

Original article

The contribution of undergraduate researches to scientific production at Fiocruz

DOI: 10.3395/reciis.v4i5.374pt

Alice Ferry de Moraes

Scientific and Technological
Information Center-Fiocruz, Rio de
Janeiro, Brazil
ferry@icict.fiocruz.br

Abstract

This article reports the results of a study on research projects presented by undergraduate research scholarship recipients from the Institutional Programs of Scientific Initiation Scholarships (PIBIC) at FIOCRUZ during the Yearly Undergraduate Research Symposiums (RAIC) between 1993 and 2007. The method involved cross-searching the names of grantees and used as basic working tools the proceedings of such symposiums and the lists of scholarships granted by both PIBIC and CNPq (National Council of Scientific and Technological Development). The grantees' names were standardized so that a search could be made on the Lattes Platform about whether or not such students had met the objectives of PIBIC, i.e. if they pursued graduate studies after completing their undergraduate degree. Moreover, the names of the universities and the areas of study undertaken by the grantees were gathered and standardized. The influence of FIOCRUZ was observed in the choice for the graduate program the grantees decided to enter and the research activities they conducted afterwards.

Keywords

science initiation – scientific production – information management – PIBIC - FIOCRUZ

1. Introduction

This study was developed at the Scientific and Technological Information Center (ICICT) from the Oswaldo Cruz Foundation (FIOCRUZ). It aimed to collect information about the names of scientific initiation scholarship recipients (hence forth grantees) who conducted research activities at FIOCRUZ between 1993 and 2007, and the results they presented at the Yearly Undergraduate Research Symposiums (hence forth RAICs) held by FIOCRUZ.

The study was centered around collecting and feeding the names of the grantees into a FIOCRUZ database named *Produtores Científicos* ('Science Producers'), developed by

the Research Laboratory for Research in Science, Technology and Innovation in Public Health (LabCiTiES).

The occurrence of the grantees' names in the aforementioned database revealed that even after their scholarships had finished, their names appeared in research presented in other events held by FIOCRUZ, particularly the symposiums geared towards graduate studies. In other words: undergraduate students, a lot of whom were former grantees from the Institutional Programs of Scientific Initiation Scholarships (henceforth PIBIC) of the National Council of Scientific and Technological Development (CNPq) continued at FIOCRUZ after their scholarships had ended, especially by entering graduate diploma, Master's and doctoral programs.

This study was conducted with the aid of the *Programa Pró-Gestão* (a program to support and foster the institutional development of science and technology), created as a result of an agreement between FIOCRUZ and FAPERJ (Research Funding Agency of the State of Rio de Janeiro), which granted two scholarships to recent graduates.

The first step of this study was the search for copies of RAIC proceedings. The names of the students who participated in each RAIC were extracted from the tables of contents. However, some details caused concern. It was observed that the names found on the inside pages of the proceedings in the abstracts of the presented studies did not always coincide with the names found in the tables of contents. A detailed, page-by-page check was made in each of the fifteen proceedings. Also, there were names of RAIC participants who were not graduate students at the time, according to the information in their curriculum in the Lattes Platform. Most of them were graduate students who worked as advisors of PIBIC grantees, and some were secondary school students who participated in the PROVOC (Scientific Vocation Program) program at FIOCRUZ.

After the initial drawbacks were overcome, this study took off.

2. The role of Fio Cruz in Brazilian education

This study, focusing on the participation of scientific initiation grantees at the RAICs, stemmed from finding the names of several grantees on the FIOCRUZ database *Pesquisadores Científicos* ('Science Producers'), which is the result of a project approved in 2003 by the Program for the Development of Research and Technological Development (PIPDT) from ICICT.

The importance of PIBIC as a provider of academic education and training is undeniable. In its turn, FIOCRUZ plays a crucial role in providing grantees with the opportunity to participate in research projects and present their work in several events held at FIOCRUZ, thus encouraging them to continue their academic career as well as offering great support to higher education in Brazil.

PIBIC, geared towards undergraduate students from all knowledge areas, aims to introduce them to the realm of scientific research and also encourage researchers-advisers to build teams and provide the institution in question with a tool for research policies [...] (MACCARIELLO et al., 1999)

FIOCRUZ is renowned for the quality of its *Lato* and *Stricto sensu* graduate programs. ICICT has recently launched a new initiative in graduate studies in Brazil. In the second half of 2009, it launched its Master's and Doctoral programs in Health Communication and Information.

In February 2006, the Anísio Teixeira National Institute for Education Studies and Research (INEP) announced that the Joaquim Venâncio Polytechnic School of Health (EPSJV) had ranked first at the National High School Examination (ENEM). EPSJV offers secondary-level technical courses in the area of health. Also at secondary level, FIOCRUZ developed the Scientific Vocation Program (PROVOC), aimed at high school students from Rio de Janeiro.

Given the acknowledged contribution of FIOCRUZ to secondary education in Brazil, this study aimed to show evidence for its contribution to the undergraduate level.

Despite its major focus on Biomedical Sciences, FIOCRUZ has significantly contributed to increased research in the field of social sciences by means of the work developed in sociology and anthropology at the *Sérgio Arouca* National School of Public Health (ENSP); in the field of applied social sciences through the work developed in communications and information science at ICICT; and the research conducted in the area of humanities with the work developed in education and history by the Oswaldo Cruz House (COC).

The contributions of FIOCRUZ reinforce its participation in the field of education and include classroom teaching and distance learning at several levels, social intervention actions to prevent diseases and promote health, and activities for the dissemination of science at *Museu da Vida* (Museum of Life).

3. Scientific initiation and research

The concern about teaching and research in higher education fostered the creation of the University of the Federal District (UFD), as conceived by Anísio Teixeira, in 1935. Since then, leading Brazilian scientists have been supportive to scientific initiation by paving the way for young students to participate in research projects at universities and research institutes such as FIOCRUZ.

As stated by Demo (2000), "research cannot be limited to the academic stereotype of a sophisticated, special activity". The author considers research to be both a scientific and educational principle and a theoretical, methodological tool to construct knowledge. Thus, participation in research is an important component of education, and the analysis of the scientific production of novice researchers can be considered a teaching activity.

In order to foster scientific and technological research and training of human resources, CNPq was created in 1951. At that time, the Student Scholarship (*Bolsa do Estudante*) started to be granted to students involved in research projects.

The University Reform Law (Law no. 5.540 of 28 November, 1968) determines (art. 2) the inseparability of teaching-research as a disciplinary norm of higher education. (MALDONADO, 1998)

The PIBIC program was created on 20 July, 1988 to grant scholarships to higher education institutions (HEI) and research institutes (RI), which should pay for a share of the scholarships. There is a program of quotas per institution, but there is also the direct grant of scholarships to researchers ('over-the-counter scholarships) through the program entitled Integrated Research Projects.

Prior to PIBIC, a small number of 'over the counter' scholarships were used. Back in 1991, only five thousand scholarships were used out of the twelve thousand which were available. (MASCUSHI, 1996). Only after 1992 the offer of PIBIC scholarships tripled and outnumbered the offer of 'over-the-counter' scholarships, which maintained the same number.

The HEIs and the RIs are supposed to hold a yearly event to present the scientific production of grantees, which is evaluated by an institutional committee.

4. Processes and steps of this research

Knowledge in the fields of library science, information science, computer science, sociology of science and education was necessary to develop actions which themselves also required patience, perseverance and a great deal of intuitive perception.

As much as possible, widely known and readily accessible technological tools were chosen to make the steps of the work easier, for example MS Excel.

Each step of the study was explained and justified to clarify the intended results, thus exercising in practice the management of scientific and technological information at FIOCRUZ. Data were collected and organized carefully to enable the possibility of analyses with different objectives in the future. In other words, data are readily available for use in other studies.

4.1 Describing how the proceedings of RAICs were formatted

The table of contents of RAIC proceedings posed an extra challenge to this research project. The organization and the format of the proceedings of every single between 1997 and 2003 were described in writing, and the following information was listed individually:

- Specifications of event: title, date and topic;
- Specifications of size: height, width and thickness of proceedings;
- Specifications of page layout: inside pages, contents pages and index pages;
- Specifications of the articles: title, author, author's name in bibliographical citation, department/laboratory/office, FIOCRUZ campus, advisor, co-advisor, title of advisor's project, area code from CNPq's knowledge area chart, major area, area, sub-area, CNPq's specialist area, abstract, keywords, address, postal code, city, state, financial aid, author's email (grantee);
- Arrangement specifications of publication: presentation in alphabetical order or not; by topics, programs, FIOCRUZ campi;
- Specifications of table of contents: alphabetical, by author's last name, by direct of author's name, by advisor, co-advisor, by single order of names; by knowledge areas.

Indexes have changed year after year and showed inconsistencies; for example, entries for the participants in RAICs (grantees and advisors) by first name or last name, whether complete or incomplete. In some proceedings, not all the names were listed in the tables of contents, hence it was necessary to write down the names found on each page to check them against the table of contents later.

Advisors and co-advisors were not always clearly identified. The list provided by CNPq revealed that only the record for advisor was entered in the tables of contents. For 'over-the-counter' grantees, the general coordinator of the project - usually a leading researcher in the field - was considered to be the advisor rather than the one who actually advised the grantee. However, actual advisors appeared in some indexes as co-advisors because of the significant contribution they gave to the grantees; they were still pursuing or had already completed their master's and doctoral degrees, and often motivated the grantees to choose their research areas.

4.2 Gathering grantees' names

The project team organized lists of grantees by checking the tables of contents and by going through every page of the volumes of the RAIC proceedings to search for names that were not recorded there.

This procedure was concomitant with a request to the PIBIC program from FIOCRUZ for an official list of grantees, which was incomplete because many researchers are granted 'over-the-counter' scholarships by means of projects approved by CNPq and forget to register them with the institutional PIBIC.

It was therefore necessary to contact CNPq and request a list of grantees between 1993 and 2007. A list was sent with all the scholarships granted, and it contained names of recipients of different types of scholarships, that is, from PIBIC grantees to visiting professors at FIOCRUZ. Two lists (by PIBIC and CNPq) were then organized in alphabetical order of names and by date so that the information could be compared to that of the RAIC proceedings.

Fortunately, because CNPq required grantees to participate in PIBIC-oriented events, there was a significant number of projects at the RAICs which were evaluated and awarded prizes by a special committee. For this reason, it could be seen that there were fewer names of grantees in the official list of PIBIC than in the table of contents of the proceedings, as shown below.

Table 1
RAIC participants by type of contract

Year	PIBIC Grantees	Over-the-counter Grantees	Total partic.
1993	66	70	136
1994	80	38	118
1995	73	33	106
1996	190	29	219
1997	171	6	177
1998	209	45	254
1999	225	35	260
2000	231	36	269
2001	150	6	156
2002	186	36	222
2003	206	54	260
2004	192	48	240
2005	207	81	288
2006	194	53	247
2007	216	92	308
Total	2596	662	3258

Source: author's data

Table 2
RAIC participants by situation

Year	New Grantees	Grantees prev yrs	Total partic.
1993	136	--	136
1994	51	67	118
1995	66	40	106
1996	172	47	219
1997	125	52	177
1998	187	67	254
1999	149	111	260
2000	143	124	267
2001	112	44	156
2002	168	54	222
2003	161	99	260
2004	165	75	240
2005	180	108	288
2006	167	80	247
2007	220	88	308
Total	2218	1040	3258

Source: author's data

4.3 Entering grantees' names into the FIOCRUZ database *Produtores Científicos*

The correct names of the grantees were inserted into the FIOCRUZ database *Produtores Científicos* after consultation to the Lattes Platform.

This database contains several spellings for the grantees' names as well as information on their participation in academic events at FIOCRUZ other than RAICs. Different proceedings emerged as a result, which was useful to search information on the scientific production at FIOCRUZ.

There were names entered under both Brazilian (ABNT) and international (Vancouver) standards. As the names of the grantees were gathered from the proceedings of academic events, several spellings and different orders were observed for their names: authors were referenced by either of the following possibilities: first name; first and last names; last name followed by initials of first and middle name(s); hyphenised compound last name determined by authors; variants of the same names according to author's marital status. The solution adopted by the aforementioned database was to record the names as they appeared in the proceedings, gathered under a main entry, determined by the *Anglo American Cataloguing Rules, 2nd Edition (AACR2)*.

The insertion of the names of the scientific initiation grantees into the database made it possible to check such

names and also identify other academic activities undertaken by the grantees and the respective areas. This procedure was particularly helpful in cases of namesakes and/or abbreviated names.

This database was installed with the MARC format in the Aleph software, available at the Library of Biomedical Sciences at FIOCRUZ. The database currently contains approximately nine thousand records, including all levels of FIOCRUZ researchers. The database was customized to be readily accessed on the Internet by users (general public and researchers) and librarians in charge of the technical processing of publications.

It can also provide information to continue the PROMAN database, which contains information on the scientific production of the early period of FIOCRUZ.

4.3.1 MARC format

The MARC format (*Machine Readable Cataloging Record*), version 21, is a product of the Library of Congress in the US, and it is used for bibliographical records. The format offers fields to describe several types of documents in different media.

It was chosen because it is a parameter for technological infrastructure, i.e. an international standard (ISO 2709/1975), which allows data exchange with any bases. MARC has a mode for authorities or authors, which enables the management of the authorship of scientific production.

The several ways to display the names of grantees and their sources of data collection were registered into the FIOCRUZ database *Produtores Científicos* under the MARC format. The fields, sub-fields and indicators of the MARC format used in the records were the following:

Field **100** was used to enter names in compliance with AACR2 standards. Indicator 1 (one) was always employed because the name to be recorded is entered by the researcher's and/or the grantee's last name.

Field **400** was used to record the other forms for the names collected in the proceedings.

Field **670** was used to record the consulted sources while field **675** was only used if the name queried were not found in the Lattes Platform.

Sub-field **c** was used for the record which refers to the grantee's advisor.

Sub-field **q** was used to indicate incomplete names, which may be completed afterwards, when the name is quoted in another source. Those incomplete names were often spelled with the initial letters of the author's full name, according to the Vancouver standards.

The indicator regarding the record by first name was always 0 (zero). This way of entering records is used by the Lattes Platform.

4.4 Queries in the Lattes Platform

The Lattes Platform, launched by CNPq in August 1999, shows the Lattes Curriculum in a standardized manner; since July 2005, a Committee for the Assessment of Lattes has conducted studies to assess, refine and improve it.

*[...] the Lattes Platform has great potential for producing knowledge that can be used to manage Science and Technology. There several possibilities to explore the stored data, which can bring insightful information on the **modus operandi** of the scientific and technological production in Brazil.* (PAULA, 2004)

Today, it is an essential tool to manage activities, define policies in the field of science and technology, and to assess scientific production. In this study, the Platform has helped to trace the academic career path of scientific initiation grantees.

It should be noted that successful retrieval of information in this Platform was directly associated with the accurate spelling and, to a large extent, the completeness of names. It no longer allows, as it did before, queries by the last name followed by the first letter of the first name, which used to result in an alphabetical list of names starting by that letter. For example, Silva, A. Successful retrieval was possible even though a long list was generated in a time-consuming process. The search scope was limited to names starting by letter A. This fact is worth of notice because many authors continue to use the Vancouver standards to present their work. Such standards prioritize the last name over the first and middle names, which are displayed as plain initials, without the use of punctuation to separate them.

As already stated, queries to the FIOCRUZ database *Produtores Científicos* became an important research tool, particularly as regards namesakes. The Lattes Platform also made it easier to identify such grantees by means of field of research and institution, thus adding relative accuracy to the retrieval of information about a grantee. There were cases where a grantee worked in a field but abandoned their academic career while their namesake developed their own brilliantly. It would be a serious mistake to associate the successful grantee with a title belonging to someone else, and hence jeopardize the reliability of this study.

It should be noted that only over the last few years registration into the Lattes Platform became mandatory, and anyway it does not always happens as desired. According to the Committee for the Assessment of Lattes, some records in

the Platform still contain incomplete and outdated personal, academic and professional data.

An increase was observed in the number of grantees' curriculums registered in the Lattes Platform over the years, as shown below:

Table 3
RAIC participants with or without Lattes Platform - overview

Year	With Lattes Plat.	Without Lattes Plat.	Total partic.
1993	55	81	136
1994	53	65	118
1995	55	51	106
1996	98	121	219
1997	83	94	177
1998	134	120	254
1999	164	96	260
2000	174	93	269
2001	119	37	156
2002	211	11	222
2003	238	22	260
2004	235	5	240
2005	283	5	288
2006	244	3	247
2007	302	6	308
Total	2448	810	3258

Source: author's data

4.5 Organization of spreadsheets

Information contained in the spreadsheets corresponding to each RAIC was obtained in the Lattes Platform from the list of grantees' names. Seven-spreadsheet Excel files were created for each year.

The first spreadsheet was named "Data" and had the following information: grantee's name, gender (male or female), year of scholarship (always related to the year of RAIC); FIOCRUZ campus where the grantee was conducting his/her activities; information on whether or not the grantee had a curriculum in the Lattes Platform.

In the first column of this spreadsheet, the grantees' names were numbered, and these numbers were used to identify the grantees in the other spreadsheets. In the "Data" spreadsheet from the RAIC held in 1993, all the names were typed in without highlights because it was the first symposium, hence there was no record of scholarships in previous years. It should be emphasized that data collection for this study started with the proceedings of RAICs, which started in 1993

at FIOCRUZ. In the other "Data" spreadsheets, the names of the PIBIC grantees were written in blue. Grantees who were not in this list had their names written in black. They were known as "over-the-counter" grantees, that is, they were hired directly by researchers according to the demands of the projects developed by the latter.

In order to identify the grantees from previous years, whose data were already known, the numbers assigned to each grantee were written in red. In every spreadsheet a choice was made to always maintain the names of the participants in each RAIC, without excluding grantees from previous years, in order to build a profile of each RAIC and record the multiple participation of some grantees.

The information from the first spreadsheet was summarized to generate tables with figures about every year in the 1993-2007 period.

The FIOCRUZ campi which hosted these grantees were the following: Biomanguinhos, CECAL, COC, DIREB, DIREH, ENSP, EPSJV, Farmanguinhos, ICC, ICICT, IFF, INCQS, IOC, IPEC, IPqAM, IPqGM, IPqLMD, IPqRR and the Presidency of FIOCRUZ. Partnerships between IOC and COC and between IOC and ENSP had their own grantees, and corresponding notes were included in the spreadsheets.

A table was created about the gender of the grantees. Information about the gender of new grantees, without repeating grantees from previous years, clearly shows the division by gender of the grantees for each year of RAIC. Six grantees whose names were incomplete could not be identified by gender.

Table 4
RAIC participants by gender

Year	Female	Male	Unidentified	Total partic.
1993	96	36	4	136
1994	36	31		67
1995	47	19		66
1996	125	45	2	172
1997	83	42		125
1998	122	65		187
1999	95	54		149
2000	98	45		143
2001	76	36		112
2002	107	61		168
2003	103	58		161
2004	115	50		165
2005	136	44		180
2006	117	50		167
2007	150	70		220
Totais	1506	706	6	2218

Source: author's data

The five subsequent spreadsheets corresponded to information on undergraduate, postgraduate diploma, master's, doctoral and postdoctoral levels, respectively. The grantees, identified by the numbers assigned to them in the spreadsheet "Data", had the following information included: year of course completion (concerning the specific course of the spreadsheet in question), institution where the course took place, and area of knowledge of the referred course.

This information was written in red in the spreadsheets for grantees from previous years. The identification number of the grantees was written in bold type to record the occurrence of more than one course at the same level of schooling. In other words, the identification number of the grantee whose curriculum had more than one undergraduate course was repeated for every extra course, and this number was written in bold type to highlight such information while preserving the grantee's identification.

Naturally, the most relevant information referred to the undergraduate level, when the PIBIC scholarships were granted. It should be noted that because FIOCRUZ does not offer undergraduate courses, it has no records for this educational level. Information on graduate courses identified whether or not the grantees had met the PIBIC objectives; it also showed, as shall be seen later in this paper, their strong preference for FIOCRUZ when they entered a graduate program.

The seventh and last spreadsheet had records about the advisors of scientific initiation grantees, according to the information gathered in the proceedings. As much as possible, the names of the advisors and the number of advisees they had at the RAIC in question were standardized.

This option was chosen to solely accept records of postgraduate diploma courses with a minimum of 360h in length. According to the Coordination for the Improvement of Higher Education Personnel (CAPES), *Master of Business Administration* (MBA) courses held in Brazil and medical residencies were considered to be postgraduate diploma courses. Extension courses were not taken into consideration.

The table below displays the numbers of advisors, year by year.

Table 5
Advisors of PIBIC grantees

Year	Advisors
1993	244
1994	237
1995	216
1996	224
1997	214
1998	294
1999	333
2000	330
2001	205
2002	290
2003	315
2004	315
2005	370
2006	334
2007	410
Total	4331

Source: author's data

4.6. Analysis of data from spreadsheets

To offer an overview of the main numbers found in the spreadsheets, pages were created containing the following information [2]:

- Year – referring to the year of the analyzed RAIC;
- RAIC – shows the number of the analyzed RAIC;
- Total grantees – absolute figures for grantees participating in the analyzed RAIC;
- New grantees – number of new grantees participating in the analyzed RAIC;
- Grantees from previous years – number of grantees from previous years, participating in the analyzed RAIC.

From this point onwards, the displayed information only referred to the new grantees from the analyzed RAIC in order to avoid repetition of data about grantees from previous years, which could yield wrong totals.

The following information was thus inserted about new grantees:

- Grantees with Lattes Platform – results obtained from the query for names of new grantees in the Lattes Platform;
- Grantees without Lattes Platform – results obtained from the query for names of grantees in the Lattes Platform, with negative results for incomplete names or namesakes without the possibility of perfect identification;
- Female grantees – identified in the spreadsheets by letter F, with information retrieved from the Lattes Platform;
- Male grantees – identified in the spreadsheets by letter M, with information retrieved from the Lattes Platform;
- Unidentified grantees – incomplete names which prevented identification of grantee's gender in the Lattes Platform.

Ways of meeting PIBIC's objectives [3] – the start and end dates of the courses were recorded with their specificities:

- Transition from undergraduate to postgraduate diploma level;
- Transition from undergraduate to master's level;
- Transition from undergraduate to doctoral level;
- Transition from doctoral to postdoctoral level.

The number of grantees with more than one course was mentioned, according to occurrence: double, triple, quadruple or quintuple occurrence. Each occurrence was registered according to the course, i.e. undergraduate, postgraduate diploma, master's, doctoral and postdoctoral levels. [4]

Finally, the number of advisors of the analyzed RAIC was shown.

4.7 Meeting PIBIC's objectives

The analysis of the curricula of the scientific initiation grantees in the Lattes Platform checked whether or not they had met PIBIC's objectives.

Lists of grantees were organized for each year of RAIC. Start and end dates were written next to the grantees' names for courses at the undergraduate, postgraduate diploma, master's, doctoral and postdoctoral levels. It was thus possible to check which graduate courses started immediately after completion of the undergraduate course. Such information was highlighted in bold type.

To obtain the accurate number of grantees who met

PIBIC's objectives, the names of grantees from previous years were italicized and thus separated from the new ones. The table below shows a summary of the grantees who met PIBIC's objectives:

Table 6

Grantees who met PIBIC's objectives

Year	New Grantees	Grantees prev. yrs
1993	16	
1994	12	6
1995	17	5
1996	35	12
1997	27	13
1998	43	19
1999	47	39
2000	50	46
2001	42	19
2002	67	24
2003	45	41
2004	44	28
2005	34	35
2006	10	15
2007	5	4
Total	494	306

Source: author's data

4.8 Queries on educational institutions

Data were gathered about the educational institutions where the grantees had studied previously, and also about those where they were going to further develop their academic careers. A six-spreadsheet Excel file was hence produced with this purpose. The spreadsheets contained information about the educational institutions where the grantees had taken courses at the undergraduate, post-graduate diploma, master's, doctoral and postdoctoral levels. There was also a general spreadsheet which consolidated these data.

Information was gathered in the Lattes Platform, and some inconsistencies were found. For example, there were cases where the grantee's curriculum had information about the college only, but not about the university that encompassed that college. There were also cases where the name of the institution was mistakenly changed or written with different spellings.

To ensure standardization, information was searched about each institution gathered in the Lattes Platform.

Whenever possible, the name of the university was chosen over the names of related schools or colleges.

If the name of the institution had changed, the most recent name was selected after being confirmed by information provided on the CAPES website.

One hundred and eighty-two teaching and research institutions were searched and recorded. Such institutions were mentioned 1687 times in undergraduate activities; 283 for post-graduate diploma activities; 665 at the master's level, 292 for doctoral activities and 39 times for postdoctoral activities.

Quantifying the institutions by type of course can provide insights for other studies about the PIBIC Program, particularly as regards Rio de Janeiro, Salvador, Belo Horizonte, Recife, Manaus and Curitiba, cities where FIOCRUZ campi are located, where the grantees developed their academic activities.

4.9 Gathering and selecting terms to identify knowledge areas within the courses

Undoubtedly, that was a time-consuming task because of the nomenclature used to name the courses gathered. For several reasons, it was not possible to establish the difference between certain undergraduate courses; for example, biomedicine and biomedical sciences. That would require consultation of their course syllabus.

As regards undergraduate courses, for example, many grantees had the following information in their curriculum: "undergraduate degree in pediatrics". To make it easier to quantify and identify the sub-area established by CNPq, the term "medicine" was selected for the field 'knowledge area'.

Post-graduate diploma courses caused the greatest difficulty for area identification, as more than one area of knowledge could be found in one single course, for example "education, culture and communication". As much as possible, topics were grouped to enable quantification of knowledge areas.

Even when grouping was possible, sometimes the grantee's description of the area was preserved because there were many grantees in such area (many of these courses were offered at FIOCRUZ). Another reason was that further specification could cause a substantial loss of information for the analysis of the obtained results.

A spreadsheet was created to make such grouping easier to understand. Areas with different designation were highlighted in blue. The grouping yielded a list of 134 topics, including ten unidentified ones.

In a similar procedure adopted for the institutions, the topics were listed generally and by course type to identify the current and future research areas of scientific initiation grantees, yielding the same number of institutions, as the number of articles were being counted.

4.10 General data about topics and institutions

Five spreadsheets were created for each type of course to clearly show the five topics and the five institutions [5] which were prominent, according to the data below:

At the undergraduate level, these were the five top knowledge areas: biological sciences (630), medicine (246), pharmacy (148), biomedicine (114) and history (88). Most grantees came from the following institutions: UFRJ (295), UFF (211), UERJ (179), UNIRIO (167) and PUC-MG (88).

At the post-graduate diploma level, given the practical nature of the courses, these were the most prominent topics: medical residency (54), education (20), clinical analysis (19), medicine (19), entomology (14) and public health (10). Former PIBIC grantees undertook their post-graduate diploma courses at the following institutions: FIOCRUZ (65), UFRJ (26), UFF (19), UERJ (10), USP (10) and UPE (09).

In master's degree courses, the main topics were molecular and cellular biology (113), parasite biology (66), public health (61), biomedicine (114), history (33) and pathology (29). Former PIBIC grantees received their master's degrees at the following institutions: FIOCRUZ (302), UFRJ (112), UFF (37), UERJ (36) and UFPE (27).

At doctoral level, the main topics were molecular/cellular biology (65), parasite biology (35), public health (15), biophysics (13), pathology (13) medicine (11), history of sciences and health sciences (11). Former PIBIC grantees pursued their doctoral degrees at the following institutions: FIOCRUZ (140), UFRJ (61), UFMG (15), USP (12) and UERJ (10).

Postdoctoral research had the following topics as the most prominent: biological sciences (16) immunology (09), pharmacology (02), history of sciences (02) and Parasitology (02). Former PIBIC grantees developed their studies at the following institutions: FIOCRUZ (10) and National Heart Lung and Blood Institute, in USA (04), University of California, in USA (04), Johns Hopkins University (02), USP (02) and Mount Sinai Medical School, in USA (02).

The overall analysis of the topics of all the courses highlighted the following areas: biological sciences (663), medicine (298), molecular/cellular biology (179), pharmacy (161) and history (137).

Among the institutions that offered undergraduate and graduate courses to grantees and former PIBIC grantees, the following can be highlighted: FIOCRUZ (517), UFRJ (496), UFF (271), UERJ (236) and UNIRIO (173).

5 Concluding remarks

The results of this study were clearly recorded as numbers and displayed in spreadsheets and tables. These numbers can provide other analyses and conclusions.

During the period of scientific initiation, FIOCRUZ is not listed as an institution for the grantees because it does not offer undergraduate courses.

Universities in Rio de Janeiro were prominent to the grantees because that is where FIOCRUZ is located. Some of the universities listed are located in states where FIOCRUZ has campi, for example Minas Gerais and Pernambuco. Some foreign universities were selected by FIOCRUZ former grantees for their postdoctoral research because that area of study could not be researched in Brazil.

The knowledge areas of courses which were frequently chosen by grantees and former grantees portray the profile of the work performed by FIOCRUZ, namely research in biological and biomedical sciences, pharmacy, social sciences and history.

The concluding remarks aim to focus on the main difficulties faced during this research and the solutions found to mitigate them.

The description of the formatting of RAIC proceedings signaled the need for greater care for graphic and editing issues in such publications. No standardization was found for size, indexing, and records of titles and authors, which could otherwise make it easier to identify such publications and perform queries or searches in written publications, in compliance with the standards currently valid in Brazil.

As regards proceedings recorded in compact discs (CDs), limitations to information search observed in the recording software shows that it needs to be reviewed. For example, authors' names can only be searched by their first name rather than last name, and queries by either advisor or topic are not available.

The institutional PIBIC program, coupled with CNPq and other funding agencies, should define standards to formalize 'over-the-counter' scholarships" by requiring the identification and registration of such scholarships with institutional scholarships.

Entering the names of scientific initiation scholarships into the FIOCRUZ database *Produtores Científicos* can be the basic tool to study the scientific production through PIBIC at FIOCRUZ. Moreover, making the different spellings of grantees' names available through the MARC format will enable, under the technical and informational point of view, the exchange of carefully described authority records, thus helping feed other information systems.

The Lattes Platform was a reliable research tool despite the remaining incompleteness of several curricula. The PIBIC at FIOCRUZ could mitigate such problems by demanding that grantees regularly update their curricula.

These standardized and updated records can be used to control and assess the grant of scholarships, ensuring quality information which can help the institution receive a greater number of scholarships.

Informing CNPq and CAPES of results like the ones found in this study will highlight the crucial role played by the institutions that participate in PIBIC and add greater value to undergraduate and graduate education in Brazil.

Notes

1. FIOCRUZ campi mentioned by their acronyms are the following: Immunobiological Technology Institute (Biomanguinhos), Laboratory Animal Breeding Center (CECAL), Brasília Advisory Board (DIREB), Human Resources Advisory Board (DIREH), Medicines and Drugs Technology Institute (Farmanguinhos), Carlos Chagas Institute (ICC), Fernandes Figueira Institute (IFF), National Institute for Health Quality Control (INCQS), Oswaldo Cruz Institute (IOC), Evandro Chagas Clinical Research Institute (IPEC), Aggeu Magalhães Research Center (IPqAM), Gonçalo Moniz Research Center (IPqGM), Leônidas & Maria Deane Research Center (IPqLMD), René Rachou Research Center (IPqRR).
2. All these data are displayed in specific spreadsheets for each RAIC, with their respective totals.
3. PIBIC aims to offer internships to motivate undergraduate students to pursue graduate studies upon completion of their degree.
4. Data for each course are displayed in separate spreadsheets.
5. The teaching institutions mentioned whose acronyms were not explained in full are the following: Federal University of Rio de Janeiro (UFRJ), Fluminense Federal University (UFF), State University of Rio de Janeiro (UERJ), Federal University of the State of Rio de Janeiro (UNIRIO), Pontifical Catholic University of Minas Gerais (PUC-MG), University of São Paulo (USP), State University of Pernambuco (UPE), Federal University of Pernambuco (UFPE), Federal University of Minas Gerais (UFMG).
6. The proceedings of each year were checked individually.
7. Team members of this study are Alexandre Medeiros Correia de Sousa, Aline Gonçalves da Silva and Leonardo de Souza Melo.

References

- AGUIAR, L.C.C. **O perfil da iniciação científica no Instituto de Biofísica Carlos Chagas Filho e no Departamento de Bioquímica Médica da Universidade Federal do Rio de Janeiro.** 1997. 119f. Dissertação (Mestrado em Química Biológica) – Centro de Ciências da Saúde, Instituto de Ciências Biomédicas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1997.
- BREGLIA, V.L.A. **A formação na graduação: contribuições, impactos e repercussões do PIBIC.** 2002. 210f. Tese (Doutorado em Educação) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2002.
- CABRERO, R.C. **Formação de pesquisadores na UFSCar e na área de educação especial: impactos do programa de iniciação científica do CNPq.** 2007. 276f. Tese (Doutorado em Educação Especial) – Centro de Educação e Ciências Humanas, Universidade Federal de São Carlos, São Carlos, 2007.
- CÓDIGO de catalogação anglo-americano – AACR2. 2. ed. Brasília: FEBAB, 2002.
- DAMASCENO, M.N. **A formação de novos pesquisadores: a investigação como uma construção coletiva a partir da relação teoria-prática.** In: CALAZANS, J. (Org.). *Iniciação científica: construindo o pensamento crítico.* São Paulo: Cortez, 1999. p.13-56.
- DEMO, P. **Pesquisa e construção de conhecimento: metodologia científica no caminho de Habermas.** Rio de Janeiro: Tempo Brasileiro, 2000.
- MACCARIELLO, M.C.; NOVICKI, V.; CASTRO, E.M. **Ação pedagógica na iniciação científica.** In: CALAZANS, J. (Org.). *Iniciação científica: construindo o pensamento crítico.* São Paulo: Cortez, 1999. p.79-115.
- MALDONADO, L.A. **Iniciação científica na graduação em nutrição: autonomia do pensar e do fazer na visão dos pesquisadores/orientadores.** 1998. 127f. Dissertação (Mestrado em Educação) – Faculdade de Educação, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 1998.
- MARC Concise Format: authority. Washington, D.C.: Library of Congress, 2004. Available at: <<http://www.loc.gov/marc/>>. Access in: 3 ago. 2004.
- MARCUSCHI, L.A. **Avaliação do programa institucional de bolsas de iniciação científica (PIBIC) do CNPq e proposta de ação.** Relatório final. Recife, 1996. Mimeografado.
- MASSI, L. **Contribuições da iniciação científica na apropriação da linguagem científica por alunos de graduação em química.** 2008. 227f. Dissertação (Mestrado em Ciências – Química Analítica) – Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, 2008.
- MORAES, A.F. **A avaliação da produção dos bolsistas de iniciação científica.** In: REUNIÃO ANUAL DE INICIAÇÃO CIENTÍFICA, 15., Rio de Janeiro, 2007. Palestra de abertura. Rio de Janeiro: Fiocruz, 2007.
- MORAES, A.F. **A gestão da informação sobre a produção dos bolsistas de iniciação científica na Fiocruz.** In: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 2., Londrina, 2007. Anais ... Londrina, UEL, 2007.
- PAULA, M.V. **Explorando o potencial da Plataforma Lattes como fonte de conhecimento organizacional em ciência e tecnologia.** 2004 1v. Dissertação (Mestrado em Gestão do Conhecimento e da Tecnologia da Informação) – Universidade Católica de Brasília, Brasília, 2004.
- PIRES, R.C.M. **A formação inicial do professor pesquisador universitário no Programa Institucional de Bolsas de Iniciação Científica – PIBIC/CNPq e a prática profissional de seus egressos: um estudo de caso na Universidade do Estado da Bahia.** 2008. 356f. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.
- REUNIÃO ANUAL DE INICIAÇÃO CIENTÍFICA, 1. a 15., 1993-2007, Rio de Janeiro. Anais ... Rio de Janeiro: Fiocruz, 1993 a 2007. [6]