

Epidemiological studies in the information and genomics era: experience of the Clinical Genome of Cancer Project in São Paulo, Brazil

V. Wünsch-Filho¹,
J. Eluf-Neto²,
P.A. Lotufo^{3,4},
W.A. da Silva Jr.⁵
and M.A. Zago⁶

¹Departamento de Epidemiologia, Faculdade de Saúde Pública,
²Departamento de Medicina Preventiva, ³Departamento de Clínica Médica,
Faculdade de Medicina, ⁴Hospital Universitário, Universidade de São Paulo,
São Paulo, SP, Brasil
⁵Departamento de Genética, ⁶Departamento de Clínica Médica,
Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo,
Ribeirão Preto, SP, Brasil

Abstract

Correspondence

V. Wünsch-Filho
Departamento de Epidemiologia
Faculdade de Saúde Pública, USP
Av. Dr. Arnaldo, 715
01246-904 São Paulo, SP
Brasil
E-mail: wunsch@usp.br

Research supported by FAPESP
(No. 01/12897-8) and Ludwig Institute
for Cancer Research.

Received June 17, 2005
Accepted March 2, 2006

Genomics is expanding the horizons of epidemiology, providing a new dimension for classical epidemiological studies and inspiring the development of large-scale multicenter studies with the statistical power necessary for the assessment of gene-gene and gene-environment interactions in cancer etiology and prognosis. This paper describes the methodology of the Clinical Genome of Cancer Project in São Paulo, Brazil (CGCP), which includes patients with nine types of tumors and controls. Three major epidemiological designs were used to reach specific objectives: cross-sectional studies to examine gene expression, case-control studies to evaluate etiological factors, and follow-up studies to analyze genetic profiles in prognosis. The clinical groups included patients' data in the electronic database through the Internet. Two approaches were used for data quality control: continuous data evaluation and data entry consistency. A total of 1749 cases and 1509 controls were entered into the CGCP database from the first trimester of 2002 to the end of 2004. Continuous evaluation showed that, for all tumors taken together, only 0.5% of the general form fields still included potential inconsistencies by the end of 2004. Regarding data entry consistency, the highest percentage of errors (11.8%) was observed for the follow-up form, followed by 6.7% for the clinical form, 4.0% for the general form, and only 1.1% for the pathology form. Good data quality is required for their transformation into useful information for clinical application and for preventive measures. The use of the Internet for communication among researchers and for data entry is perhaps the most innovative feature of the CGCP. The monitoring of patients' data guaranteed their quality.

Key words

- Multicenter studies
- Large-scale studies
- Molecular epidemiology
- Data control quality
- Cancer epidemiological studies

Introduction

Large-scale epidemiological studies have been conducted in the past. An example is the largest human experiment ever executed - the population-based study that evaluated the effectiveness of the Salk vaccine in 1954, involving almost one million children (1). However, the major risk factors for non-transmissible chronic diseases were identified as the result of epidemiological studies initiated and carried out by individual investigators belonging to relatively small research groups. Large-scale projects in the biomedical field are currently being developed, involving a large number of centers and the interaction of researchers from different fields of knowledge seeking a common goal. The first and best known of these endeavors was the Human Genome Project, an international consortium integrating scientific institutions from a number of countries (2). Several other studies with similar characteristics, such as the BioBank UK study (3), aimed at obtaining biological samples from 500 thousand individuals aged 45-69 years in order to study the role of genes and their interaction with environmental and lifestyle variables in the occurrence of a number of diseases, are still ongoing. The organization of such massive projects was made possible by the great advances in informatics which took place in the last decades and by the easy communication provided by the world computer network. The exponential dissemination of the Internet in the 1990's affected the daily routines of millions of people, and an expressive volume of data currently circulates among the computers of researchers throughout the world.

In cancer research, the demand for large-scale studies is due, at least in part, to the advances in the fields of genetics and molecular biology. Only the availability of a large number of observations will provide the statistical power necessary for an analysis of the effects of gene-gene and gene-environ-

ment interactions in neoplasm etiology (4-6).

Three major options for large-scale studies are available: meta-analysis, pooled analysis and multicenter studies with an individual base. In multicenter studies the design and conduct of the investigation and the collection of data at different centers are done according to a common study protocol. One main challenge of multicenter studies is to maintain the comparability of data in terms of exposure, outcome and confounder variables. It is also necessary to consider logistic issues in order to obtain a similar timing among the different clinical groups. Additionally, great care is necessary during the joint analysis due to the heterogeneity of results from different centers.

Practical issues regarding initiating, organizing, managing, and evaluating the studies emerge in the context of multicenter large-scale projects. The articulation of several research groups located in different regions, the volume of data generated from a large number of patients, and the manipulation and collection of biological samples from these subjects require non-conventional solutions for registration, storage, and quality control operations.

The present paper describes, from an epidemiological perspective, the methodology developed for the collection and data quality control in the multicenter study called "The relationship between the differences in gene expression and the clinical and pathological features of human cancers", or, simply, the Clinical Genome of Cancer Project (CGCP), a project initiated by FAPESP (São Paulo State Research Support Foundation) and financed by this agency and by the Ludwig Institute for Cancer Research (7). The CGCP is a large-scale project - possibly the largest currently being developed in Brazil in the field of oncology - aimed at investigating the profiles of gene expression in normal and cancerous cells and correlating these profiles with the etiology and prognos-

sis of the tumors under investigation. This knowledge may be used in the future to monitor new methods for the diagnosis and treatment of cancer.

The experience acquired by the epidemiology team while organizing and managing the CGCP data may be of use to researchers in the field of health care currently involved, or who may become involved in the future, in large-scale multicenter studies.

Material and Methods

The CGCP involves specialists in internal medicine, surgery, pathology, molecular biology, and epidemiology (participants are listed at the end of the article). The project is aimed at consolidating data on a large number of patients with well-defined diagnoses of nine types of tumors (astrocytoma, head and neck squamous-cell carcinoma, esophageal squamous-cell carcinoma, gastroesophageal junction cancer, gastric adenocarcinoma, colon and rectum carcinoma, multiple myeloma, osteosarcoma, and acute lymphoblastic leukemia) and at collecting biological samples (blood, tumor tissue, and normal tissue) from these patients.

Participant groups were selected according to FAPESP's peer-review principles. Clinical, pathology, and epidemiology groups answered a call from FAPESP, which defined evaluators from international institutions to select the groups that would participate in the study. Initial meetings were conducted in order to consolidate the plan and to define a field work strategy. Researchers maintain contact through the Internet and occasionally hold specific meetings with their groups and CGCP holds a biannual meeting for all members.

Basic designs of the collaborative study

Each clinical group is guided by specific objectives; however, three common epidemiological study designs can be identified in

the CGCP:

Cross-sectional studies for the analysis of gene expression. Using micro-array technology, the aim is to compare the prevalence of gene expression between normal and cancerous tissues. Therefore, when feasible (head and neck, esophageal, gastroesophageal junction, stomach, colon and rectum tumors and osteosarcoma), samples of normal tissue adjacent to the tumor are collected. In the case of astrocytomas, the analysis of gene expression will be compared to non-neoplastic tissue samples obtained from individuals without a diagnosis of cancer who were submitted to other neurosurgical procedures, most of them related to surgical correction of epilepsy.

Case-control studies for the analysis of etiological factors. These studies are aimed at evaluating the risk of disease according to the prevalence of specific genetic polymorphisms. Thus, DNA was extracted from the peripheral blood of cases (patients with specific tumors) and controls (patients with diseases other than cancer - except for skin cancer - which are not related to risk factors for the tumors under investigation, matched with cases by sex and age). Potential interactions between polymorphisms and lifestyle-related factors (smoking or alcohol consumption) may be studied.

Follow-up studies for the evaluation of prognosis. Five-year survival and other outcomes - such as regional and distant metastases, response to treatment, and clinical evolution of patients with the same type of cancer - will be investigated in terms of different combinations of clinical or histological variables and distinct patterns of gene expression and genetic polymorphisms.

Patient recruitment logistics

The research protocol was approved by the National Commission of Ethics in Research (CONEP, Brasília, DF, Brazil) and by the Ethics Committees of all hospitals

included in the CGCP. The recruitment of patients for the project began in the first trimester of 2002 and should continue to the end of 2005. Cases and controls come from eighteen clinical facilities in the cities of São Paulo, Ribeirão Preto, Campinas, São José dos Campos, and Botucatu, merged into twelve clinical groups linked to the nine groups of tumors under investigation (Figure 1). The clinical team of each hospital identifies cases and controls.

Throughout the year 2002, several meetings between the researchers of each clinical group and the epidemiology group were held in order to discuss the routines of patient recruitment, the format and content of the research forms, the procedures of data regis-

tration, and the transportation and storage of biological samples.

Five forms were designed to record patient data: a) the general form, containing information such as age, sex, place of birth, and previous exposure to lifestyle and environmental risk factors such as smoking and alcohol consumption; b) the clinical form, containing clinical and laboratory data; c) the pathology form; d) the follow-up form, containing data on the clinical status of the patient during the follow-up period, and finally e) a specific form for organizing the data relative to the biological samples.

We developed forms with specific questions for each type of tumor. Control patients answered the questions in the general form

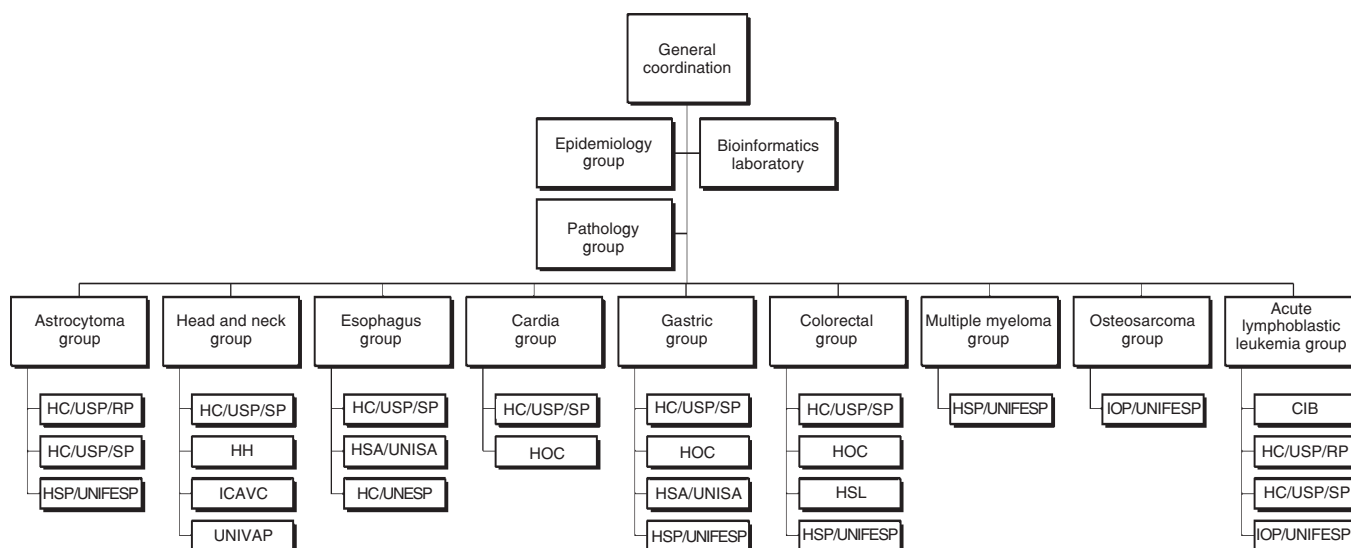


Figure 1. Flow chart of the Clinical Genome of Cancer Project.

Clinical centers:

CIB	Boldrini Child Center/Campinas
HC/UNESP	University Hospital/State University of São Paulo/Botucatu
HC/USP/RP	University Hospital/State University of São Paulo/Ribeirão Preto
HC/USP/SP	University Hospital/State University of São Paulo/São Paulo
HH	Heliópolis Hospital/São Paulo
HOC	Oswaldo Cruz Hospital/São Paulo
HSA/UNISA	Santo Amaro Hospital/Santo Amaro University/São Paulo
HSL	Sírio Libanês Hospital/São Paulo
HSP/UNIFESP	São Paulo Hospital/Federal University of São Paulo/São Paulo
ICAVC	Arnaldo Vieira de Carvalho Cancer Institute/São Paulo
IOP/UNIFESP	Pediatric Oncology Institute/Federal University of São Paulo/São Paulo
UNIVAP	Vale do Paraíba University/São José dos Campos (includes the following hospitals: Pio XII Hospital, Municipal Hospital, Do Vale Oncology Institute, São José dos Campos Gastric Clinic, Policrin, Santa Izabel Clinic, and São José Hospital).

only. Printable copies similar to the computerized forms are available on-line at the Ribeirão Preto Cell Therapy Center website (<http://ctc.fmrp.usp.br>). Following the interview, the data are entered into the system on-line.

For the centers without the infrastructure required for processing and storing their own biological samples (blood), a routine was organized for the collection and transportation of this material. At a frequency previously agreed upon with each center, blood samples are collected, packed and transferred to the Laboratory of Medical Investigation-38 of the University Hospital, School of Medicine of the University of São Paulo. These samples are then transported weekly to their final destination at the Ribeirão Preto Cell Therapy Center. Tumor and normal tissue samples are also periodically taken to this Center.

Database and system management

The different forms are completed at different times during the patient's clinical history. The general form is completed upon the patient's entry into the system, at the time of the interview and the remaining forms are frequently filled at later times.

Available through the Internet, access to the CGCP is personalized, using a login name and a security password defined by the user. Different degrees of access were established in order to ensure the privacy of patient data. The researchers from a given hospital have unrestricted access to the data of their own patients, but, in their clinical group, they can see only the consolidated quantitative data regarding the number of cases and controls. The researchers from one clinical group do not have access to any data from the other groups.

In order to facilitate database use and understanding, we established code names that distinguish each field of each form. The full code is available in the system in both

digital page (front-page on-line) and printable formats. Thus, the system has a common language which facilitates communication between the epidemiology and clinical groups. On the other hand, the code allows any new researcher that enters the project to immediately understand the meaning of the fields in each form.

Evaluation of data quality

The clinical epidemiology team is composed of three epidemiologists. The operational center is located in the Department of Epidemiology of the School of Public Health, University of São Paulo, and is staffed by a statistician, a database management technician, and two support technicians for data analysis. Two strategies were established: periodic evaluation of data quality and general evaluation of the consistency of the data entry performed at each center.

Periodic evaluation of data quality

The clinical epidemiological team issues periodical reports on the consistency of the data entered into the general form (cases and controls) and into the clinical and follow-up forms (cases). The criteria used for evaluating the data entered into the fields of the general form are discussed with the clinicians. In the reports, we describe possible problems in data consistency. For example, in the general form, confirmation was sought for all entries regarding the onset of smoking or alcohol consumption before the age of 10 years.

Data consistency reports for each form (general, clinical, and follow-up) were sent to each clinical group between September 2003 and October 2004. When necessary, alterations in fields for which there were doubts or which were left blank are carried out on-line by the centers themselves. The situation concerning these fields was reevaluated on December 31, 2004.

Evaluation of data entry consistency

In order to evaluate the data entry performed at the centers and to estimate the magnitude of the possible errors, we examined a random sample of 5% of all patients included in the system up to February 27, 2004. The selection was stratified by tumor class and into cases and controls. The final sample included 66 cases and 46 controls.

The data of the patients selected were reentered by an external computer's skilled expert specifically contracted for this task and trained in completing these forms. After telephone contact, the external expert went to the clinical center and, based on the paper copy of the completed forms or on the medical charts of the selected patients, reentered the data, also on-line. These forms received different numbers than those used for the normal entry of patients into the study. Data reentry took place between April 22, and May 10, 2004.

We subsequently identified the differences between the original entry by the clinical

group and that by the external expert. Discrepancies in the data entered into each field were detected during electronic verification. The comparison was carried out for the forms' fields and for each clinical group. We also determined whether discrepancies were due to mistakes made during the data entry procedures by the clinical group or by the external expert. Entry mistakes were estimated by dividing the total number of mistakes made by the clinical group by the total number of the electronic forms' fields evaluated. These analyses were processed using the Statistical Analysis Software (SAS)[®], version 8.02 for Windows[®].

Results

From the beginning of 2002 to the end of December 2004, 1749 cases and 1509 controls were interviewed and entered into the CGCP database. Eleven cases and 14 potential control patients refused to participate in the study. Head and neck tumors accounted for the largest number of cases (652), followed by colon and rectum tumors (328). With the exception of colon and rectum tumors, there was a predominance of males over females. The highest male/female ratios are observed for head and neck, esophagus, and cardia tumors (Table 1).

Tables 2 to 5 present the results of the evaluation of data entry consistency. In the general form, considering a total of 3639 fields of the questionnaire examined, the mean percentage of discrepant information between the data entry procedures conducted by the clinical group and that did by the external expert was 1.7%, ranging from 0 to 4.0% depending on the clinical group (Table 2). In the clinical form, discrepant information ranged from 0 to 6.7% according to the different clinical groups. The mean proportion of discrepant information in the fields of the clinical form was only 1.1% (Table 3). The pathology form showed the lowest proportion of discrepancy, with the highest value

Table 1. Number of participants by tumor group and controls recruited from January 2001 to December 2004.

Cases and controls ¹	Male	Female	Male:female ratio	Total
Cases by tumor group				
Astrocytoma	102	63	1.6	165
Head and neck	547	105	5.2	652
Esophagus	52	10	5.2	62
Gastroesophageal junction	24	4	6.0	28
Stomach	150	90	1.7	240
Colon and rectum	167	161	1.0	328
Multiple myeloma	35	21	1.7	56
Acute lymphoblastic leukemia	112	74	1.5	186
Osteosarcoma	18	14	1.3	32
All cases	1207	542	2.2	1749
Controls				
Adults (>15 years)	664	371	1.8	1035
Children (≤15 years)	187	220	0.9	407
Epilepsy ²	31	33	0.9	64
No information ³	1	2	0.5	3
All controls	883	626	1.4	1509

¹Excluding patients who refused to participate in the study (11 cases and 14 controls).

²All adults (>15 years). ³Lacking age information.

of 1.0% for the head and neck cancer group (Table 4). The follow-up form showed the greatest amplitude in the percentage of discrepancy, which ranged from 0 (zero) for clinical groups of head and neck, esophagus and cardia to 11.8% for the leukemia clinical group (Table 5).

Discussion

The basic characteristics of the CGCP were established by agreement of the researchers involved. The option for autonomy of the groups with respect to data collection and for the use of a computerized structure for data entry and storage was based on the consideration that these were the most adequate strategies given the project's circumstances and needs.

For the epidemiology group, which typically functions as a bridge between the clinical and the bioinformatics groups, the greatest challenges were related to the development of alternatives to allow communication between groups, to the determination of the levels of access of each group to the computerized system, and to the development of procedures for the evaluation of data quality aimed at preparing the data for analysis.

Compared to traditional epidemiological research, the use of the Internet for communication between research groups and for data entry into a computerized database is perhaps the most innovative aspect of the CGCP. A potential disadvantage of the use of decentralized virtual systems for biomedical data entry is the absence of printed documents for each of the study's patients, kept at a centralized storage location, as is usual for clinical and epidemiological studies. After entry into the system, the CGCP data are validated, and creating an organized filing system for keeping printed copies of the record is not a concern, even though the patients' charts are always a source of information in case of doubts. This use of

informatics in large-scale studies represents a break with some of the procedures of traditional health research, but requires adequate planning and monitoring.

The greatest positive result of the CGCP is the integration between different clinical groups in a common project. A single clinical center would not be able to contribute a suffi-

Table 2. Evaluation of data entry consistency in the general form.

Tumors	Forms evaluated		Fields examined ¹	Discrepant information
	Cases	Controls		
Astrocytoma	7	4	301	2 (0.7%)
Head and neck	23	13	1455	21 (1.4%)
Esophagus	2	0	80	0 (0%)
Gastroesophageal junction	2	0	65	1 (1.5%)
Stomach ²	7	0	245	5 (2.0%)
Colon and rectum	13	2	403	6 (1.5%)
Multiple myeloma	3	2	147	6 (4.0%)
Osteosarcoma	1	1	60	0 (0%)
Leukemia	7	10	522	8 (1.5%)
HC/USP/SP controls ³	0	13	361	13 (3.6%)
All	65	45	3639	62 (1.7%)

Data are reported as number with percent in parentheses. ¹The total number of fields examined varied among individual subjects. For example, smoking and alcohol consumption information is absent for nonsmokers and nondrinkers. ²Two questionnaires selected during sampling but not evaluated (1 case and 1 control). ³At the University Hospital, School of Medicine, University of São Paulo, controls were recruited from different inpatient units as a whole group and not separately for each tumor group, although the same control recruitment procedures as used at the other CGCP centers were followed.

Table 3. Evaluation of data entry consistency in the clinical form.

Tumors	Forms evaluated		Fields examined ¹	Discrepant information
	Cases	Controls		
Astrocytoma	7		239	4 (1.7%)
Head and neck	23		1595	8 (0.5%)
Esophagus	2		350	0 (0%)
Gastroesophageal junction	2		383	5 (1.3%)
Stomach	7		219	0 (0%)
Colon and rectum	13		751	11 (1.5%)
Multiple myeloma	3		96	0 (0%)
Osteosarcoma	1		59	1 (1.7%)
Leukemia ²	5		179	12 (6.7%)
All	63		3871	41 (1.1%)

Data are reported as number with percent in parentheses. ¹The total number of fields examined varied among individual subjects. For example, for head and neck cancer, the field concerning site of the sore was only filled if the patient reported having a sore. ²Two leukemia patients did not have completed clinical forms.

cient number of cases of a given tumor to allow for combined analyses of genetic and environmental variables and for relevant results to be obtained with respect to etiology and prognosis. A large number of patients with different tumors had already been recruited as of December 2004, and another contingent should be added by the end of 2005.

During the last decades there has been a

Table 4. Evaluation of data entry consistency in the pathology form.

Tumors	Forms evaluated ¹	Fields examined ²	Discrepant information
Astrocytoma	7	12	0 (0%)
Head and neck	22	602	6 (1.0%)
Esophagus	2	46	0 (0%)
Gastroesophageal junction	2	73	0 (0%)
Stomach	7	475	0 (0%)
Colon and rectum	12	571	2 (0.4%)
All	52	1777	8 (0.4%)

Data are reported as number with percent in parentheses. ¹There are no pathology forms for multiple myeloma and acute lymphoblastic leukemia. The pathology form was not completed for one patient of each of the following groups: head and neck cancer, colorectal cancer, and osteosarcoma. ²The total number of form fields examined varied among individual subjects. For example, form field concerning the number of affected lymph nodes in a particular anatomical region was only fulfilled if lymph nodes had previously been indicated in that region.

Table 5. Evaluation of data entry consistency in the follow-up form.

Tumors	Forms evaluated ¹	Fields examined ²	Discrepant information
Astrocytoma	7	141	1 (0.7%)
Head and neck	15	79	0 (0%)
Esophagus	1	49	0 (0%)
Gastroesophageal junction	1	6	0 (0%)
Stomach	10	131	3 (2.3%)
Multiple myeloma ³	1	-	-
Leukemia	5	271	32 (11.8%)
All	40	677	36 (5.3%)

Data are reported as number with percent in parentheses. ¹Number of patients per group without a completed follow-up form: head and neck - 8; esophagus - 1; cardia - 2; gastric - 6; colorectal - 3; multiple myeloma - 2; osteosarcoma - 1; leukemia - 2. ²The total number of fields examined varied among individual subjects. For example, for patients with astrocytoma, only fields related to treatment were completed if the patient had undergone chemotherapy or radiotherapy. ³The only patient selected was not evaluated. In this group the researchers did not open new follow-up forms each time the patient returned, but, instead, information on clinical alterations was repeatedly altered in the same form.

significant evolution in epidemiological methods, especially with respect to statistical analysis. Epidemiologists today are able to operate with mathematical models; however, mastering these technologies does not solve essential problems related to the quality of research data and the magnitude of the biases that cannot be controlled during analysis, two key elements if one wishes to establish accurate cause and effect inferences. The mean percentage of entry errors among the CGCP clinical groups was only 1.7%, which indicates the reliability of the data included in the database via the Internet. Furthermore, the continuous monitoring of data will further ensure their quality. We have properly identified the biological samples from essentially all patients and, according to preliminary analyses carried out at the Ribeirão Preto Cell Therapy Center, this material is very satisfactory. The quality of the diagnoses is also ensured, as indicated by the analysis of the pathology forms.

Large-scale projects are multidisciplinary and, in order to be effective, depend on a good informatics infrastructure. As it consolidates an expressive number of cases of different neoplasms, the CGCP in fact involves four large projects involving tumors at specific anatomical sites such as neurological, head and neck, and digestive system tumors, and the pathology group, and also into two smaller projects of multiple myeloma and osteosarcoma, each of which includes only a single clinical center. These projects are aimed at testing specific hypotheses regarding the etiology and prognosis of these diseases and have become feasible thanks to the availability of reliable and comprehensive patient data, as well as of a biological material bank well integrated to the clinical data.

Genomics is expanding the horizons of epidemiology, providing a new dimension to classical case-control, cohort, and cross-sectional studies, and is stimulating the development of large-scale multicenter studies aimed at discovering and characterizing genes related to common diseases (8). How-

ever, the principle remains that the transformation of data into information useful for clinical application and for the planning of preventive measures depends essentially on the quality of these data. Thus, study design and the implementation of a strategy based on large-scale multicenter studies for cancer research using the Internet require an objective scrutiny of the data so that valid results can be obtained in the analysis.

References

- Francis Jr T, Korn RF, Voight RB, Boisen M, Hemphill FM, Napier JA, et al. An evaluation of the 1954 poliomyelitis vaccine trials. *Am J Public Health* 1955; 45: 1-63.
- Collins FS, Morgan M, Patrinos A. The Human Genome Project: lessons from large-scale biology. *Science* 2003; 300: 286-290.
- Wright AF, Carothers AD, Campbell H. Gene-environment interactions - the BioBank UK study. *Pharmacogenomics J* 2002; 2: 75-82.
- Brennan P. Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? *Carcinogenesis* 2002; 23: 381-387.
- Caporaso NE. Why have we failed to find the low penetrance genetic constituents of common cancers? *Cancer Epidemiol Biomarkers Prev* 2002; 11: 1544-1549.
- Wunsch Filho V, Zago MA. Modern cancer epidemiological research: genetic polymorphisms and environment. *Rev Saude Publica* 2005; 39: 490-497.
- São Paulo Network for Cancer Research. The relationship between the differences in gene expression and the clinical and pathological features of human cancers. São Paulo: Fundação de Amparo à Pesquisa do Estado de São Paulo and Ludwig Institute for Cancer Research, 2001.
- Khoury MJ, Millikan R, Little J, Gwinn M. The emergence of epidemiology in the genomics age. *Int J Epidemiol* 2004; 33: 936-944.

List of CGCP participants

Coordinator: Marco Antonio Zago.

Astrocytoma group:

Alberto Alain Gabbai, Carlos Gilberto Carlotti Júnior, Suely Kazue Nagahashi Marie, Suzana Malheiros, Benedicto O. Colli, Sueli Oba.

Multiple Myeloma group:

Gisele Colleoni, José Orlando Bordin, José Salvador R. De Oliveira, Maria de Lourdes L.F. Chauffaille, Maria Regina Regis Silva, Maria Stella Figueiredo, Mihoko Yamamoto, Yuri V. Pinheiro.

Osteosarcoma group:

Antonio Sergio Petrilli, Sílvia Regina Caminada de Toledo.

Acute Lymphoblastic Leukemia group:

Antônio Sérgio Petrilli, Carlos Gilberto Carlotti Junior, Luiz Gonzaga Tone, Maria Lúcia de Martino Lee, Sílvia Regina Brandalise, Vicente Odone Filho, Vitória Régia Pereira Pinheiro.

Esophageal and Gastroesophageal Junction Cancer group:

Ivan Ceconello, José Carlos del Grande, Danilo Gagliardi, Maria Aparecida Arruda Henry, Marcelo Augusto de Oliveira, Orlando Contrucci.

Stomach Cancer group:

Joaquim Gama-Rodrigues, Laércio Lourenço, Sérgio Leonardi, Nelson Andreollo, Reginaldo Ceneviva, Fabio Lopasso, José Eduardo Krieger, Kiyoshi Iriya, Marcelo Eidi Nita, Osmar Yagi, Ulysses Ribeiro Jr., José Carlos Del Grande, Cláudio Bresciani, Carlos Eduardo Jacob, Carlos Malheiros, Fares Rahal, Shoiti Kobayasi, Nadin Safatle, Paulo Kassab.

Colon and Rectum Cancer group:

Angelita Habr-Gama, Délcio Matos, Nora Manoukian Forones, Raul Cutait, José Eduardo Krieger, Bernardo Garicochea.

Head and Neck Cancer group:

Francisco Gorgonio da Nóbrega, José Francisco de Góis Filho, Marcos Brasilino de Carvalho, Pedro Michaluart Junior, Vera Capelozzi, Patrícia M. Cury, Erica Erina Fukayama, Marina Pasetto Nóbrega, Carlos Frederico D. Pinto, Arthur C. Pereira da Silva, Abaeté Leite do Canto, João Moreira dos Santos, Paulo Vitor F. Souza Nascimento, Carlos Flavio Turci, Adriano Batista Diniz Mendes, Carlos de Oliveira Lopes.

Pathologists' group:

Kioshi Iriya, Marcelo Fabiano de Franco, Patrícia M. Cury, Sergio Rosenberg, Venâncio Avancini Ferreira Alves, Vera Luiza Capelozzi.

Clinical Epidemiologists group:

José Eluf Neto, Paulo Andrade Lotufo, Victor Wunsch Filho.