

Complexities of Conducting Cross-Disciplinary Biomedical Research

Jennifer Larsen, MD, Vice Chancellor for Research
W. Scott Campbell, PhD, Senior Director of Research and IT
University of Nebraska Medical Center

Cross disciplinary research has become routine at academic health centers because larger teams with a broader range of skills are needed to solve complex health related problems. Researchers routinely reach out to colleagues to understand how their work “at the bench” is, or could be, relevant to future clinical care, as well as, how to better incorporate what has been learned in a clinical trial into the community. To build effective teams, there are many “complexities” that must be anticipated and/or addressed.

The Complexities

Some of the common complexities involved in conducting cross disciplinary biomedical research are outlined below.

Defining the Rules of Engagement

To form an effective team requires time and discussion. Even as the team is assembled and before the data is gathered, teams should discuss rules on how the data will or will not be shared with others, could or could not be moved, and how each member will be acknowledged for their role in any published results. Who will write or lead each manuscript or grant requires frank discussions long before they are written or submitted.

Vocabulary

As simple as it sounds, to form a functional team, an environment must be created where all members can be understood, and their ideas are welcomed. This starts with encouraging everyone to speak in words most can understand by avoiding terminology and acronyms specific to one discipline that might not be understood by others. This is also important if teams want to ever attract new members, including students. Vocabulary goes beyond conversations. It also addresses how the data is captured and stored so that the data can be more easily

shared. Using a common format, preferably an established vocabulary standard (e.g., SNOMED, LOINC) that includes meta-data to allow members to understand how the fields are defined, for more consistent and reproducible data collection, as well as queries and analyses, or to combine with other data sets. Team members with terminology expertise are very valuable, and team members who can translate between disciplines are essential to an effective multidisciplinary team.

Data Transfer and Storage

Many teams require data to move from one place to another, like from a research instrument or electronic health record to a data storage space or a research database where the analysis will be performed. More teams are working with large research files, terabytes or more, such as DNA sequencing data or image files (e.g., MRI or other anatomic imaging files) or data from large populations. These data sets often have to be stored in the cloud or in large data centers able to accommodate such large data, but moving files can be time and resource consuming. Large data sets often must be stored in their entirety at the point of creation until the full copy of the dataset is transported, stored and validated for completeness in its new lo-

cation. Often discussions arise regarding “whose data is it”. Differences of opinions need to be ironed out, including what federal or commercial entity rights or patient/research subject perceptions that might be involved. The cost of data storage is often underrecognized as well. How that cost will/will not be subsidized by the team members or grants must be determined and use of data steward(s) (e.g., personnel) to maintain and distribute data sets may be necessary to include in the cost structure.

Privacy and Security

The data storage vehicle depends in part on what data is being stored. Protected health information (PHI), Protected individual information (PII), as well as other sensitive data (e.g., student data, high security data) may require special controls for who can access the data and the ability to audit who has accessed the data. Data associated with an FDA application or trial needs to meet FDA’s Title 21 Code of Federal Regulations (CFR), Part 11 requirements. Many researchers are not as knowledgeable as they should be of the eighteen PHI identifiers defined by the HIPAA legislation (<https://www.hhs.gov/hipaa/index.html> and shown in Table 1). As a result, researchers incorrectly believe their data is deidentified and attest as such, when the data is, in fact, still considered identifiable.

Special Considerations with Global Sites, Teams or Focus

There are increasing, and often changing rules, when data, samples, equipment, or team members move between or live in other countries. Countries have varying “export control” regulations concerning what is or is not allowed to cross into or out of their country. For the US, this includes interactions with specific individuals whether located in another country or in the US and specific types of equipment. For many countries, this

involves export of biologic samples or data. In particular, the new European Union Data Protection Regulation (EUD-PR) introduced in 2016 requires anyone acquiring data in an EU country, even if acquired on the property of and from citizens of another country to meet specific standards and receive EU approval. These standards further apply to data moved from an EU country.

Table 1: 18 PHI identifiers

- Names
- Dates, unless year alone
- Telephone numbers
- Geographic data (address, full zip)
- FAX numbers
- Social Security numbers
- Email addresses
- Medical record number
- Health plan beneficiary numbers
- Account numbers
- Certificate/license numbers
- Vehicle identifiers and serial numbers including license plates
- Web URLs
- Device identifiers and serial numbers
- Internet protocol addresses
- Full face photos or comparable images
- Biometric identifier (i.e. fingerprint)
- Any unique identifying number or code

Problem Solving

Teams will always encounter problems including personality disputes, intellectual property disputes, ‘I contributed more than you did’ disputes, and ‘but you promised me’ disputes, among others. Ideally, the team would have discussed potential conflicts as the team is developed, including how the team would anticipate solving conflicts and identify a structure, process, or person(s) within or outside the team to resolve disputes if

additional help is needed. Such a resolution strategy is often required in multi-PI grants by funding sponsors. These kinds of agreements are best documented and discussed for new teams so there is no misunderstanding later. If the team has not had such discussions, team leaders or members may need to reach out and find the best mediator after the fact, such as the research integrity officer or another senior leader that all parties agree to listen to for dispute resolution. Team leaders should be proactive—watching for signs of frustration or conflict and address issues before they become impossible to resolve. Michelle Bennett, PhD, of the National Cancer Institute (NCI) assembled a Field Guide for Collaboration and Team Science available on line that provides many practical approaches to common problems (<https://www.cancer.gov/about-nci/organization/crs/research-initiatives/team-science-field-guide/collaboration-team-science-guide.pdf>).

Discussing Data Sharing with the Public

As time has passed, the public has seen more and more examples of times their data has been shared or “leaked” that they were not aware could occur. Many investigators believe sharing of de-identified data is acceptable and may even assume that no one would care. In fact, many individuals are comfortable with researchers sharing their personal data, even identified data, if they are informed in advance and have given their permission, such as through informed consent or broad consent. However, others may feel differently, even if their data is deidentified, hence the recent class action lawsuit or patients who objected to the University of Chicago Hospital who gave deidentified health data to Google as part of an artificial intelligence project. These concerns can be proactively addressed through the informed

consent document, town hall meetings, or other public discussions about the importance of the study and what the study is supposed to accomplish, or having a community advisory board of community leaders to be a sounding board about the methods to be used and help the investigators disseminate the results when they are found. As researchers, if we do not have the public’s trust, we may not have funding long-term. We can all do a better job of discussing the value of the data and the project with the public and working with community leaders to implement the data into day-to-day healthcare or other outcomes.

Summary

Team science is here to stay, and curating the datasets assembled by those teams needs to be discussed in advance. This is just one aspect of the complexities of conducting cross disciplinary biomedical research. Data sharing is usually good, often mandated by some funding mechanisms, is essential to multidisciplinary collaborations, and may result in even bigger datasets which can make moving and storing the data more challenging. The nuts and bolts of achieving data sharing, which may include data deidentification, moving large data sets, or loading data into a specific website can be confusing at best, and often difficult, as well. Data sharing may require new tools that are often developed with biomedical informatics experts, who are in too short of supply. Data sharing can create new risks if those researchers who are sharing data are not aware of the pitfalls, particularly when PHI is involved. But lastly, we cannot forget the public, who needs to be part of the communication before data is shared and after, to bring them along, to understand the value, and to fully understand and make use of the results that are found.