Data Management: Why You Should Care and What You Should Know

June 19, 2013

Christine E. Chaisson
Data Coordinating Center, BUSPH
Clinical and Translational Sciences Institute (CTSI), BUMC

Why Worry About the Data?

At the end of the study, after:
• Interviews
• Clinical assessments
• Lab tests
• Other data elements...

All you have is the data

Things that can go wrong

- Data may not be saved or backed up
- Crucial data elements may be missing
- Data may be incorrect due to
  - Data collection errors
  - Data entry errors
- Data may be incorrectly identified
  - Cannot be merged
  - May be merged incorrectly
- Data files may be lost or corrupted

Real World Examples

- A few illustrations from the popular news sources

A Google Search

Data error gives Wisconsin justice big lead

- Waukesha County apologized for the error that 14,000 votes were not reported. The clerk explained that she had imported vote totals transmitted by the city of Brookfield but had not saved the data.
Forbes: Bad data hurt Haverford in college rankings

“Forbes' annual list is out, and Haverford plummeted from No. 7 to No. 27 - for no obvious reason. A College spokesman explained that the error was based on a single figure:

A zero was incorrectly entered in database instead of 108 for the graduation rate of white women who enrolled in 2004.

…But no revision is planned, since the magazine and the online list has already been published.”

Oops: Excel Error Calls Into Question...

“Serious errors that inaccurately represent the relationship between public debt and GDP growth among 20 advanced economies in the post-war period” were recently identified ‘This Time It’s Different,’ a 2009 book by Harvard researchers.

The Authors admitted they forgot to include five rows in an Excel file resulting in exclusion of data from Australia, Austria, Belgium, Canada, and Denmark — a “coding error” which they said was “a significant lapse on our part.”

Vertex stock slides over cystic fibrosis data mistake

“Shares in Vertex Pharmaceuticals have taken a hit after the company had to take the rather embarrassing step of correcting previously-announced interim mid-stage results of a combination cystic fibrosis treatment.

…the result of a misinterpretation of the denominator of the treatment group between the firm and its outside statistical vendor…”

How Bright Promise in Cancer Testing Fell Apart

Duke Cancer Center’s gene-based tests proved worthless, research behind them was discredited.

Statisticians from MD Anderson discovered errors such as columns moved over in a spreadsheet which Duke team “shrugged them off” as “clerical errors.”

Four papers were retracted

Duke shut down three trials

Center leaders resigned or were removed.

People died and relatives sued Duke.

Why Plan for Data Management?

- Standardized procedures
- Appropriate systems for
  - Data collection
  - Monitoring/auditing
  - Tracking
- Lead to higher quality data
- More reliable conclusions
With a Successful Data Management System:

Data will be:
- Complete
- Accurate
- Timely
- Answer the scientific questions

What you should consider when budgeting:
- Data entry if using paper forms
- Software (if applicable)
- Personnel (FTE) may include:
  - Data manager
  - SAS programmer
  - Web/database programmer
- Tasks include
  - Data cleaning and hecking
  - Reporting
  - Auditing

Goal: Convert Data into Electronic Format
- Get the data from report (CRFs) double-entered as soon as possible – don’t wait until the end of the study!
- Consider direct electronic capture of data
- Decide what you’re going to do before you start
- Make sure you have someone on your team that can handle the data tasks (preferably not a work study student that will change every semester)

Data Management 101:
- More than one approach to managing data.
- Consider the:
  - Environment
  - Available resources
- “Requirements analysis”
  - Planning prior to beginning the study
- Do what works for the study at hand

Data Management Stages
Using a Public Health Model

Prevention
- Design and development of data systems

Detection
- Subject accrual/data collection & cleaning

Treatment
- Correcting errors, Creation of analytic datasets, Analysis

Good Systems Prevent Errors
- Begin with:
  - Identification of tasks and timelines
  - Written, standardized procedures
  - Well designed data collection forms
  - Staff training programs
  - Documentation of systems
  - Plans for monitoring data
Where to Begin?
- Budget appropriately
- Finalize protocol and analysis plan
- Determine type/frequency of data
  - Case report forms (CRFs)
  - Biologic samples, X-Rays
- Electronic Data Capture vs. paper
- Determine key people and roles
- Establish timelines (work backwards)

Example: Example Reverse Time-line
- June 2013: Study Enrollment will begin
- May 2013: Final staff training/Investigator meeting
- April 2013: Finalize systems/pilot/testing
- Jan-March 2013: Construct data systems
- December 2012: Finalize assessment
- December 2012: Final visit protocol
- 11/12: Other final decisions

Visit Protocol: Data by Time-point
- Determine Visit Schedule and type (e.g., semi-annual in-person, 3 and 9 month phone)
- Determine data collected at each visit
  - Questionnaires
  - Labs
  - Adverse Events
  - Other data elements?
- Consider data that may not be connected to a time-point (hospitalization, death)
- Finalize the windows around time-points and how will data be associated if applicable

Sample Visit Grid

Timelines and Tasks
- Development of Protocol, analytic plan
- Creation & piloting of forms
- Design/construction of data entry and participant/data tracking systems
- Development of Manual of Operations
- Subject recruitment
- Data collection & follow-up
- Data cleaning, auditing, QA
- Analysis
- Manuscript preparation & submission

Create a Visual Timeline
- It doesn’t have to be fancy
- More detail is better but something simple is better than nothing
- Plan to review and revise it often
Simple overview Timeline

**Study Timeline**

- **Selected Activities**
  - 1-5: Design & Construct data management systems
  - 6-10: Training
  - 11-19: Planning instruments & tools
  - 20-28: Enrollment
  - 29-34: Follow-up
  - 35-42: Data collection
  - 43-51: Preparing & secondary data analysis
  - 52-54: Publications

Tools Of The Trade

- Well-designed forms
- Data management plan
- Tracking system / tracking database
- Data Capture System
- Database
- Study manuals
- Data query system

Data Management Plan*

* Required by many funders

- Outlines how data will be handled — should include descriptions of:
  - How data will be collected
  - How data will be processed (software, procedures)
  - How and where data will be stored (including software, formats, coding)
  - What QC the procedures will be and the schedule for carrying them out
  - Plan for long-term storage or archiving
  - Annotated forms

Sample Task-based Gantt

<table>
<thead>
<tr>
<th>Tasks</th>
<th>Year 1</th>
<th>Year 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Months 1-3</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 4-6</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 7-9</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 10-12</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 13-15</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 16-18</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 19-21</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 22-24</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 25-27</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Months 28-30</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

Tools Of The Trade

- Well designed forms
- Data management plan
- Tracking system / tracking database
- Data Capture System
- Database
- Study manuals
- Data query system

Information about DM Plans

Data Management Plans

- Importance of Data Management
- Background
- Data Management Plans
- Resources
- Ask a Question

Questions for a data management plan:
1. What [ ] within the dataset are the essential parts of the project?
2. What will be the role of each person?
3. How will data be managed?
Creating a Data Collection Form

Why is this topic important?

- Sloppy forms indicate sloppy research
- CRF doesn’t answer study questions
- Danger of collecting:
  - too much data
  - too little data
  - the wrong data
- Annoyed:
  - Study Coordinator
  - Participants
  - Analyst...

Successful Form: Consider ALL Functions

- Data Collection - who is completing form?
  - Study Staff (Coordinator, Clinician)
  - Participant
  - Clinician
- Data entry - who is entering data?
  - Study staff
  - Students
  - Outsourced
- Data management/cleaning
- Data analysis

Think Google

Not Yahoo
What makes a good form?

- User-friendly, uncluttered, well organized
- Provides clear instructions for completion
- Terminology familiar to person filling out
- Reading level matches study participants/evaluators
- Coded for easy data entry
- Questions only asked/data collected in one place and only one place
- Easy to refer back and clean data

Pilot Your Forms Prior to Data Collection

- Test in target population (age, gender, literacy)
  - Are items generating a high non-response rate?
    - Rework/drop question
  - Are "skip" patterns being followed correctly?
    - Train clinic personnel/revise forms
  - Are open-ended questions generating common responses?
    - Categorize/code
- Corrections made prior to start of study
- Try not to start data collection prior to finalizing forms

Avoid Open-ended & Include Response Measure

What is your date of birth? __________
2. How much do you weigh? ______
3. How tall are you? ______
4. Record subject’s temperature _____

Unit of Measurement: Clearly Defined

1. Date of Birth? ______/____/____ MM DD YYYY
2. How much do you weigh? ______ (pounds)
3. How tall are you? ______ (feet) ______ (inches)
4. Record subject’s temperature ______ (f)

Include Clear Instructions

A. What is your race/ethnicity? (Check one)
   - Caucasian
   - African American/Black
   - Asian, Pacific Islander
   - Native American
   - Other __________

B. What is your race/ethnicity? (Check all that apply)
   - Caucasian
   - African American/Black
   - Asian, Pacific Islander
   - Native American
   - Other __________

Beware: Missing data with “check all that apply”

a. High blood pressure
b. Heart disease
c. Diabetes
d. Cancer
e. Pulmonary disease
Account For Missing Data

<table>
<thead>
<tr>
<th>CBC</th>
<th>Unit</th>
<th>Value</th>
<th>Not Done</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Hemoglobin</td>
<td>g/dl</td>
<td>__ __</td>
<td>Not Done</td>
</tr>
<tr>
<td>2. Hematocrit</td>
<td>%</td>
<td>__ __</td>
<td>Not Done</td>
</tr>
<tr>
<td>3. RBC</td>
<td>M/mm³</td>
<td>__ __</td>
<td>Not Done</td>
</tr>
</tbody>
</table>

ID Assignment

- Should appear on every form (preferably page)
  - Links paper form with specific record in database
  - Multiple forms, “merge key” in database
- Must be UNIQUE for each subject
- May be a simple number 1001
- May be multi-part: 102101
  - 1 = Site
  - 02 = Language
  - 101 = ID

Inclusion of “Other (specify)”

- May cut down on items left blank
- Position “Other” last in list of possible responses to ensure all responses considered first
- Continue to monitor “Other” to ensure a common response category was not overlooked

Anticipated Responses Categorized

1. In what country were you born (check one)?
   - USA
   - Guatemala
   - Mexico
   - Dominican Republic
   - Other

Updated Due to Overwhelming Response

1. In what country were you born (check one)?
   - USA
   - Guatemala
   - Mexico
   - Dominican Republic
   - El Salvador
   - Other
In Summary, when designing questions:

- Avoid open ended responses
- Determine to whether question should be collected as "continuous" or "categorical"
- Consider all possible responses
- Make categories mutually exclusive
- Allow for unanticipated responses
- Put ID on every form/page
- Pilot your forms in the target population

Once you know what to collect...

- Decide how it will be collected
  - Paper
  - Electronic
  - Both
- If electronic, how?
- Who will:
  - Enter data
  - Handle data

Data Collection: Paper Or Paperless?

<table>
<thead>
<tr>
<th>METHOD</th>
<th>Scan/ FAX</th>
<th>CAPI* (laptop/tablet)</th>
<th>Kiosk/ Touch Screen</th>
<th>Hand Held/ Smartphone</th>
<th>ACASI† (Audio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAPER</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>PAPERLESS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

*Computer Assisted Personal Interview
†Audio Computer Assisted Self Interview
Electronic Data Capture

Advantages
- Cleaner data at entry (required fields, skips, ranges)
- More efficient for larger studies
- Electronic data in real time (or close to it)
- Eliminate data entry/shipping costs
- Data can inform next visit even for short follow up

Disadvantages
- May entail increased upfront programming costs
- Additional training of staff
- Increased equipment costs
- Infrastructure (software versions, internet connection, back-up equipment)
- Data security

Types of electronic data capture
- Local device data capture (data saved on local machine/tablet)
- Web-based data collection (data saved on central server)
  - Desktop computer
  - Laptop/Tablet/PC
- Handheld devices (HAPI, smart-phones)
- Faxed/scanned data forms system
- Others

Consider languages when selecting data entry methods and software

Dropdown Menu for Codes
Timely Submission of Urgent Data

Consider who is Completing a Form

Advantage of Web-based Randomization

Find out what may be available to you

A Word About “Canned” Software

- There are many types of “canned” or commercially available software available
- No single “best” choice
- Cost can vary widely
- Database structure can vary
- Do your homework to make sure what you get will work for your project

Consider the Database Structure

- Relational database:
  - each form is a record/row
  - Can be queried in real-time
  - Better choice for data management
- “Long Skinny” file:
  - each variable is its own short record/row,
  - difficult or impossible to query in real-time
  - Good choice for data capture
### Example of “Relational database”

<table>
<thead>
<tr>
<th>ID</th>
<th>VisDate</th>
<th>Sex</th>
<th>Race</th>
</tr>
</thead>
<tbody>
<tr>
<td>1001</td>
<td>09/09/2011</td>
<td>F</td>
<td>AA</td>
</tr>
<tr>
<td>1002</td>
<td>09/08/2011</td>
<td>M</td>
<td>W</td>
</tr>
<tr>
<td>1003</td>
<td>09/08/2011</td>
<td>F</td>
<td>AA</td>
</tr>
</tbody>
</table>

“One” row per subject

“Many” rows per subject

Best structure for real-time querying of data. Can be structured this way in most database packages including: Oracle, SQL, MS Access

### Example of “Long Skinny”

<table>
<thead>
<tr>
<th>ID</th>
<th>Date</th>
<th>VarName</th>
<th>VarType</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1001</td>
<td>09/09/2011</td>
<td>ID</td>
<td>Numeric</td>
<td>1001</td>
</tr>
<tr>
<td>1001</td>
<td>09/09/2011</td>
<td>AgeDate</td>
<td>Date</td>
<td>09092011</td>
</tr>
<tr>
<td>1001</td>
<td>09/09/2011</td>
<td>Sex</td>
<td>Numeric</td>
<td>2</td>
</tr>
<tr>
<td>1001</td>
<td>09/09/2011</td>
<td>Race</td>
<td>text</td>
<td>AA</td>
</tr>
<tr>
<td>1001</td>
<td>09/09/2011</td>
<td>Lesion_mm</td>
<td>numeric</td>
<td>6</td>
</tr>
</tbody>
</table>

... One row per variable:

Many “canned” web software packages use this structure including: Survey Monkey, REDCap, StudyTrax

### REDCap “Long Skinny”

<table>
<thead>
<tr>
<th>PI ID</th>
<th>Study ID</th>
<th>Partic. ID</th>
<th>Date</th>
<th>VarName</th>
<th>VarType</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12345</td>
<td>44556</td>
<td>1001</td>
<td>09/09/12</td>
<td>ID</td>
<td>Numeric</td>
<td>1001</td>
</tr>
<tr>
<td>12345</td>
<td>44556</td>
<td>1001</td>
<td>09/09/12</td>
<td>VisDate</td>
<td>Date</td>
<td>09092011</td>
</tr>
<tr>
<td>12345</td>
<td>44556</td>
<td>1001</td>
<td>09/09/12</td>
<td>Sex</td>
<td>Numeric</td>
<td>2</td>
</tr>
<tr>
<td>12345</td>
<td>44556</td>
<td>1001</td>
<td>09/09/12</td>
<td>Lesion_1</td>
<td>numeric</td>
<td>6</td>
</tr>
</tbody>
</table>

... Includes Investigator and Study code

<table>
<thead>
<tr>
<th>PI ID</th>
<th>Study ID</th>
<th>Partic. ID</th>
<th>Date</th>
<th>VarName</th>
<th>VarType</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12345</td>
<td>44556</td>
<td>1001</td>
<td>09/09/12</td>
<td>ID</td>
<td>Numeric</td>
<td>1001</td>
</tr>
<tr>
<td>12345</td>
<td>44556</td>
<td>1001</td>
<td>09/09/12</td>
<td>VisDate</td>
<td>Date</td>
<td>09092011</td>
</tr>
<tr>
<td>12345</td>
<td>44556</td>
<td>1001</td>
<td>09/09/12</td>
<td>Sex</td>
<td>Numeric</td>
<td>2</td>
</tr>
<tr>
<td>12345</td>
<td>44556</td>
<td>1001</td>
<td>09/09/12</td>
<td>Lesion_1</td>
<td>numeric</td>
<td>6</td>
</tr>
</tbody>
</table>

... One huge table with multiple investigators and studies

### Consider the Database Structure

- For straight electronic data capture underlying structure may not matter
- If you want an “intelligent” e-form with sophisticated checking or custom error and warning messages, database structure does matter

### Paper/Electronic Hybrid Systems: Optical Character Recognition software (scan/fax)

- Data collected on paper “TELEForm”
- Form scanned/uploaded or faxed to processing center
- Software “reads” forms and enters data into a database
- Questionable characters are set aside for manual review
- “Verifier” may be customized for each form
  - Different level can be set for various fields
  - 100% for key fields or hand written fields

### Optical Scanning/Faxes

#### Advantages
- Don’t need constant internet access
- Easy to train clinical staff
- Relatively inexpensive
- Shorter time between data collection and inclusion in database

#### Disadvantages
- Software costs, skills
- Not practical for text or handwritten data
- More sensitive to quality of forms
Technology Changes: A word of Caution on Handheld Devices

A Caution About New Technology

A Word of Caution on Smart-phones

- Encryption can be difficult (or impossible)
- Small screens make it difficult to view some question types
- Navigating around questionnaire (going back) is challenging
- Battery life is short (need to recharge frequently)
- Target for theft

Beginning A Study: Quality Control Systems

Public Health Model:
- Prevention
- Detection
- Treatment
Reports

- For the study team to manage the project
  - Screened, eligible enrolled
  - Key demographics (by randomization group)
  - Follow up rates
- For the study staff to help them manage components (e.g., call lists, follow up visit schedules)
- For Data managers to identify data problems

Tracking: Reports

- Run regular reports of information collected in tracking database
  - Subjects
  - Forms
  - Follow ups
  - Key variables (when applicable)
  - Missing:
    - Visits
    - Forms
    - Data elements

Sample Follow Up Report

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screened</td>
<td>112</td>
<td>57</td>
<td>169</td>
</tr>
<tr>
<td>Eligible</td>
<td>160</td>
<td>104</td>
<td>264</td>
</tr>
<tr>
<td>Received</td>
<td>74</td>
<td>47</td>
<td>121</td>
</tr>
<tr>
<td>Backordered</td>
<td>74</td>
<td>47</td>
<td>121</td>
</tr>
<tr>
<td>Randomized</td>
<td>60</td>
<td>61</td>
<td>121</td>
</tr>
<tr>
<td>Gender Male</td>
<td>37</td>
<td>37</td>
<td>74</td>
</tr>
<tr>
<td>Gender Female</td>
<td>23</td>
<td>24</td>
<td>47</td>
</tr>
<tr>
<td>P2O 10</td>
<td>27</td>
<td>28</td>
<td>55</td>
</tr>
<tr>
<td>P2O 11</td>
<td>13</td>
<td>53</td>
<td>66</td>
</tr>
<tr>
<td>Sex 1</td>
<td>39</td>
<td>50</td>
<td>117</td>
</tr>
<tr>
<td>Sex 2</td>
<td>13</td>
<td>8</td>
<td>21</td>
</tr>
</tbody>
</table>

Reports: Visual as well as Tabular

U19 Study: Enrollment Reports as of 10/15/2012

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>MALE Screened</td>
<td>510</td>
<td>100.0</td>
</tr>
<tr>
<td>Eligible</td>
<td>150</td>
<td>23.4</td>
</tr>
<tr>
<td>Total</td>
<td>550</td>
<td>6.8</td>
</tr>
</tbody>
</table>

Actual vs. Targeted Enrollment

Actual and Targeted Subjects by Randomization as of January 01, 2012
Tracking the Data
- Identify what data have been collected
  - For each Subject
  - For each Visit
    - Questionnaires
    - Exams, images
    - Labs results, specimen shipping
    - Other data elements

Tracking the data (continued)
Record:
- What data will not be collected
- What data have been received
- What data have been entered
Create:
- Schedules
- Reports

Form Shipment and Receipt
Track what will NOT be done

Real Time Data Cleaning
- Database updated as soon as form “Submitted”
- Reports and queries can be run in real time (as opposed to “freezing” the database and running reports at specified intervals)
- If data irregularities found, can notify sites and correct immediately
- Can identify missing data while there may be an opportunity to collect it

Data should be reviewed for accuracy prior to entry whenever possible
Look at the Data Early and Often

- You cannot fix a problem if you don’t know it exists
- Get data into electronic format ASAP so it can be more easily reviewed
- Monitor the first few and participants
- Ongoing audit percentage of forms
- Pay extra attention to key variables

Do simple checks

- Frequency (count) and distribution (range) of each and every variable
- Do crosstabs of variables where appropriate
- What is missing?
- What is out of range?
- What contradicts (e.g., pregnant males)
- Are there systematic problems?

Perform Systematic Data Audits

- Data forms and source documents are compared with database on X % of forms
- Set an “acceptable” error rate. For example:
  - 0.5% overall
  - 0.1% for key fields
- If audit yields a larger error rate, you must check and correct the database
Pay Extra Attention To Key Data

Be sure to pay particular attention to key data points where applicable.
- Query all SAE’s?
- Query all entries of crucial variables (e.g., study outcome)
- Extra attention to problematic variables (e.g., time-line-follow-back)

Data Cleaning: Essentials For Success

- Clean data in stages:
  - “Freeze” dataset for interim analysis (DSMB)
  - Subsequent cleaning of the data will be from that date forward
- The programmer must be familiar with the CRF
- The investigator or someone who really “knows” the study and the data must be involved in setting cleaning parameters and making decisions on what is invalid

Document, Document, Document!

Once you have identified errors in the data, be sure to document:
- All instances of errors
- All edits and corrections of the data
- History of manipulations, modifications, corrections to files/variables
- Location, type of media storage
- Archival procedures

Take Home Message

- Budget appropriately
- Be careful and be accurate
- Double and triple check the data
- Bring problems to the attention of study staff or PI right away
- Learn from your/other’s mistakes
- If you do things right it’s less work and you are more likely to discover the truth at the end

Data Security - General

- Keep paper records should be kept in locked cabinets and/or offices
- Store identifiers like names and addresses separate from clinical data
- Keep particularly sensitive data apart from other identifiers (e.g., SSN) – in a separate file, by ID
- Do not collect sensitive data unless you really need it
Data Security - Hardware
- Password protect all computers
- Set to automatically timeout if inactive
- Encrypt laptops, flash-drives and other storage devices when possible
- Do not put identifiable data on portable media (e.g., CDs, flash-drives) unless password protected, preferably encrypted

Data Security - Electronic Data
- Make sure web and database servers are behind firewalls
- Encrypt all data transmissions from data collection point to servers (e.g., SSL)
- If sensitive fields must be collected, (e.g., SSN) encrypt them
- System users should have own logins and be instructed not to share usernames and passwords

For more information...
- Contact the CTSI
  - See the CTSI website: [http://ctsi.bu.edu/](http://ctsi.bu.edu/)
  - Attend a CTSI drop-in session
  - Send an email: [ctsi@bu.edu](mailto:ctsi@bu.edu)

- Contact the DCC
  - See the website: [http://sph.bu.edu/DCC](http://sph.bu.edu/DCC)
  - Send an email: [chaisson@bu.edu](mailto:chaisson@bu.edu)