<table>
<thead>
<tr>
<th>What’s inside...</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Letter from Doris Rubio</td>
<td>4</td>
</tr>
<tr>
<td>- Letter from Wishwa Kapoor</td>
<td>5</td>
</tr>
<tr>
<td>- Overview</td>
<td>6</td>
</tr>
<tr>
<td>- Faculty and Staff</td>
<td>7</td>
</tr>
<tr>
<td>- Financial Summary</td>
<td>8</td>
</tr>
<tr>
<td>- Project Summary</td>
<td>9</td>
</tr>
<tr>
<td>- Services</td>
<td>10</td>
</tr>
<tr>
<td>- Custom Software Development</td>
<td>11</td>
</tr>
<tr>
<td>- Biostatistics</td>
<td>12</td>
</tr>
<tr>
<td>- Graphic Design</td>
<td>13</td>
</tr>
<tr>
<td>- Projects</td>
<td>14</td>
</tr>
<tr>
<td>- Data Coordinating</td>
<td>14</td>
</tr>
<tr>
<td>- R01 Research Studies</td>
<td>16</td>
</tr>
<tr>
<td>- Other Awards</td>
<td>18</td>
</tr>
<tr>
<td>- Data Security</td>
<td>20</td>
</tr>
<tr>
<td>- Publications</td>
<td>22</td>
</tr>
</tbody>
</table>
Over the past ten years, the Center for Research on Health Care (CRHC) Data Center has assisted University of Pittsburgh faculty in conducting high-quality clinical and translational research in over 500 studies. By providing staff with expertise in high quality statistical analysis and database development, the CRHC Data Center has become an important resource for investigators who want to prepare for grant awards and to implement data collection and statistical analysis mechanisms.

The CRHC Data Center biostatistics unit includes PhD faculty statisticians and staff with MS or BS-level degrees. Our information science unit includes staff with BS and MS-level degrees in computer or information science. The breadth of knowledge and experience of our faculty and staff allow us to provide advanced statistical analysis, current technological solutions for data management, website development and management as well as graphic design.

The CRHC Data Center takes measures to ensure the integrity and quality of data by implementing the latest security and data validation procedures. The Data Center is compliant with the FDA’s Title 21 CFR Part 11 federal regulations and the Federal Information Security Management Act (FISMA). The DC team works with researchers to make sure they understand these requirements and determines the best combination of policies to enable study teams to maintain data integrity.

For his support and dedication to our endeavors, I am grateful to Dr. Wishwa Kapoor, director of the CRHC.

Best Regards,

Doris Rubio, PhD
Director, Data Center, Center for Research on Health Care
Ten years ago, most research studies used paper forms to collect data. When databases were used, they were simplistic single-user MS Access databases. The future direction seemed clear. More researchers would want to turn to robust electronic data collection systems. That was the atmosphere when Doris Rubio became the director of the CRHC Data Center with only a handful of support staff. Not only would the investigators need someone conversant in the latest technology to develop the data collection system, but also he or she would need someone to help with design of the study and analysis of the gathered data. Dr. Rubio developed her team based on those goals. From these humble beginnings, the CRHC Data Center has grown into a highly respected resource to the University. Today, the Data Center employs numerous faculty and staff who are experienced either in statistical analysis, computing technology, data management, or web design.

Recently, government spending has suffered severe cuts, and the NIH has not been immune from the budgetary knife. Dr. Rubio has been proactively responsive to the predicament. Currently, her team is developing mechanisms to minimize database development costs. It is reassuring to know that University of Pittsburgh researchers can turn to the CRHC Data Center to receive cost effective quality database solutions and statistical analysis.

Best Regards,

Wishwa Kapoor, MD, MPH
Director, Center for Research on Health Care
I have been working with Doug Landsittel in the past two years. He and I are co-PIs of on-going 2U01DK056961-12 Consortium for Radiologic Imaging Studies in Polycystic Kidney Disease (CRISP III). We are also local co-PIs of R01DK079856-03 Identifying genetic modifiers of severity in ADPKD (Study PI: Peter Harris). Doug is great to work with. He is very reliable and pro-active. His analytic approach is terrific. His opinion and quality of work are highly regarded by the steering committee (mostly nephrologists) involved in these NIDDK sponsored studies.”

K. Ty Bae, MD, PhD, Professor and Chair of Radiology

Over the past ten years, the CRHC Data Center has served as a significant resource for University of Pittsburgh investigators. Our talented team of faculty and staff provides quality statistical analysis, database, website, and design solutions. With over 500 research projects in our portfolio, we offer both experience and innovative technology.

Mission

The Center for Research on Health Care (CRHC) Data Center’s mission is to support University of Pittsburgh faculty in conducting high-quality clinical and translational research. To meet this objective, the CRHC Data Center provides high quality data management services, statistical design and analysis, and web site design and support to all University investigators.

The specific objectives of the Data Center include the following:

- To collaborate on development of study designs across a wide range of research projects, including trials, observational analyses, and biomarker and other imaging studies;
- To develop and implement optimal statistical approaches for all types of data, including complex hierarchical or other clustered structures and complex designs;
- To provide expert statistical consultation to guide researchers through all aspects of publication and funding efforts, and offering the needed resources and facilities for managing large data coordinating centers and statistical cores of program projects;
- To implement quality control efforts in all data collection and reporting; and
- To provide all services efficiently and within budget.
**Director**  
Doris Rubio, PhD

**Associate Director**  
Doug Landsittel, PhD

**Faculty**

Kaleab Abebe, PhD  
Douglas Landsittel, PhD  
Charity Moore, PhD  
Seo Young Park, PhD  
Doris Rubio, PhD  
Dana Tudorascu, PhD  
Jonathan Yabes, PhD  
Lan Yu, PhD

**Staff**

**Administrator**  
Patrick Reitz, MA

**Data Analysts**  
Elan Cohen, MS  
Diane Comer  
Patrice Gibbs, MS  
Irina Karpov, MS  
Jie Li, MS

**Financial Manager**  
Nicole Kaefer

**Information Systems Manager**  
Tim Bragg

**Systems Administrator**  
Donald Grimm

**Systems Analysts**  
Kyle Holleran, MSIS  
Jason Kojtek, MSIT  
Desheng Li, MSEE  
Linda Quinn  
Shelley Rowe  
Terry Sefcik, MSIS  
Xiuyun Shen, MS  
Joseph Weiss  
Charlene Xie, MSIS
Financial Summary

The Data Center has seen a tremendous growth over the last 10 years. We have significantly expanded the number of projects resulting in the need to increase the number of faculty and staff. With the increase of projects, we have considerably more revenue. Also over the last 10 years, the faculty have collaborated on numerous grants, thereby increasing our grant revenue. Faculty also are principal investigators on data coordinating centers, which also contributes to the substantial increase in grant funding. The graphs below document our continued success.

Total Revenue

Revenue that the CRHC Data Center received from all sources

Total Grant Revenue

CRHC Data Center revenue specifically from grant resources.

2002-2012
Over the past ten years, the CRHC Data Center has been the go-to solution for University of Pittsburgh researchers. Our talented team of 35 faculty and staff members provides quality statistical analysis, database, website, and design solutions. With over 500 research projects in our portfolio, we offer both experience and innovative technology.

Mission

The Center for Research on Health Care (CRHC) Data Center mission is to support University of Pittsburgh faculty in conducting high-quality clinical and translational research at the University of Pittsburgh. To meet this objective, the CRHC Data Center provides top-notch data management, statistical analysis, and coordination of clinical and translation research services to all University investigators.

The specific objectives of the Data Center include the following:

- To provide University of Pittsburgh investigators with experts in data management, programming, and statistical analyses
- To assist the investigator in achieving high standards for research quality on every project
- To provide complete data analysis and data management efficiently and within budget

Project Summary

Number of Projects

Projects that the CRHC Data Center actively worked on for each calendar year.

Number of Faculty and Staff

Total number of CRHC Data Center faculty and staff for each calendar year.

2002-2012
Over the past ten years, our programming team has created data management systems that are compliant with current data security practices, including FDA’s Title 21 CFR Part 11 federal regulations. We customize these programs for the research study in order to optimize study flow. Most of our data management systems are designed to be Internet based and can be run on tablet computers. Our servers utilize 128-bit SSL security and are secured behind a firewall.

**Research Studies**
- Forms consultation and design
- Paper-based data collection including double entry verification
- Paperless data collection including dynamic validation for range and logic checks and missing values
- Automated randomization for clinical trials
- Interactive tracking for recruitment, screening and follow up
- Customized interfaces for researchers and/or participants
- Customized online and printed reports: recruitment, tracking and follow up
- Automated email reminders programmed into applications to alert the research team of pending time points or data collection needs
- Customized user documentation: data dictionary, codebook
- Customized user training

**Other Web Services**
- Interactive online application systems
- SharePoint services (online collaboration tool from Microsoft that allows groups to share files based on security access settings)
- Web site development and maintenance

*What they are saying...*
Your Web team is awesome. I love working with you. Accurate, fast, and when there is a question, it’s always so professional as to how it is handled.

Thank you for being there.

Adrienne J. Weiss
Project Coordinator
UPMC Heart and Vascular Institute
The Biostatistics Core includes experienced faculty with strong backgrounds in clinical trials, observational studies, comparative effectiveness research, biomarker and imaging studies, risk prediction modeling, survey design, and survival analysis. The PhD faculty contribute to a wide range of grant-funded research collaborations and associated pre-award development, including design of studies and development and execution of statistical analysis plans. The Core also includes MS- and BS-level statisticians who conduct a wide range of analyses, including analysis of large data sets, modeling of multilevel and correlated data, and other regression analyses. Through these and other collaborative and independent research activities, our statisticians contribute significantly to not only developing and conducting analyses, but also interpreting and writing for publications and grant applications.

Our statisticians collaborate with investigators in a variety of ways from funding proposals and study design through analysis and publication. Some expertise we provide includes:

**Design of Clinical and Translational Studies**
- Collaboratively discuss/revise research objectives
- Develop overall statistical strategies including methodology for analyzing imaging data and other complex diagnostic measurements
- Design clinical trials and other experimental studies
- Develop sampling approaches for survey/observational studies
- Calculate sample sizes and power for research studies
- Determine/implement randomization and treatment assignment
- Design surveys and measures

**Analysis of Data**
- Perform statistical programming and data management of large data sets and evaluate risk factor significance and treatment effectiveness
- Fit mixed models and other estimating approaches for longitudinal and correlated data structures
- Produce standardized reports of recruitment and data quality measures for clinical studies
- Merge electronic health record data with study data and analyze health outcomes.
The CRHC Data Center has graphic designers who have created electronic newsletters, brochures, flyers, and annual reports. We have worked with investigators to create a look and feel for their studies. We have worked with departments to advertise a new or existing program through brochures or fliers. We have created automated electronic newsletters so that department staff member can input the text for the newsletter and have it automatically emailed to the distribution list.

Our graphic design team provides a wide range of services, involving:

- Web site design and layout
- Graphs, charts, and figures
- Photography (Profile photos and events)
- Flash design and development
- Brochures
- Branding (look and feel)
- Image and graphic editing
- Video and audio editing
Projects

Over the last 10 years, the CRHC Data Center has been involved in a wide range of projects. The following examples show the breadth of experience the CRHC Data Center has with projects of all types and sizes. We serve as the Data Coordinating Center for several large studies. We work with numerous investigators who have secured R01 funding. Finally, we collaborate with investigators on other types of funding such as career development awards, program project grants, and pharmaceutical studies. Below are a few examples of our work.

Data Coordinating Centers (DCC)

CAPSIL
Principal Investigator: Nina Singh, MD, Professor of Medicine
Funded by: National Institute of Allergy and Infectious Diseases

CAPSIL is a prospective, randomized, multicenter trial of preemptive therapy vs. prophylaxis for prevention of CMV disease in R-D+ liver transplant patients. Dr. Singh leads the data coordination and statistical analysis for the project. While Dr. Singh has her own study team, she used the Data Center to build an electronic data management system. The system we built is compliant with FDA Title 21 CFR Part 11. The study is ongoing. According to the project coordinator, “The data collection system created by the CRHC Data Center for our multicenter clinical research study, CAPSIL, is second to none. Our NIH sponsor has been very pleased with all aspects of our data system. On the user end, our site coordinators have given positive feedback and are pleased with the intuitive ease of use and flow of our study forms. I have been very pleased with the attentive and punctual service provided by our Data Center Project Manager, Terry Sefcik, and our Data Center Systems Programmer, Xiuyun Shen, and the Data Center in general.”

CRISP Data Coordinating and Imaging Analysis Center
Principal Investigator: Kyongtae Bae, MD, PhD, Professor and Chair of Radiology and Doug Landsittel, PhD, Professor of Medicine and Associate Director of the CRHC Data Center.
Funded by: National Institute of Diabetes & Digestive & Kidney Diseases

CRISP is a prospective, longitudinal study to evaluate the accuracy and validity of magnetic resonance imaging to determine disease progression in Autosomal Dominant Polycystic Kidney Disease (ADPKD) defined as a change in both renal and renal cyst volumes and renal function over time. Drs. Bae and Landsittel serve as Co-Principal Investigators with Dr Bae leading the imaging component and Dr. Landsittel leading the data management and analysis component. The CRHC Data Center was responsible for building an electronic data management system—a CRISP portal for posting study documents, data, and data entry forms, and for tracking sample collection from clinical sites. The Data Center also manages all of the reports needed for the study including recruitment report, tracking report, and all of the reports for the Steering Committee and the External Expert Group. Dr. Landsittel also conducts or oversees the analyses of other Data Center statisticians, and leads the organization of monthly Steering Committee meetings, External Expert Group reviews, annual face-to-face meetings, and contact with NIDDK program officers and the NIDDK repository.
HALT-PKD

**Principal Investigator:** Charity Moore, PhD, Professor of Medicine, Biostatistics, and Clinical and Translational Science (CRHC Data Center faculty).

**Funded by:** National Institute of Diabetes & Digestive & Kidney Diseases

HALT-PKD includes two multi-site randomized clinical trials to evaluate certain FDA-approved drugs that may be effective in slowing kidney growth in persons who have polycystic kidney disease. Seven centers across the United States are participating in the two trials. The CRHC Data Center is a crucial resource for these studies as we serve as the Data Coordinating Center (DCC) for the trials, with Dr. Moore as the Principal Investigator. Dr. Moore, a faculty member of the CRHC Data Center, is responsible for coordinating all aspects of the trials and all analyses. For this project, the CRHC Data Center developed and maintains the electronic data management system, communication platform for the members of the study team, and a website for participants (haltpkd.org). The statistical team led by Dr. Kaleab Abebe (Data Center faculty) conducts all of the analyses for the studies and generates all of the required reports for the Steering Committee and Data and Safety Monitoring Board.

SPARX-DOSE PD

**Principal Investigator:** Margaret Schenkman, PT, PhD, FAPTA, Professor and Associate Dean for Physical Therapy Education at the University of Colorado.

**Funded by:** National Institute of Neurological Disorders and Stroke

SPARX-DOSE PD is a multi-site Phase II clinical trial to determine the effects of exercise on the progression of PD symptoms. The primary aim is to determine whether individuals with de novo Parkinson’s disease (naïve to drug treatment) can achieve the randomly assigned levels of mean exercise intensity and adhere to the exercise protocol. Dr. Moore is the lead statistician and oversees all of the data management for the study. The CRHC Data Center created the electronic data management system for the project. Randomization is built into the system enabling the staff to be blind to the randomization. The study has just started recruitment.
R01 research studies

CABG, RELAX, Online Treatments

Principal Investigator: Bruce Rollman, MD, MPH, Professor of Medicine, Psychiatry, Biomedical Informatics, and Clinical and Translational Science

Funded by: National Heart, Lung, and Blood Institute and National Institute of Mental Health

Dr. Rollman investigates ways to improve the quality of care for those seeking care for depression or anxiety disorders. He has had several R01 funded research projects where the CRHC Data Center has created the data management platform and provided statistical expertise.

CABG was a randomized clinical trial that studied the effect of using case managers with people with depression following cardiac bypass surgery.

RELAX was a randomized clinical trial designed to optimize treatment and 12-month clinical outcomes for primary care patients who have panic disorder, generalized anxiety disorder, or both (PD/GAD).

ONLINE TREATMENTS is another randomized clinical trial that is designed to test the comparative effectiveness of two on-line treatments for these conditions provided through the context of a Collaborative Care program. For this project, the CRHC Data Center not only created the data management system but also built a social media tool for the participants to use.

Effectiveness of PT with Knee OA

Principal Investigator: G. Kelley Fitzgerald, PhD, PT, FAPTA, Associate Professor of Physical Therapy

Funded by: National Institute of Arthritis and Musculoskeletal and Skin Diseases

Enhancing the Effectiveness of Physical Therapy for People with Knee Osteoarthritis is a multi-center, multi-national, randomized clinical trial to examine the clinical and cost effectiveness of using booster sessions in the delivery of exercise therapy with manual therapy. This comparative effectiveness trial is unique in that it combines booster sessions and manual therapy techniques with exercise therapy to improve the effectiveness of the rehabilitation of knee osteoarthritis. This study is funded by the Agency for Healthcare Research & Quality. The CRHC Data Center built the electronic data management system for this trial. Dr. Landsittel serves as the lead statistician.

What they are saying...

Charlene Xie has done a terrific job creating both the smartphone and desktop versions of our trial’s Internet support group.

Bruce L. Rollman, MD, MPH, Professor of Medicine, Psychiatry, Biomedical Informatics, and Clinical and Translational Science
T2P2 (Talking To Pregnant Patients)

Principal Investigator: Judy Chang, MD, MPH, Assistant Professor of Obstetrics, Gynecology and Reproductive Sciences

Funded by: National Institute of Drug Abuse

T2P2 is a prospective observational study that uses mixed methods to understand how obstetric care providers ask pregnant patients about alcohol and illicit drug use. Patients’ first obstetric visits are audio recorded to study the communication styles of providers with disclosure of use by patients. Women are asked to provide urine samples in order to test for the presence of alcohol, tobacco, or illicit drugs. The CRHC Data Center created an electronic data management system that enabled the study to use tablet computers without the need for paper. All data are directly entered into the computer and the data are stored securely on our data servers. Dr. Rubio serves as the lead statistician on the project.

SingleVDoubleBundle

Principal Investigator: James Irrgang, PhD, PT, ATC, Associate Professor of Medicine

Funded by: National Institute of Arthritis and Musculoskeletal and Skin Diseases

SingleVDoubleBundle is a prospective randomized clinical trial to determine if anatomic double bundle ACL reconstruction can effectively restore normal dynamic knee function. This double-blind trial compares single bundle vs. anatomic double bundle ACL reconstruction. The CRHC Data Center built the electronic data management system for this trial. Dr. Moore serves as the lead statistician.
Other Awards

FAST (Functional Assessment System Tablet)

Principal Investigator: Rachel Hess, MD, MSc

Funded by: Institutional Funds

FAST is a computerized clinical intake tool that includes patient medical history, family history, and social history; collects patient-reported outcomes such as depression and health-related quality of life using standard instruments; facilitates data collection for Medicare Annual Wellness Visits; and allows the patient to complete a review of systems. At the end of the clinical intake, patients are informed about research studies for which they may be eligible and are asked to participate in the research registry. For this project, the Data Center built the tool for the clinic, created the Division's Patient Registry, and work with the FAST team to program the system when new studies are initiated.

Late Life Depression Prevention and Treatment Center (P30)

Principal Investigator: Charles (Chip) F. Reynolds III, MD, UPMC Endowed Professor in Geriatric Psychiatry.

Funded by: National Institute of Mental Health

The Late Life Depression Prevention and Treatment Center is dedicated to the development of effective prevention and interventions in late-life mood disorders and grief for adults of all ages. Not only does the center provide treatment, but it is also home to numerous studies on depression in older adults. For this project the CRHC Data Center developed an electronic data management platform allowing multiple studies to collect data in a uniform manner and then easily cross-examine those data at any time. The platform is also scalable so that unique data management tools and instruments for future studies can be incorporated into the existing infrastructure with minimal startup cost or time investments. Finally, because some research staff work on multiple studies at any given time this platform allows them to collect data for all of them using one familiar interface rather than having to learn multiple data management processes. According to the Director of the Center, “My colleagues and I working in the NIMH sponsored center for late life depression prevention and treatment have been extremely pleased with the quality of service from the Data Center within the Center on Research in Health Care. Data Center colleagues have worked with us to develop a user-friendly and flexible web-based data entry and management system, accommodating clinicians, database administrators, and researchers. I recommend them highly”

Options NOW (Foundation funding)

Principal Investigator: Ian McGowan, MD, PhD, FRCP, Professor of Medicine

Funded by: Bill and Melinda Gates Foundation

Options NOW is a Phase 1 Open Label Safety, Acceptability, Pharmacokinetic and ex vivo Pharmacodynamic Study of TMC278 Long Acting (LA) Administered Intramuscularly to HIV-1 Seronegative Women. The primary objective is to evaluate the safety and acceptability of TMC278 LA given as an intramuscular (IM) injection. The CRHC Data Center serves as the data management and statistical core of the study. Under the direction of Dr. Kaleab Abebe, the CRHC Data Center created the electronic case report forms and study database and is responsible for all statistical analysis.
RAPlaque utilizes a novel imaging technique, microbubble contrast-enhanced B-mode ultrasonography (CU), to determine whether 1) patients with RA have an increased density of carotid artery adventitial vasa vasorum compared to control patients after controlling for traditional CV risk factors, 2) both traditional CV risk factors and inflammatory modulators of plaque vulnerability are associated with increased density of carotid artery adventitial vasa vasorum in RA patients, and 3) increased disease activity measures in RA are associated with increased density of carotid artery adventitial vasa vasorum. Dr. Landsittel conducts analyses to assess differences in imaging measures between RA patients and controls, quantify relationships between imaging measures and other risk factors with disease activity, and check data quality. He also provides expert guidance on modifying enrollment strategies and assessing reproducibility of novel imaging measures.
Data security is an important factor to consider in the data collection process. Making sure that data is secure from unauthorized changes or access is important to ensure the data’s accuracy and integrity. In this vein, the Food and Drug Administration (FDA) developed guidelines (Title 21 CFR Part 11) for electronic records to ensure the integrity and reliability of the electronic data.

The CRHC Data Center creates database systems that are compliant with the FDA’s Part 11 guidelines. We include some of our policies here.

**Database Access Procedures**

- Study personnel are assigned unique login credentials to access only the data they are permitted to manage.
- Only database programmers and statisticians assigned to a project have rights to the database, files, and directories that contain sensitive project data.
- When a programmer or statistician leaves the Data Center, his/her user account is immediately disabled and ultimately deleted.
- When new members are added to the study team, the PI or approved designate must email the Data Center with user information and security restrictions.

**Validation Procedures**

- Use of allowable data ranges for variables reduces the possibility of entering an incorrect value.
- Logical branching (skip pattern) takes the data entry specialist to the next applicable field, form, or screen, based on the entered response.
- Drop down boxes and lookup tables allow only pre-determined, pre-formatted values to be entered.
- Designation of certain fields as required and the use of error messages alerts the user when these are not completed.
- Standardized codes for commonly used variables such as gender, race, and ethnic origin generate consistent data across studies.
- Standardized codes for fields with missing data assure fields have intentionally been set as missing.
- Implementation of automated procedures such as eligibility determination, randomization, and instrument scoring eliminate human error for a variety of study processes.

**Data Backup/Disaster Procedures**

- Databases, which reside on a CRHC Data Center server, are backed up to tape every night and archived weekly. Weekly archived media is stored at an offsite secure location and maintained there for six months.
- All servers use hardware fault tolerance methods to ensure the continued availability of data.
Version Control Procedures

• The Data Center programmer uses version control software, TortoiseSVN, to manage changes made to databases and websites. All revisions are stored and can be accessed by the programmer.
• In instances where changes cause a database application to malfunction, the programmer can return to an earlier stable version of the application.
• With projects that require multiple programmers, all programmers have access to all iterations of the project’s databases and websites.
• Study personnel are given access to the latest approved version of the forms.

Server Security

• The CRHC Data Center servers are located in a secure server room, which contains an alarm system, temperature control, and a double lock on the main door with a tracking key card entry system.
• Our dedicated web server utilizes 128-bit SSL security for online real time data entry.
• Our dedicated SQL server, which is used for database and data storage, offers 128-bit SSL security and has limited access via network firewall.
• Our dedicated web site and database development server is located behind a firewall with only developers on the intranet having access.
• Servers are housed in a secure rack with dedicated UPS power sources to assure 24/7 uptime.
• Production servers are scanned on a monthly schedule by Computing Services and Systems Development (CSSD) to ensure all software and hardware are running at optimal performance and systems are secure to the latest industry standard.


