Evaluation/metrics of success Breakout Session
BD2K AHM Friday, November 13th 2015

Title: Evaluation Break Out Group – BD2K ALL-Hands Meeting, November 13, 2015
Co-Moderators: Rochelle Tractenberg, Georgetown, and Valerie Florance, NLM. Scribe: Ebony Hughes, NLM

Attendees: Of the roughly 25 attendees, about 8 were NIH program staff, the rest were BD2K awardees or colleagues. Represented were Career Awards, Centers of Excellence, Targeted software, R25 courses, training programs, LINCS-Perturbation Center, BioCaddie, and others. Incomplete list of attendee names/organizations at the end of this document.

Initial Statement of Purpose (provided by the moderators in initial draft of Breakout Session document, abstracted here): In this session we hope to initiate the development of a rubric that can be useful for applicants and awardees to gauge whether, how, and to what extent the proposed or completed work is: (1) successful based on its stated goals, and (2) aligned with the BD2K driving strategies. While it is essential to capture and communicate the extent to which funding has been successful, it is equally important to be able to identify where challenges remain (or arose), so that they can be addressed. Metrics of success should incorporate concrete and consistently evaluable (i.e., valid and reliable) indicators of whether - and specifically how - training, research, and infrastructure projects are aligned with each of these driving strategies, as potential measures of strategic impact.

Actual Breakout Session Activities: The moderators created two activities to support the stated purposes, one was to elicit a list from all participants of the actual metrics for success or evaluations that they had included/were using for their funded BD2K work. We planned to ask participants to align those actual metrics with the BD2K driving strategies in order to determine which of the strategies were best, and which less well, supported or addressed by evaluations and metrics of success that are currently being used. The second activity was to discuss the structure of a “stream”, where activities along the path from “big data” (BD) to (2) “knowledge” (K) could be classified. The purpose of this stream is to help the BD2K initiative to define what is meant by “upstream” and “downstream” effects of either the initiative, the funded projects, or some combination of these for future work. With a conceptual model of a “stream”, success and evaluation metrics could then be classified or used according to where in the stream the work was being proposed or executed. However, because of time constraints, we only achieved the first part of the first activity – the elicitation of the metrics of success and evaluation methods from the participants in the room. Those who did not have BD2K grants contributed either metrics they had used, or metrics they would like to see.

Results: Based on the actual session, the following summary was prepared by Valerie Florance and reported back to the full AHM group. The overall message from the breakout session was agreement that it would be helpful to build metric counters into every tool and resource, so as to provide automatic reports on quantitative measures to include size, scope, and use (if possible). The group was also fairly unanimous in their expressions that these measures required supplementation with qualitative assessments, which were described as equally important – but essentially nonexistent at this point. A partial list of examples of additional ways to document “success” for BD2K grant projects included:

- Replication tests
· User: surveys, demographics, geographic reach, including international, satisfaction as well as use and impact data
· Successful use cases
· How many new collaborations (are/were supported) and the level of interdisciplinarity in those collaborations
· Influence networks for cited data objects
· New non-BD2K grants BD2K PIs or new non-BD2K grants that use tools, resources or concepts developed with BD2K support
· New businesses, new therapeutics, technology transfer
· Impact on scientific methods and practices – use expert panel or scientific study of diffusion of innovation
· Improvements made by users
· Cost effectiveness of training or new process
· Qualitative as well as quantitative.

**Recommended Outcomes or Activities:** Based on the breakout session discussion and results, the moderators developed the following recommendations:

1. Go through funded projects in each strategy area to obtain a complete list of proposed/actual metrics of success and evaluation. Then, obtain feedback from the awardees about changes to these metrics that they might have implemented since writing the proposal. Use this to create a master list of metrics for success or evaluation that BD2K grantees are using.

2. Some participants seemed interested in a workgroup on evaluation, to continue to refine ideas initiated in this breakout session. Determine whether and how this would best be done – crosscut the centers and training and software groups somehow? Use the coordination centers by e.g., having them reach out to the BD2K awardees whose activities they are coordinating?

3. Foster discussion among interested parties, via the coordination centers, on the concept of the stream – what does it contribute to evaluation – is it a meta analytic tool for long term or post-program evaluation, and/or could it be useful earlier (i.e., now) in strategy areas?

4. Develop consensus around a minimal set of metrics for different strategy areas.