo Pequeno.

Com o objetivo de demonstrar visualmente a mudança nas características, é apresentada nas Figuras 29 e 30 a rede de coautoria LN no segundo período da pesquisa (1985 a 1988) e no sétimo período (2005 a 2008) respectivamente. As redes foram distribuídas utilizando-se o algoritmo *Yfan Hu* na ferramenta Gephi. Mesmo se levando em consideração a diferença entre o número de nós e o de linhas, com 1.192 nós e 4.090 linhas no segundo período e 6.892 nós e 33.967 linhas no sétimo, pode-se verificar que existe uma diferenciação na estrutura das duas redes.

**FIGURA 29. Rede leishmaniose – período 1985-1988**

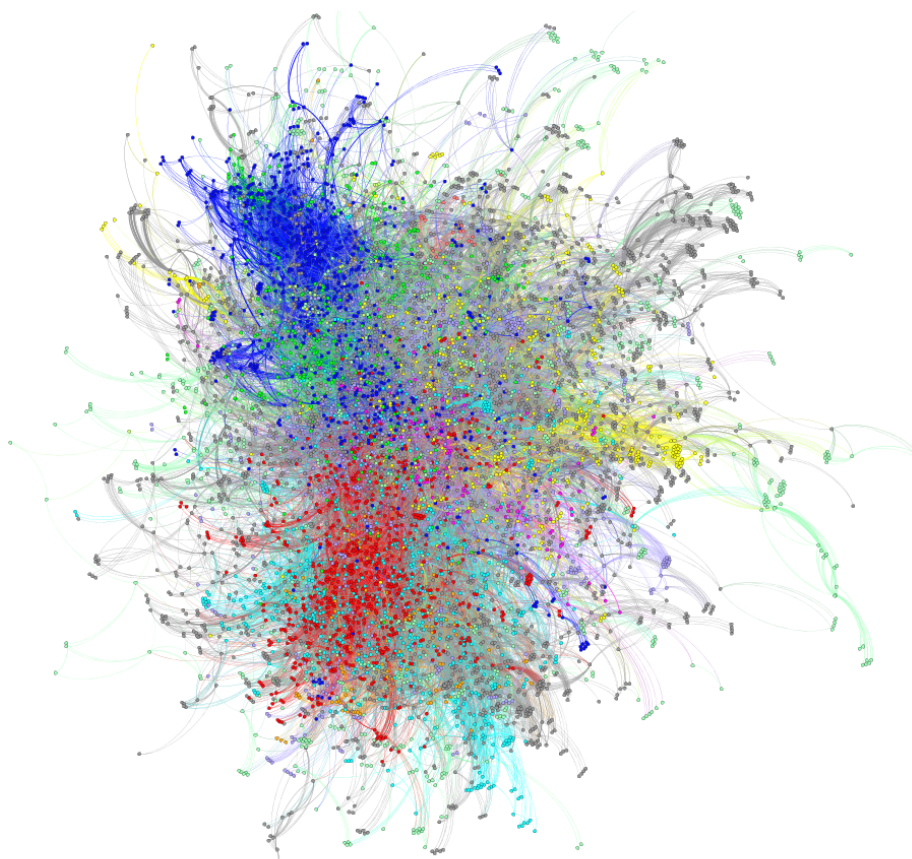
Fonte: elaboração do autor

Legenda: na rede leishmaniose no período 1985 a 1988 são apresentados os nós, que representam os pesquisadores, com as cores relativas aos seus respectivos grupos. As linhas de ligação entre dois nós representam as colaborações entre eles.

No segundo período (Figura 29) existem mais nós longe do centro e ligados a outros por alguns nós mais conectados, uma característica de redes Livre de Escala, enquanto no sétimo período (Figura 30) se nota uma grande concentração em grupos específicos, que são diferenciados pelas cores. Essa estrutura é uma consequência do coeficiente de agregação e do grau médio dos autores.

Portanto, o que se pode perceber em termos de mudança é que no início alguns pesquisadores eram o principal elo entre os diferentes autores. Com o passar do tempo, os pesquisadores centrais acabaram formando grandes centros de pesquisa nos quais vários outros pesquisadores se conectam e trocam informações de forma colaborativa em artigos científicos. Nesse sentido, os grupos de pesquisa tornam-se menos dependentes de poucos pesquisadores. Fazendo-se um paralelo com a participação dos hubs discutida no item anterior, esses pesquisadores centrais passam a tecer relações com outros pesquisadores renomados de outros grupos.

**FIGURA 30. Rede leishmaniose – período 2005-2008**

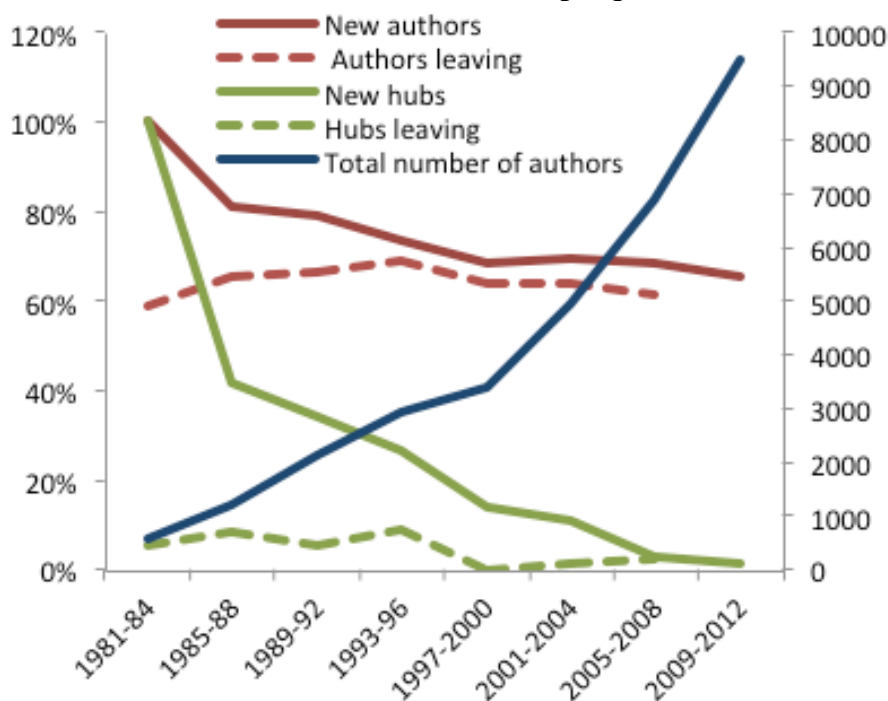


Fonte: elaboração do autor

Legenda: na rede leishmaniose no período 2005 a 2008 são apresentados os nós, que representam os pesquisadores, com as cores relativas aos seus respectivos grupos. As linhas de ligação entre dois nós representam as colaborações entre eles.

Outra forma de se avaliar a evolução da rede é por meio da investigação dos pesquisadores e sua entrada ou saída do campo de pesquisa. Na Figura 31 consta um gráfico com o número de autores que entram e saem da área de pesquisa; os hubs e o número geral de pesquisadores na área também estão representados. Embora o número de cientistas aumente ao longo do tempo, muitos deles entram e saem. Avalia-se que um pesquisador sai da área quando este passa mais de quatro anos (um período) sem publicar. Contudo, os hubs estão presentes de forma consistente ao longo do tempo. Os hubs entraram no campo logo no início do período estudado e sua taxa de abandono é bastante baixa.

**FIGURA 31. Evolução da entrada e saída de pesquisadores em leishmania**



Fonte: elaboração do autor

Legenda: *turn-over* de autores (%) e o total de autores em diferentes períodos de tempo.

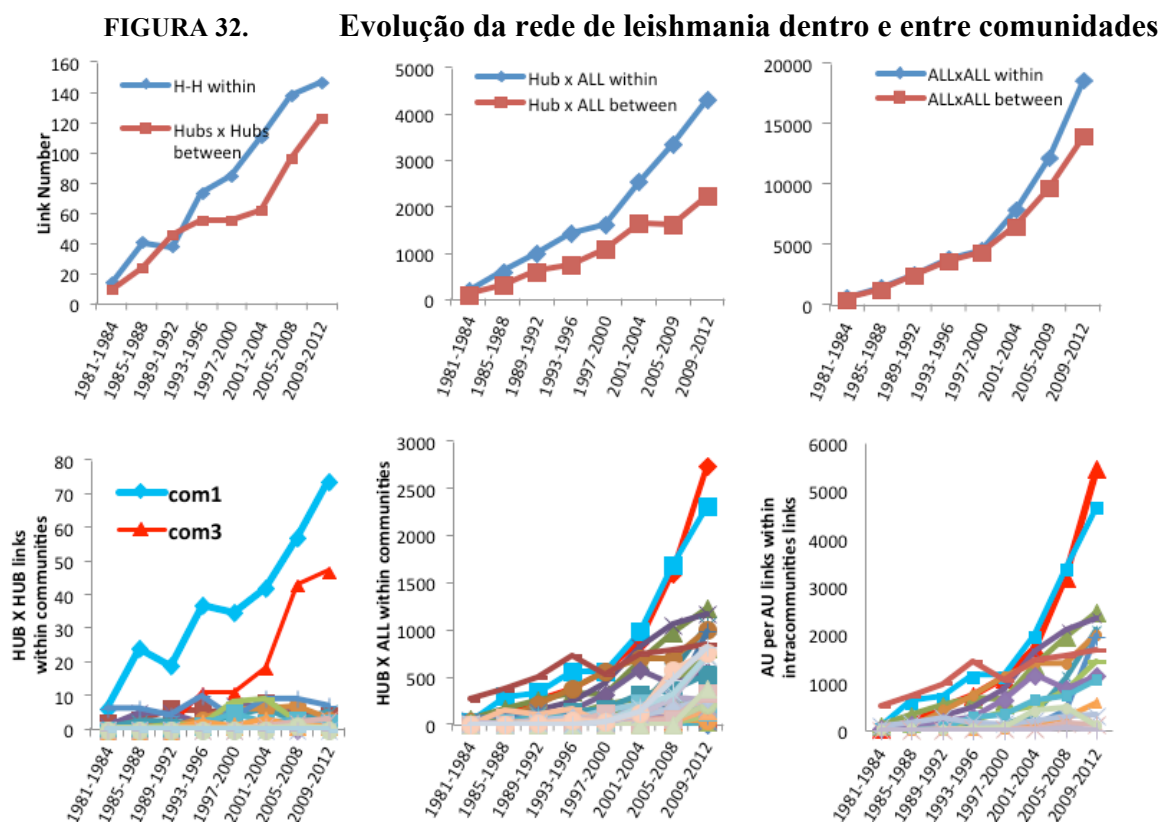
Portanto, foi encontrado que os hubs têm uma vida média de publicações no campo LN de 26 anos (mínimo = 8 anos; máximo = 32 anos) para um número significativo de publicações de 60 (mínimo = 23 e máximo = 172), cerca de 24 vezes acima da produtividade média geral, e nas quais os centros envolvem um número total médio de 156 coautores (mínimo = 111; máximo = 413).

#### 4.5 Interação entre comunidades e geografia

O próximo passo é decifrar como as comunidades, e sua estrutura de poder, evoluem. A visão quantitativa do crescimento global das comunidades e seus relacionamentos são apresentados na Figura 32. O foco principal é estabelecer um entendimento sobre o número de relações dentro e fora dos grupos. Primeiro foi avaliado o número de *links* dentro das comunidades em relação às ligações entre as comunidades. A intensidade de colaborações intracomunidades e intercomunidades aumenta ao longo do tempo. Na parte superior da Figura 32 são apresentadas primeiro as relações dos hubs dentro dos seus grupos e entre os diferentes grupos, com a proporção mantendo-se e assim corroborando o fenômeno do clube dos ricos. O segundo gráfico da figura demonstra as relações dos hubs com os demais pesquisadores da rede dentro dos grupos e entre os grupos. Neste momento pode-se notar que



existe uma tendência de relacionamentos mais focada entre os pesquisadores em geral com os hubs dos seus próprios grupos. O último gráfico da parte superior exibe as relações entre todos os nós da rede dentro dos seus próprios grupos e entre os grupos. Embora o número de ligações intracomunitárias geralmente supere o de ligações intercomunitárias, é significativamente maior o número de ligações entre hubs e outros cientistas a partir de 2000, confirmando a importância do papel dos hubs na formação dessas comunidades.



Fonte: elaboração do autor

Ainda quanto à Figura 32, na parte inferior desta são apresentadas as relações dentro de cada uma das comunidades, com um foco especial nas comunidades 1 e 3, que possuem o maior número de relações. Seguindo a lógica da parte superior da figura em questão, são apresentadas as ligações entre hubs do lago esquerdo, entre hubs e outros pesquisadores ao centro e entre todos os pesquisadores à direita. Neste caso são apresentadas apenas as ligações entre nós da mesma comunidade.

Com a análise desses resultados pode-se notar que o crescimento da rede LN está relacionado diretamente ao crescimento das comunidades 1 e 3, que consequentemente têm seu crescimento diretamente relacionado às relações entre os hubs da própria comunidade. Dessa maneira, embora muitas comunidades apareçam com grande relevância em uma análise estática do conjunto de dados LN, apenas dois grupos realmente se tornam relevantes no

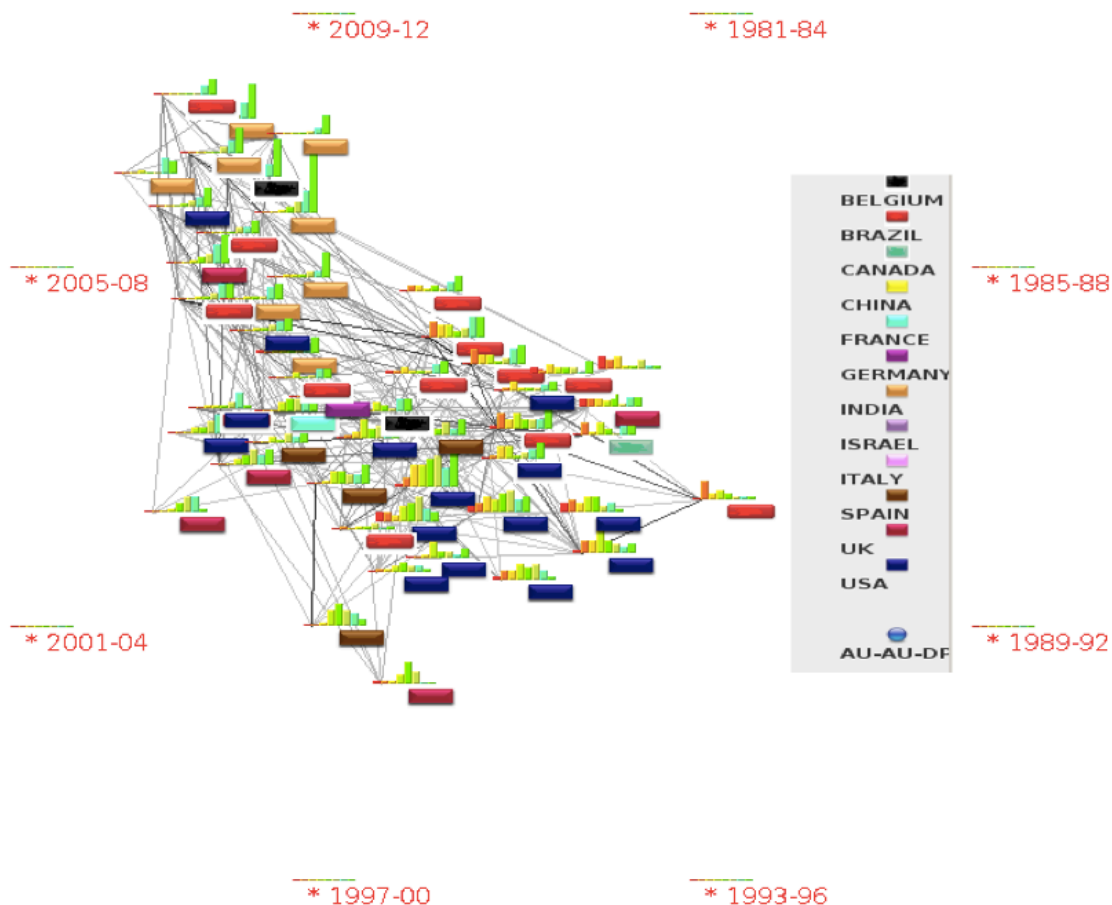
último período (2009 a 2012). Isso significa que essas duas comunidades e seus principais cientistas são os protagonistas da maior produtividade científica quanto aos números de coautorias e de publicações. Curiosamente o crescimento das comunidades 1 e 3 ocorre principalmente depois de 2000, o que corresponde ao calendário de investimentos em leishmaniose.

Para investigar a validade da estrutura da comunidade identificada e entender o que significa, examinou-se se a geografia desempenhou algum papel na formação dos grupos e se suas interações sociais dependeram dela (ONNELA et al., 2011). Uma vez que os hubs são responsáveis pela formação da estrutura de comunidade de rede, que explorou a relação entre suas posições topológicas e geográficas, constatou-se que a organização do clube dos ricos à luz da geografia é um fator significativo. As fronteiras geográficas de algumas comunidades são bem claras, como as da comunidade 4, que compreende a maioria dos hubs que trabalham no Brasil, as da comunidade 2, na França, as da comunidade 6, no Canadá, as da comunidade 5, no Reino Unido, as da comunidade 20, na Suécia, e as da comunidade 33, na Alemanha.

As comunidades mais importantes, no entanto, compartilham hubs de diferentes países. Curiosamente, a comunidade 1, muito central, conforme observado na Figura 25, consiste, não surpreendentemente, em ligações principalmente entre os centros de trabalho nos Estados Unidos, mas também inclui cientistas-chave de instituições brasileiras, enquanto a comunidade 3 é definida por ligações entre os cientistas em sua maioria provenientes do Reino Unido, da Bélgica e da Índia.

Na Figura 33 analisa-se a evolução do comportamento dos hubs e a geografia das comunidades 1 e 3 em diferentes momentos da rede LN. O resultado mais importante é que os líderes de cada comunidade foram se alternando com o passar dos anos. A comunidade 1 é por um longo tempo liderada essencialmente por pesquisadores americanos, enquanto os cientistas brasileiros começam a assumir o controle depois do ano 2000. A comunidade 3 passa a ser dominada por cientistas indianos depois de 2004.

**FIGURA 33. Interações entre os hubs dos grupos 1 e 3**



Fonte: elaboração do autor

Legenda: evolução da produtividade científica por país (esquerda). Legenda de países das comunidades 1 e 3 por país (à direita). Países com o maior volume de publicações como primeiro autor através do tempo. Posições de primeiro autor possuem mais crédito, sendo considerados cargos “chave” em artigos biomédicos. Cada barra de cor representa um autor, e a cor representa o país no qual a instituição ao qual está afiliado se encontra. O posicionamento dos nós obedece ao número de publicações no período demonstrado na parte exterior da figura, com o crescimento dos períodos seguindo o sentido horário.

A Figura 33 é distribuída usando-se um algoritmo de desenho gráfico direcionado que leva em conta a distribuição dos autores no período em que estes são mais ativos quanto às suas publicações. O grafo é posicionado como uma face do relógio analógico, com os períodos sendo alocados no sentido horário dos marcadores de horas. Os pesquisadores ativos em mais de um período estão posicionados entre esses períodos, com uma proximidade maior para o período em que são mais ativos. Pesquisadores igualmente ativos em todos os períodos são colocados exatamente no centro do gráfico. Para uma avaliação quantitativa sobre a atividade global dos hubs através do tempo são utilizados também histogramas que demonstram o quanto cada pesquisador publicou naquele período. O tamanho da barra apresenta o número de publicações, e as cores da barra de comutação variam progressivamente de vermelho no primeiro período para verde no mais recente. Hubs são também identificados pela origem geográfica das respectivas instituições de pesquisa usando-

se retângulos com códigos de cores sob histogramas, como indicado na inserção direita. A transição de azul (hubs dos EUA; comunidade 1) para vermelho (hubs brasileiros; comunidade 1) e, finalmente, para bege (hubs indianos; comunidade 3) é facilmente perceptível.

## 5 Resultados da pesquisa na Web of Science

Na segunda fase do trabalho de pesquisa decidiu-se avaliar a base de dados Web of Science (WoS) tomando como referência as mesmas doenças: leishmanioses (LN) e tuberculose (TB). Além de outra base, foi realizado também outro recorte dentro do universo de pesquisa nas duas áreas. O recorte utilizado para a pesquisa na WoS levou em consideração apenas pesquisadores de instituições brasileiras e suas relações com pesquisadores de instituições nacionais e internacionais. As relações entre os pesquisadores internacionais foram descartadas, a menos que houvesse algum pesquisador de instituição brasileira colaborando no artigo.

Ademais, foi feita uma análise mais detalhada sobre a Fundação Oswaldo Cruz (Fiocruz) e seus colaboradores. Como a WoS possui uma qualificação dos dados dos pesquisadores mais detalhada que a PubMed, esse recorte foi possível e nos permitiu identificar subgrupos de pesquisas mais refinados do que os apresentados no estudo anterior.

A Fiocruz é a principal instituição não universitária de formação e qualificação de recursos humanos para o Sistema Único de Saúde (SUS) e para a área de ciência e tecnologia no Brasil. Dessa maneira, viu-se o estudo com foco na Fiocruz como importante para demonstrar o potencial de pesquisa deste trabalho com a utilização dos métodos de análise de redes.

Para a coleta dos dados referentes às publicações científicas foram utilizadas informações obtidas por meio do serviço de indexação Web of Science (WoS). O serviço, ou base de dados, como muitos o chamam, pode ser acessado a partir do portal Periódicos Capes, disponível para a maioria das instituições federais e estaduais brasileiras de ensino e pesquisa.

A Web of Science, anteriormente conhecida como Web of Knowledge, é um serviço de indexação de citações científicas *on-line* mantido pela Thomson Reuters, que fornece um grande arcabouço de publicações nas áreas científicas. O serviço *on-line* disponibiliza acesso a vários bancos de dados, por meio de uma só plataforma, que fazem referência a pesquisas interdisciplinares, permitindo com isso a exploração em profundidade de subáreas especializadas dentro de uma disciplina acadêmica ou científica.

Um índice de citações é construído partindo do pressuposto de que as citações na ciência servem como ligações entre elementos de pesquisa semelhantes, e com isso levam a elementos científicos correspondentes, como artigos de periódicos, anais de conferências, resumos, etc. No trabalho apresentado não foi realizada pesquisa sobre a rede de citações, que é classificada como uma rede de conhecimento (NEWMAN, 2010), mas sim sobre as

colaborações entre os pesquisadores, que podem ser entendidas como uma rede social (WASSERMAN; FAUST, 1994). Independentemente do fato de a WoS ser um serviço de indexação de citações, a grande quantidade de artigos científicos indexados em sua base atende amplamente à necessidade definida para o trabalho de pesquisa proposto – avaliar as colaborações científicas.

Algumas vantagens relativas à utilização desta base de dados são apontadas a seguir:

- abrange um grande número de periódicos acadêmicos e tem alta representatividade de periódicos na área da saúde;
- é largamente utilizada para a geração de indicadores internacionais de produção científica;
- fornece informações detalhadas sobre afiliação e/ou endereços dos autores, o que permite o acompanhamento do comportamento colaborativo dos membros da população-alvo, que geralmente publicam em revistas de alto impacto;
- fornece o nome completo dos autores na maioria dos artigos. Nos estudos de análise de redes a definição correta dos nomes dos autores é fator crítico para a obtenção de resultados corretos e confiáveis no que diz respeito às ligações entre eles e conseqüentemente a toda a rede;
- permite a exportação dos dados em formato texto para importação em *softwares* de análise de dados, redes ou análises bibliométricas.

Estes recursos e vantagens que a WoS oferece estão adequados aos propósitos de identificar a participação de pesquisadores brasileiros e de pesquisadores da Fiocruz na produção de conhecimentos científicos sobre leishmaniose e tuberculose e seus comportamentos nas redes de colaboração por meio de publicações em coautoria. Dessa maneira, o primeiro passo do processo de investigação científica exploratória, proposta por Nooy et al. (2011), de “definição da rede” foi iniciado.

Outro elemento de grande relevância na análise dos dados de publicações científicas é a categorização realizada pela WoS para todos os artigos indexados em sua base. Essa categorização permite uma avaliação macro acerca de quais categorias ou tópicos são mais ou menos relevantes em determinadas áreas. As categorias são definidas de acordo com as revistas científicas e suas áreas de interesse, que podem mudar em caso de edições especiais. Um resumo das principais categorias encontradas durante a pesquisa em LN e TB é apresentado ao final deste capítulo.

O restante do capítulo Resultados da pesquisa na Web of Science é subdividido em dois subcapítulos, com foco primeiro nas leishmanioses e depois na tuberculose. No capítulo Resultados da pesquisa na PubMed, os resultados sobre LN e TB foram apresentados intercalados, diferentemente do que está sendo apresentado neste capítulo.

### 5.1 Rede de pesquisadores em leishmaniose

As atividades de “manipulação da rede” e seus dados, de acordo com a proposta de investigação científica exploratória, puderam ser realizadas a partir da coleta de informações na base WoS. A busca foi realizada em setembro de 2014 e abrangeu o período de 1984 a 2013. Utilizaram-se como parâmetros de delimitação da busca os seguintes campos e descritores: OG = (Fundação Oswaldo Cruz); OR CU = (Brasil OR Brazil); AND TS = (leish\*) OR Kala Azar.

Os campos OG e CU identificam, respectivamente, a organização com as variações de seus nomes disponíveis na base e o país de afiliação das instituições de vinculação dos autores. Por sua vez, o campo TS recupera documentos que contenham os descritores nos campos título, resumo e palavras-chave. O descritor “leish”, combinado ao caractere de truncagem<sup>10</sup> da base, que é o símbolo asterisco (\*), abarca as variações do termo leishmania, e o descritor “Kala Azar” possibilita recuperar publicações que tenham utilizado esse termo para a leishmaniose visceral.

Os tipos de publicações incluíram artigos publicados em periódicos (*articles/articles in press*), artigos de revisão (*reviews*), artigos publicados em anais de eventos (*conference papers/proceedings papers*), resumos publicados em anais de eventos (*meeting abstracts*), cartas/comunicações (*letters*) e notas/comentários (*notes/commentary*).

Após a recuperação e a compilação dos documentos, realizou-se a padronização dos nomes dos autores, evitando-se casos de grafias divergentes, sinonímia e homonímia, processo imprescindível para a atribuição correta da autoria dos documentos. Em seguida, os dados foram traduzidos em redes de coautoria autores *versus* autores. Como a colaboração de coautoria pressupõe reciprocidade entre os participantes, as ligações foram consideradas não direcionais.

---

<sup>10</sup> A truncagem é composta por um símbolo que varia conforme as bases de dados. Quando o símbolo é adicionado à raiz de uma palavra, permite recuperar variações de ortografia ou várias terminações de uma mesma raiz de palavra.

As interpretações dos resultados foram feitas tomando-se por base a avaliação dos grafos e dos indicadores considerados relevantes à análise, a saber: número de autores dos artigos no período analisado; número de conexões entre os autores; grau de centralização da rede; grupos selecionados por similaridade com base em um algoritmo; e densidade da rede – razão entre as interações efetivamente existentes entre os atores da rede e o total de ligações potenciais ou possíveis.

A busca recuperou 13.585 autores que declararam estar afiliados a instituições brasileiras. Destes foram selecionados apenas os pesquisadores que possuíam três ou mais publicações, totalizando 1.908, dos quais 332 declaram ser afiliados à Fiocruz. Estes pesquisadores publicaram um total de 4.211 artigos. É importante mencionar que a afiliação à Fiocruz é informada pelos próprios pesquisadores na ocasião da submissão da publicação, não representando, portanto, uma posição oficial da Fiocruz quanto à inclusão desses indivíduos em seu quadro profissional.

Tendo em vista o período da busca – trinta anos –, os artigos foram separados em dez períodos trienais contíguos. Essa estratégia possibilitou uma análise temporal e dinâmica da participação dos pesquisadores nos grupos das redes de colaboração.

Com base nesses dados foram analisadas duas redes: a rede brasileira de colaboração em pesquisa sobre LN e a rede de colaboração dos pesquisadores da Fiocruz que compõem esta rede. É importante ressaltar que a análise dessas redes, em função do seu recorte territorial (Brasil), representa apenas uma parte da rede de pesquisa mundial nesta área.

Nos subitens a seguir são apresentados os dados da seguinte maneira: primeiro uma avaliação bibliométrica dos resultados levando-se em consideração o crescimento temporal ou a dinâmica dos dados. No item seguinte o foco passa a ser nos resultados relacionados à análise de redes e seus indicadores, avaliando-se as redes brasileira e da Fiocruz. No próximo subitem sobre a pesquisa em LN são apresentadas as colaborações internacionais no universo de pesquisa brasileiro e no universo de pesquisa da Fiocruz. Logo em seguida são analisados os grupos ou comunidades, também chamados de *clusters*. É feita uma análise de como esses grupos são representados em termos de região geográfica ou instituição, por último são analisadas as categorias, segundo a WoS, e como estas são representativas dos grupos de pesquisa encontrados.



### 5.1.1 Análise bibliométrica das publicações em leishmaniose

O número de publicações por triênio está indicado à frente do período e logo após os parênteses. Estes contêm o número do triênio para efeito de acompanhamento nas tabelas a seguir: 1984 a 1986 (1) – 112 publicações; 1987 a 1989 (2) – 86 publicações; 1990 a 1992 (3) – 168 publicações; 1993 a 1995 (4) – 202 publicações; 1996 a 1998 (5) – 225 publicações; 1999 a 2001 (6) – 304 publicações; 2002 a 2004 (7) – 417 publicações; 2005 a 2007 (8) – 625 publicações; 2008 a 2010 (9) – 939 publicações; 2011 a 2013 (10) – 1.133 publicações.

Observe-se que o número de publicações aumentou mais de dez vezes entre o primeiro período e o último. Destaque-se que o ano de 2014 não foi incorporado em razão de a busca ter sido realizada em setembro desse ano, não correspondendo à produção total do ano.

Na Tabela 11 são expostos os 15 pesquisadores que reportaram filiação a instituições brasileiras e que possuem mais de cinquenta publicações no período. Pode-se notar que a maioria dos pesquisadores apresenta publicações em todos os triênios, apesar de haver variações. Apenas seis deles não têm publicações registradas em um ou mais períodos segundo as publicações recuperadas da WoS. No caso do pesquisador Marsden, este possui um alto número de publicações nos primeiros períodos e nenhuma publicação registrada nos últimos períodos. Isso pode ser consequência de sua aposentadoria ou de uma mudança em sua linha de pesquisa. No caso do pesquisador Meyer-Fernandes acontece o oposto do caso anterior, com nenhuma publicação na área nos primeiros períodos, mas rapidamente se tornando um pesquisador com alto número de publicações no último período. Um aprofundamento nessas duas características de autores pode ajudar a entender o ciclo de vida de um pesquisador na área e qual o caminho que outros pesquisadores devem seguir para garantir uma alta produtividade e participação na área.

**TABELA 11. Número de publicações por ano dos 15 pesquisadores com mais de 50 publicações nos triênios avaliados**

Pesquisadores	Triênios										
	Soma	10	9	8	7	6	5	4	3	2	1
CARVALHO, E. M.	153	26	18	27	13	9	13	17	13	5	12
BARRAL, A. M.	144	26	20	23	15	6	9	14	18	6	7
BARRAL-NETTO, M.	118	22	18	14	12	5	7	12	16	7	5
SHAW, J. J.	76	5	2	3	13	7	6	6	16	11	7
GRIMALDI, G.	92	8	9	5	6	9	15	9	16	6	9
LAINSON, R.	65	0	8	5	5	1	4	4	16	14	8
BRAZIL, R. P.	76	24	19	12	6	8	1	0	2	0	4
MAYRINK, W.	70	4	13	15	9	11	3	3	5	2	5
REED, S. G.	74	3	3	2	3	10	5	19	12	7	10
CUPOLILLO, E.	68	15	15	7	12	5	6	7	1	0	0
LEON, L. L. P.	65	13	13	10	18	3	3	2	2	0	1
BADARO, R.	62	2	2	4	2	5	5	12	13	7	10
SCHUBACH, A. O.	57	15	18	12	2	1	6	1	2	0	0
MARSDEN, P. D.	52				0	2	2	3	13	6	26
MEYER-FERNANDES, JR.	51	13	11	8	9	9	1	0	0	0	0

A segunda análise de temporalidade aplicada aos artigos abrangeu as categorias definidas pela WoS. Essa análise é bastante importante para se entender o quanto determinados temas são mais ou menos utilizados na área de interesse, no caso das leishmanioses, ou em que área determinado tipo de estudo é mais prevalente. Neste caso específico estão sendo analisadas somente publicações com pelo menos um pesquisador afiliado a alguma instituição brasileira. Assim, esses dados também podem servir de base para analisar o quanto certas categorias estão tornando-se mais ou menos evidentes nas pesquisas realizadas por instituições brasileiras.

Na Tabela 12 são apresentadas as vinte categorias mais relevantes ou que apareceram com maior frequência entre as publicações encontradas. Um mesmo artigo pode conter uma ou mais categorias definidas pela WoS, de acordo com seu tema, área de estudo e revista na qual foi publicado.

**TABELA 12. Avaliação temporal das categorias da Web of Science**

Categorias no WoS	Triênios										
	Soma	10	9	8	7	6	5	4	3	2	1
<i>Parasitology</i>	<b>1.286</b>	325	273	196	133	113	83	70	63	19	11
<i>Tropical Medicine</i>	<b>1.125</b>	278	243	157	75	73	66	61	86	37	49
<i>Immunology</i>	<b>575</b>	98	113	114	76	46	39	33	23	14	19
<i>Public, Environmental &amp; Occupational Health</i>	<b>480</b>	94	93	69	28	33	26	31	36	26	44
<i>Infectious Diseases</i>	<b>380</b>	109	66	71	54	29	16	16	9	6	4
<i>Biochemistry &amp; Molecular Biology</i>	<b>387</b>	91	65	60	57	53	20	27	12	1	1
<i>Veterinary Sciences</i>	<b>344</b>	116	104	70	28	16	6	2	0	2	0
<i>Microbiology</i>	<b>322</b>	61	50	70	59	32	17	19	6	2	6
<i>Pharmacology &amp; Pharmacy</i>	<b>253</b>	93	75	30	34	15	3	1	2	0	0
<i>Chemistry, Medicinal</i>	<b>191</b>	83	55	25	21	6	1	0	0	0	0
<i>Medicine, Research &amp; Experimental</i>	<b>176</b>	23	38	20	20	16	12	17	7	7	16
<i>Multidisciplinary Sciences</i>	<b>107</b>	74	15	4	5	4	1	2	2	0	0
<i>Biology</i>	<b>119</b>	21	7	5	12	14	17	13	9	7	14
<i>Cell Biology</i>	<b>118</b>	26	21	19	17	9	6	10	6	2	2
<i>Entomology</i>	<b>113</b>	26	37	19	11	11	4	2	0	1	2
<i>Dermatology</i>	<b>85</b>	31	16	3	9	3	13	6	1	1	2
<i>Plant Sciences</i>	<b>62</b>	24	24	7	3	3	1	0	0	0	0
<i>Pathology</i>	<b>59</b>	6	15	5	6	3	5	6	3	5	5
<i>Biophysics</i>	<b>60</b>	18	7	13	14	4	3	0	1	0	0
<i>Multidisciplinary Chemistry</i>	<b>54</b>	20	16	5	7	4	2	0	0	0	0

Na pesquisa em leishmaniose pode-se observar que os temas *parasitology*, *tropical medicine* e *immunology* são os mais utilizados ao longo do tempo. Como as categorias estão relacionadas às revistas científicas, pode-se supor que as revistas nessas áreas ou que contemplam esses tópicos são as que dão maior espaço para publicações sobre o tema LN. Outros temas como *veterinary sciences*, *pharmacology and pharmacy*, *chemistry, medicinal*, *plant sciences*, *biophysics*, *chemistry* e *multidisciplinary* não apareceram em um ou mais períodos. A definição de cada uma dessas categorias é apresentada ao fim deste capítulo.

No que diz respeito aos países, como o foco principal da pesquisa e a busca dos dados tiveram como base instituições brasileiras, os números demonstram as parcerias brasileiras. A Tabela 13 exhibe os países com maior frequência de colaborações com pesquisadores de instituições brasileiras. A análise de colaborações desses países com o Brasil e a Fiocruz é apresentada mais adiante neste capítulo.

**TABELA 13. Número de colaborações do Brasil com outros países**

<b>País</b>	<b>Colaborações</b>	<b>País</b>	<b>Colaborações</b>
Estados Unidos	497	Uruguai	23
Inglaterra	141	Bélgica	19
Portugal	67	Austrália	18
Espanha	59	Canadá	17
França	59	Sudão	16
Itália	52	Japão	15
Alemanha	44	Peru	14
Argentina	41	Holanda	12
Colômbia	39	Paraguai	10
Escócia	30	Bolívia	10
Suíça	25	Coreia do Sul	9
Índia	24	Tailândia	5
Venezuela	23	Irã	5

Por último apresenta-se uma análise ainda superficial das instituições que mais publicam na área de LN com exceção da Fiocruz, devido ao recorte de busca utilizado, pois poderia causar uma distorção na leitura dos dados. A Fiocruz é a instituição brasileira com maior número de publicações na área, no entanto diversas entidades de ensino federais e estaduais também estão presentes neste trabalho.

Na Tabela 14 são apresentadas as instituições com maior número de publicações na área depois da Fiocruz. Todas elas são instituições de ensino federais ou estaduais.

**TABELA 14. Instituições com publicações em leishmaniose no Brasil**

<b>Instituição</b>
Universidade Federal do Rio de Janeiro
Universidade Federal de Minas Gerais
Universidade de São Paulo
Universidade Federal de Ouro Preto
Universidade Federal da Bahia
Universidade Federal Fluminense
Universidade Federal do Mato Grosso do Sul
Universidade Estadual do Rio de Janeiro
Universidade Federal de Pernambuco
Universidade Federal do Espírito Santo
Universidade Federal do Maranhão
Universidade Federal de Juiz de Fora
Universidade Federal de Santa Catarina
Universidade Federal do Mato Grosso
Universidade Federal do Piauí
Universidade Federal de São Paulo

### 5.1.2 Análise da rede científica de leishmaniose

Os principais indicadores tanto das redes de pesquisadores brasileiros quanto da rede dos pesquisadores da Fiocruz e seus colaboradores são apresentados na Tabela 15. São mostrados, além do número de nós e *links*, os graus médio e máximo, o coeficiente de agregação, a densidade e a distância média da rede. As medidas de rede das publicações em LN foram realizadas levando-se em consideração todo o período: 1984 a 2013.

**TABELA 15. Indicadores da rede de colaboração em pesquisa sobre leishmaniose**

<b>Indicador</b>	<b>Valor – Geral</b>	<b>Valor – Fiocruz</b>
Número de nós/pesquisadores	1.908	1.184
Número de <i>links</i>	12.210	4.727
Grau médio	12,7	7,9
Grau máximo	185	185
Coeficiente de agregação	0,585	0,59
Distância média da rede	3,72	3,53
Grau de centralização da rede	0,09	0,14
Densidade da rede	0,007	0,007

Como dito anteriormente, o número de pesquisadores foi retirado das publicações científicas coletadas da base WoS com um recorte para aqueles que possuíam três ou mais publicações no período analisado. O recorte se deu pelo fato de haver um número muito grande de pesquisadores com pouca representatividade quanto ao número de publicações em LN, dificultando o trabalho de análise. Como o objetivo principal era entender como os grupos de pesquisa podiam ser classificados quanto a sua posição geográfica, áreas de pesquisa e influência na rede, o recorte com pesquisadores mais frequentes na rede pareceu adequado.

Quanto à rede de pesquisa da Fiocruz, foi levado em consideração o mesmo recorte anterior, sendo retirados os pesquisadores que não tinham nenhuma relação com pesquisadores filiados à Fiocruz segundo os dados da WoS. Da rede LN com recorte no Brasil, 724 pesquisadores não tinham relação direta com pesquisadores da Fiocruz, representando cerca de 38% da rede. Esse número indica o quanto a Fiocruz é relevante no que diz respeito a atividades de pesquisa e suas colaborações nessa área.

No que concerne à análise das redes, mesmo com um número de nós e *links* diferentes, as duas redes são bastante semelhantes em alguns de seus indicadores, como o coeficiente de agregação, a distância média e a densidade. Isso pode ser resultado da ampla participação dos pesquisadores da Fiocruz na rede. Do total de pesquisadores da rede, 28,3% ou 332 nós

correspondem a pesquisadores afiliados a essa instituição. Mesmo quando são retirados os autores não conectados aos pesquisadores da Fiocruz, a rede permanece inalterada na maioria de suas características. O grau máximo, por exemplo, é idêntico nas duas redes, uma vez que o pesquisador mais conectado está diretamente ligado a pesquisadores da Fiocruz e, portanto, está inserido em ambas as redes.

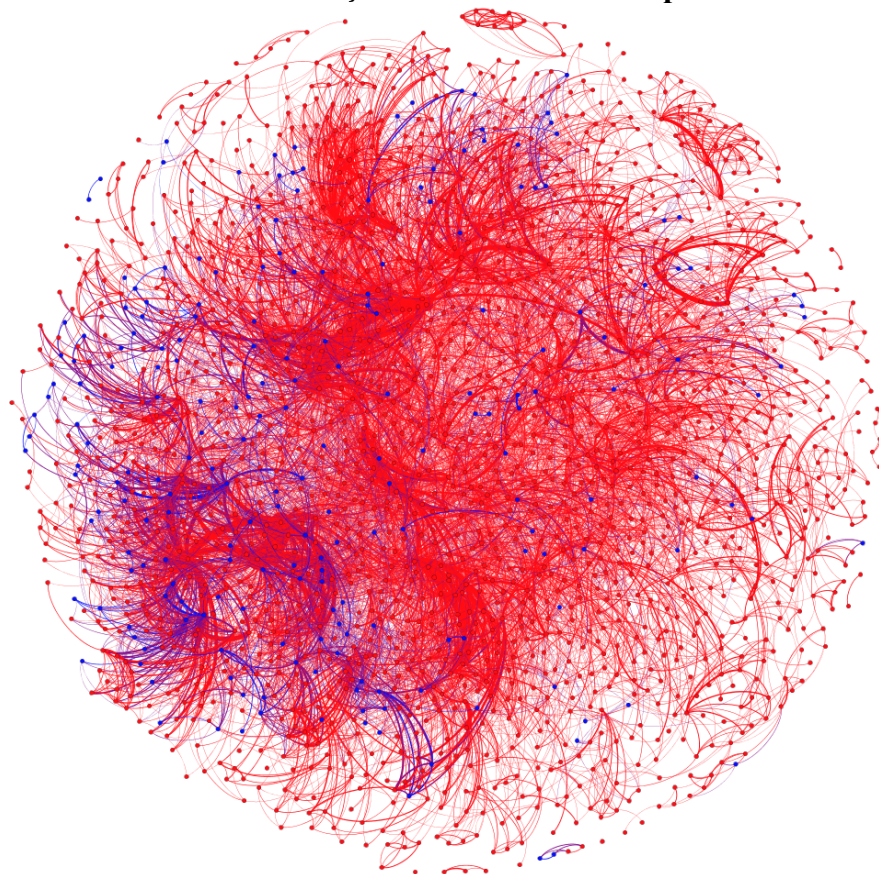
A avaliação do parâmetro de densidade, que se traduz na razão entre as interações efetivamente existentes entre os atores da rede e o total de ligações potenciais ou possíveis, mostra que as duas redes são similares. O grau de centralização complementa a avaliação da densidade, demonstrando que as redes analisadas não estão concentradas em seus nós mais centrais. De fato, o aumento da rede por si só pode levar a uma diminuição de sua densidade.

A rede também demonstra que alguns autores mantêm um alto grau de relacionamentos, principalmente com os demais pesquisadores, o que a caracteriza como rede Livre de Escala. Esse tipo de rede apresenta uma distribuição irregular de ligações entre os diferentes nós, seguindo uma lei de potência. Nelas há um grande número de ligações a um grupo pequeno de nós (instituições) denominados hubs. Barabási e Albert (1999) mostram que atores em redes com essas características, à medida que crescem e novos atores são incorporados, tendem preferencialmente a se conectar aos nós com maior grau, ou seja, aos atores que já possuem um número elevado de colaborações com outros pesquisadores.

No que diz respeito à análise do componente gigante da rede geral, há apenas 34 pesquisadores que não se conectam com o resto da rede. Isso significa que 98,2% dos pesquisadores estão conectados ao restante da rede. Uma vez definido o componente gigante, foi aplicado o algoritmo para clusterização da rede ou de formação de comunidades, a ser discutido no item 1.3 deste capítulo.

A Figura 34 ilustra a rede de colaboração em pesquisas sobre LN que envolve cientistas brasileiros. Os pesquisadores brasileiros estão representados pelos nós de cor vermelha, e os pesquisadores de outros países estão representados pelos nós de cor azul.

**FIGURA 34. Rede de colaboração em pesquisa sobre leishmaniose dos pesquisadores afiliados a instituições brasileiras com os parceiros internacionais**



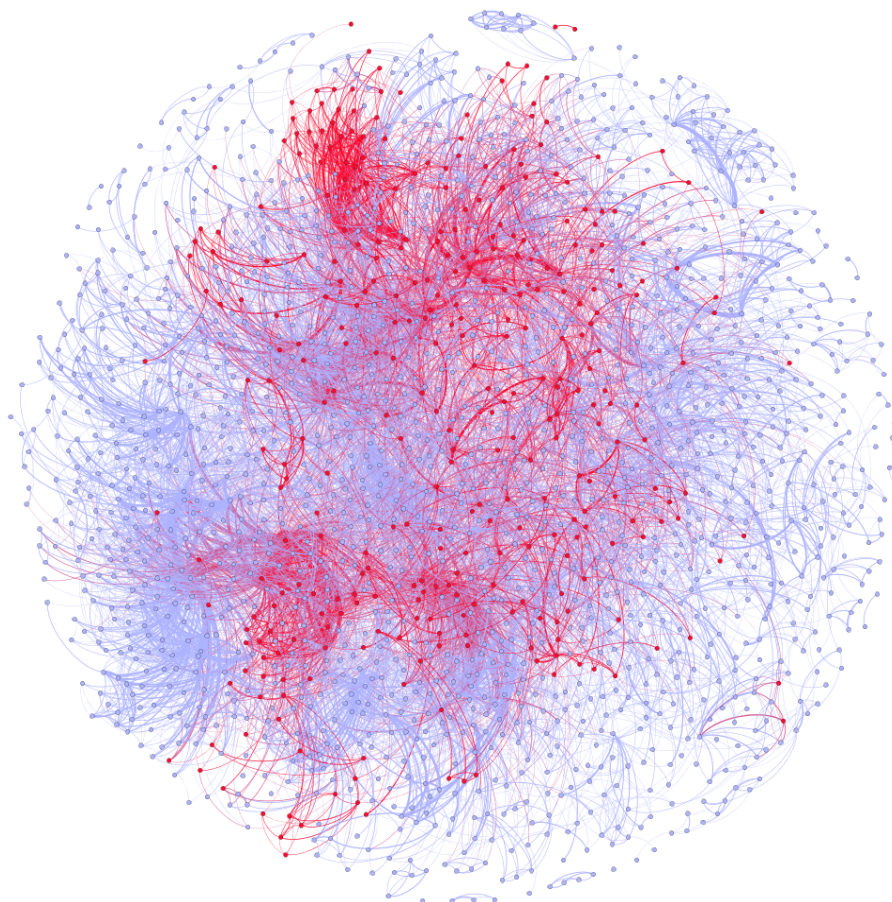
Fonte: elaboração do autor

Legenda: os nós representam os pesquisadores, e as arestas (ligações entre os nós) representam o compartilhamento da autoria de uma publicação científica. As cores dos nós indicam se o pesquisador é afiliado à instituição brasileira (vermelho) ou a instituições estrangeiras (azul).

Nessa primeira avaliação da rede de pesquisadores ligados a instituições brasileiras e seus colaboradores pode-se perceber uma forte coesão na rede. Os pesquisadores parecem se comunicar fortemente entre eles, não sendo percebidos grupos distintos nesse momento. Percebe-se também que alguns pesquisadores se encontram mais na periferia da rede, enquanto um grupo bastante ativo se encontra no centro da figura. Essa situação central e periférica está relacionada diretamente ao número de publicações e colaborações dos pesquisadores, pois a partir do momento em que as publicações científicas se tornam mais frequentes, os pesquisadores tornam-se também mais centrais na rede. A distribuição utilizada para a formação dessa figura foi a Frunchterman Reingold.

Na Figura 35 observa-se a mesma rede com os pesquisadores da Fiocruz, representados pelos nós de cor vermelha, e os demais pesquisadores, independentemente de sua nacionalidade, representados pelos nós de cor azul.

**FIGURA 35. Rede de colaboração em pesquisa sobre leishmaniose dos pesquisadores afiliados a instituições brasileiras**



Fonte: elaboração do autor

Legenda: os nós representam os pesquisadores, e as arestas (ligações entre os nós) representam o compartilhamento da autoria de uma publicação científica. As cores dos nós indicam se o pesquisador é afiliado à Fiocruz (vermelho) ou a outras instituições (azul).

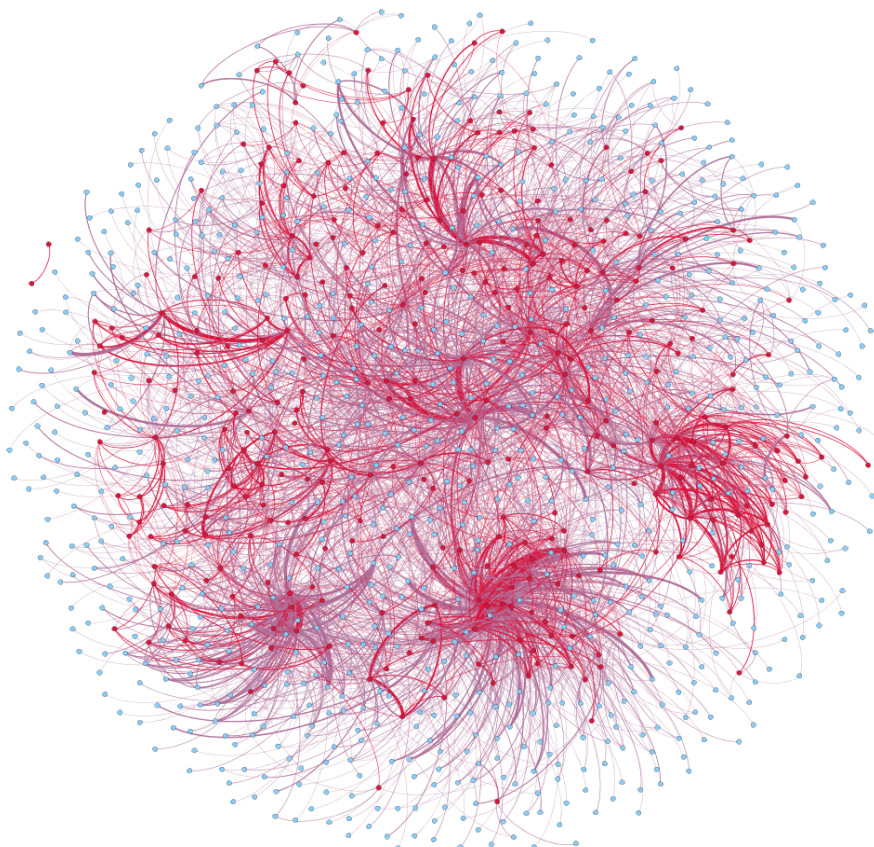
No que diz respeito às relações e aos grupos, já se pode perceber na Figura 35 que existem grupos relativamente separados dentre os pesquisadores da Fiocruz, estes se encontram um na parte superior da figura e outro na parte inferior. Quanto maior o detalhe aplicado na figura, mais informações se consegue extrair dela. Essa figura ainda retrata todos os 1.908 pesquisadores. A diferença agora é a ênfase dada aos pesquisadores da Fiocruz. Pode-se perceber a existência de grupos de pesquisadores não conectados diretamente aos nós de cor vermelha, entendendo-se com isso que estes não possuem trabalhos de coautoria com pesquisadores da Fiocruz.

Na Figura 36 é feito um recorte no qual são apresentados apenas os pesquisadores da Fiocruz e sua rede de colaboração direta. A rede agora possui 1.184 pesquisadores, conforme descrito na Tabela 15. Os pesquisadores da Fiocruz estão representados pelos nós de cor



vermelha, e os demais pesquisadores, independentemente de sua nacionalidade, representados pelos nós de cor azul. A figura demonstra que não só a participação da Fiocruz por meio de seus pesquisadores é extensiva na área da pesquisa em LN, como também sua rede de colaboração e consequente influência.

**FIGURA 36. Rede de colaboração dos pesquisadores da Fiocruz**



Fonte: elaboração do autor

Legenda: os nós representam os pesquisadores, e as arestas (ligações entre os nós) representam o compartilhamento da autoria de uma publicação científica. As cores dos nós indicam se o pesquisador é afiliado à Fiocruz (vermelho) ou a outras instituições (azul).

Com foco na rede de pesquisa da Fiocruz e seus colaboradores, pode-se avaliar como esse trabalho de colaboração ocorre. De acordo com os dados apresentados, percebe-se a existência de uma característica de pesquisa relacionando a Fiocruz à academia, uma vez que dentre as instituições brasileiras com maior número de pesquisa todas elas são entidades de ensino, com exceção da Fiocruz, ou seja, além de a Fiocruz ter relações diretas dentro do seu universo de colaboradores, ela também colabora constantemente com outros institutos de ensino. Outra análise que pode ser feita, com base na Figura 36, é que na Fiocruz se pode perceber agora uma crescente subdivisão, com três grupos aparentemente distintos na parte inferior da figura.

### 5.1.3 Rede de colaboração internacional

Quando foi realizado o recorte do universo de pesquisa na WoS foram levadas em consideração todas as publicações que tinham pesquisadores em instituições brasileiras. Dessa maneira, as publicações em que havia colaborações do Brasil com instituições internacionais puderam ser avaliadas, e um resumo dessas colaborações é apresentado a seguir.

A Tabela 16 apresenta as colaborações existentes entre os pesquisadores encontrados e os demais. No lado esquerdo da tabela são apresentadas as colaborações dos pesquisadores de instituições brasileiras com pesquisadores de outros países, do lado direito são apresentados os países, incluindo o Brasil, que colaboram com os pesquisadores da Fiocruz.

**TABELA 16. Países que mais colaboram com instituições brasileiras e com a Fiocruz no âmbito da rede de colaboração em pesquisa sobre leishmaniose**

País	Intensidade da colaboração (Brasil)	País	Intensidade da colaboração (Fiocruz)
Estados Unidos	49,64%	Brasil (exceto Fiocruz)	76,99%
Inglaterra	9,44%	Estados Unidos	7,95%
Espanha	5,15%	Inglaterra	3,64%
Itália	3,73%	Espanha	1,99%
Portugal	3,16%	Alemanha	1,31%
Índia	3,12%	França, Holanda, Índia, Uruguai, Argentina	3,58%
Outros	25,77%	Outros	4,55%

Por ser um país com alta incidência de LN e reconhecidamente um dos maiores produtores de conhecimento científico nessa área, o Brasil é um importante colaborador na comunidade científica internacional. No que tange à cooperação de pesquisadores brasileiros e da Fiocruz com colegas da América Latina, essa colaboração ocorre com poucos países, aspecto que pode ser fomentado e fortalecido. Entretanto, existe uma forte colaboração com países europeus e os EUA. Um dos motivos dessa forte colaboração é a existência de grande número de instituições de ensino nesses países, situação que estimula os pesquisadores brasileiros a procurar esses centros de pesquisa com vistas a dar continuidade aos seus estudos, acabando por criarem um vínculo de colaboração mais intenso, o que não acontece em países da América Latina. Outro elemento que se faz necessário devido à alta produtividade é a colaboração entre Brasil e Índia. Apesar de esses dois países terem o maior crescimento em termos de produtividade científica na área nos últimos anos, isso não é traduzido em colaboração científica.

Note-se ainda que os principais países colaboradores tanto do Brasil quanto da Fiocruz, à exceção da Índia, são aqueles que não possuem alta incidência de LN. Adicionalmente, os pesquisadores brasileiros em doenças negligenciadas que publicam em

coautoria com colegas de outros países privilegiam instituições de países desenvolvidos (MOREL, 2007).

Outro aspecto a ser mencionado é que dois terços dos recursos voltados ao financiamento de pesquisa e desenvolvimento tecnológico (P&D) para doenças negligenciadas entre 2008 e 2012 foram oriundos de governos. Desse montante, somente os Estados Unidos foram responsáveis por 71% do financiamento. Em ordem decrescente aparecem ainda o Reino Unido e a Comissão Europeia, com respectivamente 7% e 6%, seguidos pela França, pela Alemanha e pela Índia, com 2% cada. Os demais países contribuíram com o percentual restante (G-Finder, 2014).

A despeito de os governos financiarem diferentes tipos de iniciativas, estes apoiam prioritariamente instituições de pesquisa e projetos de pesquisa básica e desenvolvimento inicial de produtos via agências de fomento. Apesar de não haver elementos suficientes para fazer inferências, verifica-se certa convergência entre os países que mais investem em P&D para doenças negligenciadas e a relação de países com os quais os cientistas brasileiros e da Fiocruz mais colaboram.

As figuras a seguir apresentam as colaborações científicas entre os países. As informações contidas na Tabela 16 podem ser visualizadas nas Figuras 37 e 38.

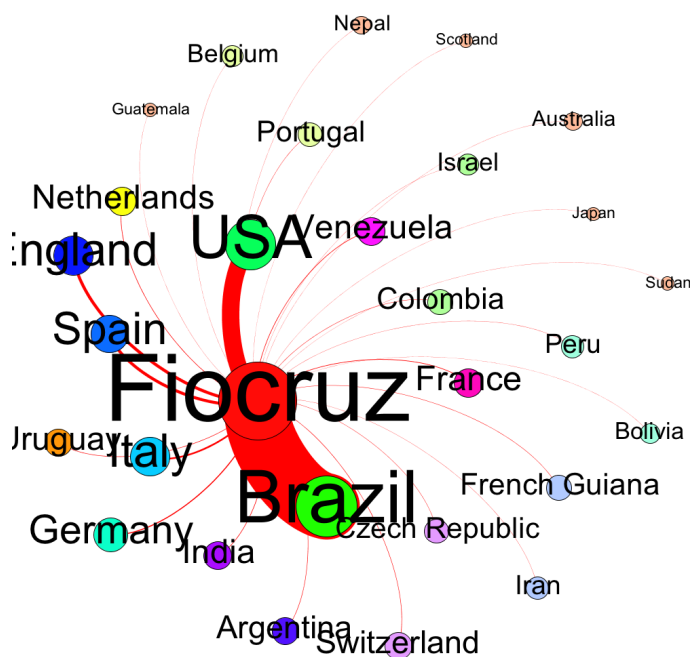
**FIGURA 37. Rede de colaboração do Brasil com outros países**



Fonte: elaboração do autor

Legenda: os nós representam países, e as arestas (ligações entre os nós) representam o compartilhamento da autoria de uma publicação científica. Quanto maior a espessura das arestas maior o número de colaborações.

**FIGURA 38. Rede de colaboração da Fiocruz com os demais países**



Fonte: elaboração do autor

Legenda: os nós representam os países, e as arestas (ligações entre os nós) representam o compartilhamento da autoria de uma publicação científica. Quanto maior a espessura das arestas maior o número de colaborações.

#### 5.1.4 Subgrupos de pesquisa em leishmaniose

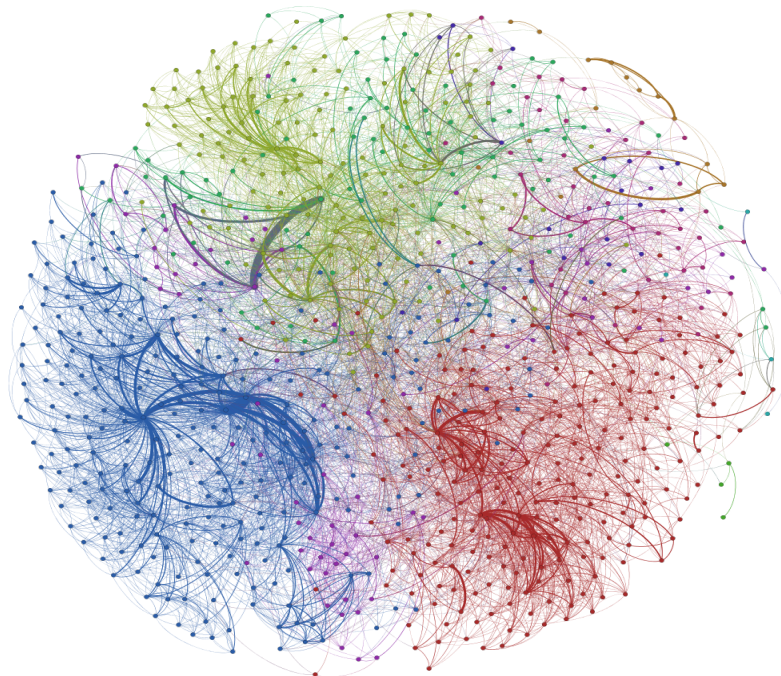
Uma vez definidos os aspectos relacionados aos dados bibliométricos, às medidas de rede e às colaborações nacionais e internacionais, decidiu-se avaliar se havia subgrupos distintos dentro da rede que pudessem ser avaliados ou mesmo caracterizados. De acordo com a pesquisa na PubMed, grande parte das divisões em grupos se dava por questões geográficas. Agora se pretendeu entender um pouco mais e avaliar as divisões em um universo mais bem definido.

No que tange à formação de comunidades, existem muitos algoritmos que buscam a divisão da rede de forma automatizada. Com base nos relacionamentos, decidiu-se usar o algoritmo *leading eigenvector*, de Newman (2006). A decisão foi tomada em razão da possibilidade de sua aplicação em redes de grande porte, o que não é possível para outros algoritmos mantendo-se um nível similar de modularidade, uma medida da qualidade da divisão da rede.

Uma vez aplicado o algoritmo foram encontradas vinte comunidades distintas. Dessas, nove tinham um número maior de pesquisadores e relacionamentos, contabilizando

aproximadamente 94% da rede. Na Figura 39 estão representados os grupos, definidos por cores, encontrados na rede geral de pesquisa em LN e suas relações.

**FIGURA 39. Rede de relacionamentos de pesquisa sobre leishmaniose, por grupos**



Fonte: elaboração do autor

Legenda: os nós representam os pesquisadores, e as arestas (ligações entre os nós) representam o compartilhamento da autoria de uma publicação científica. As cores dos nós indicam o grupo ao qual o pesquisador pertence. Foram colocados apenas os nove maiores grupos.

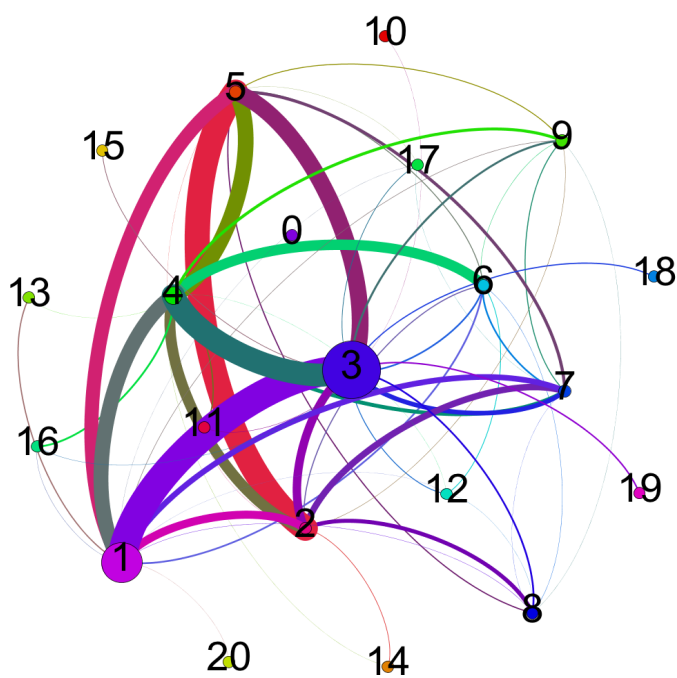
A divisão dos grupos só foi possível com a aplicação de um algoritmo que analisa se as relações em um determinado grupo de pesquisadores é mais intensa do que com pesquisadores externos a esse subgrupo. Com as aplicações de cores distintas para cada grupo, percebeu-se uma maior intensidade de colaboração na comunidade, no entanto isso não impede que pesquisadores de comunidades distintas colaborem em diferentes publicações.

Os subgrupos formados possuem características bastante regionalizadas. Tomando como referência as unidades da Fiocruz e as regiões brasileiras, notam-se grupos bem definidos entre a Universidade Federal da Bahia e a Fiocruz Bahia, representados pelos nós de cor vermelha na Figura 39. Esse mesmo padrão de colaboração entre unidades da Fiocruz e universidades repete-se em Pernambuco e em Minas Gerais, bem como no Rio de Janeiro, tendo a colaboração neste último caso menos clareza, uma vez que a cidade do Rio de Janeiro concentra um número elevado de diretorias da Fiocruz.

Com o objetivo agora de melhor entender as relações entre os diferentes grupos, fez-se uma análise similar à anterior, mas agora aglutinando todos os pesquisadores de cada grupo

em um só nó. Dessa maneira, as relações entre os diferentes grupos puderam ser mais bem percebidas. A Figura 40 apresenta essa figura com os nós da rede representando cada um dos grupos encontrados.

**FIGURA 40. Rede de relacionamentos entre grupos de pesquisa sobre leishmaniose**



Fonte: elaboração do autor

Legenda: os nós representam os grupos, e o compartilhamento da autoria de uma publicação científica entre os membros dos grupos é indicado por uma ligação entre os respectivos nós. A espessura das linhas indica a intensidade ou a frequência das relações.

Em uma análise pontual dos grupos e suas relações destacam-se os grupos ou clusters 1, 3 e 5. O grupo 1 é bastante coeso e formado por pesquisadores essencialmente da Fiocruz Bahia (Centro de Pesquisa Gonçalo Muniz) e da Universidade Federal da Bahia. O grupo 3 possui pesquisadores da Fiocruz Pernambuco (Centro de Pesquisa Ageu Magalhães) e das Universidade Federal de Pernambuco e USP, com pesquisadores dessas instituições como referências. Além destes, o grupo 3 possui vários componentes de outras universidades no Nordeste. O grupo 5 é formado por pesquisadores de unidades do Rio de Janeiro e da Federal do Rio.

O entendimento acerca da localização ou da formação dos pesquisadores do grupo auxilia no entendimento das características regionais e respectivas colaborações. Os grupos 1 e 3 têm forte relação entre eles, o que pode creditado à proximidade geográfica. Contudo, o grupo 3 possui também forte relação com outros grupos, provavelmente devido aos membros

da USP, que acabam por manter relações com pesquisadores de estados mais ao sul do Brasil. O grupo 5 demonstra sua característica articuladora, uma vez que mantém forte relação com vários grupos no Brasil e por ser o centro de várias unidades de pesquisa da Fiocruz, como por exemplo o Instituto Oswaldo Cruz, um dos maiores centros de pesquisa da Fiocruz, e a Escola Nacional de Saúde Pública, que prepara diversos pesquisadores na área.

Com base no recorte de grupos, foram calculados alguns indicadores. Na Tabela 17 estão indicados os números de pesquisadores que compõem cada grupo, os graus médio e máximo, além da análise de relações dentro e fora do grupo. Esta última análise é interessante de se avaliar para que se tenha uma ideia de quais grupos tendem a ser mais ou menos fechados no que diz respeito a colaborações externas ao seu grupo de trabalho.

**TABELA 17.** Indicadores da rede de colaboração em pesquisa sobre leishmaniose

<b>Grupo</b>	<b>Nós</b>	<b>Grau médio</b>	<b>Grau máximo</b>	<b>Relações no grupo</b>	<b>Relações com outros grupos</b>
1	420	13,87	175	2.913	1.091
2	168	7,88	32	662	687
3	439	10,89	89	2.391	1.243
4	266	11,24	79	1.495	1.165
5	226	6,96	61	787	793
6	53	5,06	25	134	252
7	111	5,95	26	330	471
8	44	4,77	9	105	52
9	61	6,95	22	212	105
10	9	8	8	36	1
11	13	0,77	2	5	19
12	33	4,61	10	76	35
13	2	1	1	1	9
14	5	4	4	10	7
15	2	1	1	1	7
16	9	2	3	9	19
17	4	2	3	4	8
18	5	2	3	5	10
19	3	2	2	3	6
20	1	0	0	0	4

Os dois maiores grupos da rede são respectivamente o 1 e o 3. Esses dois grupos são os que contêm também o maior número de pesquisadores da Fiocruz. Isso explica a forte relação entre os dois grupos, fato evidenciado pela espessura da linha que os une na Figura 41. A comunidade 5 também possui muitos pesquisadores da Fiocruz e curiosamente tem uma forte relação com a comunidade 2, que possui pesquisadores da Universidade Federal do Rio de Janeiro.

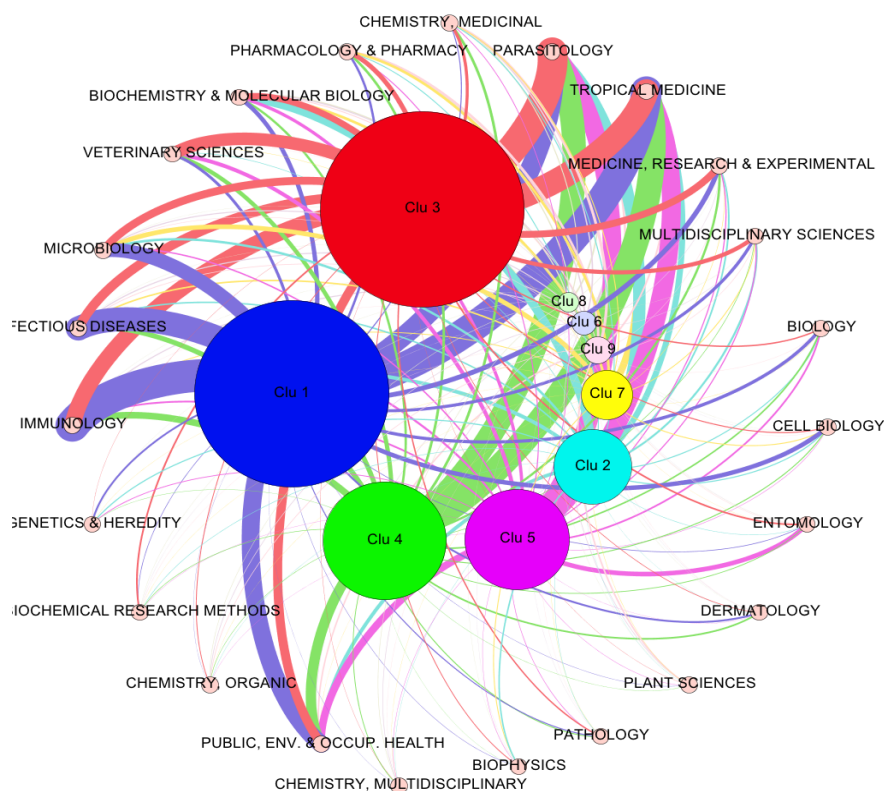
Quanto às relações dentro e fora dos seus respectivos grupos, são destacados dois grupos com características distintas. O primeiro é o grupo 1, cujo número de relações dentro

do grupo é aproximadamente três vezes maior que o número de relações com outros grupos. Esse é o maior valor dentre os grupos com maior número de pesquisadores. Isso demonstra uma alta coesão dentro do grupo, com os pesquisadores tendendo a trabalhar muito uns com os outros. No outro extremo há o grupo 5, que possui quase o mesmo número de colaborações tanto internas quanto externas. A característica desse grupo são suas fortes relações com outros grupos, podendo servir como um agregador de conhecimento entre as diferentes unidades ou regiões.

### 5.1.5 Grandes áreas do conhecimento e caracterização dos grupos

Para se ter uma dimensão sobre quais categorias são mais frequentes entre os diferentes grupos, demonstraram-se as relações de cada categoria com cada um dos grupos. A Figura 41 apresenta as categorias indexadas pela WoS para os artigos. A espessura das linhas que ligam uma categoria a um grupo está relacionada ao número de vezes que aquela categoria foi utilizada pelo grupo. O tamanho dos nós ao centro está relacionado ao número de pesquisadores do grupo. Pode-se notar, por exemplo, que o grupo 1 tem uma relação mais frequente com as categorias *Immunology*, *Infection Diseases* e *Tropical Medicine*.

**FIGURA 41. Relações dos grupos com as categorias**



Fonte: elaboração do autor



Legenda: os nós no centro representam os grupos, e os nós na periferia representam as categorias da WoS. A inclusão de um artigo publicado pelo grupo em determinada categoria da WoS é indicada por uma ligação entre os respectivos nós. As cores dos nós no centro da rede indicam o grupo ao qual o pesquisador pertence, e o tamanho desses nós reflete o número de pesquisadores que compõem o grupo. A espessura das linhas mostra a frequência com a qual determinada categoria foi utilizada nas publicações do grupo.

Outra forma de visualização das áreas de maior interesse de cada um dos grupos encontra-se na Tabela 18. Na parte superior da tabela são apresentados os números de cada grupo, conforme a Figura 41. Do lado esquerdo da tabela são apresentadas as áreas de conhecimento. Portanto, os valores de cada coluna indicam o quanto aquela área de conhecimento é representativa daquele grupo. No caso de *Parasitology*, por exemplo, 11% das pesquisas realizadas pelo grupo 1 utilizam essa área de conhecimento, enquanto nos grupos 2, 4 e 7 essa área de conhecimento é utilizada em 23% dos artigos publicados. Essa análise demonstra o quanto cada um dos grupos é mais ou menos focado em cada uma das áreas de conhecimento. O percentual de cada um dos grupos não necessariamente totaliza 100%, uma vez que um mesmo artigo pode abarcar mais de uma área de conhecimento, além do fato de não haver representação na Tabela 18 de todas as 109 categorias encontradas nos artigos.

**TABELA 18. Indicadores da rede de colaboração em pesquisa sobre leishmaniose**

Grupos/ Áreas do conhecimento	1	2	3	4	5	6	7	8	9
<i>Parasitology</i>	11%	23%	19%	23%	25%	15%	23%	15%	12%
<i>Tropical medicine</i>	17%	11%	16%	24%	25%	9%	4%	4%	10%
<i>Immunology</i>	18%	7%	11%	7%	2%	7%	3%	1%	0%
<i>Public, environmental and occupational health</i>	10%	7%	6%	10%	10%	3%	0%	0%	8%
<i>Infectious diseases</i>	13%	3%	6%	5%	3%	7%	5%	2%	4%
<i>Biochemistry and molecular biology</i>	2%	16%	4%	2%	4%	7%	9%	12%	5%
<i>Veterinary sciences</i>	2%	1%	8%	4%	5%	2%	1%	4%	1%
<i>Microbiology</i>	8%	5%	5%	5%	2%	7%	15%	2%	2%
<i>Pharmacology and pharmacy</i>	1%	2%	3%	2%	1%	8%	9%	11%	13%
<i>Chemistry, medicinal</i>	1%	2%	2%	3%	1%	9%	3%	9%	15%
<i>Medicine, research and experimental</i>	3%	3%	5%	2%	1%	3%	2%	1%	2%
<i>Multidisciplinary sciences</i>	2%	4%	3%	1%	2%	2%	3%	0%	3%
<i>Biology</i>	2%	2%	1%	2%	2%	0%	1%	0%	1%
<i>Cell biology</i>	3%	2%	1%	1%	1%	1%	3%	1%	0%
<i>Entomology</i>	0%	1%	1%	1%	6%	0%	0%	1%	2%
<i>Dermatology</i>	1%	0%	0%	2%	0%	1%	0%	0%	0%
<i>Plant sciences</i>	0%	0%	0%	0%	0%	4%	1%	2%	9%
<i>Pathology</i>	1%	1%	1%	1%	1%	0%	1%	0%	0%
<i>Biophysics</i>	0%	3%	1%	0%	0%	0%	3%	1%	0%
<i>Chemistry, multidisciplinary</i>	0%	0%	0%	1%	0%	3%	1%	4%	1%
<i>Chemistry, organic</i>	0%	0%	1%	0%	0%	4%	0%	4%	4%
<i>Biochemical research methods</i>	0%	1%	1%	1%	0%	0%	0%	3%	1%
<i>Genetics and heredity</i>	1%	2%	0%	0%	1%	0%	0%	1%	0%

Tanto a Figura 41 como a Tabela 18 demonstram, de forma visual e por números respectivamente, o quanto cada um dos grupos pode ser identificado devido às suas áreas de interesse. Como as categorias estão diretamente relacionadas às revistas científicas, percebe-se a tendência maior ou menor por certas publicações, que normalmente estão mais alinhadas com uma ou outra área de conhecimento ou forma de pesquisa. O entendimento dessas nuances pode apoiar o direcionamento de pesquisas para um ou outro grupo ou mesmo propor colaborações ou linhas de pesquisa para este ou aquele grupo.

## 5.2 Rede de pesquisadores em tuberculose

A coleta de dados para a rede TB na WoS foi feita no dia 11 de setembro de 2014 e abrangeu o período de 2003 a 2014. As consultas foram feitas no modo *advanced search* direcionada ao tópico da pesquisa, que abrange título, resumo e palavras-chave, e também à organização e ao país de interesse (Fiocruz e Brasil). A fim de obter as publicações em tuberculose de autores brasileiros e também de autores afiliados à Fiocruz, a consulta foi feita utilizando-se os seguintes termos: TS = (TUBERCULOS\*); AND OG = (Fundação Oswaldo Cruz); OR CU = (Brasil OR Brazil). Foi utilizado o caractere de truncagem (\*) para compreender as variações do termo “tuberculose” e, assim, recuperar o máximo possível de documentos relevantes. Também foram utilizados os campos “Organização – Consolidada” (OG) para recuperar as possíveis variações da Fiocruz na base. Os tipos de publicação incluíram artigos publicados em periódicos (*articles/articles in press*), artigos de revisão (*reviews*), artigos publicados em anais de eventos (*conference papers/proceedings papers*), resumos publicados em anais de eventos (*meeting abstracts*), cartas/comunicações (*letters*) e notas/comentários (*notes/commentary*).

Após a compilação de todos os documentos, um processo de padronização foi cuidadosamente realizado para reunir os vários nomes de um autor, evitando casos de sinonímia e homonímia. Este processo de desambiguação é extremamente importante para a atribuição da autoria de um artigo ao autor correto.

Os dados obtidos por meio da base de dados foram traduzidos em redes de coautoria autores x autores. Como a colaboração de coautoria pressupõe reciprocidade entre os participantes, todas as ligações foram consideradas não direcionais.

A coleta de dados recuperou 2.326 publicações, com um total de 11.219 pesquisadores. Após o processo de desambiguação foram identificados 10.701 autores. Considerando esse universo foi feito um recorte para a avaliação das colaborações apenas dos

pesquisadores que tinham mais de uma publicação. Essa abordagem foi adotada por entender-se que os pesquisadores que apresentavam uma única publicação na área não representavam parcela significativa da comunidade científica que ativamente realiza pesquisa em tuberculose.

Após essa limpeza dos dados foram analisados 2.181 pesquisadores, sendo 1.800 pesquisadores brasileiros e 477 deles afiliados à Fiocruz. com base nesses dados foram analisadas duas redes: a rede de colaboração em pesquisa sobre tuberculose e a rede de colaboração dos pesquisadores da Fiocruz. É importante ressaltar que a análise dessas redes incidiu sobre uma estrutura parcial da rede de pesquisa mundial em tuberculose, definida aqui pela rede brasileira de pesquisa neste tema. Portanto, as redes aqui apresentadas não representam todas as relações possíveis e as realmente existentes na pesquisa mundial em TB, mas somente o recorte considerado foco do estudo: pesquisadores afiliados a instituições brasileiras e suas correlações com outros pesquisadores.

### **5.2.1 Análise bibliométrica das publicações em tuberculose**

A análise temporal das publicações mostra que o número de artigos científicos na área de tuberculose no Brasil quase triplicou de 2006 a 2009, mas desde então tem ficado relativamente estável. São os seguintes os números de publicações durante os anos de 2003 a 2013: 2003 (289); 2004 (97); 2005 (95); 2006 (125); 2007 (186); 2008 (223); 2009 (274); 2010 (266); 2011 (282); 2012 (282); 2013 (289). O ano de 2014 não foi apresentado devido ao fato de não corresponder à produção total do ano.

Na Tabela 19 são mostrados os 25 pesquisadores com maior número de publicações no período. Pode-se notar que alguns pesquisadores são constantes quanto ao número de publicações por ano, enquanto outros, como Pavan ou Noia Maciel, passaram a publicar artigos na área a partir de 2007. Essa avaliação de temporalidade ajuda no entendimento do ciclo de vida de produção dos pesquisadores e pode colaborar na definição de políticas públicas quanto à colaboração de pesquisadores em suas carreiras.

**TABELA 19. Número de publicações por ano dos 25 pesquisadores com mais publicações no período 2003-2013**

Pesquisador	Total	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
BASSO, L. A.	79	5	8	7	10	6	8	8	6	5	9	7
SANTOS, D. S.	77	5	8	6	10	6	7	8	5	5	9	8
KRITSKI, A. L.	70	4	5	3	1	11	7	10	7	9	9	4
SCATENA VILLA, T. C.	53	0	0	0	0	3	6	6	8	11	11	8
TRAJMAN, A.	39	1	1	2	1	3	5	5	8	4	3	6
SUFFYS, P. N.	39	2	4	1	2	2	5	2	2	8	5	6
SILVA, C. L.	39	0	0	2	1	5	2	6	9	5	4	5
DIETZE, R.	33	2	1	1	2	1	6	10	3	3	2	2
RUFFINO- NETTO, A.	33	0	0	0	2	5	5	8	3	7	2	1
CHAISSON, R.	32	0	0	1	3	7	5	5	4	2	1	4
SARNO, E. N.	30	3	4	4	1	3	3	2	2	3	1	4
RODRIGUES, L. C.	28	4	3	5	1	3	3	1	2	2	3	1
LEITE, C. Q. F.	27	0	0	0	0	3	2	3	6	4	3	6
GOLUB, J. E.	27	0	0	0	0	2	5	3	6	5	1	5
PAVAN, F. R.	25	0	0	0	0	0	2	5	6	4	2	6
FERES SAADA, M. H.	23	0	1	3	2	3	3	1	5	1	3	1
PALACI, M.	23	1	3	0	1	2	3	5	2	1	2	3
FUJIMURA LEITE, C. Q.	23	0	0	0	0	2	2	1	3	5	1	9
PESSOLANI, M. C. V.	23	1	2	4	0	4	2	3	0	2	3	2
NOIA MACIEL, E. L.	22	0	0	0	0	0	3	8	6	2	0	3
GOMES, H. M.	21	1	1	0	2	1	2	1	1	3	5	4
LILEMBAUM, W.	21	1	0	0	2	1	0	4	4	1	5	3
ZAHA, A.	21	2	1	1	3	0	3	3	0	3	2	3
AZEVEDO JR, W. F.	21	0	0	0	0	2	8	6	2	1	2	0
JUNQUEIRA- KIPNIS, A. P.	18	0	0	0	0	1	4	2	3	4	2	2

A segunda análise de temporalidade aplicada aos artigos abrangeu as categorias definidas pela WoS (Tabela 20). Um mesmo artigo pode conter uma ou mais categorias definidas pela WoS, de acordo com seu tema e sua área de estudo. Neste caso específico estão sendo analisados somente os artigos com pelo menos um pesquisador afiliado a alguma instituição brasileira. Esses dados também podem servir de base para analisar o quanto certas categorias estão se tornando mais ou menos evidentes nas pesquisas.

TABELA 20. Avaliação temporal das categorias da Web of Science.

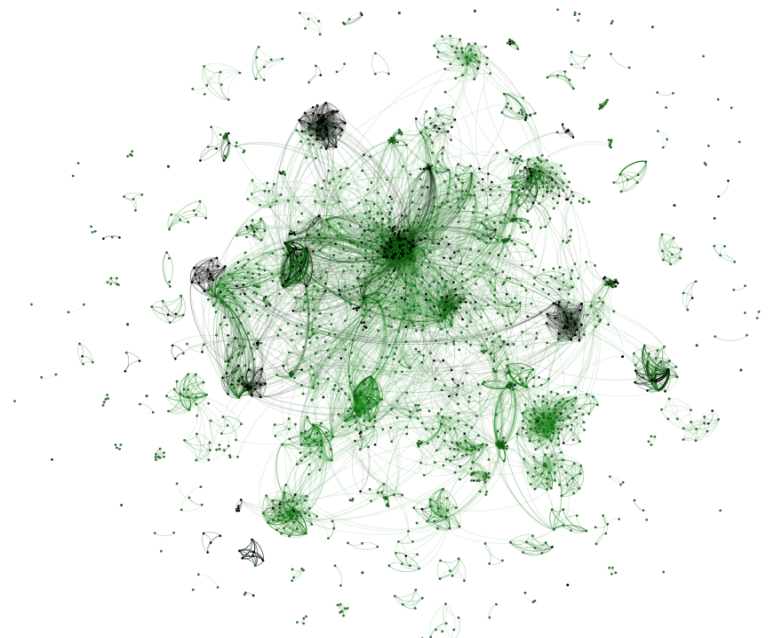
Categoria	Total	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
<i>Infectious diseases</i>	383	15	19	21	28	42	42	51	41	34	41	49
<i>Respiratory system</i>	297	10	9	17	20	16	19	63	40	36	36	31
<i>Immunology</i>	291	13	21	14	19	41	26	34	31	27	32	33
<i>Microbiology</i>	239	16	15	11	17	28	33	25	15	29	28	22
<i>Public, environmental and occupational health</i>	204	6	6	13	11	33	18	15	18	39	18	27
<i>Tropical medicine</i>	190	6	11	10	11	14	25	23	24	22	24	20
<i>Biochemistry and molecular biology</i>	162	8	11	8	14	7	22	24	15	25	17	11
<i>Pharmacology and pharmacy</i>	134	3	6	6	13	9	17	15	13	26	18	8
<i>Chemistry, medicinal</i>	132	1	2	3	8	5	14	20	16	29	14	20
<i>Parasitology</i>	85	4	7	4	6	5	6	11	10	12	11	9
<i>Medicine, research and experimental</i>	79	7	4	6	5	9	5	10	9	6	8	10
<i>Multidisciplinary sciences</i>	63	0	0	2	1	2	4	0	5	6	14	29
<i>Veterinary sciences</i>	71	4	3	2	6	2	4	9	7	9	15	10
<i>Nursing</i>	75	0	0	0	0	4	7	8	14	11	21	10
<i>Biophysics</i>	59	3	7	4	6	3	8	7	4	5	7	5
<i>Biotechnology and applied microbiology</i>	54	3	3	4	2	1	3	9	5	7	4	13
<i>Chemistry, multidisciplinary</i>	50	0	2	3	2	4	6	7	9	1	7	9
<i>Cell biology</i>	40	2	3	2	7	3	3	1	4	7	5	3
<i>Medicine, general and internal</i>	41	0	2	1	0	3	3	8	5	10	7	2
<i>Chemistry, organic</i>	38	0	0	0	1	4	9	9	5	1	5	4
<i>Biochemical research methods</i>	36	2	7	3	3	2	3	3	4	2	6	1
<i>Rheumatology</i>	33	0	1	1	4	1	3	4	8	5	3	3
<i>Genetics and heredity</i>	33	2	3	2	2	2	3	6	4	4	3	2
<i>Critical care medicine</i>	34	5	2	5	4	2	1	11	2	0	2	0
<i>Surgery</i>	32	1	2	0	1	6	0	6	3	3	4	6
<i>Transplantation</i>	27	0	1	1	0	4	1	5	2	4	2	7
<i>Dermatology</i>	26	1	1	0	1	1	1	2	4	8	4	3
<i>Chemistry, inorganic and nuclear</i>	22	0	0	1	0	1	5	2	3	2	4	4
<i>Crystallography</i>	22	0	2	0	1	2	0	1	6	3	2	5
<i>Pediatrics</i>	22	0	2	1	4	1	1	4	1	2	3	3

Na pesquisa em tuberculose pode-se observar que o tema *Infectious Diseases* é o mais frequente nos artigos, seguido pelos temas *Respiratory System*, *Immunology* e *Microbiology*. O tema *Biochemistry and molecular biology* passou a integrar boa parte das pesquisas a partir de 2008, e a categoria *Critical Care Medicine* foi tema de vários artigos no ano de 2008, mas não de tantos em outros anos. As publicações nos temas *Nursing*, *Biotechnology and applied microbiology* e *Multidisciplinary sciences* têm crescido nos últimos anos.

### 5.2.2 Rede de colaboração em pesquisa sobre tuberculose

O grafo de rede de colaboração tem como objetivo demonstrar como estão representadas as relações entre os diferentes pesquisadores. Dependendo da distribuição de rede utilizada pode-se ter uma visualização com características diferentes que enfatizam mais ou menos certos padrões. A Figura 42 ilustra a rede de coautoria em pesquisa sobre tuberculose que envolve cientistas afiliados a instituições brasileiras. Estes estão representados pelos nós de cor verde, e os pesquisadores de outros países estão representados pelos nós de cor preta.

**FIGURA 42. Rede de colaboração em pesquisa sobre tuberculose de pesquisadores afiliados a instituições brasileiras**



Fonte: elaboração de Bruna Fonseca – Relatório GT de Redes Fiocruz, 2014

Legenda: os nós representam os pesquisadores, e o compartilhamento da autoria de uma publicação científica é indicado por uma ligação entre os respectivos nós. As cores dos nós indicam se o pesquisador é de nacionalidade brasileira (verde) ou estrangeira (preto).

No grafo da Figura 42 foi feita uma distribuição com base no algoritmo *Force Atlas*, que tem como característica o adensamento de nós com maiores relações e o distanciamento de grupos menos coesos. Essa representação é diferente da distribuição de nós das figuras apresentadas na leishmaniose, quando foi utilizada a distribuição Fruchterman Reingold. Pode-se notar pela Figura 42 que os grupos mais ao centro, além de terem uma forte conexão entre os nós do mesmo grupo, possuem também uma forte conexão com outros grupos ou nós próximos. Com essa distribuição fica mais fácil a identificação de subgrupos de pesquisa dentro da rede.

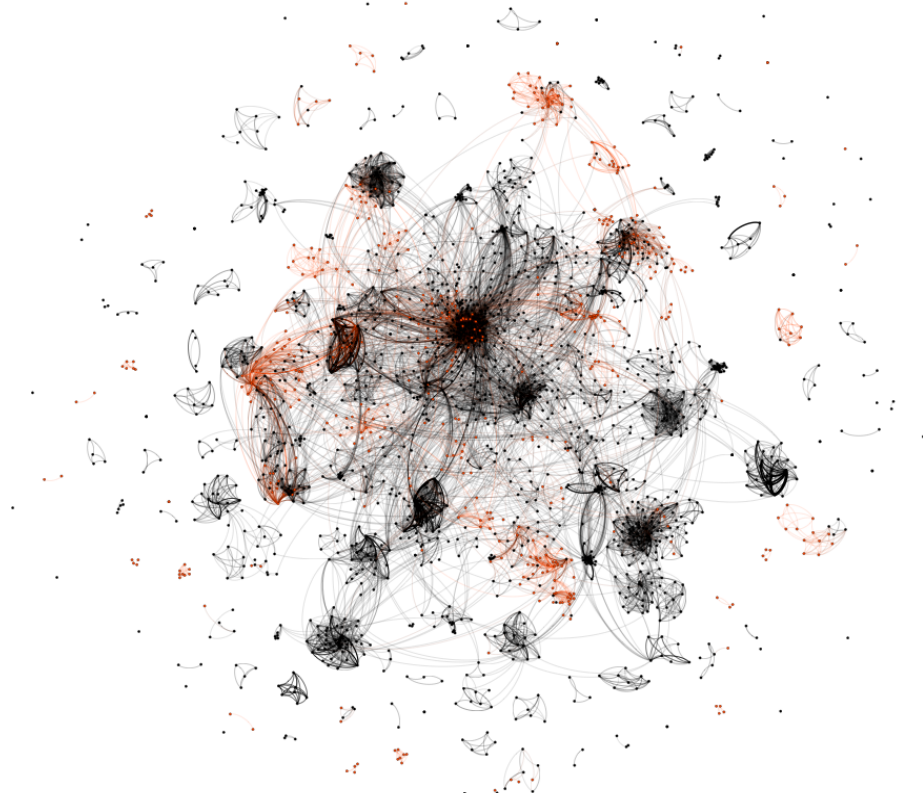
O que se avalia nessa figura são as colaborações internacionais e onde se encontram esses colaboradores quanto aos agrupamentos percebidos na figura. Pode-se notar que as instituições internacionais colaboram muito com o grupo que está ao centro, pois normalmente ele é constituído por pesquisadores com maior número de publicações e consequentemente com maior número de colaborações. Isso implica dizer que existe uma relação direta entre colaborações na rede e colaborações internacionais.

Outros três focos de nós pretos estão acima, ao lado esquerdo e ao lado direito do centro da figura. No grupo da parte superior existe uma concentração maior de instituições estrangeiras em relação às brasileiras, o que demonstra a participação de um grupo de pesquisadores em algum grupo de pesquisa específico e bastante forte estrangeiro. Nas concentrações do lado esquerdo o número de nós de pesquisadores de instituições brasileiras parece ser maior, dessa maneira imagina-se uma relação inversa, na qual pesquisadores estrangeiros se aproximam de instituições brasileiras. Na última concentração internacional com nós pretos do lado direito da figura parece haver poucos nós verdes, o que pode ser um ou poucos pesquisadores brasileiros se relacionando com um grupo estrangeiro específico.

Essas análises são baseadas na visualização do grafo, que, segundo a proposta de pesquisa, se relaciona com a “inspeção visual”. Uma vez feita essa análise visual exploratória, é necessário um aprofundamento da pesquisa para se avaliar os motivos e as características dessas relações. Neste momento as análises realizadas estão apenas no nível macro, e não nos detalhes das relações.

Na Figura 43 observa-se a mesma rede com os pesquisadores da Fiocruz, representados pelos nós de cor laranja, e os demais pesquisadores, independentemente de sua nacionalidade, representados pelos nós de cor preta. Os pesquisadores da Fiocruz representam 21,87% de toda a rede.

**FIGURA 43. Rede de colaboração em pesquisa sobre tuberculose de pesquisadores afiliados a instituições brasileiras**



Fonte: elaboração de Bruna Fonseca – Relatório GT de Redes Fiocruz, 2014

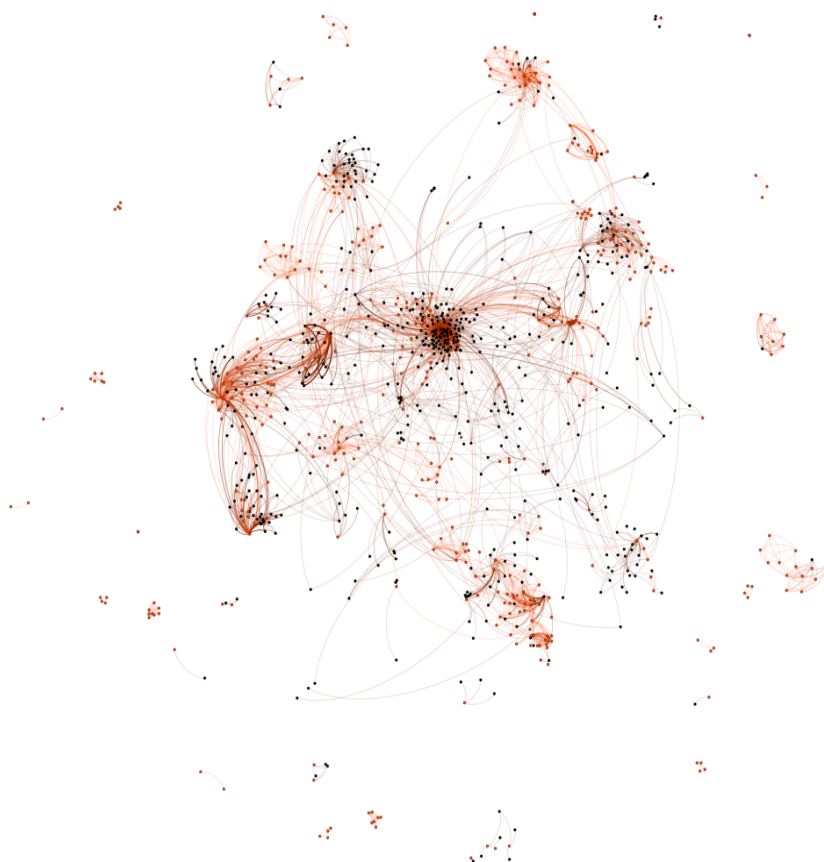
Legenda: os nós representam os pesquisadores, e o compartilhamento da autoria de uma publicação científica é indicado por uma ligação entre os respectivos nós. As cores dos nós indicam se o pesquisador é afiliado à Fiocruz (laranja) ou a outras instituições (preto).

O que se pode notar nessa figura é a dispersão dos nós laranja, que representam a Fiocruz. Eles estão espalhados por todo o espectro do grafo, o que implica a participação em diferentes grupos de pesquisa no território nacional. Sendo a Fiocruz uma referência na área de pesquisa em saúde, essa constatação encontra-se alinhada com a realidade. Os demais pesquisadores tendem a estar vinculados a institutos de ensino e pesquisa.

Na Figura 44 observa-se a rede apenas dos pesquisadores da Fiocruz, representados pelos nós de cor laranja, e os demais pesquisadores que possuem relações diretas com esses pesquisadores da Fiocruz, independentemente de sua nacionalidade, representados pelos nós de cor preta.



**FIGURA 44. Rede de colaboração em pesquisa sobre tuberculose de pesquisadores afiliados à Fiocruz**



Fonte: elaboração de Bruna Fonseca – Relatório GT de Redes Fiocruz, 2014

Legenda: os nós representam os pesquisadores, e o compartilhamento da autoria de uma publicação científica é indicado por uma ligação entre os respectivos nós. As cores dos nós indicam se o pesquisador é afiliado à Fiocruz (laranja) ou a outras instituições (preto).

A Figura 44, por fim, apresenta os pesquisadores da Fiocruz e suas relações com outros pesquisadores, assim como foi feito na LN. Os dados da rede com todos os nós e com os nós somente da Fiocruz são mostrados adiante. Os principais indicadores dessas duas redes estão apresentados na Tabela 21. As medidas de rede das publicações em tuberculose foram realizadas levando-se em consideração todo o período analisado (2003-2014). Na tabela são mostrados os graus médio e máximo, o coeficiente de agregação e a distância média da rede.

A avaliação do parâmetro de densidade, que se traduz na razão entre as interações efetivamente existentes entre os atores da rede e o total de ligações potenciais ou possíveis, aponta que a rede em questão é extremamente esparsa, com apenas 0,5% (de um total de 100%) do potencial de interação da rede sendo utilizado. A baixa densidade evidencia a subutilização do potencial da rede, com muitas possibilidades de relacionamento ainda inexploradas. O grau de centralização da rede complementa a avaliação da densidade, demonstrando que a rede analisada não está concentrada em seu nó mais central. Esses valores diferenciam-se dos resultados encontrados na rede LN na WoS devido ao recorte de

pesquisadores utilizado: pesquisadores com duas ou mais publicações em TB e pesquisadores com três ou mais publicações em LN.

**TABELA 21. Indicadores da rede de colaboração em pesquisa sobre tuberculose**

Indicador	Valor – Geral	Valor – Fiocruz
Número de nós	2.181	1.011
Número de <i>links</i>	11.369	2.984
Grau médio	10,28	5,90
Grau máximo	208	168
Coefficiente de agregação	0,730	0,757
Distância média da rede	13	12
Grau de centralização da rede	0,09	0,16
Densidade da rede	0,005	0,006

No que diz respeito à análise do componente gigante – maior grupo de nós que estão conectados uns aos outros, mas não estão conectados a outros nós na rede –, dos 2.181 pesquisadores identificados há apenas 365 que não se conectam com o resto da rede. Isso significa que 83% da rede está conectada quanto às colaborações. Uma vez definido o componente gigante, foi aplicado o algoritmo para a clusterização da rede ou a formação de comunidades.

A Tabela 22 mostra os países do mundo, com exceção do Brasil, com os quais os pesquisadores brasileiros mais colaboram, assim como os países com os quais os pesquisadores da Fiocruz mais colaboram.

**TABELA 22. Países que mais colaboram com os cientistas brasileiros e da Fiocruz no âmbito da rede de colaboração em pesquisa sobre tuberculose**

País	Intensidade da colaboração (Brasil)	País	Intensidade da colaboração (Fiocruz)
Estados Unidos	34,9%	Estados Unidos	40,2%
França	16,8%	França	16,3%
Suíça	6,5%	Inglaterra	4,3%
Inglaterra	4,9%	Holanda, Itália e Espanha	3,8%
Espanha	3,4%	Portugal	3,2%

A análise do padrão de cooperação tanto do Brasil quanto da Fiocruz mostra que seus principais países colaboradores não são aqueles com alta incidência de tuberculose, e sim países cuja incidência é de menos de cinquenta novos casos por cada 100 mil habitantes. Isso revela certa desarticulação da pesquisa em tuberculose realizada por pesquisadores brasileiros, tendo em vista que os países que mais se beneficiariam do avanço científico na área ou que poderiam fornecer dados e materiais importantes para a pesquisa não estão inseridos de maneira significativa na rede de cooperação neste tema.

### 5.2.3 Subgrupos de pesquisa em tuberculose

O estudo sobre os agrupamentos formados com base nas relações tem sido de grande valia para o estudo de grupos de pesquisa, pois demonstra de forma mais objetiva onde se encontram os pesquisadores e quais as áreas nas quais estes tendem a trabalhar com maior frequência. O que se tem encontrado até o momento tem sido uma forte relação entre colaboração e posicionamento geográfico. Com a rede de tuberculose não tem sido diferente.

Na Tabela 23 são apresentados os agrupamentos formados e o número de nós em cada rede. São apresentados também o grau médio de cada grupo, o grau máximo, que reflete o autor com maior número de relações dentro do grupo, o número de relações dentro da comunidade e, por último, o número de relações dos autores de um grupo com outros grupos.

**TABELA 23. Comunidades e relações existentes na rede de pesquisa em tuberculose**

<b>Grupo</b>	<b>Nós</b>	<b>Grau médio</b>	<b>Grau máximo</b>	<b>Relações no grupo</b>	<b>Relações com outros grupos</b>
1	366	14,2	168	2595	560
2	80	10,9	49	437	178
3	99	11,1	74	553	69
4	112	8,8	39	498	387
5	25	10,8	24	136	10
6	60	8,8	35	264	184
7	52	7,8	35	205	238
8	11	8,7	10	48	27
9	26	6,3	16	82	83
10	32	21,2	30	339	10
11	136	9,7	55	663	219
12	108	9,2	45	498	80
13	39	9,5	34	187	42
14	32	9,8	22	158	22
15	106	9,1	44	485	27
16	61	9,6	45	294	56
17	257	4	18	517	187
18	56	6,25	14	175	110
19	8	7	7	28	12
20	90	7,7	26	346	73
21	11	2,7	7	15	24
22	38	17,6	28	336	60

Foram encontradas 22 comunidades na rede, mas uma análise qualitativa faz-se necessária para avaliar o grau de coerência dos agrupamentos gerados pelo algoritmo com os grupos de pesquisa ou laboratoriais que existem na realidade.

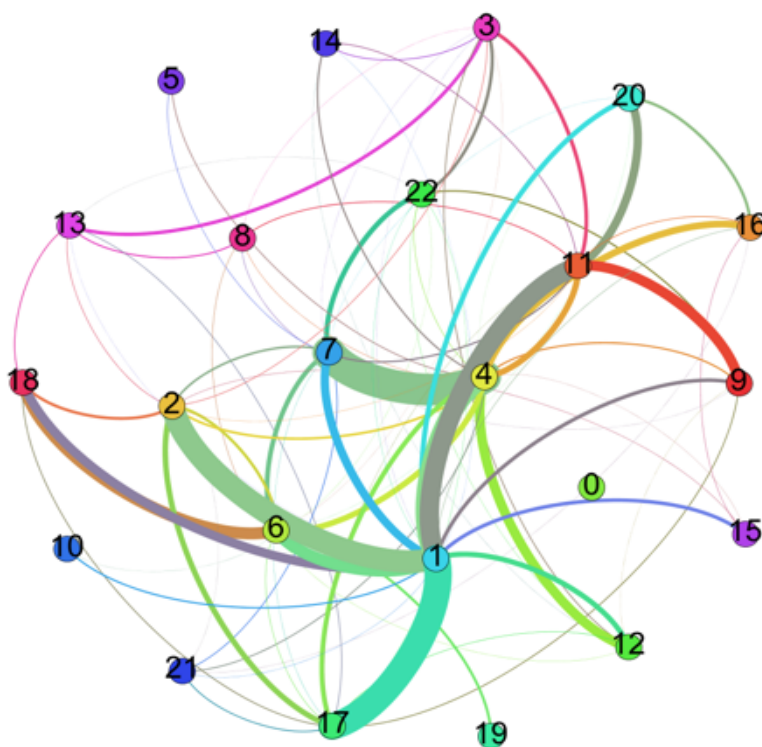
No que diz respeito aos graus médio e máximo, percebe-se o quanto um grupo é altamente conectado e o quanto pessoas dentro do grupo exercem forte influência em decorrência das suas relações com os demais. No caso do grupo 3, por exemplo, o grau máximo ou o pesquisador com o maior número de relações possui contatos com 75% do

grupo ao qual pertence. Isso significa que muito do que é realizado ou pesquisado por este autor é caracterizado ou seguido pelo grupo. Normalmente, nesses casos o pesquisador apresenta-se como o responsável por um laboratório ou grupo de pesquisa. No caso do grupo 19, todos os pesquisadores estão conectados entre si, uma vez que os graus médio e máximo têm o mesmo valor, provavelmente em razão de um ou mais artigos terem sido elaborados por todos conjuntamente.

No que diz respeito à avaliação de relações dentro e fora do grupo, percebe-se também características diferentes entre os grupos. O grupo 1 é mais coeso, apesar de seu grande número de relações externas ao grupo. Ele possui cerca de cinco vezes mais relações dentro do próprio grupo do que com pesquisadores externos. Essa é uma característica da maioria dos grupos, nos quais os relacionamentos externos estão a cargo de alguns poucos pesquisadores, também conhecidos como *gatekeepers*. No entanto, também existem grupos como o grupo 9, cujo número de relações internas e externas é similar, diferenciando-se dos demais.

As relações entre os grupos podem ser mais bem percebidas na Figura 45. A espessura das linhas demonstra a intensidade ou o número de relações entre os grupos. Os grupos 1 e 17, por exemplo, possuem um número de relações entre si muito maior do que os grupos 1 e 12.

**FIGURA 45. Rede de relacionamentos entre grupos de pesquisa sobre tuberculose**

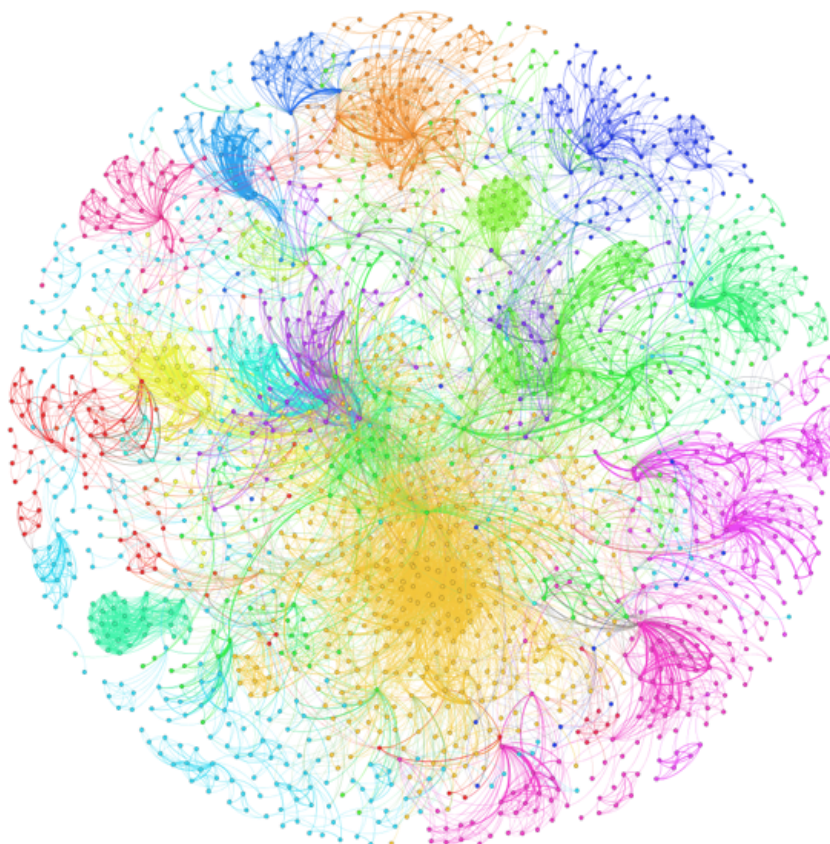


Fonte: elaboração do autor

Legenda: os nós representam os grupos, e o compartilhamento da autoria de uma publicação científica entre os membros dos grupos é indicado por uma ligação entre os respectivos nós. A espessura das linhas indica a intensidade ou a frequência das relações.

Na Figura 46 está demonstrado como esses grupos se apresentam na rede, exibindo todos os seus pesquisadores, e não somente um nó por grupo, como na figura anterior. Por meio de uma análise visual é possível perceber como os pesquisadores de um mesmo grupo tendem a se relacionar entre si. Isso na verdade é uma característica do próprio algoritmo utilizado, que leva em consideração as relações similares entre os autores. A distribuição utilizada na figura foi a Fruchterman Reingold, que tende a deixar os nós com maior grau no centro da figura, enquanto os nós com menor grau ficam mais na periferia. Com essa distribuição pode-se perceber também os grupos de maior colaboração e, conseqüentemente, maior influência na rede.

**FIGURA 46. Rede de relacionamentos entre grupos de pesquisa sobre tuberculose, separada por grupos**



Fonte: elaboração do autor

Legenda: os nós representam os pesquisadores, e o compartilhamento da autoria de uma publicação científica é indicado por uma ligação entre os respectivos nós. As cores dos nós indicam o grupo ao qual o pesquisador pertence.

Além do recorte sobre os pesquisadores brasileiros, foi feita uma análise levando-se em consideração as instituições às quais os pesquisadores são afiliados segundo sua definição nos artigos indexados pela WoS. Para avaliar a influência, o relacionamento e o posicionamento dos pesquisadores da Fiocruz nos respectivos grupos, a Figura 47 representa a rede de tuberculose original, mas com os nós em vermelho representando os pesquisadores da Fiocruz. A figura em questão revela-nos que esses pesquisadores estão “pulverizados” por toda a rede. Existe uma concentração na parte superior direita na qual quase todo o grupo é representado por pesquisadores da instituição, no entanto quase todos os demais grupos congregam pesquisadores da Fiocruz.

**FIGURA 47. Rede de tuberculose e pesquisadores Fiocruz – grupos**



Fonte: elaboração do autor

Legenda: os nós representam os pesquisadores, e o compartilhamento da autoria de uma publicação científica é indicado por uma ligação entre os respectivos nós. As cores dos nós indicam o grupo ao qual o pesquisador pertence. Os pesquisadores da Fiocruz estão representados pelos nós de cor vermelha.

A Figura 48 apresenta a mesma rede que a anterior, mas apenas com os pesquisadores da Fiocruz, representados pelos nós em vermelho, e seus colaboradores diretos. São ao todo 985 pesquisadores na rede. Pode-se perceber que ao centro existe um grupo bastante denso e de cor amarelada, no qual apenas um pequeno percentual de nós é da Fiocruz. Esse grupo está relacionado a instituições de pesquisa internacional, e uma pequena parcela dos colaboradores é proveniente de entidades brasileiras. Outro grupo interessante de ser analisado está no lado esquerdo inferior, em cor esverdeada. Esse grupo é extremamente coeso, podendo inclusive ser um clique, que é um subgrupo no qual todos os nós se conectam entre si. Esse grupo foi formado muito provavelmente pela construção de um ou mais artigos com elevado número de coautores.

**FIGURA 48. Rede de relacionamentos dos pesquisadores Fiocruz – grupos**



Fonte: elaboração do autor

Legenda: os nós representam os pesquisadores, e o compartilhamento da autoria de uma publicação científica é indicado por uma ligação entre os respectivos nós. As cores dos nós indicam o grupo ao qual o pesquisador pertence. Os pesquisadores da Fiocruz estão representados pelos nós de cor vermelha.

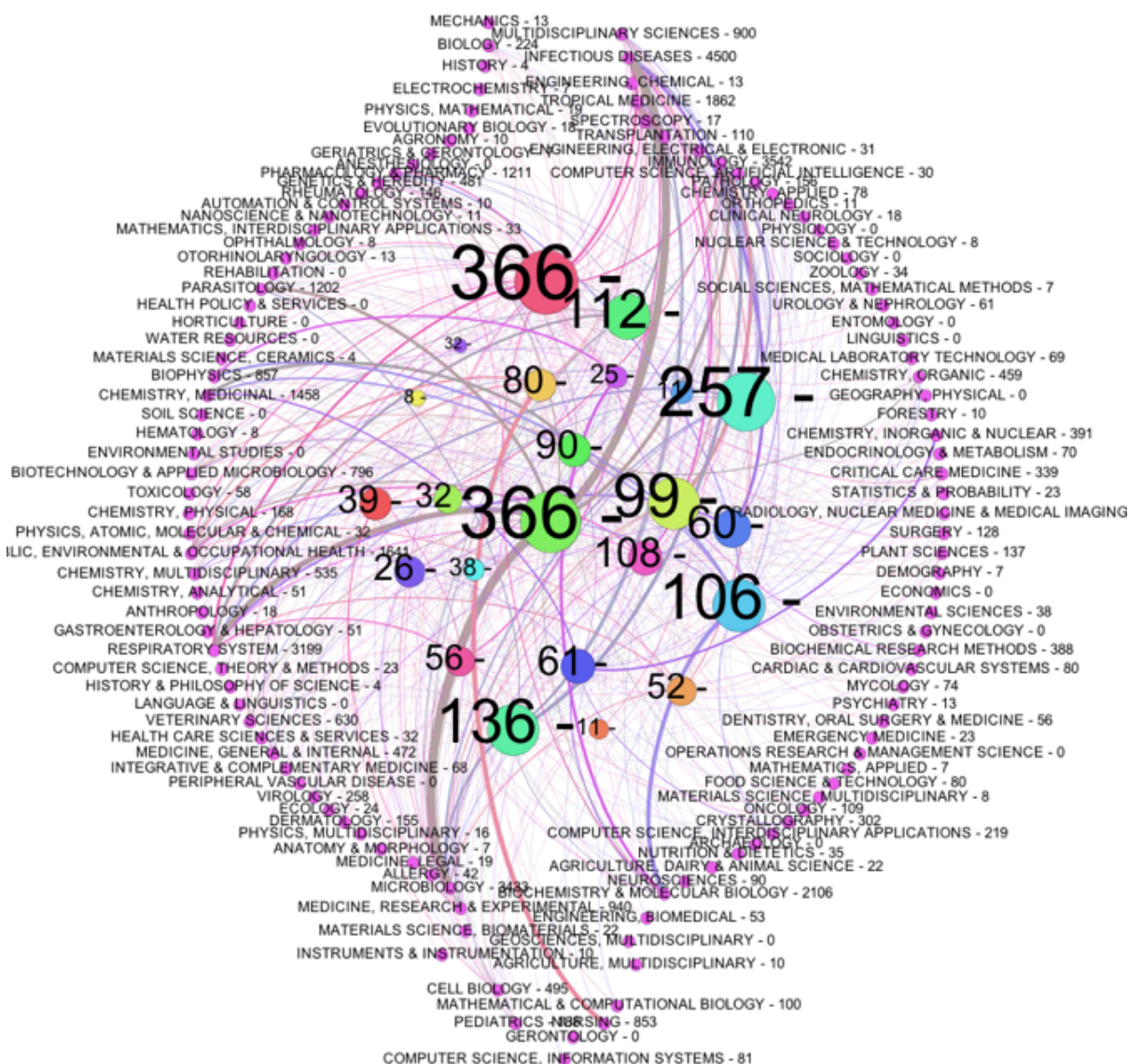
#### **5.2.4 Grandes áreas do conhecimento e caracterização dos grupos**

A WoS utiliza uma forma de classificação de seus periódicos indexados chamada de *Web of Science Subject Categories* ou *Areas*. Essas categorias totalizam 250. Nos artigos foco deste estudo foram encontradas 121 categorias, estando as trinta categorias mais frequentes representadas na Tabela 13, no início desta subseção.

Como cada artigo está inserido em uma ou mais dessas classificações, os pesquisadores podem ser ordenados segundo essas áreas bem como segundo os grupos de pesquisa que eles representam. Algumas categorias, como *Infectious diseases*, *Respiratory system* ou *Immunology*, são utilizadas pela maioria dos pesquisadores e consequentemente pelos grupos. Todavia, outras categorias tendem a ser mais evidentes em grupos específicos. Devido a essa divisão, propõe-se uma análise de classificação dos grupos por meio das categorias que podem ser representantes de cada grupo. Neste momento, restringiu-se a apresentação visual de como alguns dos grupos se relacionam com essas categorias e de como essas categorias representam alguns grupos (Figura 49).



FIGURA 49. Relações dos grupos com as categorias



Fonte: elaboração do autor

Legenda: os nós no centro representam os grupos, e os nós na periferia representam as categorias da WoS. A inclusão de um artigo publicado pelo grupo em determinada categoria da WoS é indicada por uma ligação entre os respectivos nós. As cores dos nós no centro da rede indicam o grupo ao qual o pesquisador pertence, e o tamanho desses nós reflete o número de pesquisadores que compõem o grupo. A espessura das linhas apresenta a frequência com a qual determinada categoria foi aplicada a uma publicação do grupo. Os números acima dos nós representam o número de pesquisadores que compõem o grupo. O círculo vermelho superior com o número 366 é formado pelos pesquisadores que não fazem parte do componente gigante.

Na Figura 49 são apresentadas todas as categorias relacionadas ao número de vezes em que foram utilizadas nos artigos. A espessura das linhas que ligam uma categoria a um grupo está relacionada ao número de vezes que aquela categoria foi utilizada pelo grupo. O tamanho dos nós ao centro está relacionado ao número de pesquisadores dentro do grupo, que é também indicado pelo número à frente do nó. Por sua vez, o círculo verde ao centro com 366

nós está relacionado ao grupo 1, segundo a Tabela 23. O nó vermelho na parte superior, também com 366 nós, não se encontra na Tabela 23, pois são os nós que estão fora do componente gigante. A Figura 49 representa uma rede de dois modos, na qual os nós são os grupos de pesquisadores de um lado e as categorias do outro, com cada elemento, seja dos grupos seja das categorias, só podendo se relacionar com membros do outro tipo.

O que se pode notar na Figura 49 é que todos os grupos possuem uma grande diversidade de temas ao qual cada grupo está relacionado. O número de relações entre as categorias e os grupos é muito alto, o que dificulta uma melhor compreensão da estrutura de relacionamentos. Pode-se notar também que alguns temas têm uma maior frequência, com linhas de maior calibre saindo para grupos específicos. Essas linhas e áreas do conhecimento estão relacionadas diretamente aos dados bibliométricos quanto ao número de aparições para cada um dos temas. Outra diferença perceptível na figura diz respeito aos dois grupos com 366 pesquisadores. O grupo do centro tem fortes relações com temas específicos; por sua vez, o grupo ou nó da parte superior não tem uma referência perceptível com nenhum dos temas apresentados.

Nas figuras seguintes buscou-se demonstrar como cada grupo é caracterizado por categorias específicas, ou seja, quais as categorias em que os artigos dos seus pesquisadores foram enquadrados.

No caso dos dois maiores grupos, com 366 pesquisadores cada, há uma diferença clara quanto ao número de relações com categorias específicas. No caso do grupo 1, apresentado na Figura 50, o número de publicações por autor é maior que no grupo de pesquisadores fora do componente gigante, o que pode ser observado pelo número de relações com categorias específicas, como *Microbiology* ou *Infectious disease*, e a espessura da linha. Entretanto, o grupo ainda aparenta ser bastante pulverizado quanto às categorias e consequente escolha de revistas para publicação. Um aprofundamento quanto aos pesquisadores e suas áreas de trabalho nesse grupo pode trazer informações que não foram percebidas nesse momento.

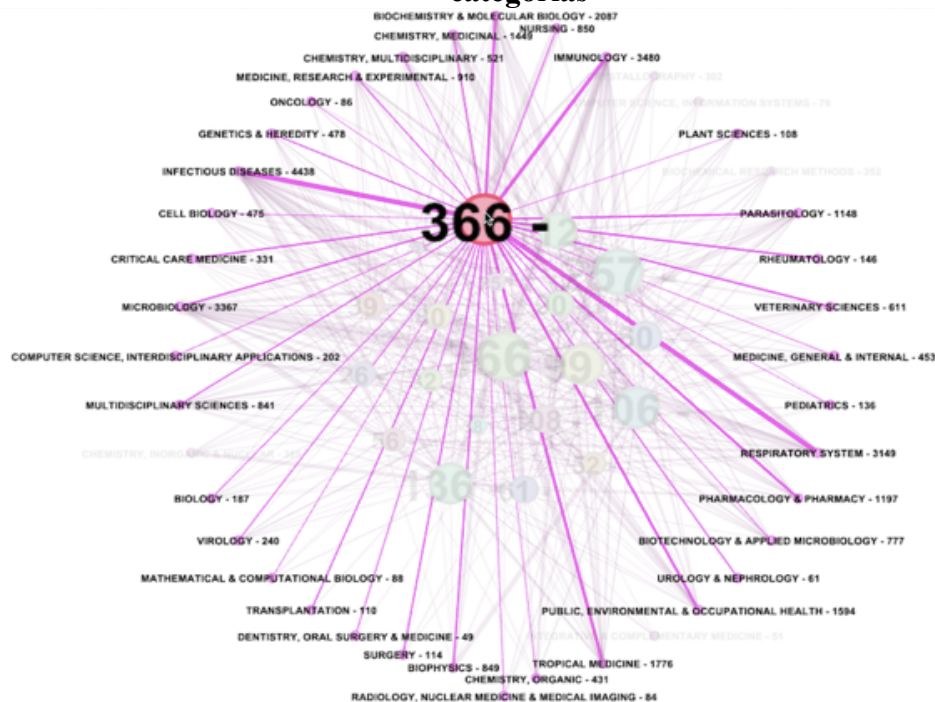
**FIGURA 50. Relações do grupo 1 com as categorias**



Fonte: elaboração do autor

Legenda: o nó destacado no centro representa o grupo 1, e os nós na periferia representam as categorias da WoS. A inclusão de um artigo publicado pelo grupo em determinada categoria da WoS é indicada por uma ligação entre os respectivos nós. A espessura das linhas mostra a frequência com a qual determinada categoria foi aplicada a uma publicação do grupo.

**FIGURA 51. Relações do grupo localizado fora do componente gigante com as categorias**



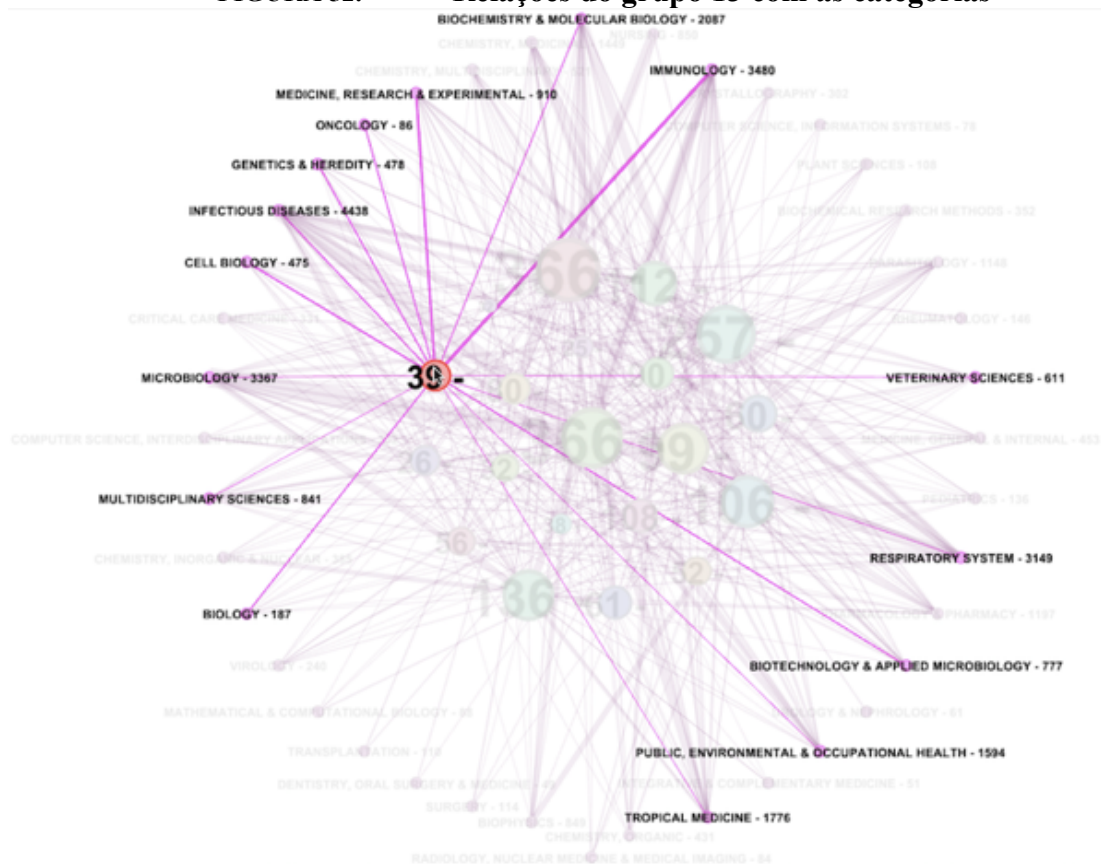
Fonte: elaboração do autor

Legenda: o nó destacado no centro representa o grupo de pesquisadores que não está incluído no componente gigante da rede (desconectados), e os nós na periferia representam as categorias da WoS. A inclusão de um artigo publicado pelo grupo em determinada categoria da WoS é indicada por uma ligação entre os respectivos nós. A espessura das linhas mostra a frequência com a qual determinada categoria foi aplicada a uma publicação do grupo.

A Figura 51 mostra que existe um número ainda maior de pulverização dos temas nesse grupo. Como este é formado por autores de diversos subgrupos que não estavam conectados entre si, esta característica é facilmente observada no grafo. Nenhum dos temas específicos possui forte relevância, além de estar abrangendo quase todas as áreas encontradas na pesquisa.

As Figuras 52 e 53 são apresentadas como referência para demonstrar que alguns grupos têm focos mais específicos.

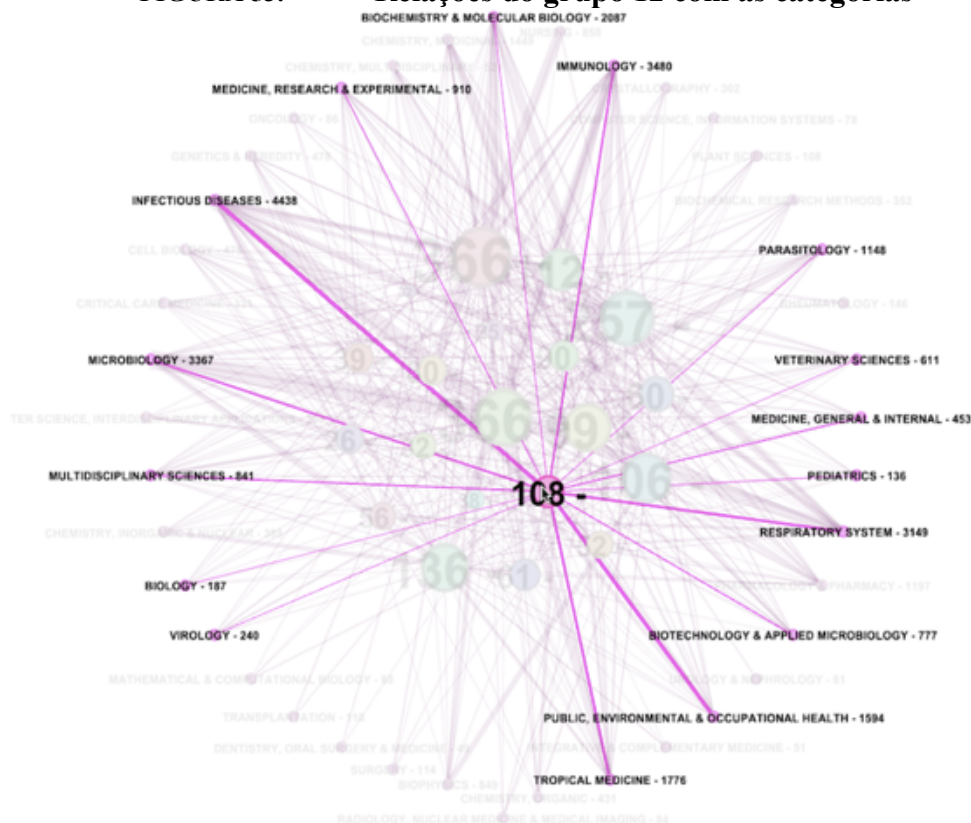
**FIGURA 52. Relações do grupo 13 com as categorias**



Fonte: elaboração do autor

Legenda: o nó destacado no centro representa o grupo 13, e os nós na periferia representam as categorias da WoS. A inclusão de um artigo publicado pelo grupo em determinada categoria da WoS é indicada por uma ligação entre os respectivos nós. A espessura das linhas mostra a frequência com a qual determinada categoria foi aplicada a uma publicação do grupo.

**FIGURA 53. Relações do grupo 12 com as categorias**

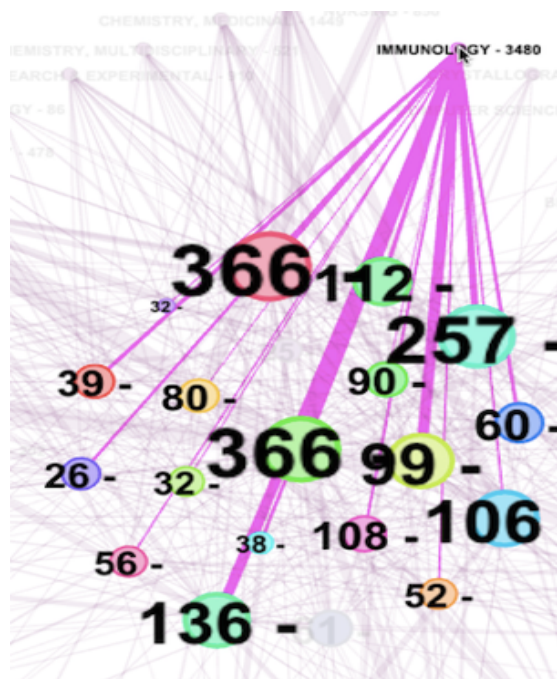


Fonte: elaboração do autor

Legenda: o nó destacado no centro representa o grupo 12, e os nós na periferia representam as categorias da WoS. A inclusão de um artigo publicado pelo grupo em determinada categoria da WoS é indicada por uma ligação entre os respectivos nós. A espessura das linhas mostra a frequência com a qual determinada categoria foi aplicada a uma publicação do grupo.

As Figuras 54 e 55 trazem as mesmas informações que as figuras anteriores, mas com foco nas categorias, demonstrando quais os grupos que possuem pesquisadores atuantes naquelas áreas específicas. Algumas áreas abrangem todos os grupos, como evidenciado na Figura 54, enquanto outras abarcam grupos de pesquisa mais específicos, conforme exposto na Figura 55.

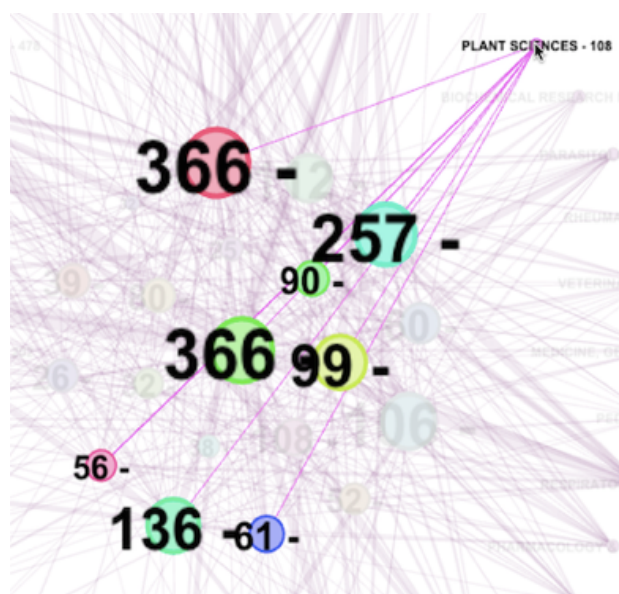
FIGURA 54. Relações da categoria *Immunology* com os grupos



Fonte: elaboração do autor

Legenda: o nó destacado na periferia representa a categoria *Immunology* da WoS, e os nós mais ao centro representam os grupos de pesquisadores. A inclusão de um artigo publicado pelo grupo na categoria *Immunology* é indicada por uma ligação entre os respectivos nós. A espessura das linhas mostra a frequência com a qual a categoria foi aplicada a uma publicação do grupo.

FIGURA 55. Relações da categoria *Plant Sciences* com os grupos



Fonte: elaboração do autor

Legenda: o nó destacado na periferia representa a categoria *Plant Sciences* da WoS, e os nós mais ao centro representam os grupos de pesquisadores. A inclusão de um artigo publicado pelo grupo na categoria *Plant Sciences* é indicada por uma ligação entre os respectivos nós. A espessura das linhas mostra a frequência com a qual a categoria foi aplicada a uma publicação do grupo.

### 5.3 Descrição das categorias Web of Science<sup>11</sup>

#### *Biochemical Research Methods* (métodos de investigação bioquímica)

Os métodos de investigação bioquímica baseiam-se em técnicas específicas utilizadas nas pesquisas biológicas e bioquímicas, incluindo-se os procedimentos para a purificação e a análise de biomoléculas, a observação da estrutura ou da função de organismos e tecidos vivos (exclusivos de microscopia) e a alteração de biomoléculas para aplicações em investigações específicas. Esta categoria não engloba aplicações clínicas ou o desenvolvimento e a concepção de ferramentas de diagnóstico.

#### *Biochemistry and molecular biology* (bioquímica e biologia molecular)

Esta categoria abrange a bioquímica e a biologia molecular em temas gerais como carboidratos, lipídios, proteínas, ácidos nucleicos, genes, drogas, substâncias tóxicas e outros produtos químicos ou componentes moleculares de células, microrganismos, plantas e animais, incluindo seres humanos. Excluem-se os temas ligados à bioquímica de células, a tecidos e a órgãos. Excluem-se também os temas cujo principal foco é o organismo de estudo, por exemplo, plantas, micróbios, etc.

#### *Biology* (biologia)

A biologia inclui temas de pesquisa cuja abordagem é ampla ou interdisciplinar. Além disso, inclui materiais que cobrem uma área específica da biologia não abrangida em outras categorias, como biologia teórica, biologia matemática, biologia térmica, *cryobiology* e pesquisa de ritmo biológico.

#### *Biophysics* (biofísica)

Nesta categoria estão incluídos assuntos como a transferência e os efeitos das forças físicas e da energia de luz, som, eletricidade, magnetismo, calor, frio, pressão, forças mecânicas e radiação dentro e sobre as células, tecidos e organismos inteiros.

#### *Biotechnology and applied microbiology* (biotecnologia e microbiologia aplicada)

A biotecnologia e a microbiologia aplicada cobrem uma ampla gama de temas sobre a manipulação de organismos vivos para fabricar produtos ou resolver problemas para atender às necessidades humanas. Os tópicos incluem a engenharia genética; técnicas diagnósticas e terapêuticas moleculares; mineração de dados do genoma; bioprocessamento de alimentos e medicamentos; controle biológico de pragas; biorremediação ambiental; e produção de bioenergia. Nesta categoria também estão inclusos temas como desenvolvimento social, negócios relacionados e questões regulatórias.

<sup>11</sup> <[http://ip-science.thomsonreuters.com/mjl/scope/scope\\_scie/](http://ip-science.thomsonreuters.com/mjl/scope/scope_scie/)>.

*Cell biology* (biologia celular)

Esta categoria trata de todos os aspectos da estrutura e da função das células eucarióticas. A principal característica dos temas nesta categoria é a ênfase na integração em nível celular da bioquímica molecular, da genética fisiológica e das informações patológicas. Nesta categoria também são considerados relevantes tecidos específicos, diferenciados e embrionários.

*Critical care medicine* (medicina de cuidado crítico)

A medicina de cuidado crítico abarca especialidades de saúde concentradas no atendimento de pacientes com doenças que trazem risco de vida ou provocam lesão aguda. Nesta categoria estão incluídos assuntos como ataque cardíaco; intoxicação; queimaduras; pneumonia; complicações cirúrgicas; nascimento prematuro; traumas, abrangendo traumatismo craniano; acidente vascular cerebral e outras lesões neurais; anestesia; cuidados intensivos; e reanimação.

*Crystallography* (cristalografia)

A cristalografia compreende o estudo da formação, da estrutura e das propriedades dos cristais. Esta categoria também inclui meios de cristalografia de raios-X e o estudo da estrutura interna dos cristais por meio da utilização da difração de raios-X.

*Dermatology* (dermatologia)

A dermatologia é uma área que abrange a anatomia, a fisiologia e as patologias da pele. Abarca assuntos como dermatologia investigativa e experimental, dermatite de contato, cirurgia dermatológica, patologia dermatológica e oncologia dermatológica. Nesta categoria há também temas específicos como queimaduras, feridas e hanseníase.

*Entomology* (entomologia)

A entomologia abarca aspectos do estudo de insetos, incluindo entomologia geral, entomologia aplicada, entomologia regional, insetos aquáticos, bioquímica e fisiologia de insetos, entomologia econômica, manejo integrado de pragas, entomologia ambiental e ciência pesticida aplicada.

*Medicine, general and internal* (medicina geral e interna)

A medicina geral e interna compreende especialidades médicas, como clínica geral, medicina interna, fisiologia clínica, gestão da dor, medicina militar e hospitalar. Temas relacionados à medicina de família e aos serviços de cuidados de saúde primários são colocados na categoria cuidados de saúde primários.



*Genetics and heredity* (genética e hereditariedade)

Genética e hereditariedade abarcam assuntos ligados à estrutura, às funções e às propriedades de genes bem como às características de herança. Esta categoria também considera as características hereditárias, a genética da população, a frequência e a distribuição do polimorfismo, bem como as doenças hereditárias e os distúrbios do processo de replicação. A categoria é distinguível da Bioquímica e da Biologia Molecular por sua ênfase específica sobre o gene como uma única unidade funcional e sobre o efeito do gene no organismo como um todo.

*Immunology* (imunologia)

A imunologia relaciona-se a todos os aspectos da resposta imunológica, à regulação e aos níveis celular, molecular e clínico. Outros tópicos incluem estudos sobre a interação entre os patógenos e a imunidade do hospedeiro, a imunologia clínica, as imunoterapias emergentes, bem como sobre a contribuição imunológica para o curso da doença.

*Infectious diseases* (doenças infecciosas)

Doenças infecciosas dizem respeito a todos os aspectos da patogênese de doenças virais ou bacterianas clinicamente significativas, incluindo doenças sexualmente transmissíveis (DST), como HIV e Aids. Nesta categoria encontram-se também os assuntos ligados às interações patógeno-hospedeiro, à prevenção, ao diagnóstico, ao tratamento e à epidemiologia de doenças infecciosas.

*Chemistry, inorganic and nuclear* (química inorgânica e nuclear)

A química inorgânica e nuclear abrange tanto a química inorgânica como a nuclear. Na química inorgânica estão os elementos não carbono e a preparação, as propriedades e as reações de seus compostos. Também inclui estudos de determinados compostos de carbono simples, incluindo os óxidos de carbono, os dissulfuretos, os halogenetos, o cianeto de hidrogênio e sais, tais como cianetos, cianatos, carbonatos e hidrogenocarbonatos. Recursos relacionados à química de coordenação e a compostos organo-metálicos (que contêm uma ligação metal-carbono) são também abarcados nesta categoria. A química nuclear inclui estudos do núcleo atômico, compreendendo a fissão e as reações de fusão e seus produtos. Esta categoria compreende também a área de radioquímica com foco em temas como a preparação de compostos radioativos, a separação de isótopos por reações químicas, o uso de marcadores radioativos em experiências sobre as reações químicas e os compostos de elementos transurânicos.

*Medicinal chemistry* (química medicinal)

A química medicinal abrange temas como o isolamento e o estudo de substâncias com potencial terapêutico. São também temas de interesse as relações quantitativas estrutura-função, a caracterização estrutural, as sínteses orgânicas de compostos naturais e químicos e as técnicas de análise utilizadas no desenho racional de drogas.

*Microbiology* (microbiologia)

A microbiologia inclui recursos relacionados a todos os aspectos de estudos fundamentais e aplicados de microrganismos, incluindo bactérias, vírus e fungos. Esta categoria também considera os recursos que englobam os aspectos clínicos da ocorrência, o tratamento de patógenos microbianos, os estudos de ciência básica e da bioquímica microbiana, a microbiologia ambiental e os usos bacterianos/virais em biotecnologia.

*Multidisciplinary chemistry* (química multidisciplinar)

A química multidisciplinar abarca recursos que têm uma abordagem geral ou interdisciplinar das ciências químicas. Recursos de química em tópico especial que tenham relevância para muitas áreas da química também estão incluídos nesta categoria. Recursos com um foco primário em química analítica, inorgânica e nuclear, biológica, física ou polímero são colocados em suas próprias categorias.

*Multidisciplinary sciences* (ciências multidisciplinares)

Nas ciências multidisciplinares encontram-se os assuntos de caráter amplo ou geral nas ciências. Incluem-se aí as principais disciplinas científicas, como física, química, matemática, biologia, etc. Natureza e ciência são os temas mais proeminentes nesta categoria e servem como exemplos típicos. O site da National Science Foundation é um bom exemplo de um assunto web incluído nesta categoria. Alguns temas especializados com uma ampla gama de aplicações nas ciências podem também se enquadrar nesta categoria.

*Nursing* (enfermagem)

A enfermagem acolhe todos os aspectos da ciência e da prática da enfermagem, como administração, economia, gestão, educação, aplicações tecnológicas e todas as especialidades de cuidados clínicos.

*Organic chemistry* (química orgânica)

A química orgânica inclui questões ligadas a compostos orgânicos sintéticos e naturais e sua síntese, estrutura, propriedades e reatividade. A pesquisa sobre hidrocarbonetos, uma importante área da química orgânica, está incluída nesta categoria.

*Parasitology* (parasitologia)

A parasitologia abrange o estudo de parasitas e organismos que vivem dentro ou sobre outros organismos vivos e disso se beneficiam, mas muitas vezes causam danos a seus anfitriões.

*Pathology* (patologia)

A patologia engloba as técnicas, as causas e/ou os efeitos no desenvolvimento da doença em tecidos vivos. Esta categoria também considera as aplicações médicas e biomédicas de métodos histológicos e citogenéticos, o desenvolvimento e a utilização de novas técnicas de diagnóstico e o estudo patológico de tecidos ou doenças específicas.

*Pediatrics* (pediatria)

Nesta categoria estão a pesquisa básica e a clínica em pediatria. Várias especialidades pediátricas são cobertas, incluindo sistema respiratório, cardiologia, odontologia, dermatologia, comportamento do desenvolvimento, gastroenterologia, hematologia, imunologia e doenças infecciosas, neurologia, nutrição, oncologia, psiquiatria, cirurgia, medicina tropical, urologia e nefrologia. A cobertura também inclui perinatologia, neonatologia e medicina do adolescente.

*Pharmacology and pharmacy* (farmacologia e farmácia)

Farmacologia e farmácia abrangem questões referentes à descoberta e à experimentação de substâncias bioativas, incluindo investigação animal, experiência clínica, sistemas de entrega e distribuição de drogas. Esta categoria também inclui bioquímica, metabolismo e efeitos tóxicos ou adversos das drogas.

*Plant sciences* (ciências vegetais)

As ciências vegetais abrangem o estudo das plantas, incluindo tópicos sistemáticos, bioquímicos, agrícolas e farmacêuticos. Esta categoria inclui materiais sobre plantas superiores e inferiores, plantas terrestres e aquáticas, células vegetais, plantas inteiras e *assemblages* de plantas.

*Public, environmental and occupational health* (saúde pública, ambiental e ocupacional)

A saúde pública, ambiental e ocupacional trata de assuntos relacionados à epidemiologia, à higiene e à saúde; doenças parasitárias e parasitologia; medicina tropical; medicina da indústria; medicina ocupacional; controle de infecção; e medicina preventiva. Aqui também estão incluídos a saúde ambiental, as causas do câncer e seu controle, a aviação, o aerosol e a medicina da floresta.

*Research and experimental medicine* (pesquisa médica e experimental)

Na pesquisa médica e experimental está incluída a pesquisa médica geral, com especial destaque para técnicas extremamente novas e intervenções clínicas em uma ampla gama de especializações e aplicações médicas, adicionando-se o desenvolvimento de vacinas, a substituição de tecidos, imunoterapias e outras estratégias terapêuticas experimentais. Aqui se encontram também intervenções clínicas em estágios iniciais de desenvolvimento, utilizando modelos *in vitro* ou com animais e ensaios clínicos de pequena escala.

*Respiratory system* (sistema respiratório)

O sistema respiratório abrange todos os aspectos de doenças respiratórias e pulmonares, incluindo sua relação com a cirurgia e as doenças cardiovasculares e torácicas.

*Rheumatology* (reumatologia)

A reumatologia abarca a pesquisa clínica, terapêutica e de laboratório sobre artrite, reumatismo e doenças autoimunes, crônicas, degenerativas e inflamatórias que afetam principalmente as articulações e o tecido conjuntivo.

*Surgery* (cirurgia)

Na cirurgia encontram-se os temas cirúrgicos gerais, incluindo os diferentes tipos de cirurgia (cardiovascular, neurocirurgia, ortopedia, pediatria, ou vascular); disciplinas afins da cirurgia (oncologia cirúrgica, patologia, radiologia); e técnicas cirúrgicas (artroscopia, microscopia, endoscopia).

*Transplantation* (transplantação)

Transplantação relaciona-se à assimilação do tecido enxertado e à reconstituição de órgãos ou partes de órgãos removidos. Concentra-se em procedimentos de transplante e na manutenção de tecidos ou órgãos transplantados, bem como em transplantes específicos, tais como de coração, pulmão, rim e medula óssea.

*Tropical medicine* (medicina tropical)

Na medicina tropical estão concentrados o estudo e o tratamento de doenças parasitárias e outras condições médicas únicas ou originárias de regiões tropicais.

*Veterinary sciences* (ciências veterinárias)

As ciências veterinárias dizem respeito tanto à pesquisa como aos aspectos clínicos da saúde animal, doenças, lesões, nutrição, reprodução e saúde pública. Nesta categoria também estão materiais de fazenda, jardim zoológico, laboratório e animais aquáticos selvagens.

## 6 Resultados da pesquisa no SciELO

O modelo exploratório de análise de redes com foco nas relações de coautoria em publicações científicas foi aplicado pela terceira vez neste trabalho de pesquisa em uma base de dados brasileira, a Scientific Eletronic Library Online (SciELO). Foram utilizadas as duas áreas de pesquisa previamente estudadas, leishmaniose e tuberculose, agora buscando se fazer uma comparação direta entre elas com o mesmo recorte temporal.

No estudo das colaborações científicas previamente realizado, o foco principal eram as redes de coautoria, com o intuito de aprimorar o entendimento sobre a estrutura da rede e a formação de grupos baseada em fatores geográficos ou de conhecimento. Como as duas bases, PubMed e Web of Science, possuem um amplo espectro na área biomédica, acredita-se que as análises nos níveis mundial e regional de colaboração científica foram contempladas quanto ao entendimento das diversas especificidades quando utilizados os métodos de análise de redes.

A base de dados SciELO tem como foco principal a indexação de publicações científicas brasileiras e demais países associados. “O SciELO foi concebido como um projeto e uma estratégia para superar o fenômeno conhecido como ‘ciência perdida’, causado pela presença muito fraca dos periódicos de países em desenvolvimento nos índices internacionais” (PACKER et al., 2014). Dessa maneira, o que se pretendeu com este estudo foi saber primeiro quais as áreas de conhecimento e revistas científicas mais utilizadas em LN e TB indexadas no SciELO e como podem ser utilizados os métodos de análise de redes para compreender as colaborações científicas entre pesquisadores e países que possivelmente tiveram presença menos expressiva em bases internacionais.

O funcionamento regular do SciELO teve início no Brasil em 1998, depois de um projeto-piloto de um ano em parceria com a Fapesp e o Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde (Bireme/Opas/OMS). O SciELO tinha dois objetivos simultâneos: o primeiro era desenvolver competência e infraestrutura para indexar e publicar na internet um conjunto selecionado de periódicos brasileiros, de diferentes disciplinas, que adotassem a avaliação pelos pares, e lidar com textos em diversos idiomas; o segundo foi aumentar a visibilidade, o uso e o impacto dos periódicos indexados e das pesquisas que publicam (PACKER et al., 2014).

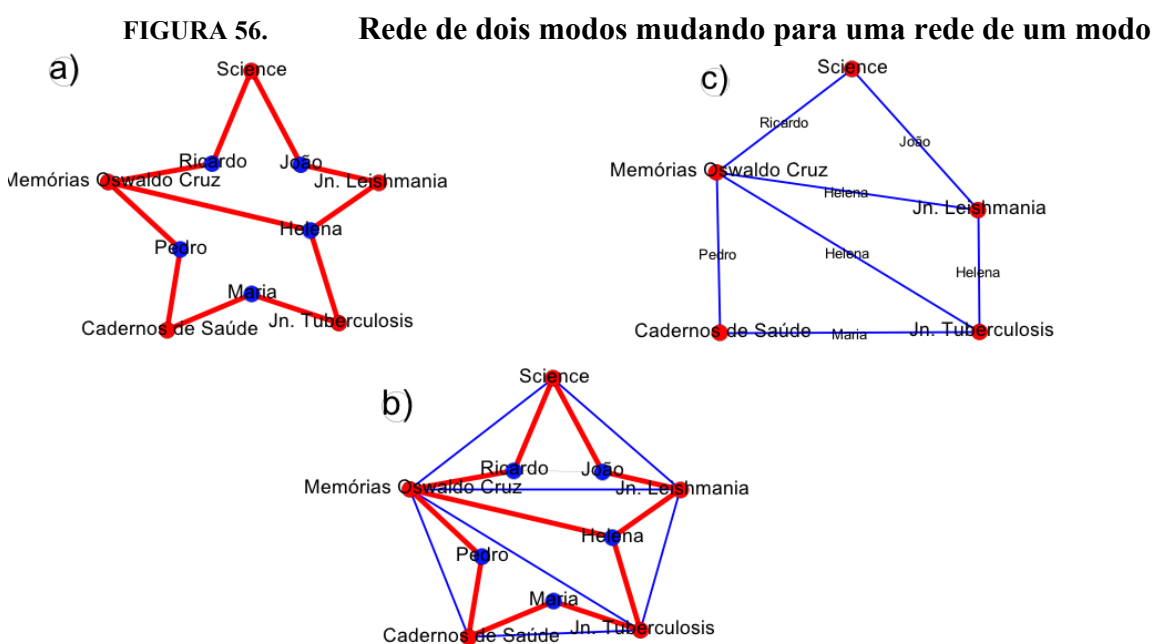
### 6.1 Definição e manipulação das redes

A fase de definição da rede no estudo da base SciELO foi dividida em duas etapas distintas, nas quais os nós e as linhas das redes estudadas possuíam características diferentes. Na primeira etapa seguiu-se a mesma estrutura dos estudos anteriores de colaboração

científica. Foram analisadas as relações entre os pesquisadores nas duas áreas de forma distinta, e os resultados de LN e TB foram comparados entre si.

Na segunda etapa foi feita uma análise entre os pesquisadores e as revistas nas quais estes pesquisadores tiveram seus artigos publicados. Esse tipo de rede é conhecido como rede de dois modos ou rede de afiliação, em que os pesquisadores representam um conjunto de nós e as revistas, outro conjunto de nós ao qual os pesquisadores se “afiliam” por meio de seus artigos. Nesse tipo de rede só existem relações de um conjunto para o outro, ou seja, dos pesquisadores para as revistas e vice-versa. Não há relação dos pesquisadores entre si ou das revistas entre si.

Contudo, existe a possibilidade de se transformar uma rede de dois modos em uma rede de um modo, sendo a relação estabelecida entre os nós da nova rede o conjunto de nós retirados da rede anterior. Para melhor exemplificar essa estrutura, exhibe-se na Figura 56 uma rede de dois modos e sua transformação para uma rede de um modo.



Fonte: elaboração do autor

Legenda: a rede (a) mostra todos os nós e as linhas de uma rede de dois modos, na qual os nós vermelhos são as revistas e os azuis, os pesquisadores. Só existem relações entre revistas e pesquisadores. A rede (b) mostra todas as linhas na transformação entre uma rede de dois modos e uma rede de um modo. A rede (c) mostra uma rede de um modo, os nós são as revistas e suas relações, os pesquisadores.

Na Figura 56 busca-se exemplificar o que seria uma rede de dois modos e sua mudança para uma rede de um modo, a rede (a) está no estado inicial de dois modos e a rede (c), no estado final de um modo. Na rede inicial existem dois grupos diferentes de nós, sendo um deles os pesquisadores, na cor azul, e o outro as revistas, na cor vermelha. Note-se que não há relação das revistas entre si ou dos pesquisadores entre si, apenas dos pesquisadores

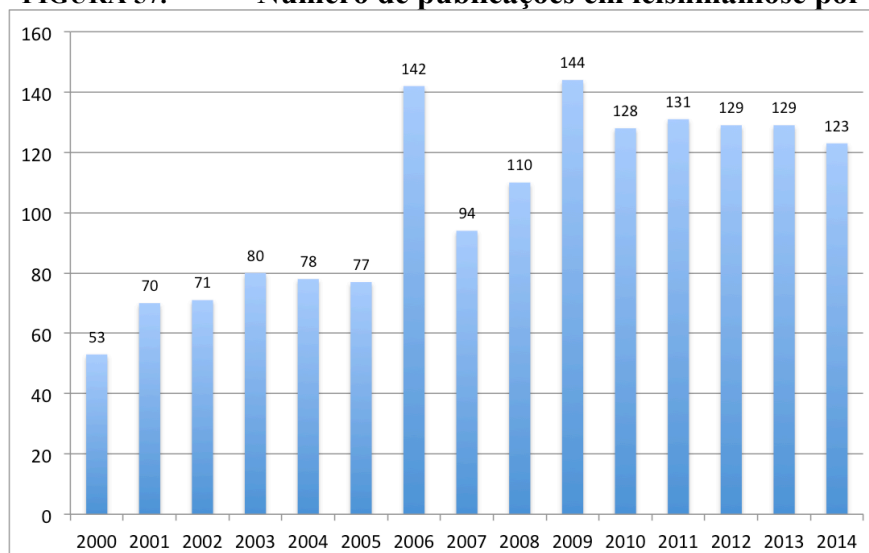
com as revistas. Na rede (c) é apresentada uma figura após a transformação da rede. Pode ser visto na rede (c) que as relações entre a revista *Memórias Oswaldo Cruz* e as revistas *Jn. Leishmania* e *Jn. Tuberculosis* ocorreram em virtude da “Helena”, que possuía relações com as três revistas. A rede (b) mostra apenas a transição entre as duas redes para efeito de ilustração.

### 6.1.1 Dados das redes

O levantamento dos dados da base SciELO ocorreu em novembro de 2014 usando-se os termos de pesquisa (leish\*) e (tubercul\*) para as respectivas áreas. O período analisado foi de 2000 a 2014 para as duas áreas, e o número de artigos foi de 1.559 para LN e 3.578 para TB. Isso demonstra que o número de publicações em TB é mais que o dobro que em LN para o período, mesmo se analisando somente as publicações indexadas no SciELO, corroborando os números encontrados em outras bases quanto à diferença entre as duas áreas. No entanto, apesar de o número de publicações em TB ser maior que em LN, o número de pesquisadores que publicaram apenas uma única vez em TB é muito maior que em LN. Esse fator é relevante quando se analisa a rede de dois modos ou a colaboração entre as revistas. Neste caso não são levados em consideração os pesquisadores com apenas uma publicação, pois não apresentam relevância para o estudo, uma vez que não poderiam ser o *link* entre quaisquer duas revistas. Mais detalhes a esse respeito são descritos quando do estudo dos dados da rede de dois modos.

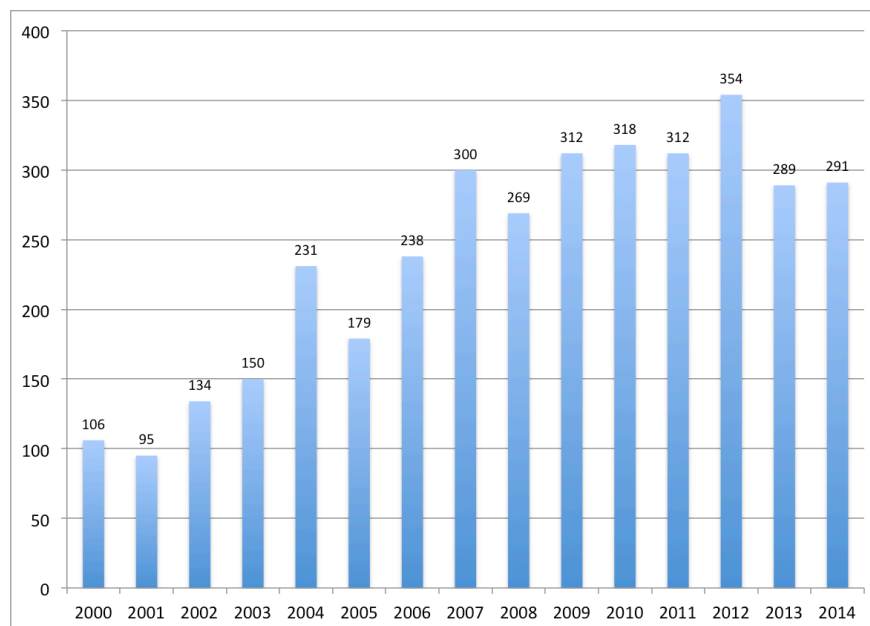
As Figuras 57 e 58 mostram no formato de gráfico de barras o número de publicações por ano para cada uma das áreas.

**FIGURA 57. Número de publicações em leishmaniose por ano**



Fonte: elaboração do autor

**FIGURA 58. Número de publicações em tuberculose por ano**



Fonte: elaboração do autor

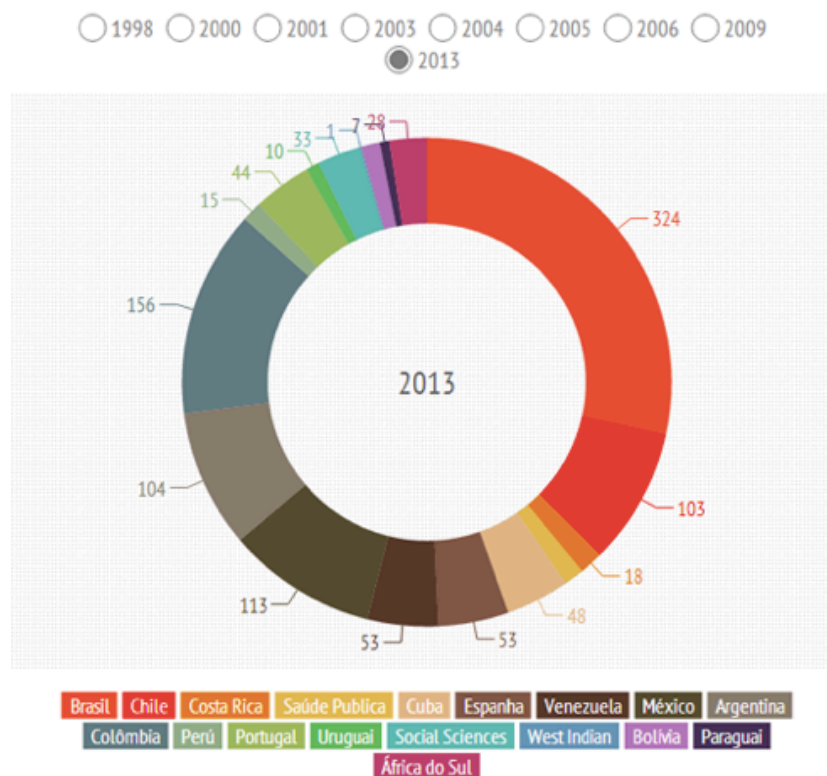
Nas Figuras 57 e 58 percebe-se um volume crescente de publicações nos últimos anos, com um acréscimo aproximado de três vezes o número de publicações do ano de 2000 para 2014 nas duas áreas. Entretanto, verifica-se que em certos anos houve um pico ou aumento elevado de publicações que se diferencia dos padrões de crescimento de outros anos. Isso pode ser percebido nos anos de 2006 e 2009 para a LN e nos anos de 2004, 2007 e 2012 para a TB. Vários podem ser os fatores para esse crescimento, como incentivo financeiro, tanto nacional como internacional, além da possibilidade de entrada de novas revistas na base indexada. Como o SciELO vem agregando revistas e parcerias internacionais nos últimos anos, a entrada de uma nova coleção pode criar essas distorções. Para este estudo não foi avaliado o motivo desses crescimentos acima da média.

### 6.1.2 Coleções, revistas e áreas do SciELO

A base de dados SciELO vem crescendo quanto ao número de títulos indexados a cada ano, incorporando diferentes países e coleções. Coleções são qualquer tipo de produção científica registrada como anais de congressos, revistas científicas e outros. Na Figura 59 são mostrados no ano de 2013 os países e o número das respectivas coleções de cada país conforme o infográfico disponibilizado pelo próprio SciELO.



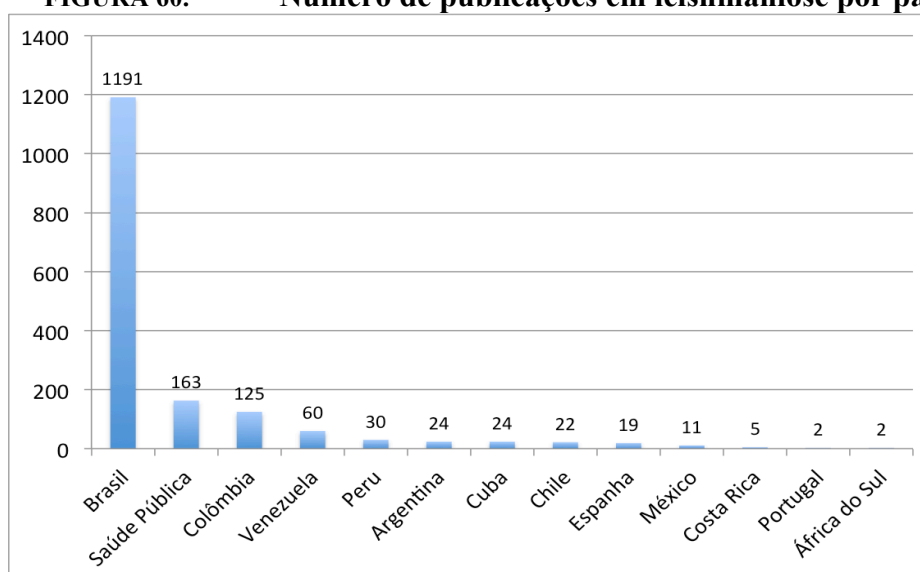
**FIGURA 59. Número de títulos indexados no SciELO por país**  
**Evolução da Rede SciELO - Número de títulos**



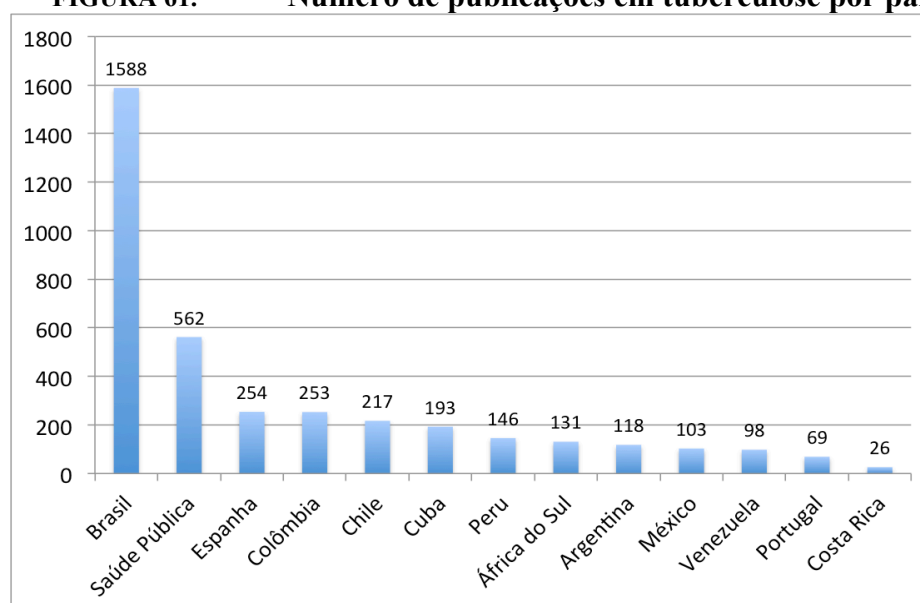
Fonte: PACKER et al., (2014)

Os resultados da Figura 59 demonstram como está distribuído o número de títulos de coleções (revistas e outras) por país, incluindo-se duas áreas não relacionadas a países específicos: Saúde Pública e Social Sciences. Essa distribuição pode ajudar a identificar como certos pesquisadores tendem a colaborar mais ou menos em determinadas áreas.

Como o Brasil é o país com o maior número de coleções indexadas, na distribuição de revistas e pesquisadores por região há uma grande vantagem para o país em números tanto na área da LN como na da TB em comparação com os demais países. As Figuras 60 e 61 mostram o número de publicações por país para cada uma das áreas. Cada publicação científica avaliada se relaciona a uma coleção, com algumas coleções tendo um número elevado de artigos a elas relacionados.

**FIGURA 60. Número de publicações em leishmaniose por país**

Fonte: elaboração do autor

**FIGURA 61. Número de publicações em tuberculose por país**

Fonte: elaboração do autor

Tanto no caso da LN como da TB o número de publicações relacionadas a coleções brasileiras é muito maior que nos demais países. Como dito antes, esse número de publicações está diretamente relacionado ao número de coleções. O que não foi avaliado nesta pesquisa é qual o percentual de coleções de cada um desses países que está indexado na base. Independentemente de haver um percentual mais elevado de coleções brasileiras, acredita-se que nas áreas de LN e TB o Brasil está bem representado em termos de publicações e coleções. Outro fator a ressaltar é que a afiliação do artigo com o país é feita por meio da coleção, e não do pesquisador, ou seja, um pesquisador brasileiro pode muito bem ter

publicado em revistas de outros países e vice-versa. Essa relação entre as revistas será analisada mais adiante.

Dando prosseguimento à análise das publicações por país, nota-se uma diferença entre LN e TB quanto à ordem dos países que publicam com mais frequência em cada uma dessas áreas. No caso da LN, os países com maior número de publicações em suas coleções depois do Brasil são Colômbia e Venezuela, com 7,4% e 3,6% do número total de publicações, respectivamente. Os demais países em ordem decrescente no número de publicações são Peru, Argentina, Cuba, Chile, Espanha, México, Costa Rica, Portugal e África do Sul. Se for analisada a TB, os países com maior número de publicações depois do Brasil são Espanha, Colômbia, Chile, Cuba, Peru, África do Sul, Argentina, México, Venezuela, Portugal e Costa Rica, sendo os quatro primeiros dessa lista responsáveis por mais de 5% do número total de publicações na área. A Tabela 24 demonstra o percentual de publicações por área e por país, além do percentual do número de coleções por país. O percentual na tabela é calculado com base no número total de títulos ou publicações.

**TABELA 24. Percentual de títulos e publicações por país**

	<b>Títulos</b>	<b>Títulos %</b>	<b>LN</b>	<b>LN %</b>	<b>TB</b>	<b>TB %</b>
Brasil	324	30,1%	1191	71,0%	1588	42,3%
Colômbia	156	14,5%	125	7,4%	253	6,7%
México	113	10,5%	11	0,7%	103	2,7%
Argentina	104	9,7%	24	1,4%	118	3,1%
Chile	103	9,6%	22	1,3%	217	5,8%
Espanha	53	4,9%	19	1,1%	254	6,8%
Venezuela	53	4,9%	60	3,6%	98	2,6%
Cuba	48	4,5%	24	1,4%	193	5,1%
Portugal	44	4,1%	2	0,1%	69	1,8%
África do Sul	28	2,6%	2	0,1%	131	3,5%
Saúde Pública	18	1,7%	163	9,7%	562	15,0%
Costa Rica	18	1,7%	5	0,3%	26	0,7%
Peru	15	1,4%	30	1,8%	146	3,9%

Fonte: elaboração do autor

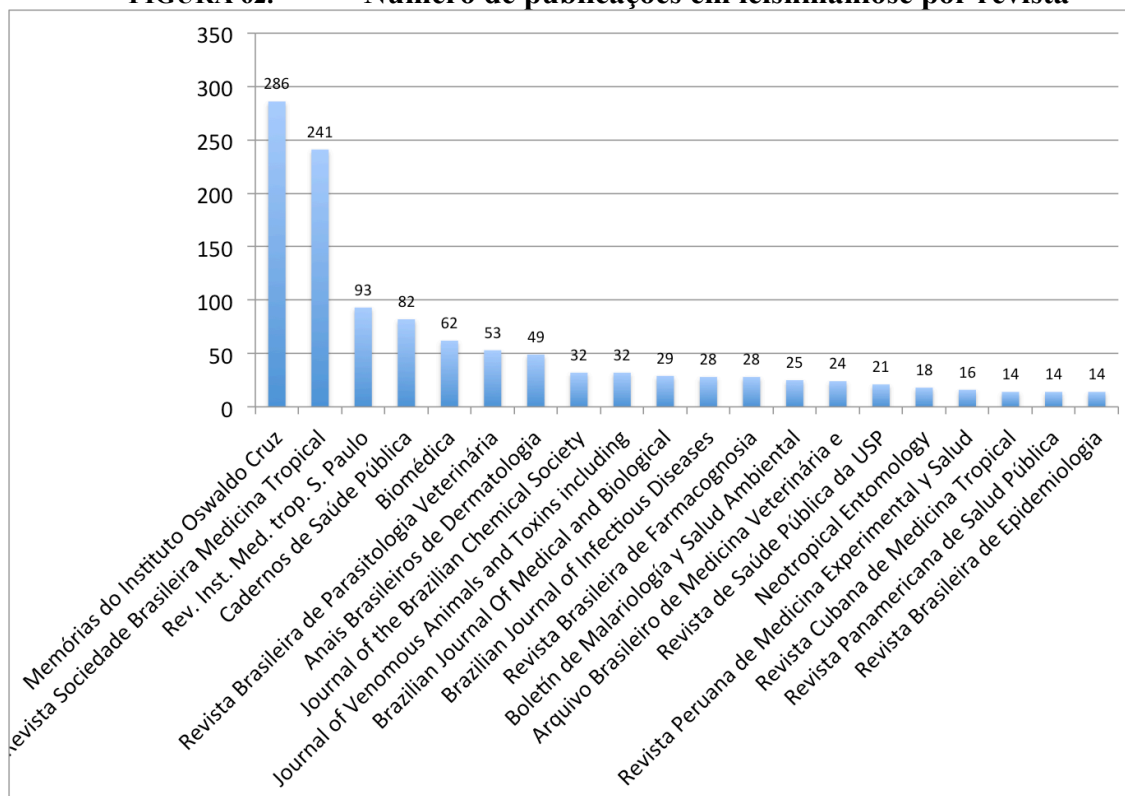
A mudança de posição no número de publicações de países como Espanha e África do Sul entre as redes de LN e TB demonstra uma clara diferença em interesses de pesquisa, tendo a TB uma abrangência territorial maior que a LN. Outro fator importante nos resultados está relacionado à hegemonia de pesquisa no caso da LN para o Brasil em detrimento de outros países da região ou indexados pelo SciELO. Essa hegemonia não é tão presente em TB, tendo em vista que vários outros países têm números elevados de publicações.

Ao se levar em consideração a proporção do número de coleções disponíveis por país e o número de publicações em cada área, surgem dados também interessantes, como no caso do Peru, de Cuba e da África do Sul, que detêm um percentual mais elevado de publicações

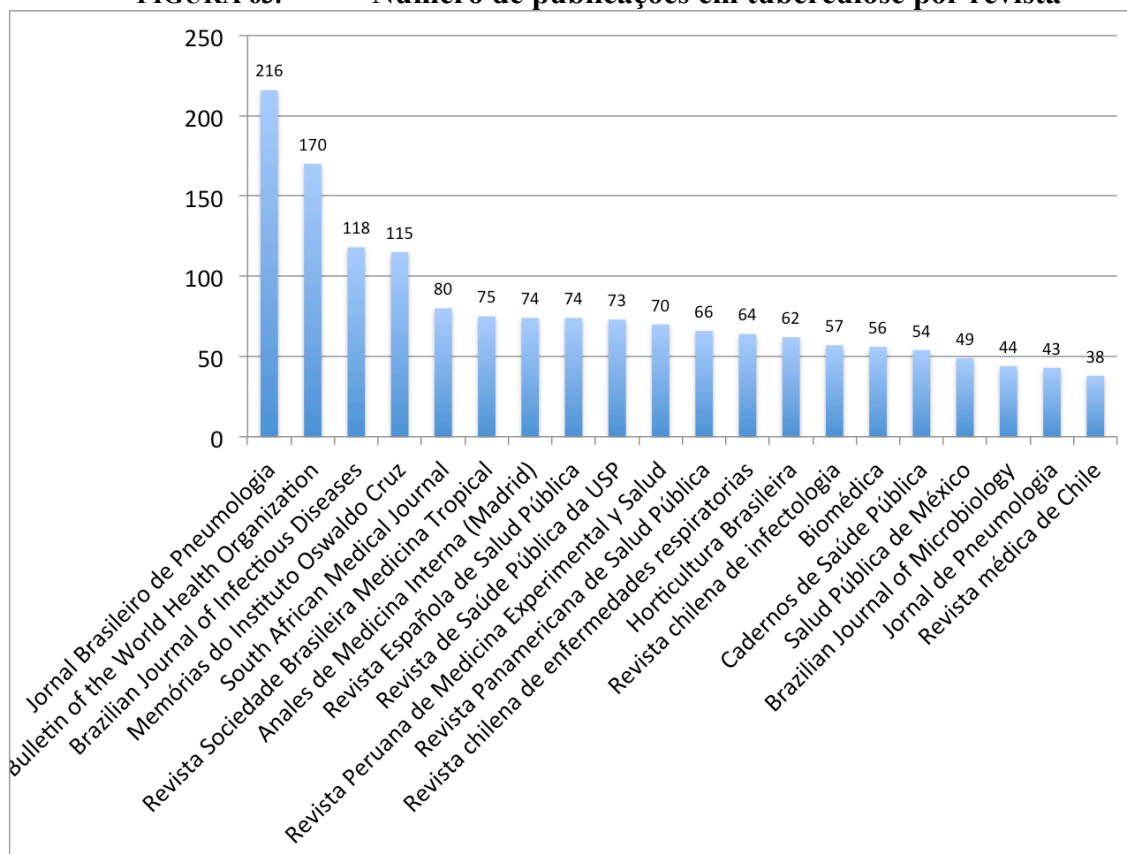
em TB que seu percentual de revistas indexadas. Esse fato não é percebido em nenhum país no caso da LN além do Brasil.

As coleções estão relacionadas aos países, e isso é o que foi demonstrado nas figuras anteriores. O que pode ser analisado agora com maiores detalhes é quais as revistas mais utilizadas na publicação de artigos científicos. As Figuras 62 e 63 exibem as revistas com maior número de publicações para LN e TB respectivamente.

**FIGURA 62. Número de publicações em leishmaniose por revista**



Fonte: elaboração do autor

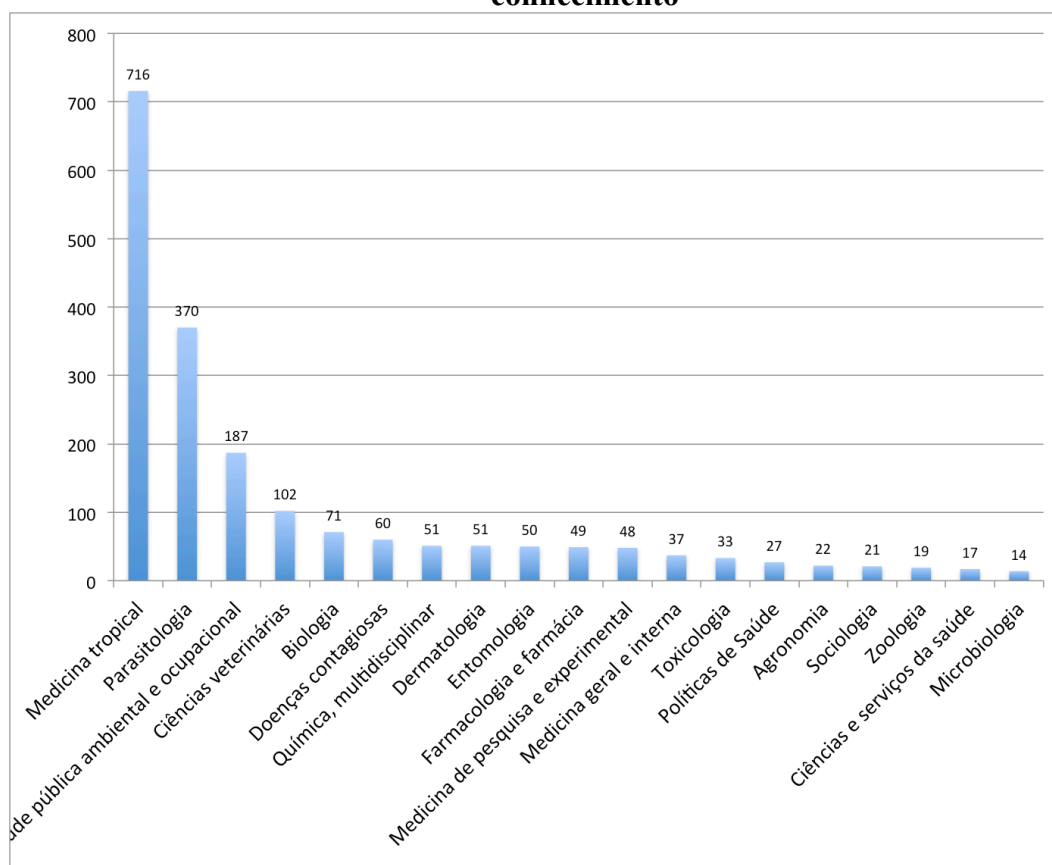
**FIGURA 63. Número de publicações em tuberculose por revista**

Fonte: elaboração do autor

Uma vez que as revistas apresentadas nas Figuras 62 e 63 são apenas um maior detalhamento das coleções vistas anteriormente, não é de se estranhar um padrão similar na distribuição das publicações. Uma primeira análise está no formato dos grafos nas duas figuras. Para LN existe uma queda mais acentuada entre as revistas com maior publicação para as demais revistas do que em TB. Esse mesmo padrão foi visto quanto às publicações por país. No entanto, pode-se perceber também uma mudança nas revistas e a frequência de utilização destas. No caso da LN, das dez revistas com maior número de publicações na área nove são brasileiras. Porém, para TB, das dez primeiras revistas em número de publicações cinco são brasileiras e as outras cinco de outros países como Espanha, África do Sul e Peru. Essa diferença é muito importante para se entender a estrutura da pesquisa nas respectivas áreas e como as colaborações internacionais são importantes.

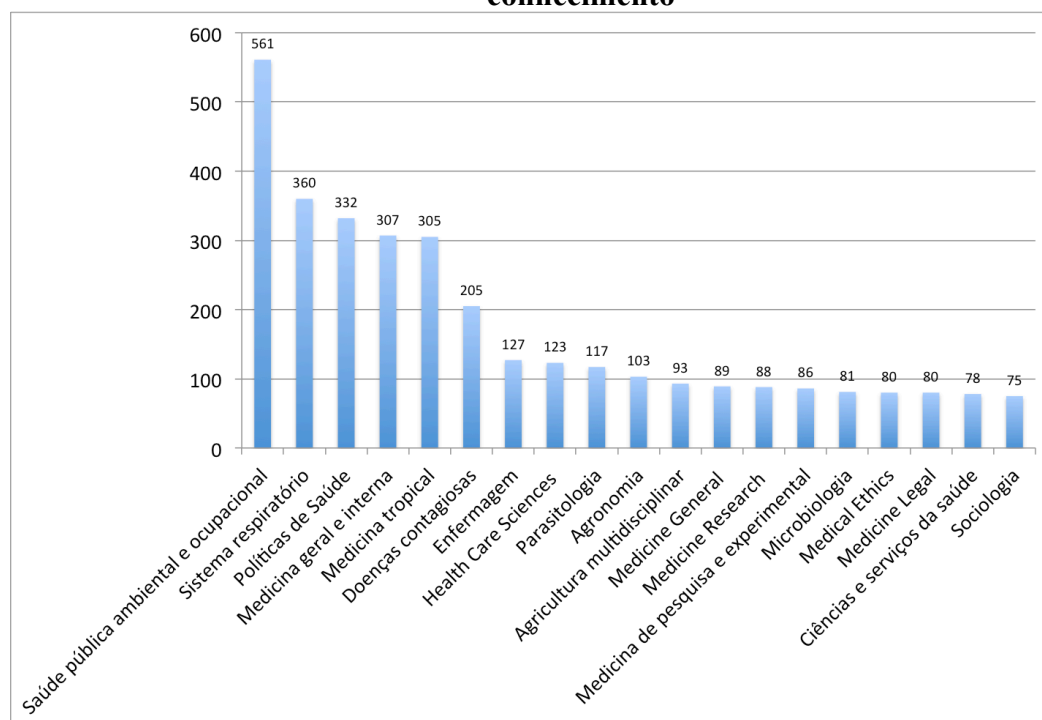
Surge agora um terceiro elemento passível de análise: as áreas de conhecimento em LN e TB. Essas áreas de conhecimento estão relacionadas à caracterização definida pela base Web of Science e utilizada pelo SciELO como referência. As Figuras 64 e 65 exibem os números por área de conhecimento.

**FIGURA 64. Número de publicações em leishmaniose associadas às áreas de conhecimento**



Fonte: elaboração do autor

**FIGURA 65. Número de publicações em tuberculose associadas às áreas de conhecimento**



Fonte: elaboração do autor

Conforme explicado no capítulo da Web of Science, as áreas de conhecimento estão relacionadas diretamente às revistas científicas, e não necessariamente às publicações nelas contidas, lembrando que uma revista pode conter uma ou mais áreas de conhecimento, e por isso os números podem divergir do total de artigos encontrados.

Uma vez que as áreas de conhecimento estão relacionadas às revistas, e as revistas às coleções dos respectivos países, a estrutura dos gráficos apresentados segue o mesmo padrão, com uma queda mais acentuada nas primeiras áreas em LN para as demais áreas em comparação à TB. Além disso, pode-se ver um entendimento interessante de como cada uma das áreas de conhecimento é percebida nos seus respectivos temas.

Enquanto a LN tem como foco principal a Medicina Tropical e a Parasitologia, na TB as áreas de conhecimento abrangem assuntos mais gerais, como Saúde Pública Ambiental e Ocupacional, Políticas de Saúde e Medicina Geral e Interna. Outras áreas são também importantes tanto para a LN quanto para a TB, mas essa característica mais pública da TB pode ser responsável pela efetividade da batalha contra a doença, segundo números de avanço das doenças, em TB mais do que em LN.

## **6.2 Determinação de características estruturais e a inspeção visual das redes**

De acordo com o modelo exploratório de análise de redes seguido neste estudo, devem ser analisadas as características estruturais e depois a ser feita a inspeção visual da rede por meio dos grafos. Como para a base do SciELO foram estudadas duas estruturas de redes, deve-se apresentar o resultado dos dois modelos, com o segundo modelo contendo duas etapas.

A seguir são apresentados os resultados em tabelas no subitem “Dados das redes” e em sociogramas nos subitens “Redes de colaboração” e “Redes de afiliação”. Nesses subitens são tratados os resultados de três redes, sendo a primeira relacionada com os dados de colaboração dos pesquisadores, a segunda, a rede de dois modos, com pesquisadores e revistas, e, por fim, a rede de um modo, com revistas se relacionando com outras revistas.

### **6.2.1 Dados das redes**

Para que se possa avaliar a estrutura das redes é necessária a avaliação de algumas das suas medidas. As medidas da estrutura de colaboração dos artigos, autores com autores, são apresentadas na Tabela 25.

**TABELA 25. Percentual de títulos e publicações por país**

	LN	TB
Número de nós	1038	2045
Número de linhas	2708	5796
Componente gigante	61,20%	37,40%
Grau médio	6,4	8,3
Densidade	0,01	0,01
Coeficiente de clusterização	0,654	0,699
Caminho mais curto	6	5,6
Diâmetro	18	14
<i>Clusters</i>	22	26

Fonte: elaboração do autor

O resultado da tabela demonstra a diferença em tamanho das duas redes, tendo o número de nós e linhas aproximadamente o dobro da rede LN para a rede TB. Uma grande diferença, no entanto, encontra-se na análise do componente gigante, ou maior componente, com todos os seus nós conectados. No caso da rede LN, esse elemento é de 61,20% da rede, o que demonstra uma forte coesão. Para TB, o componente gigante é de 37,40% do total da rede. Levando-se em consideração os dados levantados anteriormente sobre as duas estruturas quanto às revistas, às coleções e às áreas do conhecimento, esse resultado corrobora os dados anteriores, uma vez que a maioria das publicações em LN está concentrada em algumas áreas, enquanto em TB a maioria está pulverizada por diferentes países e revistas.

No que diz respeito ao grau médio ou ao número de colaborações de cada pesquisador, essa medida é maior em TB do que em LN, com 8,3 e 6,4 respectivamente. Isso demonstra que em TB, apesar de o grupo componente gigante ser relativamente menor, eles tendem a trabalhar mais de forma colaborativa. Os demais indicadores das duas redes são similares, e a densidade da rede nos dois casos é 0,01, havendo espaço para mais colaborações. O Coeficiente de Clusterização, que é a medida de colaboração dos nós, é alto para as duas redes, e o caminho médio e o diâmetro da rede estão alinhados com outras redes de colaboração. Em LN, o diâmetro, ou a distância mais longa entre dois nós na rede, é 18, enquanto em TB é 14. Para esse elemento, quanto mais densa a rede e quanto maior o número de colaborações menor é a distância em média.

O número de subgrupos ou *clusters* na rede mostrou-se similar nas duas redes. No entanto, para esse resultado vale ressaltar que os demais subgrupos não conectados não foram levados em consideração.

Quanto à rede de dois modos ou rede contendo os pesquisadores e as relações com as revistas, tem-se os resultados apresentados na Tabela 26.



**TABELA 26. Dados das redes de dois modos**

	Número de nós	Número de linhas	Pesquisadores	Revistas	Grau máximo
LN	1.093	1.903	928	165	351
TB	1.545	1.674	1.296	249	184

Fonte: elaboração do autor

A diferença no número de nós das duas redes diverge da diferença apresentada na rede de colaboração. Isso ocorre devido à existência de um percentual maior de pesquisadores em TB que publicaram apenas uma vez. Como neste estudo está sendo avaliada a capacidade de um ator relacionar duas revistas, os pesquisadores com uma só publicação não foram levados em consideração.

Outro elemento interessante e divergente da rede de colaboração analisada anteriormente é o número de linhas entre as redes LN e TB, que está diretamente relacionado ao grau máximo da rede. Em LN, o número de linhas é muito maior do que em TB, levando-se em consideração o número de nós das duas redes. Isso demonstra a grande concentração de pesquisadores que publicam com regularidade e em diversas revistas, diferentemente de TB, que possui muitos pesquisadores que publicaram em uma única revista uma ou poucas vezes. Isso pode ser visto quanto ao número de revistas e quanto à diferença entre as duas redes, em que TB tem 84 revistas a mais que LN, o que é mais da metade do número total de revistas em LN.

Por último é analisada a rede de dois modos transformada em um modo. Nessa rede são consideradas apenas as revistas e as relações entre elas. Diz-se que uma revista tem relação com outra revista quando um pesquisador escreve artigos para as duas. A Tabela 27 expõe os números das redes de LN e TB.

**TABELA 27. Dados das redes de dois modos**

	Número de nós	Número de linhas	Componente gigante	Demais revistas	Densidade	Grau médio
LN	165	340	53%	78	0,09	7,724
TB	249	387	48,70%	126	0,064	4,7

Fonte: elaboração do autor

Na Tabela 27, o que se pode perceber é que a rede LN possui um adensamento maior que a rede TB. Isso significa que entre as redes, quando se analisam as relações entre as revistas, se percebe que em LN há uma grande concentração nas revistas mais comuns, enquanto na TB existe uma diversidade maior, pois esta possui um número maior de revistas e mais espalhadas por diferentes países. Dessa maneira se explica esse adensamento em LN, pois algumas poucas revistas, como *Memórias do Instituto Oswaldo Cruz* e *Revista Brasileira*

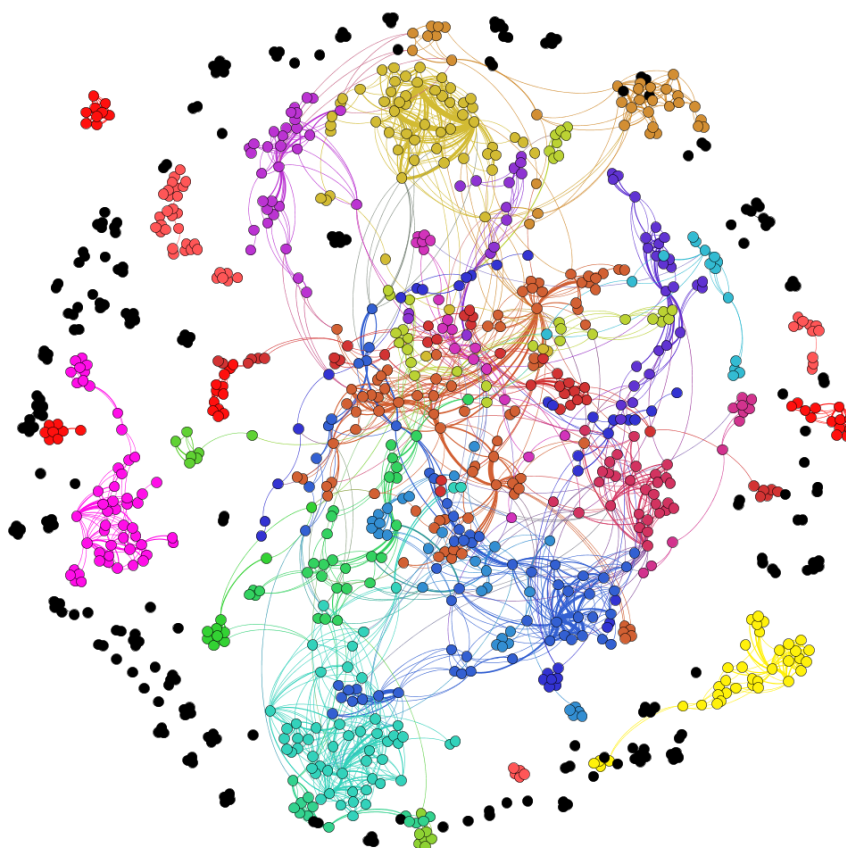
de *Medicina Tropical*, acabam por concentrar a maior parte das publicações e, consequentemente, o maior número de colaborações entre as demais revistas.

Todos os resultados apresentados neste subitem corroboram os sociogramas do subitem seguinte para cada uma das redes, seja de colaboração, de dois modos ou de um modo.

### 6.2.2 Redes de colaboração

As redes ou sociogramas a seguir são essenciais para a inspeção visual, e sua distribuição ocorreu de forma que fossem ressaltados os resultados que se pretendia demonstrar na discussão destas. A Figura 66 mostra a rede de colaboração em LN, na qual cada nó representa um pesquisador, e as linhas entre dois nós, a colaboração entre eles. As cores nos nós centrais demonstram os subgrupos aos quais estes nós pertencem.

FIGURA 66. Rede de colaboração em leishmaniose



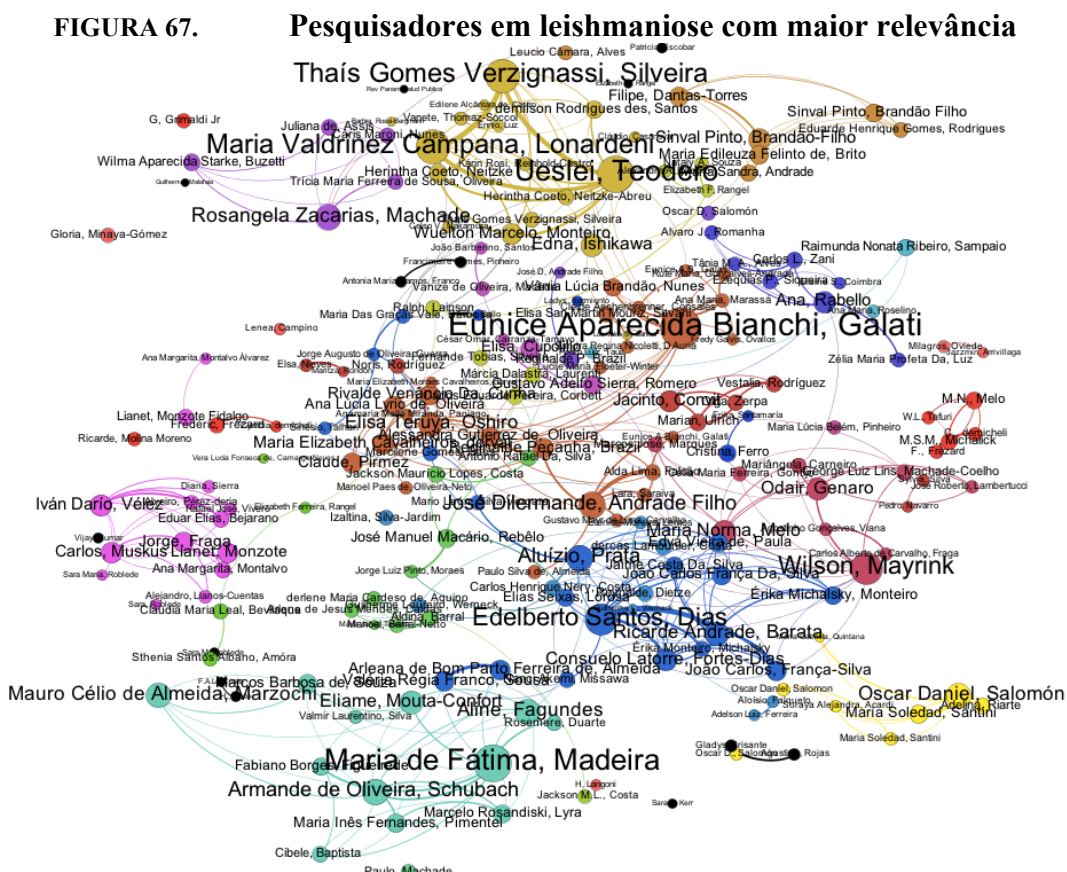
Fonte: elaboração do autor

Legenda: rede de coautoria em leishmaniose representando todos os pesquisadores da rede. No centro da rede estão os nós do componente gigante, todos conectados entre si. Na periferia estão os subgrupos ou nós não conectados ao resto da rede. As cores dos nós representam seus subgrupos tanto no componente gigante quanto externo a eles. Os nós em preto são nós com três colaborações ou menos.

A Figura 66 demonstra o componente gigante e os demais componentes separados da rede. É possível notar que existem grupos relativamente grandes não conectados ao grupo central. Esse fator é fácil de entender, uma vez que as coleções são de diferentes países, que não necessariamente possuem colaboração nesse tema específico. Os pesquisadores ou nós em preto são aqueles com poucas relações que podem ter publicado uma ou poucas vezes e por isso possuem poucas relações com os demais membros da rede.

No que diz respeito ao componente gigante, no centro da rede é possível notar, pelos diferentes subgrupos definidos pelas cores, que mesmo estando conectados eles tendem a se concentrar em grupos menores. Isso pode estar diretamente relacionado a uma ou mais revistas específicas, que tendem a reforçar um ou outro assunto específico.

Na Figura 67 são apresentados os nós da mesma rede anterior, com ênfase no componente gigante e em alguns nós específicos. O tamanho de cada nó está relacionado ao número de colaborações ou grau do pesquisador nessa rede.



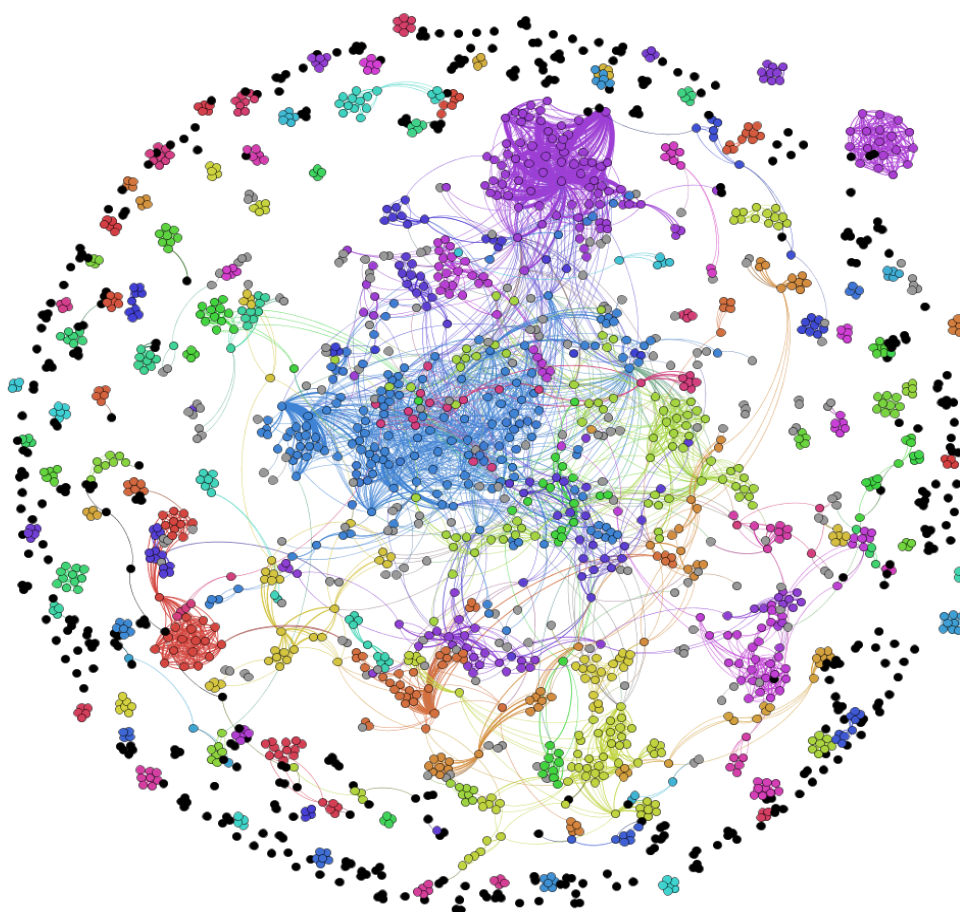
Fonte: elaboração do autor

Legenda: sub-rede da rede de colaboração em leishmaniose. As cores dos nós representam os subgrupos formados pelo algoritmo de clusterização e caracterizam uma maior afinidade com os elementos daquele grupo do que com elementos fora do grupo. O tamanho dos nós está relacionado ao número de colaborações do pesquisador na rede. O nome apresentado é o nome do pesquisador segundo disponibilizado pelo SciELO. As linhas entre dois representam a colaboração entre os pesquisadores.

No sociograma da Figura 67 é apresentado um recorte da rede de colaboração anterior com o objetivo de avaliar o nome dos pesquisadores e seus colaboradores. Como o SciELO tem o foco da sua indexação nas revistas dos países afiliados, vários dos pesquisadores brasileiros encontrados com grande frequência em revistas internacionais não são tão evidentes nessa análise. Quando se faz uma análise superficial dos pesquisadores que aparecem, vê-se que há uma concentração muito grande de pesquisadores brasileiros, uma vez que a maioria das revistas em LN que apareceu nos resultados era brasileira. Caso fossem analisados os grupos menores, é possível que pesquisadores de outros países ficassem em maior evidência.

Na Figura 68 aparece agora o mesmo tipo de análise feita em LN, mas agora em TB. A rede é uma rede de colaboração científica entre os pesquisadores.

**FIGURA 68. Rede de colaboração em tuberculose**



Fonte: elaboração do autor

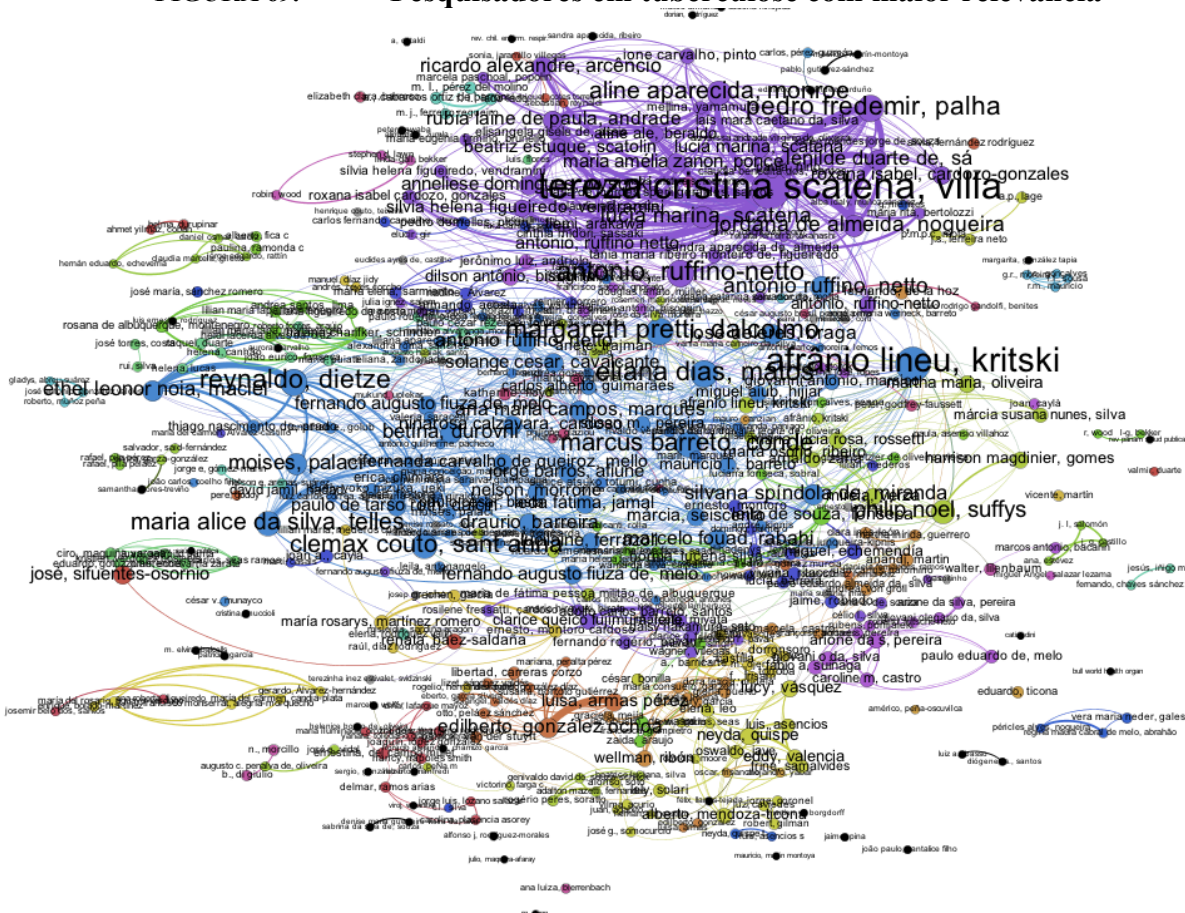
Legenda: rede de coautoria em tuberculose representando todos os pesquisadores da rede. No centro da rede estão os nós do componente gigante, todos conectados entre si. Na periferia estão os subgrupos ou nós não conectados ao resto da rede. As cores dos nós representam seus subgrupos tanto no componente gigante quanto externo a eles. Os nós em preto são nós com três colaborações ou menos.

A Figura 68 mostra a rede de colaboração em TB com um grande grupo central e vários grupos menores na periferia. Uma clara diferença visual na rede TB comparada com LN é a quantidade de pequenos subgrupos, assim como foi visto na Tabela 25. Como existe uma grande quantidade de revistas e países no estudo da TB, essas diferenças aparecem de forma clara na comparação das duas figuras. Outro fator relevante é que o componente gigante ao centro possui dois grandes grupos bem definidos, sendo um azul mais ao centro e um roxo no topo da figura, além de um terceiro em verde logo à direita do azul e bem conectado com ele. Os grupos são bastante densos e demonstram um maior grau de colaboração dos pesquisadores, assim como foi exposto na Tabela 25.

Pode-se descrever a característica dessa rede de colaboração científica como tendo vários subgrupos espalhados em diversas áreas e um grupo central muito conectado e bem dividido entre si.

A Figura 69 demonstra a mesma rede de colaboração em TB, com ênfase no componente gigante e seus atores. Como a rede é mais densa no centro, a visualização dos nomes acaba por ficar um pouco prejudicada.

FIGURA 69. Pesquisadores em tuberculose com maior relevância



Fonte: elaboração do autor

Legenda: sub-rede da rede de colaboração em tuberculose. As cores dos nós representam os subgrupos formados pelo algoritmo de clusterização e caracterizam uma maior afinidade com os elementos daquele grupo do que com elementos fora do grupo. O tamanho dos nós está relacionado ao número de colaborações do pesquisador na rede. O nome apresentado é o nome do pesquisador segundo disponibilizado pelo SciELO. As linhas entre dois nós representam a colaboração entre os pesquisadores.

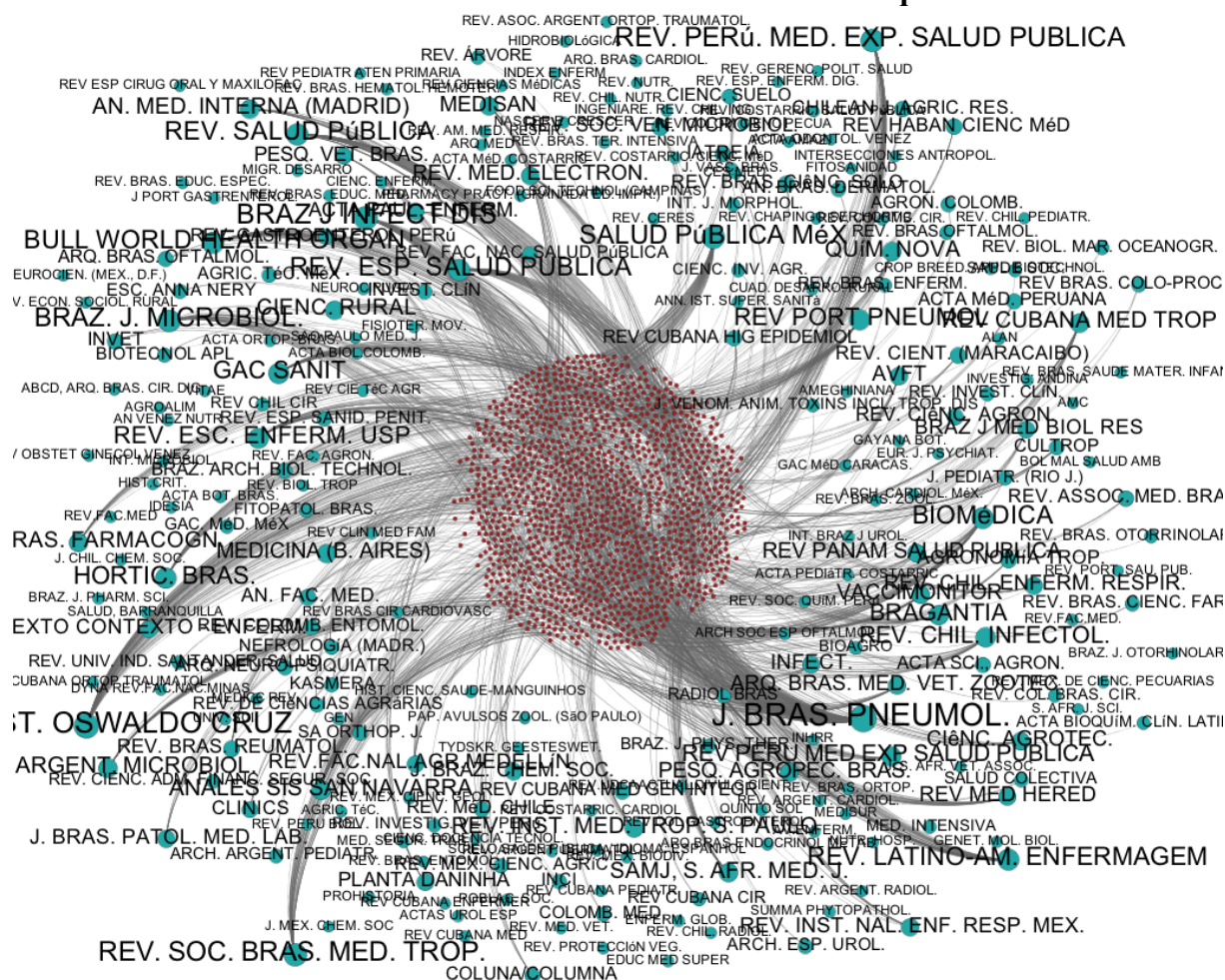
Na sub-rede apresentada na Figura 69 pode-se ver de forma mais evidente os dois grandes grupos e alguns menores na periferia. No caso do grupo em roxo, pode-se perceber que existe uma pessoa que controla as relações com os demais membros do grupo cujo nome é Tereza Cristina Scatena Villa, o que diferencia um pouco do grupo azul, em que os pesquisadores tendem a ter um grau similar de colaboração, com exceção de um pesquisador, Afrânio Lineu Kritski, que tem forte colaboração com pesquisadores de outros grupos também.

### 6.2.3 Redes de afiliação

Neste item é feita uma análise visual da rede de dois modos, ou rede de afiliação, e de sua mudança para a rede de um modo. O objetivo desta análise é buscar o entendimento de como ocorrem as colaborações entre as revistas e se existe maior ou menor tendência de



**FIGURA 71. Rede de dois modos em tuberculose. Pesquisadores e revistas**



Fonte: elaboração do autor

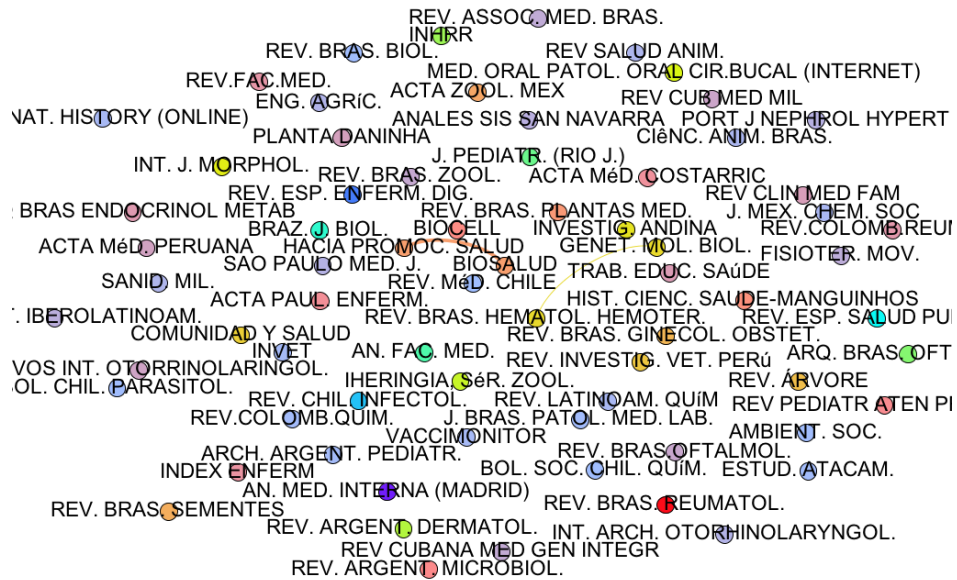
Legenda: rede de dois modos da tuberculose. Os nós em vermelho no centro da figura representam os pesquisadores, e os nós em azul na periferia, as revistas. As linhas identificam se um autor publicou um artigo na referida revista. Algumas revistas possuem um número de publicações muito maior que outras, o que pode ser visto pela quantidade de linhas indo para elas. As revistas com número acentuado de publicações têm um ligeiro aumento de tamanho.

O que se pode perceber de diferente na Figura 71 para a mesma rede de LN é a diversidade de revistas com relevância em TB. Além do maior número de revistas também existe uma diversidade em países com revistas mexicanas, peruanas, portuguesas e outras tantas. O fato de a TB ser muito mais pulverizada por diferentes países fica evidente na comparação das duas figuras. O que resta saber é o quanto essa diversidade pode ter apoiado ou não o tratamento da doença. Pode-se esperar pelo menos uma maior opção ou tipos de pesquisa diferenciados.

As figuras seguintes concentram-se na rede de um modo após a transformação da rede de pesquisadores e revistas. As Figuras 72 e 73 exibem, para efeito de ilustração, as revistas cujos pesquisadores que nelas publicaram não o fizeram em outras revistas, com exceção de dois casos para LN e cinco para TB.



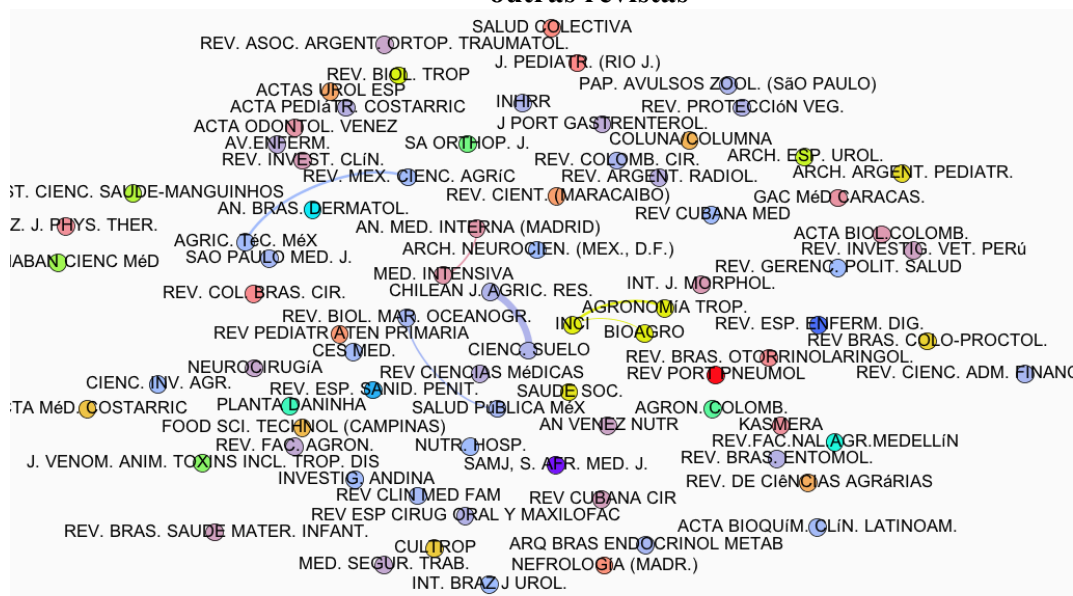
FIGURA 72. Revistas em leishmaniose sem relacionamentos expressivos com outras revistas



Fonte: elaboração do autor

Legenda: rede de um modo de leishmaniose transformada. As relações entre as revistas são estabelecidas quando um autor escreve artigos para duas ou mais revistas. No caso desta figura, apenas em dois casos os autores escreveram para duas revistas distintas. Em todos os outros casos os autores que publicaram nessas revistas não publicaram em nenhuma outra.

**FIGURA 73. Revistas em tuberculose sem relacionamentos expressivos com outras revistas**



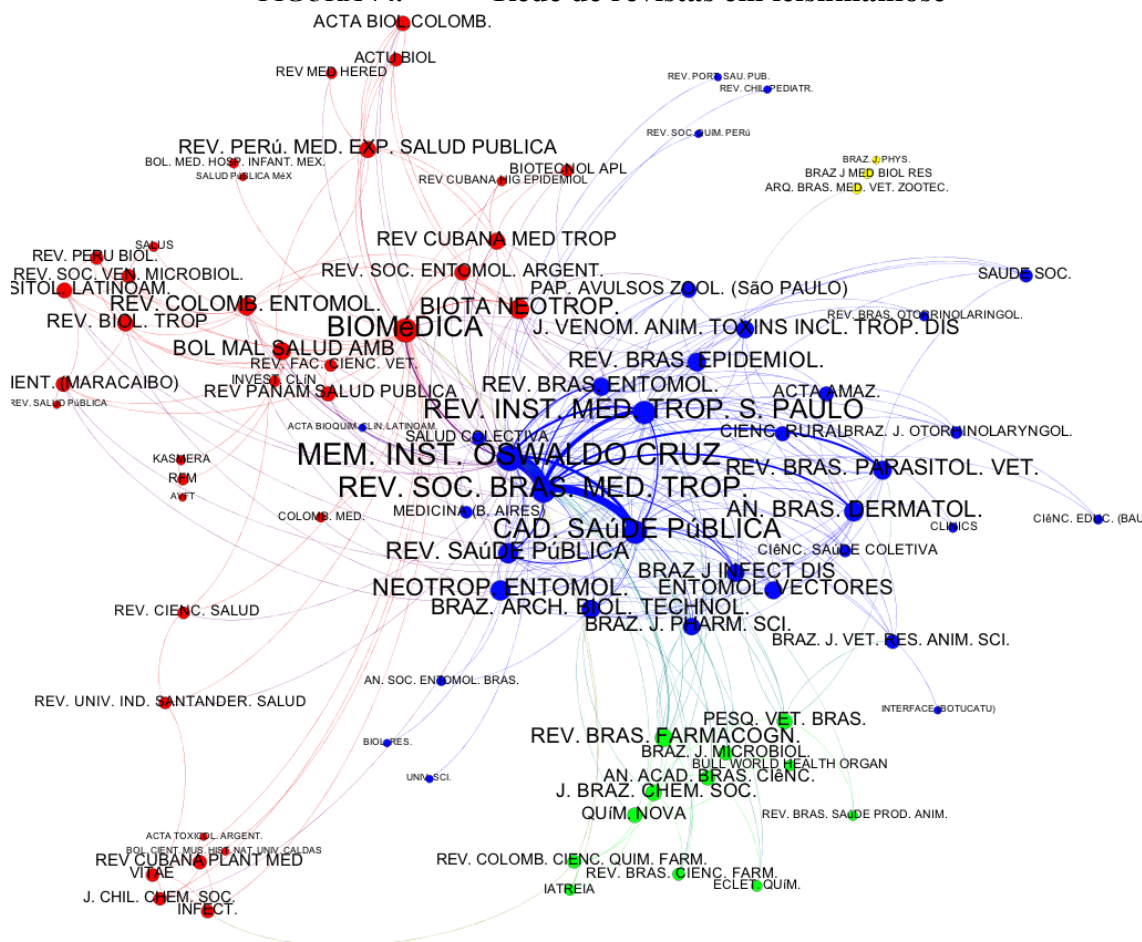
Fonte: elaboração do autor

Legenda: rede de um modo de tuberculose transformada. As relações entre as revistas são estabelecidas quando um autor escreve artigos para duas ou mais revistas. No caso desta figura, apenas em cinco casos os autores escreveram para duas revistas distintas. Em todos os outros casos os autores que publicaram nessas revistas não publicaram em nenhuma outra.

As Figuras 72 e 73 mostram a existência de várias revistas em que seus autores não publicam em outras revistas. Isso pode se dar pelo alto nível de especificidade dessas revistas ou pelo público diferenciado quanto ao nível de estratificação das revistas. Entender o público que essas revistas atendem é de grande importância para divulgar as possibilidades de trabalho com o objetivo de se estender a abrangência de trabalhos científicos.

Por último, as Figuras 74 e 75 mostram a rede de um modo transformada da LN e da TB com as revistas que possuem colaborações intensas com outras revistas. Os subgrupos dessas redes são apresentados pelas diferenças em cores, e em alguns casos acabam por representar regiões geográficas ou áreas de conhecimento diferentes.

**FIGURA 74. Rede de revistas em leishmaniose**



Fonte: elaboração do autor

Legenda: rede de um modo de leishmaniose transformada. As relações entre as revistas são estabelecidas quando um autor escreve artigos para duas ou mais revistas. As cores representam os diferentes subgrupos, que podem estar relacionados por questões geográficas ou temas específicos. O tamanho dos nós possui uma pequena diferença para as revistas com maior colaboração com outras revistas.

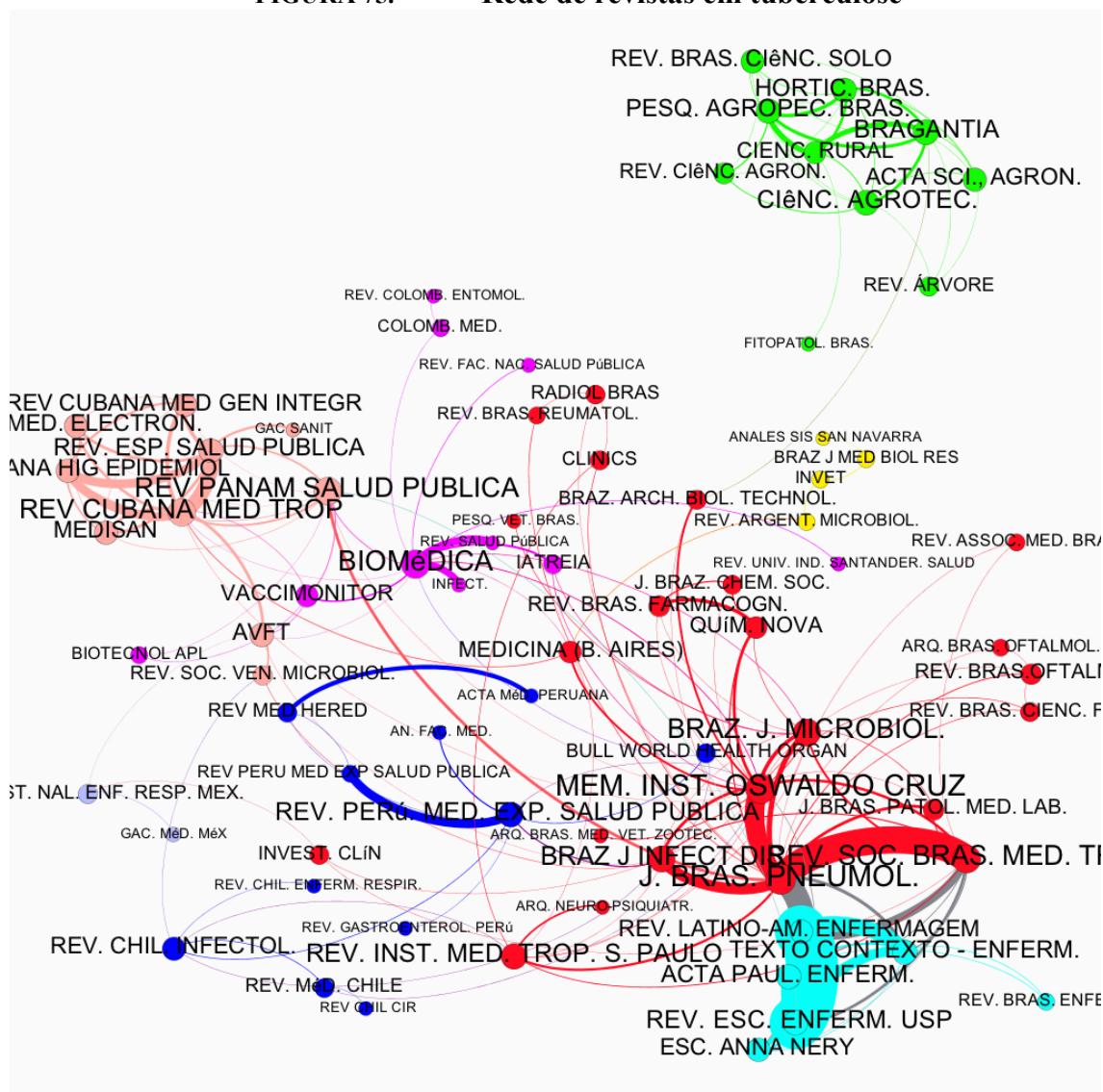
A Figura 74 apresenta as relações das revistas de LN, podendo-se perceber quatro grupos, definidos pelas cores, bem distintos, com o grupo de cor vermelha possuindo também um subgrupo um pouco afastado no lado inferior esquerdo, e o grupo amarelo do lado superior direito ser bastante restrito. Como já foi visto nos resultados anteriores, a área de LN é bastante concentrada em algumas poucas revistas e prioritariamente no Brasil. Esse grupo é representado na cor azul ao centro, sendo a distribuição do grafo utilizada a Force Atlas, que apresenta uma maior ênfase nos agrupamentos. *Memórias do Instituto Oswaldo Cruz*, *Revista da Sociedade Brasileira de Medicina Tropical* e *Cadernos de Saúde* são as revistas com maior colaboração. Isso significa que os pesquisadores que publicam em uma dessas revistas tendem a publicar também nas outras duas. Isso é previsível pelos resultados anteriores, uma vez que essas são as revistas com maior número de publicações.

O grupo de cor vermelha, na parte superior da figura, representa os países latino-americanos, com exceção do Brasil. Essa clara diferença de grupos, estabelecida com base no número de relações dentro e entre os grupos, aponta a necessidade de haver uma maior troca de saberes com países próximos ao Brasil que sofrem de problemas similares. Revistas cubanas, peruanas e colombianas são as que mais prevalecem neste grupo, uma vez que estes são os países que mais produzem artigos nesta área depois do Brasil. As duas revistas que demonstram ser o elo entre os pesquisadores do Brasil e demais regiões são *Memórias do Instituto Oswaldo Cruz* e *Biomédica*. Falta saber se são os pesquisadores brasileiros que publicam nas revistas latino-americanas ou se os pesquisadores desses países é que publicam na *Memórias do Instituto Oswaldo Cruz*. Esse tipo de estudo mais aprofundado pode apoiar o direcionamento de políticas públicas cujo objetivo é aumentar a interação e a colaboração de países na região.

Na parte inferior da Figura 74 aparece o grupo de cor verde, que possui tendências para pesquisa em química e veterinária na área de LN. Essas não são necessariamente muito bem conectadas entre si, uma vez que abordam assuntos diferentes, mas se destacam do grupo central. Coincidentemente, o subgrupo vermelho, também na parte inferior da figura, contém revistas de países como Argentina, Cuba e Chile, sendo evidente em um dos casos o foco em química, assim como o grupo da cor verde.

A Figura 75 mostra as relações entre as revistas com foco agora em TB. Sua estrutura mostra-se diferente de LN, com um número mais elevado de subgrupos, apesar de o componente gigante de TB ser mais restrito, conforme dados anteriores.

FIGURA 75. Rede de revistas em tuberculose



Fonte: elaboração do autor

Legenda: rede de um modo de tuberculose transformada. As relações entre as revistas são estabelecidas quando um autor escreve artigos para duas ou mais revistas. As cores representam os diferentes subgrupos, que podem estar relacionados por questões geográficas ou temas específicos. O tamanho dos nós possui uma pequena diferença para as revistas com maior colaboração com outras revistas.

A Figura 75 exhibe sete grupos bem definidos e concentrados em suas respectivas áreas, definidos pelas cores de seus elementos. O sociograma não aparenta apresentar uma centralidade muito bem definida, conforme visto em LN. Apesar de haver dois grupos na parte inferior esquerda de cores vermelho e azul claro com fortes relações e número de colaborações, estes não se encontram totalmente centrais no grafo.

O grupo vermelho, que aparenta ser o maior grupo, possui as revistas brasileiras mais utilizadas na área de TB, com *Memórias do Instituto Oswaldo Cruz* se repetindo em LN, mas as outras quatro não se repetem, sendo elas *Brazilian Journal of Microbiology*, *Brazilian Journal of Infectious Diseases*, *Revista da Sociedade Brasileira de Medicina Tropical* e por

último, dentre as revistas com maior número de relações no grupo, *Jornal Brasileiro de Pneumologia*. Esta última aparenta ser a revista com maior relacionamento com as demais revistas do grupo de acordo com as linhas que a conectam com as outras quatro revistas. Além disso, o *Jornal Brasileiro de Pneumologia* também estabelece o maior relacionamento com outro grupo por meio da linha cinza de larga espessura. Neste caso, vários pesquisadores que publicam nesta revista também publicam na *Revista Latino-Americana de Enfermagem*.

Um segundo grupo de destaque é o grupo azul claro na parte inferior esquerda, com revistas brasileiras na área de enfermagem, que por sua vez possuem uma alta relação entre elas. Isso quer dizer que os pesquisadores dessa área específica que tratam do assunto TB publicam com frequência nas outras revistas do grupo.

No lado superior direito aparece o grupo de cor verde, com revistas relacionadas à agronomia e às ciências do solo. É importante ressaltar que o único motivo de esse grupo se manter conectado às demais revistas é um pesquisador apenas que publicou em *Bragantia* bem como em *Brazilian Archives of Biology and Technology*.

Outros três grupos representados na Figura 75 estão relacionados a outros países latino-americanos que não o Brasil. A Colômbia está representada pela revista *Biomedica*, em destaque na cor rosa no centro superior com outras revistas. Revistas do Panamá e de Cuba estão bem conectadas entre si no lado superior esquerdo, enquanto revistas definidas pela cor azul escuro entre o centro e o lado esquerdo inferior representam outros países da região, como Chile e Peru.

## PARTE III – DISCUSSÃO

### 7 Discussão dos resultados

Os dados analisados nos capítulos anteriores buscaram revelar “As estruturas globais e regionais do campo de pesquisa, desenvolvimento e inovação das doenças negligenciadas leishmaniose e tuberculose sob a ótica das redes complexas”. O que será discutido doravante é de que forma esses resultados ajudaram a melhor compreender essas estruturas globais e regionais.

As estruturas de rede mais comuns e estudadas neste trabalho são denominadas Livre de Escala, Mundo Pequeno e Centro Periferia. Além dessas estruturas, elementos e teorias da Ciência de Redes foram essenciais para o entendimento da evolução das redes e das tendências no universo pesquisado. Dentre esses elementos podem ser citados os hubs, as sub-redes, o componente gigante, a modularidade e os *clusters*, o conceito de clube dos ricos e outros.

Outro fator de grande relevância no trabalho é a evolução das redes e dos elementos e estruturas definidas nesse intervalo de tempo. O estudo dessas mudanças nas redes pode colaborar com o entendimento sobre as características do universo de pesquisa com vistas a melhor apoiar as políticas públicas.

Dentre as estruturas encontradas, as redes com características Livre de Escala possuem um grupo de pesquisadores com um número muito maior de publicações e conseqüentemente mais colaborações do que os demais membros da rede. Essa característica já havia sido investigada por outros pesquisadores (NEWMAN, 2004; BARABÁSI et al., 2002; CLAUSET; SHALIZI; NEWMAN, 2009), no entanto a mudança nessas estruturas não é algo tão discutido (GAY, 2012). Podem-se notar na pesquisa realizada mudanças da estrutura Livre de Escala para Mundo Pequeno.

Essa mudança demonstra que no início do universo temporal estudado alguns poucos pesquisadores eram o centro de todas as pesquisas realizadas, mas com o passar dos anos esses pesquisadores evoluíram para centros de pesquisa onde dezenas de pesquisadores se tornaram relevantes, e não apenas alguns.

Nas redes com estrutura Mundo Pequeno, os grupos são relativamente separados entre si, com a presença de pesquisadores que conseguem fazer o intercâmbio entre os diferentes grupos, reduzindo com isso a distância média da rede. Esses pesquisadores são os novos

membros dos grupos de pesquisa que foram emergindo em consequência da existência dos pesquisadores mais relevantes do início do estudo.

Esses pesquisadores relevantes, seja na estrutura Livre de Escala, seja na estrutura Mundo Pequeno, são denominados hubs e são os elementos de maior evidência na rede. Eles possuem o maior número de publicações e colaborações, possuem a capacidade de interagir com outros grupos, servindo de intermediadores do conhecimento, além da possibilidade de divulgar novas tecnologias com maior rapidez que os demais membros devido à sua proximidade com o resto da rede.

Essas características dos hubs levam à discussão de outra estrutura de rede denominada Centro Periferia, encontrada nos grupos de pesquisa avaliados. Essa estrutura definida por Borgatti (2000) demonstra que a rede é composta por elementos, sejam eles grupos ou pesquisadores, muito bem conectados entre si, mas que também têm relações com todo o resto da rede. Nessa estrutura identifica-se uma clara relação de poder, na qual os membros da periferia da rede, apesar de em maior número, dependem de certa forma dos conhecimentos gerados pelos grupos do centro, uma vez que estes são os mais citados e seus trabalhos, os mais difundidos. A possibilidade de mudança do pesquisador de sua posição periférica para uma central vai depender de sua capacidade de publicação e da aceitação dos seus conhecimentos pelos seus pares e eventualmente de sua participação em grupos de pesquisa já consolidados ou da criação de novos grupos de pesquisa que sejam aceitos pelos “membros” do centro.

A característica de se manter no centro da rede e com boas relações com outros grupos também no centro é conhecida como clube dos ricos. Essa teoria define que as pessoas mais relevantes tendem a se manter relacionadas com outras pessoas também relevantes, e dessa maneira todo o conhecimento gerado por esse “clube” é controlado e difundido para toda a rede.

Outro elemento importante discutido anteriormente é o componente gigante, que é o maior elemento conectado da rede. Na Ciência de Redes normalmente se analisa somente este componente, uma vez que medidas como distância média da rede ou estruturas como Mundo Pequeno só podem ser calculadas de forma apropriada nesse elemento.

Para as áreas de pesquisa bem definidas, como é o caso das leishmanioses e da tuberculose, percebeu-se que o componente gigante chega a mais de 90% dos elementos da rede em alguns casos. No entanto, para áreas de pesquisa mais difusas quanto as áreas de conhecimento ou áreas de pesquisa mais recentes, essa forte relação entre os pesquisadores pode não ser tão evidente. Ou seja, pesquisadores de uma área de conhecimento como



Ciências Sociais Aplicadas quando do estudo de colaborações científicas não estarão relacionados com pesquisadores de Ciências da Computação estudando o mesmo tema. Dessa maneira, o entendimento sobre a existência desse componente gigante quanto a sua abrangência pode também ajudar a entender melhor a ciência em seus diferentes recortes.

Quando do estudo dos diferentes componentes na pesquisa realizada, pode-se notar uma clara divisão dos elementos por meio do uso da modularidade e da formação dos *clusters*. Esses *clusters*, que em alguns casos podem ser chamados de sub-redes, ajudaram a identificar a formação de grupos de pesquisa com base, muitas vezes, em questões geográficas tanto para um universo mundial quanto para recortes regionais.

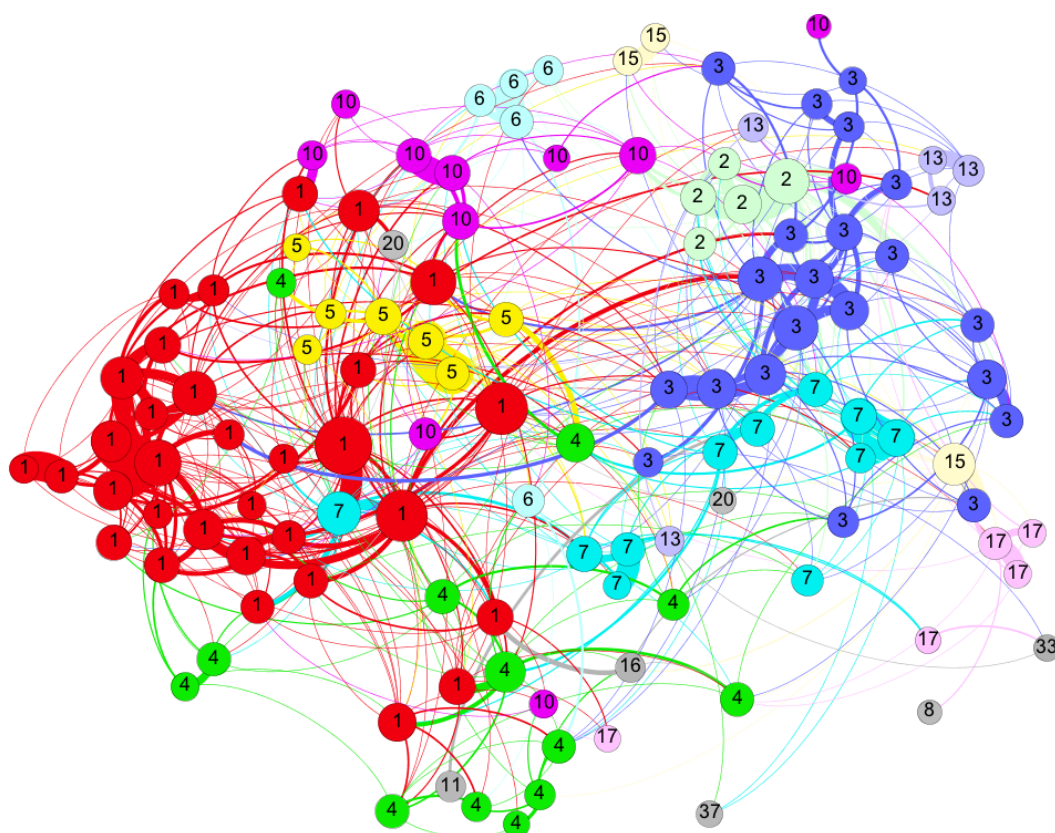
Todos esses elementos e estruturas, portanto, foram essenciais na pesquisa para o entendimento das particularidades das redes pesquisadas.

## **7.1 Discussão dos resultados PubMed**

O estudo apresentado foi tido como o ponto de partida para uma investigação sistemática da estrutura da ciência sobre as doenças negligenciadas e sua evolução, especialmente na última década. Buscou-se demonstrar neste estudo que, embora os hubs sejam poucos em números (menos de 5% do número total de cientistas conectados), eles não deixam de ser o principal fator na determinação da topologia complexa da rede de coautoria em leishmaniose e tuberculose, incluindo a definição das estruturas em comunidades.

Comunidades distintas são resultados claros e definidos a partir de interações entre os pesquisadores denominados hubs, em espaços geográficos específicos ou por temas de interesse comum. De que forma os resultados do PubMed revelaram estruturas globais e regionais? A Figura 76 demonstra a interação entre os hubs da rede leishmaniose.

FIGURA 76. Rede de relações entre hubs



Fonte: elaboração do autor

Legenda: organização estrutural dos hubs. Cada nó e cada linha são coloridos de acordo com a comunidade da qual participam na rede, o que também é demonstrado pelo número dentro do nó. A distribuição de comunidades foi determinada utilizando-se o algoritmo Leading Eigenvector. O tamanho dos nós representa o grau na rede. A espessura das linhas representa o número de colaborações entre dois nós, e a cor determinada para a linha segue a de um dos nós ao qual está ligada. A distribuição da rede segue o algoritmo Force Atlas 2, que demonstra aproximações estruturais de forma visual.

O recorte realizado para a pesquisa da base PubMed teve como objetivo abranger o universo de pesquisa mundial nas respectivas áreas. De forma global, observou-se a existência de *clusters* com características geográficas e colaborações bem estabelecidas entre países. Os *clusters* com maior número de pesquisadores tinham também o maior número de hubs na maioria dos casos, demonstrando onde se inseriam as estruturas de pesquisa por localização geográfica.

No caso das leishmanioses, duas comunidades atualmente dominam o campo, com cientistas dos EUA e do Brasil (grupo 1 da Figura 76) de um lado e cientistas do Reino Unido, da Bélgica e da Índia do outro lado (grupo 3 da Figura 76). Todavia, embora os cientistas americanos ainda mantenham uma posição forte no campo de pesquisa, cientistas do Brasil e da Índia, agora as grandes economias nacionais emergentes, tornaram-se centrais nestas duas comunidades depois do ano 2000. Essa grande mudança na pesquisa nesses dois

países, que corresponde também a um salto substancial na produtividade científica global, está em consonância com o aumento nos investimentos e com as novas políticas públicas para LN, que começaram no início da década e parecem indicar o quanto esses esforços deram frutos.

No caso da tuberculose, existe uma interação entre os hubs e entre os grupos que vai além das fronteiras geográficas e possui, em alguns casos, temas de pesquisa como elemento agregador. Os dois grupos mais bem conectados da tuberculose possuem pesquisadores situados ao norte da parte continental europeia, incluindo países como Holanda, Bélgica, França e Alemanha. O outro grupo possui pesquisadores dos EUA, da China e alguns poucos do Reino Unido. Outros grupos que se destacaram têm pesquisadores da Índia, da Itália e da Suíça em um deles e do Canadá e da França em outro. Há um quinto grupo no qual os pesquisadores são prioritariamente chineses, mas com forte relação com o grupo que possui pesquisadores dos EUA e da China. Este último grupo foi o único, dentre os analisados com hubs, que se mostrou concentrado em uma única região geográfica. Os pesquisadores brasileiros não foram encontrados entre os hubs, ou pesquisadores com o maior número de publicações na área de tuberculose, apesar do crescente número de publicações nos últimos anos.

A pesquisa demonstra que a rede está organizada em torno de um “clube dos ricos” bastante estável, com interações duradouras entre seus membros, o que levanta a questão da adequação desse tipo de estrutura para o futuro da pesquisa nessas áreas. Os pesquisadores centrais ou hubs mantêm relações diretas com grande parte da rede, caracterizando com isso as estruturas centro periferia. Dessa maneira, esses hubs possuem a capacidade de intervir ou direcionar os rumos da pesquisa com base nas suas próprias tendências ou interesses. Além disso, eles possuem forte relação entre si, formando com isso uma estrutura hierárquica e com grande capacidade de influência entre os demais pesquisadores da rede. Essa estrutura rígida e bastante hierárquica é mais bem percebida na rede leishmaniose em virtude do menor número de subgrupos e maior regionalização.

Outro elemento analisado na pesquisa foi a evolução temporal das redes. Além do crescimento do número de publicações, novas estruturas de poder foram se formando com o passar dos anos. No caso da LN, o Brasil e a Índia tiveram um crescimento bastante superior aos demais países, especialmente quando comparado com períodos anteriores. Para a TB, o crescimento elevado comparado com períodos anteriores é percebido para os EUA, mas também para os países do chamado Brics, com exceção da Rússia. A China e a Índia tiveram

o maior crescimento percentual no número de publicações no último período de quatro anos, seguidas da África do Sul e do Brasil.

A evolução das redes não se deu apenas quanto aos países, mas também quanto à estrutura de colaboração. Houve um aumento no número de parcerias científicas que subiu de uma média de 3,5 para seis colaboradores por artigo nos últimos 32 anos. Essa mudança é percebida nas estruturas das redes com o aumento da densidade destas. Esse aumento na densidade deve-se a vários fatores, como maior número de publicações por ano, maior número de colaboradores por artigo e também maior número de pesquisadores na área. Dessa maneira, as relações mudaram de uma centralidade baseada em alguns poucos pesquisadores para uma centralidade em grupos de pesquisa, que por sua vez possuem pesquisadores chave que mantêm esse grupo coeso. Para a LN, os centros ou grupos de pesquisa concentram-se em dois grandes grupos; para a TB, eles se concentram em seis grandes grupos.

A hipótese com a qual se trabalha é que os grupos possuem pessoas-chave, definidas pelas estruturas de Hubs ou *k-core*, que são responsáveis pelo direcionamento das pesquisas, uma vez que sua influência sobre os demais representantes da rede é muito grande. Além disso, acredita-se que cada comunidade possui um universo mais ou menos definido de pesquisa, e que sua relação com outros grupos compõe um elemento de definição da sua área.

No caso das estruturas regionais, o estudo dos *clusters* e de seus elementos trouxe evidências sobre os grupos de pesquisa. Ficou clara a forte relação entre os pesquisadores dos respectivos *clusters*, e as barreiras geográficas mostraram-se um elemento importante na definição dessa estrutura. No entanto, o crescimento da importância dentro de um grupo pareceu diretamente relacionado à capacidade de interagir com pesquisadores de outros países, característica identificada nos hubs.

## 7.2 Discussão dos resultados Web of Science

O estudo da LN e da TB na Web of Science teve um recorte voltado para as publicações científicas que tinham dentre seus autores pesquisadores de instituições brasileiras. Assim, este estudo teve um foco mais restrito que o realizado na base PubMed. Também foi realizada uma análise mais aprofundada sobre a participação de uma instituição específica, a Fiocruz, que tem como foco de suas ações a atenção à saúde pública e sua influência e colaboração com as demais entidades.

Tendo em vista a crescente complexidade e o caráter sistêmico do processo de geração de inovações na área da saúde, a colaboração científica desponta como um importante

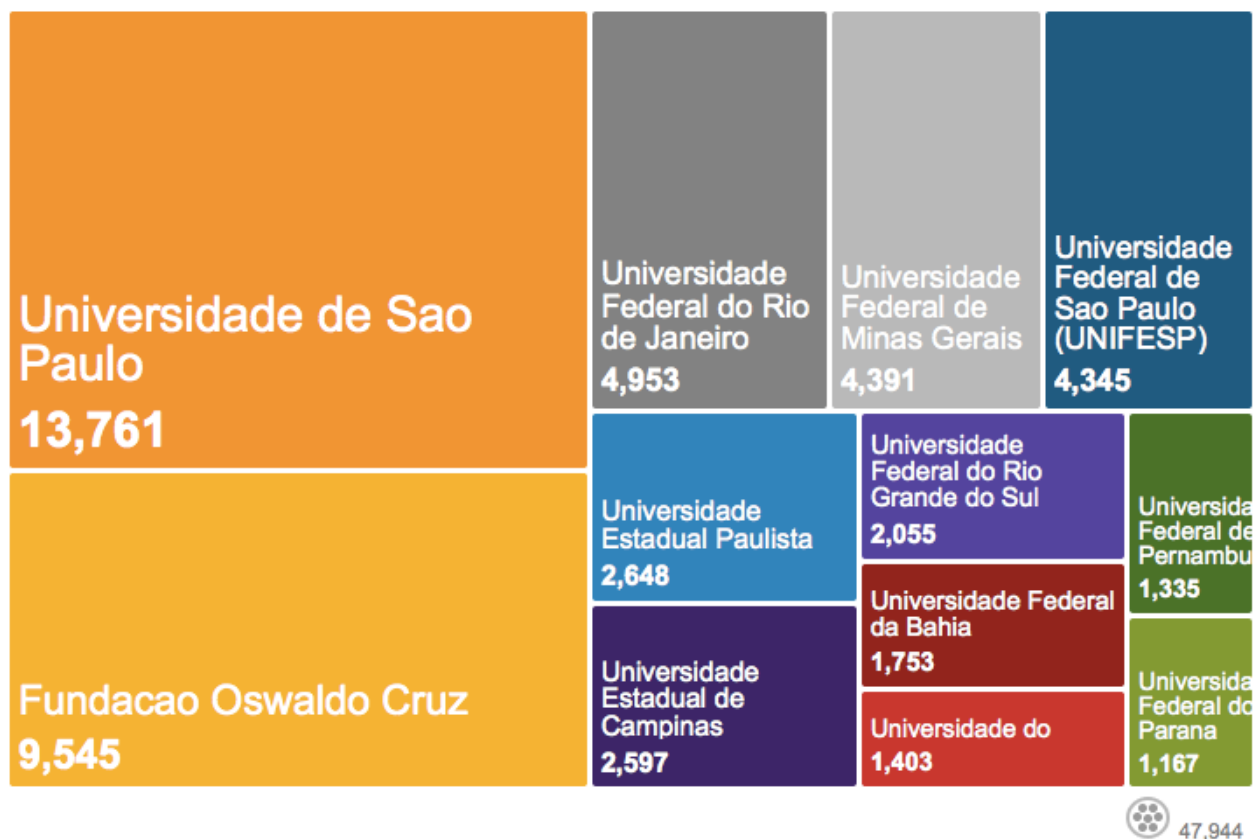
mecanismo de promoção de inovações. Dessa maneira, o mapeamento das redes de colaboração científica no Brasil facilitou a busca por informações estratégicas que podem apoiar a construção de um futuro alinhado às demandas sociais do país.

Para a construção da rede LN foram utilizadas informações obtidas nas publicações em leishmaniose de autores em instituições brasileiras no período de 1984 a 2013. A análise foi feita em um recorte específico de 1.908 pesquisadores, vinculados a 4.211 publicações, dos quais 332 declararam ser afiliados à Fiocruz. Os pesquisadores da Fiocruz representam 17,5% de toda a rede de pesquisa em leishmanioses no universo pesquisado.

Quanto à análise em TB, a coleta de dados recuperou 2.326 publicações, com um total de 10.701 autores. Após o tratamento e o recorte dos dados, a rede ficou com 2.181 pesquisadores, sendo 1.800 pesquisadores brasileiros e 477 deles afiliados à Fiocruz. Os pesquisadores da Fiocruz representam 21,87% de toda a rede de colaboração em pesquisa sobre tuberculose.

O número de pesquisadores da Fiocruz nas duas áreas de pesquisa e o número de publicações desses pesquisadores nas categorias mais comuns para LN e TB são semelhantes. A Fiocruz representa 17,5% na rede LN e 21,87% na rede TB, e sua produção é a segunda nas áreas de medicina tropical, imunologia, doenças infecciosas, microbiologia, parasitologia, sistemas respiratórios e saúde pública, ambiental e ocupacional, que são as categorias mais comuns para publicação em LN e TB. A Figura 77 aponta os números para as entidades brasileiras com maior quantidade de publicações englobando todas as áreas citadas. O total de publicações nessas áreas para as entidades brasileiras totaliza 47.944.

**FIGURA 77. Número de publicações de entidades brasileiras em áreas específicas na Web of Science**



Fonte: elaboração do autor – ferramenta In Cites Thomson Reuters

Legenda: número de publicações nas áreas de medicina tropical, imunologia, doenças infecciosas, microbiologia, parasitologia, sistemas respiratórios e saúde pública, ambiental e ocupacional para as 12 instituições brasileiras com maior número de publicações na área.

A Fiocruz é um importante elemento no universo de pesquisa nessas áreas no Brasil, primeiro pelo número de publicações e consequentes colaborações científicas, segundo pela sua abrangência territorial, que, ao contrário das universidades federais, possui uma grande capilaridade por meio das suas unidades e diretorias nos diferentes estados brasileiros. Em virtude dessa situação privilegiada na área da pesquisa e da colaboração científica, a rede de pesquisadores da Fiocruz assemelha-se à rede de pesquisadores brasileiros, resultado da ampla participação dos pesquisadores da Fiocruz em ambas as redes.

Essa similaridade entre as redes nacionais e a da Fiocruz pode ser mais bem percebida por meio das colaborações internacionais. O que se percebeu na pesquisa é que os principais países colaboradores não são aqueles com alta incidência de LN ou TB, isso demonstra que os países que mais se beneficiariam do avanço científico na área ou que poderiam fornecer dados e materiais importantes estão pouco inseridos na rede de cooperação neste tema no Brasil.

No caso da LN, o principal país pesquisador e com acometimento da doença é a Índia, que não aparece como um grande parceiro científico. No caso da TB, os países da América Latina possuem grandes problemas na área, no entanto não são evidenciados nas colaborações.

Os principais países colaboradores, tanto da Fiocruz quanto do Brasil, estão entre aqueles que mais investem em P&D para doenças negligenciadas: Estados Unidos e Inglaterra. A avaliação do número de colaborações internacionais da rede de pesquisa sobre leishmaniose no Brasil aponta um aumento da cooperação a partir dos anos 2000, provavelmente ligada ao aumento do número de artigos publicados em coautoria que vêm se estabelecendo nas últimas décadas.

No que diz respeito às colaborações nacionais, percebeu-se uma grande participação da Fiocruz nas mais diferentes áreas do Brasil. A questão geográfica é um importante fator na colaboração científica, situação evidenciada pela formação dos grupos de pesquisa. Em razão de sua capacidade de se relacionar com as diferentes entidades nacionais, a Fiocruz torna-se o elemento agregador na pesquisa em saúde, com a vocação de colaborar com a grande maioria das pesquisas realizadas em território nacional bem como de influenciá-las.

Além da colaboração científica e da capacidade de influenciar dos atores, foram estudadas as categorias de pesquisa realizadas pelos diferentes grupos encontrados na rede. Conforme dito antes, as categorias definidas pela Web of Science estão relacionadas às revistas científicas, que, por sua vez, podem ser mais ou menos específicas. Para LN, os temas mais comuns foram medicina tropical e parasitologia. No entanto, um dos grupos teve uma tendência um pouco maior de publicar em temas como imunologia, doenças infecciosas e saúde pública, ambiental e ocupacional. O entendimento sobre essas diferenças pode apoiar o direcionamento de políticas públicas para uma ou outra área específica com o objetivo de fortalecer assuntos menos recorrentes, caso essa seja a decisão.

Para TB, as categorias estudadas tiveram uma pequena diferença do que para LN, incluindo áreas como sistema respiratório e microbiologia. No que diz respeito aos grupos, percebeu-se uma maior diversidade e interesse, demonstrando que o estudo da TB se encontra em diversas áreas do conhecimento e da saúde.

Ao longo deste estudo, pôde-se observar que a análise aprofundada das redes de cooperação conseguiu identificar seus principais interlocutores, além de mostrar a maneira como se relacionam, o que discutem e o que produzem. O estudo pôde determinar quais são os pesquisadores e como eles se relacionam, quais os temas mais frequentes e quais as

possíveis estratégias para articular atividades de pesquisa que poderiam atender às demandas nacionais.

O estudo permitiu identificar os relacionamentos, suas estruturas de poder e as relações centrais e periféricas desempenhadas pelos pesquisadores brasileiros. Ao mesmo tempo, a pesquisa contribuiu para identificar características que auxiliam no desempenho da rede brasileira de pesquisa, minimizando os riscos de fragmentação de suas ações e identificando possibilidades de intervenção.

Com uma análise mais aprofundada dos vários resultados obtidos neste estudo, associando-os a informações qualitativas, pode-se ter informações para subsidiar a tomada de decisão e estimular as colaborações externas ou integrar mais a rede interna, onde concentrar os recursos de suas pesquisas para a produção de conhecimento capaz de gerar novas tecnologias, onde atuar mais fortemente na integração e na estimulação de redes temáticas de interesse.

### 7.3 Discussão dos resultados SciELO

O SciELO, por ser uma base científica brasileira, possui um número maior de coleções (revistas, congressos, anais e outros) do próprio país do que de outros países associados. Independentemente dessa diferença, a pesquisa mostrou que na área da saúde, especialmente em LN, o Brasil se destaca quanto ao número de publicações ante os demais países, sendo eles África do Sul, Argentina, Bolívia, Chile, Costa Rica, Cuba, Espanha, Índias Ocidentais, México, Paraguai, Peru, Portugal, Uruguai e Venezuela. A Tabela 28 é aqui novamente apresentada para ressaltar essa diferença.

**TABELA 28. Percentual de títulos e publicações por país**

	Títulos	Títulos (%)	LN	LN (%)	TB	TB (%)
<b>Brasil</b>	324	30,1	1191	71,0	1.588	42,3
<b>Colômbia</b>	156	14,5	125	7,4	253	6,7
<b>México</b>	113	10,5	11	0,7	103	2,7
<b>Argentina</b>	104	9,7	24	1,4	118	3,1
<b>Chile</b>	103	9,6	22	1,3	217	5,8
<b>Espanha</b>	53	4,9	19	1,1	254	6,8
<b>Venezuela</b>	53	4,9	60	3,6	98	2,6
<b>Cuba</b>	48	4,5	24	1,4	193	5,1
<b>Portugal</b>	44	4,1	2	0,1	69	1,8
<b>África do Sul</b>	28	2,6	2	0,1	131	3,5
<b>Saúde pública</b>	18	1,7	163	9,7	562	15,0
<b>Costa Rica</b>	18	1,7	5	0,3	26	0,7
<b>Peru</b>	15	1,4	30	1,8	146	3,9

Fonte: elaboração do autor



A Tabela 28 demonstra que no Brasil, apesar de ter 30,1% das coleções na base, o número de publicações em LN é de 71% e o de TB é de 42,3%, caracterizando a importância desse tema para o Brasil comparado aos demais países associados ao SciELO. Esse levantamento diz respeito apenas às coleções indexadas nesta base, não sendo uma visão geral de cada um desses países.

Decidiu-se também por fazer uma análise comparativa das duas áreas, LN e TB, quanto ao perfil de publicações. A primeira grande diferença encontrada diz respeito ao percentual de publicações mencionado de 71% e 42,3% apresentados na Tabela 28. Para o caso da LN, o Brasil é um dos principais pesquisadores na área, com níveis similares e até maiores em alguns casos que os EUA e alinhado com a Índia. Para a TB, mesmo dentro do SciELO, outros países mostraram-se envolvidos com o tratamento da doença e a pesquisa na área.

Esse mesmo perfil e abrangência de países podem também ser percebidos quanto às revistas utilizadas em cada área. No caso da LN, existem duas revistas que possuem um número muito mais elevado de publicações que todas as outras juntas: *Memórias do Instituto Oswaldo Cruz* e a *Revista da Sociedade Brasileira de Medicina Tropical*, ambas brasileiras. Para TB, essa diferença não é tão grande, com revistas de diferentes países com números similares de publicação na área. Essa diversificação é um fator relevante e também necessário na ciência, pois diferentes linhas de conhecimento e competências são necessárias para se alcançar resultados mais promissores no tratamento da doença.

O terceiro elemento estudado na pesquisa diz respeito à colaboração entre países por meio da publicação em diferentes revistas. Alguns podem argumentar que existe uma barreira na língua, com pesquisadores de países de língua espanhola ou inglesa tendo uma maior dificuldade de publicar no Brasil ou em Portugal e vice-versa. Entretanto, todas as revistas aceitam a publicação na língua inglesa, que pode ser considerada padrão na ciência em algumas áreas. Acredita-se, portanto, que essa barreira vai muito além da questão da língua e exige uma intervenção por meio de políticas de incentivo e troca científica.

Para a LN, percebeu-se uma diferença clara entre as revistas brasileiras e as revistas de outros países latino-americanos. Dentre os países latinos, especialmente Cuba, Panamá e Peru, existe uma grande colaboração. Contudo, para o Brasil existe pouca colaboração ou diversificação nas escolhas por revistas. Como a maior parte das revistas de grande impacto está nos EUA e nos países da Europa Ocidental, este pode ser um elemento importante na escolha das revistas para publicação.

A pesquisa em TB mostrou outro aspecto interessante quanto à escolha das revistas para publicação pelos pesquisadores. Além das diferenças regionais percebidas em LN, evidenciou-se uma diferença temática, demonstrando a abrangência das áreas em TB. Os pesquisadores da área da enfermagem e farmácia, por exemplo, focam sempre nas suas revistas, tendo um *link* forte com uma revista específica, o *Jornal Brasileiro de Pneumologia*. Outro grupo bastante diferenciado é a área de ciências agrônômicas, pois os pesquisadores que publicam em revistas nessa área não publicam em revistas da saúde.

De maneira geral, percebeu-se uma clara diferença entre os padrões de publicação para LN e para TB, além de uma separação geográfica na escolha por revistas para publicação.

## 8 Conclusão

Neste trabalho decidiu-se realizar a análise de publicações científicas em duas áreas da saúde relacionadas às doenças leishmaniose e tuberculose. Para entender como ocorreu a evolução científica nesse universo e quais elementos podem ser utilizados para apoiar decisões futuras, foi elaborado um *corpus* sobre publicações acadêmicas em diferentes bases científicas nesses dois temas. Na pesquisa, analisou-se a estrutura de colaboração mundial e regional e a conseqüente relação de poder entre os pesquisadores e seus respectivos grupos de pesquisa e colaboração. Publicações sobre leishmaniose e tuberculose foram recuperadas por meio das bases PubMed, Web of Science e SciELO, processadas com ferramentas de mineração e tratamento de dados e analisadas com pacotes de *software* de análise de redes sociais.

Apresentou-se que a ciência nas áreas de leishmaniose e tuberculose vem crescendo exponencialmente desde a década de 1980, crescimento medido pelo número de artigos publicados a cada ano, pela taxa de influxo de novos autores e pelo aumento na colaboração científica. Constatou-se ainda que existe uma estrutura de poder revelada pela existência de um “clube dos ricos”, constituído por um grupo de cientistas interligados entre si que possuem uma relação direta com o resto da rede com pelo menos dois graus de separação e que controlam a estrutura da rede em todos os momentos.

Foi observada também uma mudança no perfil dos pesquisadores chave que têm o potencial de controlar a rede. Pesquisadores dos países chamados emergentes ou do bloco dos BRICS a cada ano vêm ocupando mais espaço. No caso da leishmaniose, Brasil e Índia se destacam, enquanto na tuberculose, China, Índia, África do Sul, Brasil e Rússia se mostram cada vez mais atuantes.

Na pesquisa mais focada na situação regional percebeu-se a participação ativa e importante da Fiocruz como instituição agregadora nas duas áreas de pesquisa e com relações fortes e bem estabelecidas com outras instituições de ensino e pesquisa nos diferentes estados brasileiros. Esse perfil mostrou-se crítico para o direcionamento de trabalhos de pesquisa, e o entendimento mais profundo dessas relações e a capacidade de influência são essenciais para a determinação de um futuro adequado às necessidades de saúde brasileiras.

Por último analisou-se a capacidade de interação de pesquisadores na região da América Latina e de outros países por meio da publicação em revistas desses países. Notou-se que existe pouca interação entre os pesquisadores e que existe uma necessidade de

intervenção mais clara e objetiva para a aproximação desse bloco, uma vez que os países deste bloco compartilham uma série de problemas similares na área da saúde.

Quanto ao modelo de pesquisa e ao universo utilizado, defende-se neste trabalho a necessidade de se aprofundar no uso de análises de redes sociais dentro da Ciência da Informação como forma de complementar o entendimento por meio de análise de citações ou análise bibliográfica. Os resultados de análises bibliométricas e cientométricas podem ser extrapolados com o aprofundamento de análises de colaboração científica. O que se buscou neste trabalho foi apresentar uma série de resultados com a aplicação da análise de colaboração científica como forma de definir uma base para trabalhos futuros.

Segundo Glanzel (2002), a autoria de um artigo é o principal descritor bibliométrico de uma publicação científica. As tendências e os padrões percebidos na autoria podem caracterizar a estrutura social e até mesmo cognitiva dos diferentes campos de pesquisa. Nos últimos anos tem havido uma intensificação das colaborações científicas, trazendo novos insumos para o entendimento da ciência. Dessa maneira, faz-se necessário o aprofundamento de métodos mais eficazes que possam apoiar o estudo dessas autorias e colaborações. Defende-se que o uso dos métodos de análise de redes para avaliação de publicações científicas deve obrigatoriamente ir além dos métodos clássicos de análise bibliométrica, de maneira que sejam abarcadas e entendidas essas novas tendências.

Os resultados de estudos de coautoria podem ser usados em uma perspectiva política de investigação da ciência. A colaboração científica fornece uma visão geral das principais características do processo de comunicação científica e estruturas de poder de uma área de pesquisa. A colaboração pode ser analisada sob a perspectiva de um artigo e suas áreas foco, de uma revista científica e suas participações, de um recorte temporal incentivado ou não por uma política pública ou de uma região geográfica e seus grupos formais ou não formais de trabalho. A análise dos padrões e das tendências em uma evolução temporal possibilita a avaliação de vários pressupostos e políticas de ciência, na medida em que eles se relacionam com a colaboração científica.

## **8.1 Pergunta de pesquisa, objetivos e hipóteses**

A pergunta de pesquisa apresentada no início do trabalho é: “Como ou de que forma os métodos e as técnicas de análise de redes complexas descrevem as estruturas e os processos globais e regionais da pesquisa a partir da investigação das redes de coautoria e

outras informações encontradas em publicações científicas no campo das doenças negligenciadas leishmaniose e tuberculose?”

Os resultados dos trabalhos e dos métodos aplicados demonstram que existe uma grande quantidade de informações que podem ser analisadas tomando como base as colaborações científicas. As métricas da análise de redes utilizadas ajudam a entender as estruturas das redes e seu potencial de crescimento, evolução e disseminação de informações e conhecimento. Essas relações sociais associadas ao uso de certas palavras ou à escolha por publicações em certos meios têm um poder de influência maior do que pode ser percebido com a utilização dos métodos clássicos de análise bibliométrica e cientométrica.

### **8.1.1 Objetivos**

Quanto ao objetivo geral do trabalho, “Estudar a aplicabilidade de um conjunto de métodos para análise de redes complexas sobre publicações científicas a respeito de doenças negligenciadas”, acredita-se que o resultado foi satisfatório no intuito de se abrir novos horizontes no campo da pesquisa das colaborações científicas. De forma alguma o trabalho procurou ser exaustivo quanto aos diferentes métodos. O foco principal foi a validação da capacidade de se produzir insumos que pudessem apoiar a tomada de decisão sobre áreas específicas da ciência.

Quanto aos objetivos específicos foram apresentados métodos e ferramentas para a realização das diferentes tarefas no processo de pesquisa. Buscou-se ser o mais claro possível na descrição das medidas e etapas do processo de maneira a possibilitar o entendimento por indivíduos de diferentes áreas do conhecimento, uma vez que o método se apresenta como transdisciplinar.

Quanto à análise das diferentes bases de dados existentes, o objetivo foi atendido com a apresentação dos resultados em cada um dos três capítulos de discussão. Não foi realizada uma comparação direta entre as bases, mas sim uma discussão quanto ao seu uso e seus principais atributos. A escolha de uma base em detrimento de outra deve ser de discernimento do pesquisador e não foi objeto deste trabalho trazer à tona essa discussão. O que se pretendeu foi demonstrar que existe a possibilidade de se realizar a análise de diferentes maneiras para atender aos diferentes interesses e necessidades informacionais. Quando da pesquisa em cada uma das bases se buscou a aplicação de métodos de certa forma complementares, e não idênticos entre si. Essa mudança ocorreu não só quanto aos métodos, mas também quanto aos recortes temporais e regionais e quanto ao número de publicações.

Dentre os resultados mais relevantes percebidos na pesquisa, o entendimento sobre a existência de uma estrutura de poder entre os participantes na ciência foi o que mais chamou a atenção. O estudo dos padrões e das tendências nas relações entre os pesquisadores chave pode revelar uma capacidade de direcionamento das pesquisas com o potencial de influenciar as políticas públicas que ainda não está claro.

No que diz respeito à inovação, acredita-se que ainda existe muito espaço para pesquisa que não foi realizado neste trabalho. Um dos elementos principais na área da saúde é o entendimento profundo quanto ao uso das MeSH associado às colaborações e à cognição.

### **8.1.2 Pressupostos**

Acredita-se portanto com a pesquisa realizada que a existência de um arcabouço de informações relevantes no estudo de publicações científicas que não são exploradas pelos métodos convencionais de análises bibliométricas e cientométricas foi corroborada. Muitos dos resultados encontrados e apresentados foram validados por especialistas em diversas áreas da ciência, incluindo as áreas pesquisadas. Os resultados encontrados foram considerados relevantes e inovadores quanto ao melhor entendimento da ciência e suas potencialidades.

O segundo pressuposto apresentado foca nas estruturas de colaboração existentes nas redes. Acredita-se que em parte foi confirmado, especialmente quanto à existência de grupos coesos com a capacidade de influência direta na rede. No entanto, não se conseguiu comprovar o real fator que leva esses elementos ao poder, que vai além do número de publicações e consequente elevado número de colaborações. Elementos financeiros não puderam ser avaliados quanto ao potencial alavancador dessas estruturas.

Por fim o trabalho descreve que apesar de haver uma grande coesão nas redes de publicação científica, existe uma estrutura centrada em poucos grupos de pesquisa tradicionais que normalmente estão estabelecidos em virtude das suas colaborações recorrentes. No caso da leishmaniose, esses grupos tradicionais estão dando espaço para os novos grupos emergentes, mas não se sabe se a custo de políticas de incentivo financeiro ou outro fator. Portanto, os elementos que podem contribuir para a criação ou o desmembramento de estruturas formais de poder não foram encontrados.

Quanto à capacidade de entrada de novos pesquisadores nas estruturas já existentes de poder, percebeu-se a dificuldade desse avanço nas estruturas regionais. Na escala mundial, novos grupos se formaram. O espaço temporal analisado manteve-se em menos de uma geração, e com isso os elementos chave se mantiveram na cadeia produtiva.

Por último, confirmou-se a necessidade de aumento dos centros de pesquisa para propiciar avanços na ciência. Em uma comparação entre as áreas de leishmaniose e tuberculose, verificou-se que na segunda existe uma diversidade maior de centros de pesquisa e conseqüente maior colaboração internacional. Partindo do pressuposto de que o aumento da colaboração é uma tendência e uma necessidade para se transpor barreiras de conhecimento, é de entendimento deste trabalho que é necessária uma maior disseminação e conseqüente colaboração de centros de pesquisa para atender às necessidades da saúde.

## **8.2 Trabalhos futuros**

Conforme discutido anteriormente, a área de Análise de Redes Complexas é transdisciplinar e pode atender diversas demandas informacionais. Além de estar alinhada a novas oportunidades provenientes de áreas das ciências exatas, por meio da Ciência de Redes, existe uma tendência quanto ao uso dos métodos de análise de redes, com especial atenção na área da Saúde. Esse crescente interesse abre portas para muitos trabalhos de pesquisa.

Considerando-se as atividades recém-realizadas, trabalhos futuros utilizando essa pesquisa devem analisar as relações colaborativas avaliando relações sociais, por meio das parcerias científicas, e cognitivas, por meio das atividades fim ou dos resultados alcançados. Outra linha de pesquisa está relacionada à capacidade de controle ou à influência dentro da rede.

## Referências

ADAMS, James; GRILICHES, Zvi. Measuring science: an exploration. *Proceedings of the National Academy of Sciences*, v. 93, n. 23, p. 12664-12670, 1996.

ALBERT, Réka; BARABÁSI, Albert-László. Statistical mechanics of complex networks. *Reviews of Modern Physics*, v. 74, n. 1, p. 47-97, 2002. doi:10.1103/RevModPhys.74.47.

ALVAR, Jorge. Leishmaniasis worldwide and global estimates of its incidence. *PloS One*, v. 7, n. 5, 2012. e35671. doi:10.1371/journal.pone.0035671.

BARABÁSI, Albert-Laszlo. *Linked: how everything is connected to everything else and what it means*. New YORK, Plume, 2003.

BARABÁSI, Albert-László; RÉKA, Albert. Emergence of scaling in random networks. *Science*, v. 286, n. 5439, p. 509-512, 1999. doi:10.1126/science.286.5439.509.

BARABÁSI, Albert-László; ZOLTÁN, N. Oltvai. Network biology: understanding the Cell's Functional Organization. *Nature Reviews Genetics*, v. 5, n. 2, p. 101-113, 2004. doi:10.1038/nrg1272.

BARABÁSI, A. L. et al. Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and its Applications*, v. 311, n. 3-4, p. 590-614, 2002. doi:10.1016/S0378-4371(02)00736-7.

BARRETO, M. L. et al. Successes and failures in the control of infectious diseases in Brazil: social and environmental context, policies, interventions, and research needs. *The Lancet*, v. 377, n. 9780, p. 1877-1889, 2011.

BARBASTEFANO, Rafael Garcia. Impactos dos nomes nas propriedades de redes sociais: um estudo em rede de coautoria sobre sustentabilidade, *Perspectivas em Ciência da Informação* v. 18, n. 3, p. 78-95, jul. 2013.

BASU, Aparna; VINU KUMAR, B. S. International collaboration in Indian scientific papers.



*Scientometrics*, v. 48, n. 3, p. 381-402, 2000. doi:10.1023/A:1005692505687.

BEAVER, Donald Deb. Reflections on scientific collaboration (and its study): past, present, and future. *Scientometrics*, v. 52, n. 3, p. 365-377, 2001. doi:10.1023/A:1014254214337.

BOLLOBÁS, Béla. *Random graphs*. 2. ed. Cambridge; New York: Cambridge University Press, 2001.

BONACICH, Phillip. Power and centrality: a family of measures. *American Journal of Sociology*, v. 92, n. 5, p. 1170-1182, 1987.

BORDIN, Andréa Sabedra; GONÇALVES, Alexandre Leopoldo; TODESCO, José Leomar. Análise da colaboração científica departamental através de redes de coautoria. *Perspectivas em Ciência da Informação*, v. 19, n. 2, p. 37-52, abr. 2014.

BORGATTI, Stephen P.; EVERETT, Martin G. Models of core/periphery structures. *Social Networks*, v. 21, n. 4, p. 375-395, 2000. doi:10.1016/S0378-8733(99)00019-2.

BORNMANN, L. & MUTZ, R. (2014). From P100 to P100': Conception and improvement of a new citation-rank approach in bibliometrics. *Journal of the Association for Information Science and Technology*, 65(9), 1939-1943.

BRASIL. . Leishmaniose tegumentar americana; leishmaniose visceral. *Doenças infecciosas e parasitárias*. Brasília: Ministério da Saúde, 2012. p. 261-273.

BRITO, Monique Araújo de. Investment in drugs for neglected diseases: a portrait of the last five years. *Revista da Sociedade Brasileira de Medicina Tropical*, v. 46, n. 1, p. 1-2, 2013. doi:10.1590/0037-8682201200212013.

CAINELLI, Giulio et al. The strength of strong ties: how co-authorship affect productivity of academic Economists?. *Scientometrics*, v. 102, n. 1, p. 673-699, 2014. doi:10.1007/s11192-014-1421-5.

CHOO, Chun Wei. *The knowing organization: how organizations use information to*

construct meaning, create knowledge, and make decisions. 2 ed. Oxford: Oxford University Press, 2005.

CHUNG-YUAN, Huang; CHUEN-TSAI, Sun; HSUN-CHENG, Lin. Influence of local information on social simulations in small-world network models. Text.Article. 31 out. 2005. <http://jasss.soc.surrey.ac.uk/8/4/8.html>.

CLAUSET, Aaron; SHALIZI, Cosma Rohilla; NEWMAN, M. E. J. Power-law distributions in empirical data. *SIAM Review*, v. 51, n. 4, p. 661-703, 2009. doi:10.1137/070710111.

COADIC, Yves-François Le. *A ciência da informação*. Brasília, Briquet de Lemos/Livros, 1996.

COLIZZA, V. et al. Detecting rich-club ordering in complex networks. *Nature Physics*, v. 2, n. 2, p. 110-115, 2006. doi:10.1038/nphys209.

COSTA, Benedita Marta Gomes; PEDRO, Edilson da Silva; MACEDO, Gorete Ribeiro de. Scientific collaboration in biotechnology: the case of the Northeast Region in Brazil. *Scientometrics*, v. 95, n. 2, p. 571-592, 2013. doi:10.1007/s11192-012-0924-1.

CROSS, Robert L.; PARKER, Andrew. *The hidden power of social networks: understanding how work really gets done in organizations*, 2004. 4.1.2004 edition. Boston, Mass: Harvard Business Review Press.

DAVENPORT, Thomas H.; PRUSAK, Laurence. *Working knowledge: how organizations manage what they know*. 2. ed. rev. Boston, Mass.: Harvard Business School Press, 2000.

DESJEUX, P. Leishmaniasis: current situation and new perspectives. *Comparative Immunology, Microbiology and Infectious Diseases*, v. 27, n. 5, p. 305-318, 2004. doi:10.1016/j.cimid.2004.03.004.

DOROGOVSTSEV, S. N., A. V. GOLTSEV, e J. F. F. MENDES. 2007. "Critical phenomena in complex networks". *arXiv:0705.0010*, abril. doi:10.1103/RevModPhys.80.1275.

DUARTE, Emeide Nóbrega. Conexões temáticas em gestão da informação e do conhecimento no campo da ciência da informação: proposta de redes humanas, *Perspectivas em Ciência da Informação*, v. 21, n. 1, 159-173, jan. 2011.

DUARTE, Emeide Nóbrega. Conteúdos temáticos como subsídios para indicação de colaboração na ciência da informação: Ppgci/Unesp e Ppgci/Ufpb em evidência, *Perspectivas em Ciência da Informação*, v. 22, n.1, p. 189-201, 2012.

ERDÖS, P.; RÉNYI, A. On random graphs. *Publicationes Mathematicae*, v.6, p. 290-297, 1959.

FERREIRA, Gonçalo Costa. Redes sociais de informação: uma história e um estudo de caso, *Perspectivas em Ciência da Informação*, v. 16, n. 3, p. 208-231, jul. 2011.

FREEMAN, Linton C. Centrality in social networks conceptual clarification. *Social networks*, v. 1, n. 3, p. 215-239, 1979.

GLOBAL FUNDING OF INNOVATION FOR NEGLECTED DISEASES (G-FINDER). Government funding for neglected diseases: why it doesn't add up, 2014. Disponível em: <[http://polycycures.org/downloads/Government\\_Funding\\_for\\_NDs.pdf](http://polycycures.org/downloads/Government_Funding_for_NDs.pdf)>. Acesso em: 10 dez. 2014.

GAY, Brigitte. Universal dynamics on complex networks, really?. *Social Network Mining, Analysis and Research Trends: Techniques and Applications*, Taiwan, IGI Global, 2012.

GLÄNZEL, Wolfgang. National characteristics in international scientific co-authorship relations. *Scientometrics*, v. 51, n. 1 p. 69-115, 2001. doi:10.1023/A:1010512628145.

GLÄNZEL, Wolfgang. Coauthorship patterns and trends in the sciences (1980-1998): a bibliometric study with implications for database indexing and search strategies. *Library Trends*, v. 50, n. 3, p. 461, 2002.

GLÄNZEL, Wolfgang; LANGE, Cornelius de. A distributional approach to multinationality measures of international scientific collaboration. *Scientometrics*, v. 54, n. 1, p. 75-89, 2002.

doi:10.1023/A:1015684505035.

GLÄNZEL, Wolfgang; LETA, Jacqueline; THIJS, Bart. Science in Brazil. Part 1: a macro-level comparative study. *Scientometrics*, v. 67, n. 1, p. 67-86, 2006. doi:10.1007/s11192-006-0055-7.

GLÄNZEL, Wolfgang; SCHUBERT, András. Double effort = double impact? a critical view at international co-authorship in chemistry”. *Scientometrics*, v. 50, n. 2, p. 199-214, 2001. doi:10.1023/A:1010561321723.

GLÄNZEL, Wolfgang; SCHUBERT, András. “Analysing Scientific Networks Through Co-Authorship”. In *Handbook of Quantitative Science and Technology Research*, organizado por Henk F. Moed, Wolfgang Glänzel, e Ulrich Schmoch, 257–76. Springer Netherlands, 2004.. [http://link.springer.com/chapter/10.1007/1-4020-2755-9\\_12](http://link.springer.com/chapter/10.1007/1-4020-2755-9_12).

GOH, Kwang-Il et al. The human disease network. *Proceedings of the National Academy of Sciences*, v. 104, n. 21, p. 8685–8690, 2007. doi:10.1073/pnas.0701361104.

GOYAL, Sanjeev; LEIJ, Marco J. van der; GONZÁLEZ, José Luis Moraga. Economics: an emerging small world. *Journal of Political Economy*, v. 114, n. 2, p. 403-412, 2006. doi:10.1086/500990.

GRAEML, Alexandre Reis. Redes sociais e intelectuais na área de pesquisa em administração da informação: uma análise cientométrica do período 1997-2006, v. 20, n. 1 p. 95-110, jan. 2010.

GRANOVETTER, Mark S. The strength of weak ties. *American Journal of Sociology*, v. 78, n. 6, p. 1360–1380, 1973. doi:10.2307/2776392.

GUIMARÃES, Francisco; MELO, Elisete. *Diagnóstico utilizando redes sociais*. Monografia, Universidade Federal do Rio de Janeiro, 2005. [http://portal.crie.coppe.ufrj.br/portal/data/documents/storedDocuments/%7B93787CAE-E94C-45C7-992B-9403F6F40836%7D/%7BE5F077FE-704C-44EA-9B70-30BC4665277F%7D/RJ11\\_Projeto01.pdf](http://portal.crie.coppe.ufrj.br/portal/data/documents/storedDocuments/%7B93787CAE-E94C-45C7-992B-9403F6F40836%7D/%7BE5F077FE-704C-44EA-9B70-30BC4665277F%7D/RJ11_Projeto01.pdf).

HANSEN, Morten; NOHRIA, Nitin; TIERNEY, Thomas. What's your strategy for managing knowledge?. *Harvard Business Review*, 1999. <http://hbr.org/1999/03/whats-your-strategy-for-managing-knowledge/ar/1>.

HOU, Haiyan; KRETSCHMER, Hildrun; LIU, Zeyuan. The structure of scientific collaboration networks in scientometrics. *Scientometrics*, v. 75, n. 2, p. 189-202, 2007. doi:10.1007/s11192-007-1771-3.

KATZ, J. Sylvan; MARTIN, Ben R. What is research collaboration?. *Research Policy*, v. 26, n. 1, p. 1-18, 1997. doi:10.1016/S0048-7333(96)00917-1.

KIM, Jongkwang; WILHELM, Thomas. What is a complex graph?. *Physica A: Statistical Mechanics and its Applications*, v. 387, n. 11, p. 2637-2652, 2008. doi:10.1016/j.physa.2008.01.015.

KRETSCHMER, Hildrun. Author productivity and geodesic distance in bibliographic co-authorship networks, and visibility on the web. *Scientometrics*, v. 60, n. 3, p. 409-420, 2004. doi:10.1023/B:SCIE.0000034383.86665.22.

LAUDEL, Grit. What do we measure by co-authorships?. *Research Evaluation*, v. 11, n. 1, p. 3-15, 2002.

LEMKE, T. L. Antiparasitic drugs. *Foye's principles of medicinal chemistry*, 7. ed. New York: Lippincott Williams & Wilkins, 2012.

LEYDESDORFF, Loet; PARK, Han Woo; WAGNER, Caroline. International coauthorship relations in the social sciences citation index: is internationalization leading the network? *Journal of the Association for Information Science and Technology*, abr. 2014. doi:10.1002/asi.23102.

LEYDESDORFF, Loet et al. Animating the development of social networks over time using a dynamic extension of multidimensional scaling, 2008. *arXiv:0809.4655 [physics]*, setembro. <http://arxiv.org/abs/0809.4655>.

LEYDESDORFF, Loet; WAGNER, Caroline. International collaboration in science and the formation of a core group, 2009 *arXiv:0911.1438 [physics]*, novembro. <http://arxiv.org/abs/0911.1438>.

LIANG, Liming et al. Age structures of scientific collaboration in chinese computer science. *Scientometrics*, v. 52, n. 3, p. 471-486, 2001. doi:10.1023/A:1014252017971.

LUNDBERG, Jonas et al. Collaboration uncovered: exploring the adequacy of measuring university-industry collaboration through co-authorship and funding. *Scientometrics*, v. 69, n. 3, p. 575-589, 2006. doi:10.1007/s11192-006-0170-5.

MÄHLCK, Paula; PERSSON, Olle. Socio-bibliometric mapping of intra-departmental networks. *Scientometrics*, v. 49, n. 1, p. 81-91, 2000. doi:10.1023/A:1005661208810.

MAIA, Maria de Fátima S.; CAREGNATO, Sônia Elisa. Coautoria como indicador de redes de colaboração científica. *Perspectivas em Ciência da Informação*, v. 13, n. 2, p. 18-31, ago. 2008.

MATHEUS, Renato F.; SILVA, Antonio Braz de Oliveira e. Análise de redes sociais como método para a Ciência da Informação, *Perspectivas em Ciência da Informação*, v. 7, n. 2, 2006.

MELIN, G.; PERSSON, O. Studying research collaboration using co-authorships. *Scientometrics*, v. 36, n. 3, p. 363-377, 1996. doi:10.1007/BF02129600.

MOREL, Carlos M. et al. Health innovation networks to help developing countries address neglected diseases. *Science*, v. 309, n. 5733, p. 401-404, 2005. doi:10.1126/science.1115538.

MOREL, Carlos M. et al.. Co-authorship network analysis: a powerful tool for strategic planning of research, development and capacity building programs on neglected diseases. *PLoS Negl Trop Dis*, v. 3, n. 8, p. 501, 2007. doi:10.1371/journal.pntd.0000501.

MORENO, Jacob Levy. 1953. *Who Shall Survive?: Foundations of Sociometry, Group Psychotherapy and Sociodrama*. Beacon House.

MOURA, Ana Maria Mielniczuk de; CAREGNATO, Sonia Elisa. Co-autoria em artigos e patentes: um estudo da interação entre a produção científica e tecnológica. *Perspectivas em Ciência da Informação*, v. 16, n. 2, p. 153-167, 2011.

MUELLER, Silva P. M. *Métodos para pesquisa em Ciência da Informação*. Brasília: Thesaurus, 2007.

NAGPAUL, P. S. Visualizing cooperation networks of elite institutions in India. *Scientometrics*, v. 54, n. 2, p. 213-228, 2002. doi:10.1023/A:1016036711279.

NATIONAL RESEARCH COUNCIL. *Network Science*, 2005. <http://www.nap.edu/catalog/11516/network-science>.

NEWMAN, M. E. J. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, v. 98, n. 2, p. 404-409, 2001.

NEWMAN, M. E. J. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences*, 101 (suppl\_1), 5200-5205, 2004. doi:10.1073/pnas.0307545100.

NEWMAN, M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, v. 103, n. 23, p. 8577-8582, 2006. doi:10.1073/pnas.0601602103.

NEWMAN, M. E. J. *Networks: an introduction*. Oxford, Oxford University Press, 2010.

NEWMAN, Mark E. J.; GIRVAN, Michelle. Finding and evaluating community structure in networks. *Physical Review e*, v. 69, n. 2, p. 0261, 2004.

NONAKA, Ikujiro; TAKEUCHI, H. *The knowledge-creating company: how japanese companies create the dynamics of innovation*. Oxford, Oxford University Press, 1995.

NOOY, Wouter de; MRVAR, Andrej; BATAGELJ, Vladimir. *Exploratory social network*

*analysis with Pajek*. New York: Cambridge University Press, 2011.

OLIVEIRA, Nivaldo; SOUZA, Donizeti Leandro de; CASTRO, Cleber Carvalho de. Análise sociométrica da rede de relacionamento das bibliotecas que constituem o consórcio das universidades federais do sul-sudeste de Minas Gerais, *Perspectivas em Ciência da Informação*, v. 19, n. 1, p. 130-148, mar. 2014.

OLIVEIRA, Sandra Cristina de et al. Inferência estatística clássica para a confiabilidade de rede de coautoria com enfoque nos vértices, v. 19, n. 4, p. 202-225, out. 2014.

ONNELA, Jukka-Pekka, et al. 2011. “Geographic Constraints on Social Network Groups”. *PLoS ONE* 6 (4): e16939. doi:10.1371/journal.pone.0016939

PACKER, Abel L. et al. *SciELO – 15 anos de acesso aberto: um estudo analítico sobre acesso aberto e comunicação científica*. Unesco, 2014. <http://scielo.org/php/level.php?lang=pt&component=56&item=61>.

PERSSON, Olle; GLÄNZEL, Wolfgang; DANELL, Rickard. Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, v. 60, n. 3, p. 421-432, 2004. doi:10.1023/B:SCIE.0000034384.35498.7d.

PINTO, Adilson Luiz; IGAMI, Mery P. Zamudio; BRESSIANI, José Carlos. Visibilidade e monitoramento científico na área nuclear e ciências relacionadas: uma perspectiva a partir da produtividade do Ipen-Cnen/Sp. *Perspectivas em Ciência da Informação*, v. 15, n. 2, p. 198-218, 2010.

ROMBACH, M. Puck et al. Core-periphery structure in networks. *Physics*, fev. 2012. <http://arxiv.org/abs/1202.2684>.

ROYAL SOCIETY. *Knowledge, networks and nations: global scientific collaboration in the 21st century*. [S.l.]: Royal Society, 2011.

SANTOS, Priscila. *Redes de patentes e publicações em vacinas para dengue e papilomavírus*



*humano*: implicações para políticas públicas de inovação em saúde. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2012. <http://www.escavador.com/pessoas/377056>.

SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, v. 1, n. 1, 2008. <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235>.

SEGLEN, Per O.; AKSNES, Dag W. Scientific productivity and group size: a bibliometric analysis of norwegian microbiological research. *Scientometrics*, v. 49, n. 1, p. 125-143, 2000. doi:10.1023/A:1005665309719.

SILVA, Alzira K. A. da; BARBOSA, Ricardo Rodrigues; DUARTE, Emeide Nóbrega. Rede social de coautoria em Ciência da Informação: estudo sobre a área temática de "organização e representação do conhecimento", *Perspectivas em Ciência da Informação*, v. 2, n. 22, p. 63-79, maio 2012.

SILVA, Antonio B. de Oliveira e; FERREIRA, Marta A. T. Gestão do conhecimento e capital social: as redes e sua importância para as empresas. *Informação & Informação*, v. 12 (1 esp), 2007. doi:10.5433/1981-8920.2007v12n1espp.

SILVA, Edna L. da; PINHEIRO, Liliane Vieira; REINHEIMER, Frederico Maragno. Redes de conhecimento em artigos de comunicação científica: estudo baseado em citações bibliográficas de artigos de periódicos na área de Ciência da Informação no Brasil, *Perspectivas em Ciência da Informação*, v. 23, n. 1, p. 145-160, jan. 2013.

SONNENWALD, Diane H. Scientific collaboration. *Annual Review of Information Science and Technology*, v. 41, n. 1, p. 643-681, 2007. doi:10.1002/aris.2007.1440410121.

STEVENS, Philip. Diseases of poverty and the 10/90 gap. *New England Journal of Medicine*, v. 359, n. 14, p. 1530-1531, 2008. doi:10.1056/NEJMbkev0804694.

TIEMERSMA, E. W. et al. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. *PLoS One*, v. 6, n. 4, e17601, 2011.

TOMAÉL, Maria Inês; MARTELETO, Regina Maria. Redes sociais de dois modos: aspectos conceituais Tr-ansinformação, Campinas, v. 25, n. 3, p. 245-253, set. 2013.

VALENTE, Thomas W. *Social networks and health: models, methods, and applications*. Oxford; New York: Oxford University Press, 2010.

VANZ, Samile A. de Souza; STUMPF, Ida Regina C. Colaboração científica: revisão teórico-conceitual, v. 15, n. 2, p. 42-55, 2010.

VASCONCELLOS, Alexandre Guimarães; MOREL, Carlos Medicis. Enabling policy planning and innovation management through patent information and co-authorship network analyses: a study of tuberculosis in Brazil. *PLoS ONE*, v. 7, n. 10, p. e45569, 2012. doi:10.1371/journal.pone.0045569.

VILAN FILHO, Jayme Leiro; SOUZA, Held Barbosa de; MUELLER, Suzana. Artigos de periódicos científicos das áreas de informação no Brasil: evolução da produção e da autoria múltipla, *Perspectivas em Ciência da Informação*, v. 13, n. 2, p. 2-17, 2008.

WANG, Yan et al. Scientific collaboration in China as reflected in co-authorship. *Scientometrics*, v. 62, n. 2, p. 183198, 2005. doi:10.1007/s11192-005-0013-9.

WASSERMAN, Stanley; FAUST, Katherine. *Social network analysis: methods and applications*. Boston, Cambridge University Press, 1994.

WATTS, Duncan J.; STROGATZ, Steven H. Collective dynamics of "small-world" networks". *Nature*, v. 393, p. 440-442, jun. 1998. doi:10.1038/30918.

WHITE, J. G., E. et al. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 314, n. 1165, p. 1-340, 1986. doi:10.1098/rstb.1986.0056.

WORLD HEALTH ORGANIZATION (WHO). Accelerating work to overcome the global impact of neglected tropical diseases: a roadmap for implementation [Geneva]: Who , 2012.

WORLD HEALTH ORGANIZATION (WHO). Global Tuberculosis Report 2013. [Geneva]: Who, 2013

WORLD HEALTH ORGANIZATION (WHO). Work to overcome the global impact of neglected tropical diseases, [Geneva]: Who, 2014.

WUCHTY, Stefan; JONES, Benjamin F.; UZZI, Brian. The increasing dominance of teams in production of knowledge. *Science*, v. 316, n. 5827, p. 1036-1039, 2007. doi:10.1126/science.1136099.

YAN, Erjia; DING, Ying. Scholarly network similarities: how bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and cword networks relate to each other. *Journal of the American Society for Information Science and Technology*, v. 63, n. 7, p. 1313-1326, 2012. doi:10.1002/asi.22680.

YAN, Erjia; DING Ying; ZHU, Qinghua. Mapping library and information science in China: a coauthorship network analysis. *Scientometrics*, v. 83, n. 1, p. 115-131, 2009. doi:10.1007/s11192-009-0027-9.