

PERFIL DOS BOLSISTAS PQ DA ÁREA DE QUÍMICA BASEADO NA PLATAFORMA LATTES

Alexandre Donizeti Alves^{*a}, Horacio Hideki Yanasse^b e Nei Yoshihiro Soma^c^aLaboratório de Computação Aplicada, Instituto Nacional de Pesquisas Espaciais, Av. dos Astronautas, 1.758, Jardim da Granja, 12227-010 São José dos Campos – SP, Brasil^bInstituto de Ciência e Tecnologia, Universidade Federal de São Paulo, Rua Talim, 330, Vila Nair, 12231-280 São José dos Campos – SP, Brasil^cDivisão de Ciência da Computação, Instituto Tecnológico da Aeronáutica, Praça Marechal Eduardo Gomes, 50, Vila das Acácias, 12228-900 São José dos Campos – SP, Brasil

Recebido em 17/07/2013; aceito em 18/11/2013; publicado na web em 03/02/2014

PROFILE OF CNPq PRODUCTIVITY RESEARCH FELLOWS IN THE CHEMISTRY AREA BASED ON THE LATTES PLATFORM. The Lattes Platform, an information system maintained by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), is the main source of information on Brazilian researchers. This paper presents a scientific output profile of the CNPq Productivity Research Fellows in the Chemistry area based on the information extracted automatically from Lattes curricula in the 2002-2011 period using the language “LattesMiner”. This paper also provides a comparison with the results of Santos *et al.* (2010). The findings confirmed that the majority of the researchers are male (67.9%), classified as category 2 (63.2%) and working in the Southeast region (60.7%).

Keywords: Lattes Platform; information extraction; Chemistry.

INTRODUÇÃO

A Plataforma Lattes (PL) é um sistema de informação mantido pelo CNPq para gerenciar informações relacionadas a pesquisadores e instituições no Brasil. A PL é considerada um banco de dados completo e altamente qualificado.¹ A qualidade e a fidelidade das informações contidas na PL são monitoradas pela própria comunidade científica, uma vez que os currículos de cada profissional são consultados e comparados na disputa por recursos das agências de fomento. Portanto, a PL é hoje uma das principais fontes de informações sobre pesquisadores brasileiros.

O principal componente da PL é o sistema Currículo Lattes, que é um sistema de informação curricular. Atualmente, o sistema armazena mais de dois milhões de currículos de pesquisadores, docentes, estudantes e profissionais das diversas áreas do conhecimento que atuam em ciência, tecnologia e inovação, principalmente do Brasil. Esse sistema tem, portanto, um elevado potencial para extração de informação, embora isso não seja uma tarefa simples e imediata, notadamente quando muitos pesquisadores necessitam ser considerados.

Diversos estudos foram realizados por pesquisadores das mais diversas áreas do conhecimento usando a base de dados da PL. Nestes estudos foram analisados o perfil dos bolsistas de Produtividade em Pesquisa (PQ) do CNPq em diversas áreas como Saúde Coletiva,^{2,3} Odontologia,^{4,5} Medicina,⁶⁻¹⁰ Química¹¹ e Pesquisa Operacional.¹²

Santos *et al.*¹¹ em 2010 analisaram o perfil dos 604 bolsistas PQ da área de Química com bolsa vigente em março de 2009, com base nas informações da PL. Nesse estudo, pioneiro na área de Química, é possível ter uma visualização do perfil desses bolsistas, sua distribuição geográfica, por instituições acadêmicas, por gênero, por idade científica¹³ (número de anos decorridos desde o ano de publicação do primeiro artigo em periódico indexado) e por subáreas de atuação. Também foram considerados os índices numéricos de produtividade extraídos dos currículos Lattes, tais como índice de orientação¹¹ (IO), índice H,¹³ número de artigos e somatório dos impactos.

O objetivo deste artigo é reavaliar o perfil dos bolsistas PQ do CNPq da área de Química no Brasil em 03/12/2012 de acordo com as informações contidas na PL, considerando a produção científica desses bolsistas nos últimos 10 anos (2002 a 2011). O diferencial deste estudo está no fato que todo o processo de aquisição e extração dos dados foi feito automaticamente usando a linguagem LattesMiner,¹⁴ gastando um tempo bem menor. Este estudo também permite comparar dados mais recentes com o que foi observado há alguns anos, além de considerar atributos não utilizados no trabalho anterior.

Este artigo está organizado da seguinte maneira: na Seção 2 é descrito como foram coletados os dados de interesse. Na Seção 3 são apresentados os resultados, destacando as principais informações obtidas. Finalmente, na Seção 4 são apresentadas as conclusões e sugestões para trabalhos futuros.

COLETA DE DADOS

No estudo realizado foi utilizada a relação de bolsistas PQ, com bolsas ativas em 3 de dezembro de 2012, disponível no sítio do CNPq da área de Química. Os bolsistas considerados foram apenas os listados com indicação de estarem em “Em folha de Pagamento”; os demais, por exemplo, com bolsas suspensas, não foram considerados. O número de bolsistas encontrados nesta condição foi 695.

A partir da identificação dos bolsistas, foram obtidos os seus currículos Lattes usando a linguagem LattesMiner que foi desenvolvida pelos autores do presente artigo. LattesMiner é uma linguagem de domínio específico que permite a extração automática de informações de currículos Lattes e a identificação e análise de redes sociais acadêmicas.¹⁴ É composta por um conjunto de classes escritas em Java que permite a desenvolvedores implementar suas próprias aplicações com alto nível de abstração e poder de expressão.

LattesMiner permite extrair informações de um pesquisador individual ou de um conjunto deles usando o nome ou número (ID) do pesquisador cadastrado na PL. Com as informações extraídas pode-se identificar, por exemplo, redes sociais acadêmicas, competências

*e-mail: alexdonizeti@gmail.com

regionais, perfil de grupos de diferentes áreas de pesquisa e conhecimento, etc.

Inicialmente, foi criado um arquivo texto com o nome dos bolsistas PQ da área de Química. Em seguida, a linguagem LattesMiner obteve automaticamente o número identificador (ID) de cada um deles, permitindo com isso que os seus currículos fossem baixados e armazenados como arquivos HTML. Foi feita a extração automática dos dados, sendo os mesmos armazenados, também de forma automática, em um banco de dados.

As seguintes informações foram extraídas automaticamente: dados pessoais, endereço profissional, formação acadêmica, produção bibliográfica em periódicos, índice H e número de citações na base ISI (*Institute for Scientific Information*) Web of Science, nomes em citações bibliográficas, orientações concluídas, contatos (é considerado um contato todo “link” identificado no currículo Lattes de um bolsista para outro currículo Lattes), idiomas e áreas de atuação. A linguagem LattesMiner permite extrair outras informações, mas para este estudo apenas essas foram consideradas. Todo esse processo foi realizado no dia 3 de dezembro de 2012 em menos de 3 h.

Os seguintes atributos foram considerados neste trabalho: gênero, categoria do bolsista, instituição de origem, ano de conclusão do doutorado, distribuição geográfica, artigos completos publicados em periódicos, índice H, número de citações no ISI, idiomas, tempo decorrido após conclusão do doutorado, orientações concluídas de iniciação científica, mestrado e doutorado, áreas de atuação, periódicos utilizados para publicação, classificação de periódicos segundo o Qualis Periódicos¹⁵ da CAPES de 2012, fator de impacto dos periódicos de acordo com o JCR^{®16} (Journal Citation Reports[®]) de 2011, somatório do fator de impacto dos periódicos, número de autores por artigo publicados em periódicos, relacionamentos entre os bolsistas e a população segundo o censo do IBGE¹⁷ de 2010.

RESULTADOS E DISCUSSÃO

Dos 695 bolsistas PQ da área de Química, 674 (97,0%) atualizaram o currículo Lattes em 2012, sendo que 535 (77,0%) deles nos últimos 3 meses considerando a data de coleta dos dados. 690 (99,3%) dos bolsistas declaram a área de Química como a sua principal área de atuação. Os demais bolsistas (5), 2 declaram atuar na área de Física, 1 em Bioquímica, 1 em Farmácia e o outro não informou. Dentre as mais de 200 diferentes subáreas de atuação informadas nos currículos Lattes dos bolsistas, as que mais se destacaram foram: físico-química (40,1%), química orgânica (38,8%), química analítica (30,6%) e química inorgânica (25,2%).

Na Tabela 1 é apresentada a distribuição dos bolsistas em cada uma das categorias e por gênero. Pode-se observar que a maioria é do sexo masculino (67,9%) e se concentra na categoria 2 (63,2%). Também é possível observar que há um predomínio masculino em

todas as categorias, principalmente nas categorias SR e 1A. Isso é esperado tendo em vista o predomínio masculino na titulação e na carreira universitária há algumas décadas passadas.

O tempo médio decorrido após conclusão do doutorado dos bolsistas é de 18,4 anos, variando de 3 a 54 anos e a grande maioria (45,2%) com tempo de 11 a 20 anos. Até 10 anos (146), a grande maioria (95,9%) é da categoria 2. Na faixa de 11 a 20 anos (314), 69,4% são da categoria 2 e 18,5% da categoria 1D. Na faixa de 21 a 30 anos (167), 43,7% são da categoria 2 e 16,8% da categoria 1C. Na faixa de 31 a 40 anos (52), 26,9% são da categoria 1A e 25,0% são da categoria 1B. E entre os bolsistas com tempo maior que 40 anos (16), 50,0% são da categoria 1A e 25,0% da categoria SR.

Dos 695 bolsistas, 692 (99,6%) informaram o seu endereço profissional. A Figura 1 apresenta a distribuição geográfica deles segundo seus endereços profissionais e o número de bolsistas por milhão de habitantes (entre []). O mapa foi gerado usando a ferramenta GPS Visualizer (acessível em <http://www.gpsvisualizer.com>), sendo destacados na cor “cinza” todos os estados em que o número de bolsistas é menor do que 10. O Sudeste é a região do Brasil com o maior número de bolsistas tendo o estado de São Paulo como principal destaque, com mais de um terço dos bolsistas. Há bolsistas em quase todos os estados do país. As exceções ocorrem na região Norte, em que os estados do Acre, Amapá, Roraima e Tocantins não possuem bolsistas e na região Centro-Oeste, em que o estado de Mato Grosso não possui bolsista.

Analisando o número de bolsistas por milhão de habitantes, o grande destaque é o estado do Rio Grande do Sul (6,8) que possui

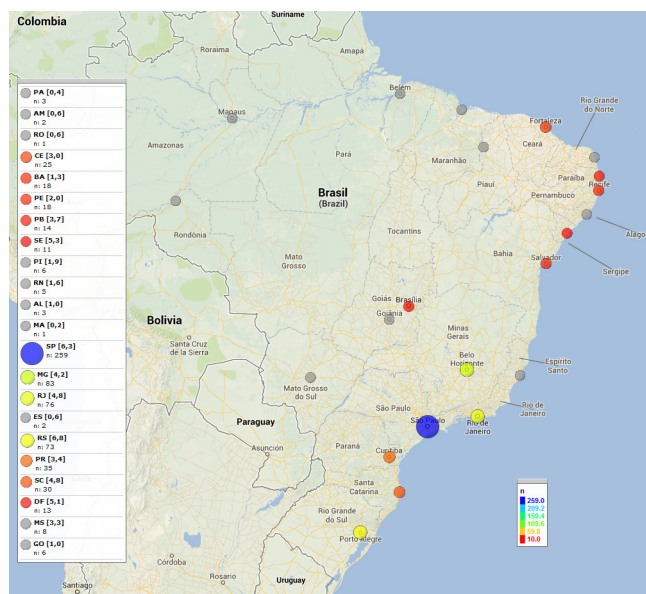


Figura 1. Distribuição geográfica dos bolsistas PQ da área de Química

Tabela 1. Distribuição dos bolsistas PQ da área de Química segundo categoria e gênero

Categoria	n	%	Tempo médio decorrido após o doutorado (anos)	Masculino	%	Feminino	%
SR	7	1,0	43,4	6	85,7	1	14,3
1A	45	6,5	31,4	42	93,3	3	6,7
1B	46	6,6	27,8	38	82,6	8	17,4
1C	63	9,0	23,3	45	71,4	18	28,6
1D	93	13,4	19,6	66	71,0	27	29,0
2	439	63,2	14,7	273	62,2	166	37,8
2F	2	0,3	7,5	2	100	0	0
Total	695	100	18,4	472	67,9	223	32,1

quase o dobro de bolsistas por milhão de habitantes quando comparado com a média do país que é de 3,6. Ainda na região Sul, o estado de Santa Catarina (4,8) também está acima da média. Na região Sudeste, é interessante observar que todos os estados, com exceção do Espírito Santo, estão acima da média nacional de bolsistas por milhão de habitantes. Na região Nordeste, os estados de Sergipe (5,3) e da Paraíba (3,7) também estão acima da média nacional. Na região Centro-Oeste, o Distrito Federal (5,1) também se encontra nessa condição.

Os bolsistas da categoria SR estão concentrados na região Sudeste, sendo que apenas 1 (14,3%) do total de 7 está na região Sul. Os bolsistas da categoria 1A estão um pouco mais distribuídos. Apesar da grande maioria (71,1%) estar na região Sudeste, também há bolsistas nas regiões Sul (17,8%) e Nordeste (11,1%). Os bolsistas da categoria 1B também se concentram na região Sudeste (76,1%). O mesmo ocorre com os bolsistas da categoria 1C (69,8%) e com os bolsistas da categoria 1D (60,2%). Já os bolsistas da categoria 2 estão distribuídos em 20 estados mais o Distrito Federal e a maioria (56,3%) também se encontra na região Sudeste. Os bolsistas da categoria 2F estão na região Nordeste (50,0%) e Sul (50,0%).

No endereço profissional os bolsistas também informam a instituição em que trabalham. São 77 instituições diferentes sendo que as 5 mais informadas são: Universidade de São Paulo (104), Universidade Estadual de Campinas (56), Universidade Federal de Minas Gerais (43), Universidade Federal do Rio de Janeiro (42) e Universidade Estadual Paulista Júlio de Mesquita Filho (40). Todas essas instituições estão na região Sudeste e representam pouco mais de 40% dos bolsistas PQ da área de Química. Na região Sul, a Universidade Federal do Rio Grande do Sul está na sexta posição com 37 bolsistas. Na região Nordeste, a Universidade Federal do Ceará é a instituição melhor colocada ocupando a nona posição com 23 bolsistas. Na região Centro-Oeste, a instituição melhor colocada é a Universidade de Brasília que ocupa a décima quarta posição com 13 bolsistas. E na região Norte, a Universidade Federal do Pará é a instituição melhor colocada ocupando a trigésima sexta posição com 3 bolsistas. É interessante destacar também que mais da metade (55,6%) dos bolsistas da categoria 1A trabalha na Universidade de São Paulo ou na Universidade Estadual de Campinas.

Dos 695 bolsistas, 688 (99,0%) informaram ter conhecimento de idiomas (total de 19 diferentes), sendo que Inglês (98,8%), Espanhol (77,8%) e Francês (49,9%) foram os três principais. Dos 687 bolsistas que declararam ter conhecimento no idioma “Inglês”, 358 (52,1%) declararam ler, escrever, falar e compreender bem. Desse total, 203 (56,7%) são da categoria 2 e 32 (71,1%) dos 45 bolsistas da categoria 1A também declaram ter esse conhecimento.

O número de artigos publicados em periódicos no período de 2002 a 2011 foi de 32.873 artigos, conforme destacado na Tabela 2. Dividindo esse período em dois períodos de 5 anos, percebe-se que houve um aumento de 23,9% no número de artigos publicados em periódicos entre o período de 2002 a 2006 e o período de 2007 a 2011. A média de artigos por ano de cada um dos 695 bolsistas foi de 4,7.

Do total de publicações em periódicos no período de 2002 a 2011, 31.876 (97,0%) artigos possuíam ISSN. A Tabela 3 apresenta a lista dos periódicos com mais de 200 artigos no período de 2002 a 2011 mais utilizados para publicação nesse período, obtidos a partir dos ISSN indicados. Nesse período, os bolsistas publicaram em 1.979 periódicos diferentes. Assim também foi possível obter o estrato do periódico de acordo com o Qualis/CAPES de 2012, considerando a área de avaliação da “Química”. Também foi possível obter o fator de impacto dos periódicos correspondentes no JCR® de 2011.

É possível perceber que há uma prevalência de publicação em periódicos nacionais, que ocupam as duas primeiras posições da lista. O primeiro periódico da lista é “Química Nova”, que teve em média

Tabela 2. Número de artigos publicados em periódicos pelos bolsistas PQ da área de Química no período de 2002 a 2011

Período	Artigos	PQ	Artigos / PQ
2002	2.608	585	4,5
2003	2.493	596	4,2
2004	2.797	633	4,4
2005	3.310	648	5,1
2006	3.474	666	5,2
2002 a 2006	14.682	692	21,2
2007	3.476	667	5,2
2008	3.683	670	5,5
2009	3.645	664	5,5
2010	3.756	664	5,7
2011	3.631	659	5,5
2007 a 2011	18.191	695	26,2
2002 a 2011	32.873	695	47,3

187,2 artigos publicados nele por ano. Dos 695 bolsistas, 489 (70,4%) publicaram pelo menos uma vez nesse periódico no período de 2002 a 2011. E o segundo periódico da lista é o “Journal of the Brazilian Chemical Society”, com média de 180,3 artigos publicados nele por ano. Dos 695 bolsistas, 498 (71,7%) publicaram nesse periódico. Dos periódicos da lista, apenas 2 (Revista Brasileira de Farmacognosia e Eclética Química) não possuem fator de impacto no JCR® de 2011, sendo que o primeiro fazia parte do JCR® quando a maioria dos artigos considerados foi publicado.

Vale salientar que no total informado de artigos dos bolsistas, alguns desses são contabilizados mais de uma vez devido a coautorias, sendo esta a forma correta e usual quando se considera a produção científica de pesquisadores de mais de uma instituição.¹⁸

A média de autores nos artigos publicados no periódico “Química Nova” foi de 4,5 e no periódico “Journal of the Brazilian Chemical Society” foi de 5,3. Entre os periódicos com mais de 200 artigos no período de 2002 a 2011, o que tem a maior média de autores (8,0) é o “Bioorganic & Medicinal Chemistry”. Existe uma diferença nas médias de autores por artigo publicado em periódicos diferentes. A razão destas diferenças mereceria uma análise mais profunda por especialistas da área.

A Tabela 4 apresenta alguns índices numéricos de produtividade dos bolsistas no período de 2002 a 2011 de acordo com as suas categorias e pelo tempo decorrido após a conclusão do doutorado, sendo que para cada índice é apresentado o valor mínimo, médio e máximo. É importante observar que quando se consideram valores médios por categoria é necessário levar em conta que há uma quantidade muito maior de bolsistas da categoria 2 em comparação com as demais.

Analisando o número de artigos publicados em periódicos no período de 2002 a 2011, percebe-se que os bolsistas da categoria 1A foram os que mais publicaram nesse período, sendo que um desses bolsistas publicou 332 artigos, com uma média de 33,2 artigos por ano e com tempo decorrido após a conclusão do doutorado de 21 a 30 anos. Os bolsistas com tempo decorrido após a conclusão do doutorado até 10 anos publicaram 4.587 (14,0%) artigos, com média de 31,4 artigos por bolsista nesse período. Os bolsistas na faixa de 11 a 20 anos publicaram 13.917 (42,3%) artigos, com média de 44,3. Na faixa de 21 a 30 anos, os bolsistas publicaram 9.838 (29,9%) artigos, com média de 58,9. Na faixa de 31 a 40 anos foram publicados 3.225 (9,8%) artigos, com média de 62,0. E os bolsistas com tempo decorrido após a conclusão do doutorado acima de 40 anos publicaram

Tabela 3. Distribuição dos periódicos mais utilizados para publicação pelos bolsistas PQ da área de Química no período de 2002 a 2011

ISSN	Periódico	n	PQ	n / PQ	Qualis	FI
0100-4042	Química Nova	1.872	489	3,8	B2	0,763
0103-5053	Journal of the Brazilian Chemical Society	1.803	498	3,6	A2	1,434
0039-9140	Talanta (Oxford)	610	168	3,6	A2	3,794
0040-4039	Tetrahedron Letters	534	117	4,6	B1	2,683
1600-5368	Acta Crystallographica. Section E	361	69	5,2	B5	0,347
0102-695X	Revista Brasileira de Farmacognosia	347	87	4,0	B5	-
0003-2670	Analytica Chimica Acta	318	127	2,5	A1	4,555
0013-4686	Electrochimica Acta	313	89	3,5	A2	3,832
0584-8547	Spectrochimica Acta. Part B, Atomic Spectroscopy	285	56	5,1	A2	2,876
0022-2860	Journal of Molecular Structure	279	107	2,6	B2	1,634
0026-265X	Microchemical Journal	256	90	2,8	B1	3,048
0021-9797	Journal of Colloid and Interface Science	254	97	2,6	A2	3,070
0100-4670	Eclética Química (Unesp)	251	95	2,6	B5	-
0021-9673	Journal of Chromatography	251	77	3,3	A1	4,531
1040-0397	Electroanalysis (New York, N.Y.)	246	69	3,6	B1	2,872
0277-5387	Polyhedron	237	100	2,4	B1	2,057
1089-5639	The Journal of Physical Chemistry. A	234	79	3,0	B1	2,946
0020-1693	Inorganica Chimica Acta	233	92	2,5	B2	1,846
0926-860X	Applied Catalysis. A, General	229	74	3,1	A2	3,903
1388-6150	Journal of Thermal Analysis and Calorimetry	221	56	3,9	B2	1,604
0009-2614	Chemical Physics Letters	221	75	2,9	B1	2,337
1381-1169	Journal of Molecular Catalysis. A, Chemical	212	82	2,6	B1	2,947
0021-8995	Journal of Applied Polymer Science	205	81	2,5	B3	1,289
0968-0896	Bioorganic & Medicinal Chemistry	201	79	2,5	B1	2,921

n: número de artigos publicados no periódico; PQ: número de bolsistas PQ da área de Química; Qualis: estrato no Qualis Periódicos da CAPES de 2012 considerando a área de avaliação da "Química"; FI: fator de impacto do JCR® de 2011.

1.306 (4,0%) artigos, com média de 81,6 artigos por bolsista. É interessante observar que esses bolsistas, proporcionalmente, foram os que mais publicaram.

A média de autores por artigo em publicações em periódicos no período de 2002 a 2011 é 5,3. A maior média foi de 10,4 autores por artigo de um bolsista da categoria 2 e com tempo decorrido após a conclusão do doutorado até 10 anos. A menor média foi de 2,2, também de um bolsista da categoria 2. É interessante observar que a média de autores por artigo diminui à medida que o tempo decorrido após a conclusão do doutorado desses aumenta. Dos 32.873 artigos publicados em periódicos nesse período, em 8.352 (25,4%) tinham como primeiro autor um dos bolsistas e em 10.191 (31,0%) deles um bolsista como último autor.

Todos os 695 bolsistas possuem pelo menos um artigo publicado em periódico indexado no JCR® no período de 2002 a 2011, sendo possível calcular o somatório do fator de impacto de todos os periódicos que esses bolsistas publicaram. Dos 1.979 periódicos utilizados, 1.225 (61,9%) possuem fator de impacto no JCR® de 2011. Do total de 32.873 publicações nesses periódicos, 27.697 (84,3%) possuem fator de impacto. O maior valor encontrado foi de um bolsista da categoria 1A, com somatório de 1.143,2 com o fator de impacto de suas publicações variando de 0,493 a 40,197. O menor valor foi de 5,0 de um bolsista da categoria 2 e a média geral foi de 98,6. Exceto pela categoria SR, observa-se que o valor médio diminui significativamente de uma categoria para outra, com os maiores valores nas principais categorias. Observa-se como esperado que quanto maior o tempo decorrido após a conclusão do doutorado, maior o valor médio do somatório do fator de impacto.

Dividindo o somatório do fator de impacto dos periódicos pelo número de artigos publicados no período de 2002 a 2011, é possível calcular o fator de impacto por artigo publicado do bolsista. É

interessante observar que os bolsistas da categoria 2 publicam em periódicos em que a média do fator de impacto desses periódicos é maior que a média dos bolsistas da categoria SR, sendo que o mesmo ocorre com os bolsistas da categoria 1C também em relação aos da categoria 1B. É curioso observar que quanto maior o tempo decorrido após a conclusão do doutorado, menor o valor médio do fator de impacto por artigo publicado do bolsista. Isso dá um indicativo que os bolsistas com menor tempo de doutoramento estão direcionando suas publicações para periódicos com maior fator de impacto, possivelmente, motivados pela classificação desses no Qualis/CAPES da área de avaliação da "Química". Todavia, é imperativo notar que uma simples comparação do valor médio do fator de impacto com o tempo decorrido após a conclusão do doutorado é insuficiente para inferir o aumento ou o declínio na produtividade da publicação de artigos em periódicos indexados no JCR®, uma vez que são considerados períodos de gerações distintas.

Em relação ao número de citações no ISI dos artigos dos bolsistas no período de 2002 a 2011, percebe-se que há uma grande variação. O maior número de citações no ISI foi de um bolsista da categoria 1A, com 13.368 citações. O menor valor foi de um bolsista da categoria 1C, com 11 citações. Dos 695 bolsistas, 680 (97,8%) informaram o número de citações no ISI em seus currículos Lattes, com valor médio de 821,9 nesse período. Considerando o tempo decorrido após a conclusão do doutorado, o valor médio aumenta significativamente em cada faixa. É interessante observar que o maior valor ocorreu na faixa de 21 a 30 anos, tendo ocorrido o mesmo em relação ao somatório do fator de impacto.

Dividindo o total de citações pelo número de artigos publicados é possível calcular o número de citações no ISI por artigo de cada bolsista. Percebe-se que o número médio de citações por artigo está fortemente correlacionado com a categoria do bolsista, pois os valores

Tabela 4. Índices numéricos de produtividade dos bolsistas PQ da área de Química por categoria e por tempo de titulação do doutorado no período de 2002 a 2011

Categorias		SR	1A	1B	1C	1D	2	2F
Número de artigos publicados em periódicos	Min	29	31	23	24	18	9	9
	Média	78,1	99,1	80,8	65,4	49,2	35,1	20,5
	Max	179	332	174	187	173	184	32
Média de autores de artigos publicados em periódicos	Min	3,6	3,0	3,0	3,3	2,6	2,2	5,3
	Média	4,8	4,8	5,2	5,2	5,1	5,4	5,8
	Max	6,6	7,5	7,7	7,8	7,5	10,4	6,2
Somatório do fator de impacto dos periódicos	Min	22,5	32,4	37,1	57,6	13,1	5,0	17,0
	Média	110,1	251,3	170,1	141,3	100,9	68,9	39,6
	Max	233,6	1.143,2	335,5	296,3	230,6	284,7	62,3
Fator de impacto por artigo publicado em periódico	Min	0,5	1,0	1,2	1,1	0,1	0,3	1,9
	Média	1,4	2,5	2,1	2,2	2,0	2,0	1,9
	Max	3,2	4,7	3,5	3,6	3,5	6,1	1,9
Número de citações no ISI	Min	763	479	447	11	207	22	48
	Média	2.299,0	3.523,4	1.584,8	1.140,3	855,2	397,6	369,5
	Max	4.362	13.368	3.346	2.389	5.933	1.964	691
Número de citações no ISI por artigo publicado em periódico	Min	4,5	5,4	4,4	0,2	3,4	1,0	5,3
	Média	44,9	40,3	21,6	20,2	19,5	12,1	13,5
	Max	123,0	121,7	45,2	60,3	109,9	64,1	21,6
Índice H	Min	13	13	14	1	9	4	4
	Média	22,7	29,3	21,6	18,4	15,7	11,1	9,5
	Max	33	51	33	28	28	26	15
Índice de Orientação (IO)	Min	7,0	13,5	10,5	12,0	5,3	0,5	5,0
	Média	30,9	52,6	48,5	40,8	33,7	21,6	6,5
	Max	57,5	124,5	96,0	121,0	80,8	101,3	8,0
Tempo decorrido após conclusão do doutorado (anos)		0 a 10	11 a 20	21 a 30	31 a 40	> 40		
Número de artigos publicados em periódicos	Min	9	10	9	27	23		
	Média	31,4	44,3	58,9	62,0	81,6		
	Max	97	184	332	135	225		
Média de autores de artigos publicados em periódicos	Min	2,2	2,2	2,9	2,8	3,3		
	Média	5,5	5,3	5,2	4,8	4,5		
	Max	10,4	8,6	8,7	7,7	6,6		
Somatório do fator de impacto dos periódicos	Min	17,0	5,0	19,5	23,1	22,5		
	Média	71,6	91,7	121,1	124,6	162,2		
	Max	284,7	643,3	1.143,2	289,5	480,6		
Fator de impacto por artigo publicado em periódico	Min	0,7	0,1	0,7	0,5	0,6		
	Média	2,3	2,1	2,1	2,0	2,0		
	Max	6,1	4,4	4,7	3,6	3,3		
Número de citações no ISI	Min	45	11	59	273	251		
	Média	363,0	659,7	1.153,1	1.578,7	2.593,4		
	Max	1.763	5.933	13.368	5.360	8.022		
Número de citações no ISI por artigo publicado em periódico	Min	1,2	0,2	1,9	5,4	4,5		
	Média	11,9	15,0	19,3	26,8	32,2		
	Max	37,7	109,9	121,7	123,0	86,3		
Índice H	Min	4	1	5	9	8		
	Média	10,9	13,3	16,9	19,4	24,0		
	Max	26	37	51	37	50		
Índice de Orientação (IO)	Min	0,5	3,8	5,3	7,5	7,0		
	Média	10,9	29,4	39,5	39,8	33,0		
	Max	40,5	121,0	124,5	96,0	57,0		

são maiores nas principais categorias. Considerando o tempo decorrido após a conclusão do doutorado, percebe-se também que quanto maior o tempo de doutorado maior o número médio de citações por artigo. Analisando estes dados com os de valor médio dos fatores de impacto dos periódicos em que o bolsista publica temos uma indicação interessante: aparentemente, a maioria das citações de artigos publicados pelos bolsistas com mais anos de doutorado não vem de artigos publicados em periódicos com fatores de impacto elevados.

O índice H foi informado em 676 (97,3%) dos currículos Lattes dos bolsistas, com valor médio de 14,3. O maior valor informado foi de um bolsista da categoria 1A, com índice H de 51. O menor índice H informado foi 1 de um bolsista da categoria 1C. Percebe-se que o maior valor médio também ocorre na faixa de 21 a 30 anos.

Outro atributo considerado neste estudo foi o número de orientações concluídas de iniciação científica, mestrado e doutorado. Dos 695 bolsistas, 694 (99,9%) informaram as orientações concluídas no período de 2002 a 2011. Na modalidade iniciação científica, 648 (93,4%) bolsistas orientaram em média 12,1 alunos nesse período. Na modalidade mestrado, 646 (93,1%) orientaram em média 6,3 alunos e na modalidade doutorado, 477 (68,7%) orientaram 5,8 alunos nesse período.

Outra forma de analisar as orientações é usando o IO.¹¹ O IO é calculado pelo somatório do número de alunos de cada modalidade de orientação multiplicado por seus respectivos pesos: iniciação científica (0,5), mestrado (1,5) e doutorado (3,0). Às co-orientações foram atribuídos pesos na metade do valor das orientações concluídas

como orientador principal. O bolsista com maior IO no período de 2002 a 2011 é da categoria 1A, com IO de 124,5 e tempo decorrido após a conclusão do doutorado de 21 a 30 anos. A média do IO dos bolsistas nesse período foi de 28,8.

Considerando o tempo decorrido após a conclusão do doutorado, foi possível verificar que os bolsistas na faixa de 31 a 40 anos (7,5%) foram os que mais contribuíram para a formação de recursos humanos, com a média do IO de 39,8. O IO também foi significativo para os bolsistas PQ com o tempo de doutorado de 21 a 30 anos (24,0%), pois sua média do IO foi de 39,5. Apenas após os 40 anos (2,3%) de conclusão do doutorado observa-se uma queda no IO, com a média de 33,0. Isso talvez se deva ao fato que não havia muitos programas de doutorado antigamente e mais recentemente, os bolsistas nesta faixa estão se aposentando e não orientam mais. Os bolsistas na faixa de 11 a 20 anos (45,2%) possuem média de 29,4. E os bolsistas com até 10 anos (21,01%) possuem uma média bem menor (10,9). Esse resultado era de se esperar, pois recém-doutores não atuam imediatamente na pós-graduação.

A linguagem LattesMiner permite identificar os relacionamentos nas orientações tidas entre os bolsistas. A Figura 2 ilustra a rede de orientações concluídas de mestrado e doutorado entre os bolsistas de acordo com as suas categorias. Essa rede foi gerada a partir do sistema SUCUPIRA,¹⁹ que foi desenvolvido utilizando a linguagem LattesMiner. As cores dos vértices representam a categoria do bolsista, sendo que “roxo” é para a categoria SR, “azul” para a categoria 1A, “verde” para a categoria 1B, “amarelo” para 1C, “laranja” para 1D, “vermelho” para a categoria 2 e “rosa” para a categoria 2F. A rede apresenta a relação orientador-orientado por categoria, do centro para as extremidades. Estas identificam as orientações concluídas de mestrado (M) e doutorado (D) entre os bolsistas. As cores das arestas representam o número de orientações entre as categorias, indicando quão intenso é um relacionamento.

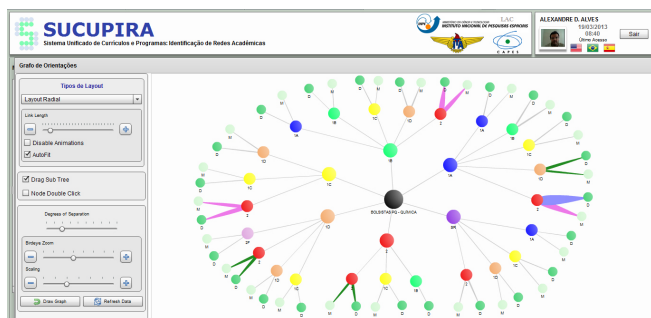


Figura 2. Rede de orientações concluídas de mestrado (M) e doutorado (D) entre os bolsistas PQ da área de Química de acordo com a categoria

Há relacionamentos nessa rede que chamam a atenção e estão destacados em “azul”, representando que o relacionamento ocorreu mais de 50 vezes; em “rosa”, representando que o relacionamento ocorreu entre 25 e 50 vezes e em “verde escuro”, representando que o relacionamento ocorreu entre 10 e 25 vezes. O principal relacionamento de orientador-orientado ocorreu entre os bolsistas da categoria 1A e os da categoria 2 nas orientações concluídas de doutorado, sendo que essa relação ocorreu 81 vezes.

A linguagem LattesMiner também permite identificar os contatos (todos os “links” identificados para outros currículos Lattes) contidos no currículo Lattes de um determinado pesquisador. Todo contato contém o ID do pesquisador, o que permite identificar os relacionamentos entre os pesquisadores. Deve-se salientar que nem todo relacionamento de um pesquisador possui “link” para outro no currículo Lattes. Estes relacionamentos não “certificados” no currículo Lattes não são contabilizados.

A Figura 3 ilustra a rede de contatos nos artigos publicados em periódicos no período de 2002 a 2011 entre os bolsistas por categoria. Ao todo foram identificados 27.328 contatos nesta condição, sendo que 5.672 (20,8%) ocorreram entre os bolsistas da categoria 2 com eles mesmos. Esse relacionamento é destacado em “azul”, representando que o relacionamento ocorreu mais de 5.000 vezes. Também são destacados os relacionamentos em “rosa”, representando que o relacionamento ocorreu entre 1.000 e 5.000 vezes e em “verde escuro”, representando que o relacionamento ocorreu entre 100 e 1.000 vezes.

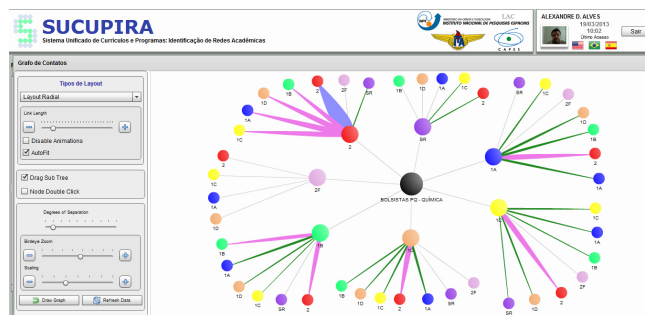


Figura 3. Rede de contatos identificados nos artigos publicados em periódicos no período de 2002 a 2011 entre os bolsistas PQ da área de Química de acordo com a categoria

Também é importante destacar que os bolsistas de todas as categorias se relacionam entre eles em praticamente todas as categorias, o que denota que a desejável cooperação acadêmica é alcançada pela área de Química.

CONCLUSÕES

Este estudo apresentou um perfil dos bolsistas PQ do CNPq da área de Química baseado em informações extraídas dos currículos Lattes de dezembro de 2012. Houve um aumento em torno de 15% no número de bolsistas desde o estudo realizado por Santos *et al.*¹¹ com os pesquisadores com bolsas ativas em 2009. A grande maioria dos bolsistas ainda é do sexo masculino, uma vez que a porcentagem continua praticamente a mesma, em torno de 68%. A porcentagem de bolsistas da categoria 1 era de 36,8 e agora está em 35,5. A porcentagem de bolsistas da categoria 2 era de 62,3 e agora está em 63,2. A região Sudeste possuía 63,7% do total de bolsistas e atualmente é de 60,7%. São Paulo continua sendo o estado com o maior número de bolsistas. Entretanto, a porcentagem diminuiu de 41,2 para 37,3. A média nacional do número de bolsistas por milhão de habitantes aumentou de 3,2 para 3,6 (aumento de 12,5%). O estado com a maior razão é o Rio Grande do Sul com 6,8 bolsistas por milhão de habitantes (o estado de São Paulo tem 6,3). Nesse ponto houve uma mudança, pois no estudo de Santos *et al.*¹¹ o estado de São Paulo possuía a maior razão com 6,0 bolsistas por milhão de habitantes e o estado do Rio Grande era o segundo (5,1).

É interessante destacar que as cinco instituições com mais bolsistas continuam sendo as mesmas e todas da região Sudeste, possuindo juntas 285 bolsistas. Porém, a porcentagem diminuiu de 45,9 para 41,0. O predomínio ainda continua sendo da Universidade de São Paulo com 15,0% dos bolsistas, tendo diminuído a porcentagem que era de 17,4. Outra questão interessante é que a Universidade de São Paulo e a Universidade Estadual de Campinas continuam com mais da metade dos bolsistas da categoria 1A.

Também é interessante destacar que houve um predomínio de publicações em dois periódicos nacionais: “Química Nova” e “Journal of the Brazilian Chemical Society”. Nesses dois periódicos, o número de artigos publicados corresponde a 36,8% das publicações

dos bolsistas considerando apenas os periódicos mais utilizados, conforme Tabela 3.

Em relação às publicações em periódicos, notamos que não houve uma queda significativa nos valores das médias quando consideramos os bolsistas por categoria, exceto em uma. No estudo realizado por Santos *et. al.*¹¹ a média de publicações por ano em cada categoria era a seguinte: SR (8,0), 1A (9,0), 1B (8,3), 1C (6,9), 1D (5,6) e 2 (4,2). A média neste estudo foi a seguinte: SR (7,8), 1A (9,9), 1B (8,1), 1C (6,5), 1D (4,9) e 2 (3,5). Comparando, percebemos que houve um aumento (10,0%) na categoria 1A e uma queda mais acentuada na categoria 2 (16,7%).

No período de 2002 a 2011, os artigos dos bolsistas foram publicados em periódicos de 149 categorias diferentes de um total de 226 constantes no JCR® de 2011. As categorias mais utilizadas foram: “Química Multidisciplinar” em 5.247 (18,9%) artigos, “Físico-Química” em 4.641 (16,8%) artigos, “Química Analítica” em 3.951 (14,3%) artigos, “Química Orgânica” em 2.718 (9,8%) artigos e “Farmacologia e Farmácia” em 1.761 (6,4%) artigos. A categoria “Química Inorgânica e Nuclear” ocupa a sétima posição com 1.727 (6,2%) artigos publicados nessa categoria. Vale lembrar que um mesmo periódico pode ser classificado em mais de uma categoria do JCR®.

É interessante observar que a distribuição dos artigos entre categorias não é igual às subáreas de atuação informadas pelos bolsistas, pois 40,1% declaram atuar em “Físico-Química”, 38,8% em “Química Orgânica”, 30,6% em “Química Analítica” e 25,2% em “Química Inorgânica”. Percebe-se, então, que os bolsistas publicam mais em “Química Analítica” do que em “Química Orgânica” apesar de declararem justamente o contrário. Isso decorre da atuação interdisciplinar dos bolsistas PQ da área de Química.

O JCR® também calcula a mediana do fator de impacto dos periódicos de uma determinada categoria, *i.e.*, subáreas de atuação, com os seguintes valores de fator de impacto: “Físico-Química” (1,997), “Química Orgânica” (2,341), “Química Analítica” (1,92) e “Química Inorgânica” (1,846).

Considerando o somatório do fator de impacto por ano, percebe-se também que houve em geral uma queda nos valores médios de acordo com a categoria dos bolsistas. No estudo realizado por Santos *et. al.*¹¹ o valor médio do somatório por ano em cada categoria era a seguinte: SR (13,5), 1A (19,8), 1B (17,1), 1C (12,8), 1D (11,1) e 2 (7,3). Neste estudo, o valor médio do somatório por ano em cada categoria é o seguinte: SR (11,0), 1A (25,1), 1B (17,0), 1C (14,1), 1D (10,1) e 2 (6,9). Houve um aumento no valor médio nas categorias 1A e 1C. O maior valor médio por ano no estudo realizado por Santos *et. al.*¹¹ era 57 e de um bolsista da categoria 1B. Neste estudo, o maior valor foi de 114,3 de um bolsista da categoria 1A.

No caso do índice H, como era de se esperar, ocorreu um aumento em todas as categorias dos bolsistas. No estudo realizado por Santos *et. al.*¹¹ o valor médio do índice H era o seguinte: SR (17,5), 1A (24), 1B (17,4), 1C (13,5) e 1D (12,3). Para a categoria 2 não foi informado o valor médio do índice H. No nosso estudo, o valor médio do índice H foi o seguinte: SR (22,7), 1A (29,3), 1B (21,6), 1C (18,4) e 1D (15,7). O maior índice H era 47 e passou a ser 51.

Em relação ao IO, verificou-se que os bolsistas com tempo decorrido após a conclusão do doutorado de 21 a 30 anos e de 31 a 40 anos foram os que mais contribuíram para a formação de recursos humanos. Segundo Santos *et. al.*¹¹, essa faixa ficava entre os 10 e 30 anos. Percebe-se que há uma mudança no perfil dos bolsistas considerando-se diferentes janelas de tempo, pois o estudo realizado por Santos *et. al.*¹¹ considerou os pesquisadores com bolsas PQ ativas em 2009.

Uma última informação que merece ser destacada é o fato de

que 34 (4,9%) dos bolsistas PQ da área de Química serem membros titulares da Academia Brasileira de Ciências, o que representa 7,6% do total de membros da Academia.

Vale ressaltar que os dados levantados para este estudo foram obtidos automaticamente usando a linguagem LattesMiner, que permite extrair as informações dos currículos Lattes e armazená-las em um banco de dados. Com a linguagem LattesMiner é possível repetir todo esse processo a qualquer momento, permitindo que seja feita uma análise temporal com pouco esforço. Isso é importante para acompanhar a evolução no tempo de uma determinada área.

Uma sugestão para trabalhos futuros seria analisar se há relação entre o número de citações de um artigo e o fator de impacto do periódico. Este último também poderia ser analisado em relação ao índice H. Outra sugestão seria analisar se há relação entre o número de artigos publicados em periódicos e o IO.

AGRADECIMENTOS

Os autores agradecem o apoio financeiro da CAPES e CNPq. Também agradecem ao Prof. Luiz Carlos Dias (Instituto de Química, Universidade Estadual de Campinas e Coordenador da Área de Química na CAPES) por discussões e contribuições ao presente manuscrito. Os autores agradecem aos assessores anônimos da revista por suas importantes contribuições para a melhoria deste manuscrito.

REFERÊNCIAS

1. Lane, J.; *Nature* **2010**, *464*, 488.
2. Barata, R. B.; Goldbaum, M.; *Cad. Saúde Pública* **2003**, *16*, 1863.
3. Santos, S. M. C.; Lima, L. S.; Martelli, D. R. B.; Martelli-Júnior, H.; *Physis* **2009**, *19*, 761.
4. Cavalcante, R. A.; Barbosa, D. R.; Bonan, P. R. F.; Pires, M. B. O.; Martelli-Júnior, H.; *Rev. bras. epidemiol.* **2008**, *11*, 106.
5. Scarpelli, A. C.; Sardenberg, F.; Goursand, D.; Paiva, S. M.; Pordeus, I. A.; *Braz. Dent. J.* **2008**, *19*, 252.
6. Martelli-Júnior, H.; Martelli, D. R. B.; Quirino, I. G.; Oliveira, M. C. L. A.; Lima, L. S.; Oliveira, E. A.; *Rev. Ass. Med. Bras.* **2010**, *56*, 478.
7. Mendes, P. H. C.; Martelli, D. R. B.; Souza, W. P.; Filho, S. Q.; Martelli-Júnior, H.; *Rev. Bras. Educ. Med.* **2010**, *34*, 535.
8. Oliveira, E. A.; Pécoits-Filho, R.; Quirino, I. G.; Oliveira, M. C.; Martelli, D. R.; Lima, L. S.; Martelli-Júnior, H.; *J. Bras. Nefrol.* **2011**, *33*, 31.
9. Oliveira, E. A.; Ribeiro, A. L. P.; Quirino, I. G.; Oliveira, M. C. L.; Martelli, D. R.; Lima, L. S.; Colosimo, E. A.; Lopes, T. J.; Simões e Silva, A. C.; Martelli-Júnior, H.; *Arq. Bras. Cardiol.* **2011**, *97*, 186.
10. Oliveira, E. A.; Colosimo, E. A.; Martelli, D. R.; Quirino, I. G.; Oliveira, M. C. L.; Lima, L. S.; Simões e Silva, A. C.; Martelli-Júnior, H.; *Scientometrics* **2012**, *90*, 429.
11. Santos, N. C. F.; Cândido, L. F. O.; Kuppens, C. L.; *Quim. Nova* **2010**, *33*, 489.
12. Alves, A. D.; Yanasse, H. H.; Soma, N. Y.; *XLIII Simpósio Brasileiro de Pesquisa Operacional*, Ubatuba, Brasil, 2011.
13. Hirsch, J. E.; *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16569.
14. Alves, A. D.; Yanasse, H. H.; Soma, N. Y.; *SPLASH/OOPSLA XI Workshop on Domain-Specific Modeling*, Portland, Estados Unidos, 2011.
15. <http://qualis.capes.gov.br/webqualis/>, acessada em Dezembro 2012.
16. <http://www.webofknowledge.com/>, acessada em Dezembro 2012.
17. <http://www.ibge.gov.br/home/>, acessada em Dezembro 2012.
18. Almeida, E. C. E. e Guimarães, J. A.; *Scientometrics* **2013**, *97*, 287.
19. Alves, A. D.; Yanasse, H. H.; Soma, N. Y.; *6th Iberian Conference on Information Systems and Technologies (CISTI)*, Chaves, Portugal, 2011.