

Main Objectives of the Statistical Center

1. To participate in the development of proposed protocols, particularly as regards experimental design, sample size and feasibility. Biostatisticians work with the study coordinator on the statistical aspects for each protocol, including a specification of the major objectives and the number of patients required to meet those objectives. Data coordinators and biostatisticians comment on concept sheets for proposed studies and work with protocol coordinators in the Operations Office to produce protocols that are concise and clear.
2. To provide for registration of all patients on all studies, and for randomization of patients where appropriate. Registrations for all studies are available by direct Internet Web-based connection to the WebReg program.
3. To develop Web and software technology for paperless submission of data to the Statistical Center.
4. To provide for review and quality control of data collected during studies. Data coordinators screen all incoming data and query institutions regarding any incompleteness or inconsistency. Further range and logical checks are made at data entry. Data are collected on timeliness and accuracy of data submitted by each institution, as part of an effort to improve quality throughout the Group.
5. To provide for data entry, and for computer processing, and storage and retrieval of data. Data entry is performed by trained data control technicians. There are three data entry submission routes: paper, fax or the Internet.
6. To work with the Group Chair and other investigators in the Group to improve the quality of clinical trials through the use of improved data forms, uniform and reproducible data definitions and economical data flow, and efficient use of Statistical Center resources.
7. To assist the Group Chair in the administration of the Group.
8. To analyze and publish the results of studies in conjunction with the study coordinators. Improvements are continually being made in the analytic and design tools available to the biostatisticians. The main analytic tool is SAS™, with locally developed software to extract SAS™ files from Oracle™ and to produce tables and survival curves for the semi-annual Report of Studies and for publication.
9. To use the data collected to try to find new leads regarding prognostic factors and late effects.
10. To perform statistical research on the efficient design, conduct and analysis of cancer clinical trials and cancer control research. In particular, research is being done on the analysis of survival data, on design and analysis strategies for clinical trials, on monitoring strategies for Phase III studies, on analysis of longitudinal data subject to non-ignorable missingness, and on methods for the analysis of high-dimensional cytogenetic, microarray and proteomic data.

11. To educate investigators, nurse oncologists and CRAs in statistical analysis, research design and the utilization of the most advanced scientific and data management strategies.

Major Accomplishments

In fulfilling these objectives, major accomplishments during the years 1998-2003 were:

1. Developed a new Web based system to allow direct registration by institutions. Expanded this system to allow use by members of other Cooperative Groups, and registration to studies conducted by other Groups.
2. Developed a Web page providing access to the Report of Studies, accrual reports, training manuals and other research information.
3. Offered a training program for experienced clinical research associates on the use of the new Web tools.
4. Improved training. Developed a formal training program and a Data Operations Procedure Manual for new data coordinators. Revamped the training program and updated the manual for institutional clinical research associates. Significantly improved Standard Operating Procedures (SOP) documentation.
5. Contributed to the development and conduct of a Young Investigators Workshop, an intensive training program designed to teach clinical trials principles to new researchers.
6. Revised the program for training new biostatisticians, and instituted a biostatisticians' meeting to discuss procedures, policies, and methodologic issues.
7. Developed an improved system for documenting and tracking serious adverse events (SAEs), formerly referred to as adverse drug reactions (ADRs).
8. Simplified all forms in response to the need for efficiencies in the clinical trials process.
9. Participated in a national cooperative effort to streamline data collection through the Common Data Element project initiated by the Cooperative Group Chairs and Group Statisticians, and supported by the NCI.
10. Served on external committees in the form of faculty participation on data monitoring committees, NIH review panels, American Joint Commission on Cancer task forces, development of CTC and RECIST definitions, editorial boards, etc.
11. Adopted a new data submission and entry system that allows data to be scanned, faxed or submitted via the Web for direct entry into the database.
12. Developed an electronic patient chart system.
13. For intergroup communication, initiated collaboration with other cooperative group statistical centers to allow direct data submission on intergroup trials. Worked with CTSU on regulatory and data submission procedures.
14. Revised guidelines for the design and assessment of quality of life on Southwest Oncology Group studies.

15. Continued development of the database management system, including more flexibility in handling multiple registrations, cycle-specific data collection, extensive additional logic checks at the time of entry of patient evaluations, and enhanced capability for managing double blind studies in anticipation of planned prevention studies.
16. Met responsibilities with respect to the conduct, design, and analysis of Southwest Oncology Group studies.
17. Made advances in statistical methods for therapeutic and prevention trials, for quality of life studies, and for the analysis of high-dimensional data, as evidenced by presentations at national meetings and publications. Developed and improved tools for the design of trials and made them available on the Web.
18. Continued the development of analytic software, including the Statistician's Report Worksheet (SRW) for producing the Report of Studies, and programs for exploratory and longitudinal data analysis.
19. Managed follow up of multiple large intergroup trials, including SWOG 8814, 8897, 9313, and 9304. Managed a number of successful Phase III trials, including an intergroup sarcoma trial, S0033, testing two doses of Geevac in GIST tumors.
20. Completed accrual to a large chemoprevention study, the Prostate Cancer Prevention Trial (PCPT). Study results were published in the New England Journal of Medicine in July 2004.
21. Opened a second cancer prevention study, the Selenium and Vitamin E Cancer Prevention Trial (SELECT), which accrued a total of 35,534 men between activation in August 2001 and the end of accrual in June 2004.
22. Participated with a group of Japanese investigators in a series of clinical trials workshops.
23. Published a book on clinical trials, Clinical Trials in Oncology, co-authored by three Statistical Center faculty. Offered short courses based on this text. A second edition of this text appeared in 2002. Edited a Handbook of Statistics in Clinical Oncology, with contributions from five Statistical Center faculty.
24. Joined a consortium of SWOG researchers in the development of an Early Therapeutic Program for the Group. Developed a new Web program to allow rapid data transmission for this initiative.
25. In conjunction with the Group Operations Office, expanded the Web security systems to require individual password access to the Group Web site, thus allowing more detailed Web-based communications between institutions and the statistical center.
26. Initiated development of a system for collection and transmission of data from off-site laboratories and institutions to enable better tracking of pathology specimens and more consistent reporting of results.
27. Moved the cytogenetics business office to the Statistical Center.

28. Developed an internal Quality Assurance program to audit the data coding quality of Data Coordinators, and to provide continuing education as needed.
29. Established a correlative sciences statistical team to support the increasingly complex analyses involving high dimensional data.
30. Managed 26,213 registrations and 31,613 patients in follow-up.

Computing Infrastructure

Information Technology Environment

The Statistical Center computing resources are based on the Microsoft Windows 2000/2003 network. The key services include the patient databases, E-mail, WEB, application and file sharing, electronic document and image management, Cardiff Teleform[®] data forms design and submission, batch application processing, disaster recovery, virus protection services, Citrix-based terminal services, remote access, desktop configuration and network monitoring/management.

Desktop Systems and Network

Each staff at FHCRC and CRAB has an Intel Pentium III or higher system. Each desktop PC runs MS Windows 2000 Professional or MS Windows XP Professional. Over 35 desktop applications are supported. All workstations access the Statistical Center servers and the Internet. The Statistical Center database, WEB, file and related services are housed at CRAB. Statistical Center staff at FHCRC access CRAB resources (over a dedicated network connection to improve performance and security) using Citrix terminal services and Internet Explorer. Sensitive information is transmitted using encryption. In addition to the dedicated network link between CRAB and FHCRC, each organization has independent connection to the Internet.

MS Windows Servers

The Southwest Oncology Group Statistical Center uses a modified, distributed architecture with the majority of its MS 2000/2003 Servers, i.e. different servers perform different functions. There are over 50 network servers including a small number of development and test servers. Production servers are Pentium III or higher servers with multiple processors, a minimum of 1 Gigabyte (GB) of memory, have fault tolerant disk subsystems and carry on-site maintenance. Several tape library systems are used for backup.

Key network server services offered at the Statistical Center include:

- **Oracle[®]** Database
- **Microsoft[®]** SQL Server
- **Microsoft[®]** Exchange Post Office
- **Microsoft[®]** IIS WEB Servers (Internet and Intranet)
- **Lyris[™]** Listserv
- **Ecora[©]** Auditor
- **TNT[©]** ELM Event Monitoring
- **Cardiff[™]** Teleform[®]
- **Citrix[®]** Terminal Services
- **Camellia[©]** C/S Batch Services
- **Veritas[™]** Backup Exec Enterprise
- **Trend Micro[™]** Virus Protection
- **Symantec[™]** Ghost Enterprise and WinInstall
- **Microsoft[®]** System Management Services
- **SAS[®]** and **Splus[®]** Statistical Packages