Data Managers Interest Group Meeting Minutes
March 25, 2014, 12:30-1:30pm
Wood Basic Science Building, West Lecture Hall

Content: Vice Dean Daniel Ford, ICTR PI and Data Management Advisory Board Chair, opened the meeting by reiterating Vice Provost for Research, Denis Wirtz's desire to establish University-wide data standards. Currently, there is an intense need to improve both data security within Hopkins, as well as data that travels outside of the Institution. It is imperative to receive input from the data management community, from the people who work with data here at Hopkins.

Next, Scott Carey, Data Management Advisory Board member, began an open discussion by highlighting areas we might want to consider when creating some type of data standards or formal policy: data organization; storage and backup; security and access controls; documentation; and metadata. These labels do not encompass all of the areas that need to be addressed at a University level, but it's a good starting point for a conversation. Scott also shared links to other institutions' data policies and practices, which appear at the end of this document. It's important to look at these other institutions, so we don't replicate work that has been already done elsewhere.

When Scott opened the floor for discussion, the group began talking about data sharing policies. It's clear that there are many ways of going about this at Hopkins. There are secure methods, like using locked databases and locked share drives. Likewise, some people like using Redcap for its high level of security. However, there are instances where this means having to put every study team member in eIRB. Somewhat related, one interest group members commented that having the denotation

of "data analyst" on eIRB has been very encouraging. That being said, not all PIs list a data analyst on every protocol. Should this be a standard? Should this require a policy? It should be noted that not all protocols require explicit roles to be named. For example, imaging studies don't require a radiologist to be listed on eIRB.

Next, Sharon Ghazarian, Data Management Advisory Board Co-Chair, asked the group about training. Interest group members expressed a definite need for it. Some type of data management best practices certificate would certainly be helpful for hiring purposes. In general, junior faculty may be less aware of current data management practices. In turn, these faculty members are training recent college graduates, who also lack knowledge of this important information. Likewise, many research coordinators are uninformed about data management. If training became required by the IRB, how would it be enforced – random audits; monitoring; peer review/reporting?

Continuing, the conversation turned to the creation of training modules. PIs are already required to know so much about the protocol and specific grant mechanism; should they really take this on too? Should this even be their responsibility? Good data management practices start from the ground up. Perhaps the most we can strive for is high standards for comprehensive data management plans, which should be written before beginning research.

One area that could use attention, is the issue of petty cash vouchers, and the required social security number. Currently, petty cash voucher PDFs are uploaded to SAP, which could easily be hacked. New practices, such as getting an electronic signature, could improve this area of concern.

Next, Kay Dickersin asked about the recent developments with the IOM, and open access to people's data, which will soon be available by request. The IOM is currently working on the principles around this new development on their website. Jon Petters and Betsy Gunia from the Sheridan Libraries, have begun to work on some training materials regarding the de-identification of data while retaining coordinated utility. This might be a good topic for a subgroup, or workgroup to focus on. Investigators might be more willing to invest in data management at the beginning of their research, if they know their data is going to be released at the conclusion of the study.

Other areas to focus on might be the data management practices involved in mult-site studies; sharing data in secure methods of communication and the risks involved with "data in motion"; and CRMS. These efforts must involve all of the Hopkins Schools, not just Medicine and Public Health. Dan Ford reminded the interest group that they have been brought together in an effort to "professionalize" data management. We want to empower data managers. If there are certain unethical or unsafe practices going on, data managers have the right to halt research, and to do so anonymously, although retaining anonymity might be impossible in many cases. This interest group should help to inform the greater Hopkins community about data management practices and procedures, and to encourage others to have high standards when it comes to handling data.

Once workgroups are established, a meeting will be scheduled for them to report back to the larger interest group. An open dialogue should exist between advisory board, subgroup, and interest group members. In the coming weeks, a survey link will

be sent to interest group members, which will contain questions used to acquire basic information from individuals concerning job responsibilities and utilities; feedback from the March 25th discussion; and to collect volunteer information.

Links to Data Management Best Practices and Policies:

• DataONE (Data Observation Network for Earth) Best Practices Data Life Cycle: www.dataone.org/best-practices

• Oxford University diagram on Good Practices in Research Data Management: http://www.databaseanswers.org/data_models/research_lifecycle/index.htm

• California Digital Library

Manage Your Data: http://www.cdlib.org/services/uc3/datamanagement/index.html

• Colorado Clinical and Translational Sciences Institute (University of Colorado Denver) Managing Your Research

Data: http://cctsi.ucdenver.edu/RIIC/Pages/DataManagement.aspx