BIOTRAIN 753 "Data Management and Quality for Biomedical PhD Students"

Overview:

This course is developed to engage Biomedical Ph.D. trainees in continued Responsible conduct of research, Rigor and Reproducibility RCR/R&R training, an emerging mandate from the National Institution of Health (NIH) and future requirement for NIGMS T32 Pre-Doctoral Training grants. The course is directed by the RCR Curriculum Manager in the Office of Biomedical Graduate Education (OBGE) in Duke School of Medicine and is offered via Duke LMS (Learning Management System). In Years 2 and 3, PhD students in the School of Medicine are required to take this course comprised of 3 online interactive modules. Each module is accompanied by an assessment. This course is presented with interactive graphics, text-based activities, short videos, and discipline specific scenarios. Data Management and Quality for Biomedical PhD Students includes 3 modules:

1. **Research Quality and Reproducibility**
   Learning outcomes:
   - Discuss the issue of research reproducibility from your personal perspective, from the perspective of your peers, and from the perspective of research stakeholders.
   - Explain how meta-science contributes to our understanding of research quality and reproducibility.
   - Discuss the terminology used to describe research metrics.
   - Understand the perspective and role of research stakeholders as strategies are developed to address research reproducibility.
   - Articulate why training in research quality and reproducibility is important.
   - Identify general risks to research quality and reproducibility.
   - Explain how understanding the basis of irreproducible research can inform and advance science.
   - Summarize the current concerns (causes, costs, consequences) associated with research rigor, quality, and transparency.

2. **Data Management**
   Learning outcomes:
   - Recognize the complexities of the research data ecosystem, including a wide range of data types, sizes, sources, and formats.
   - Describe the stages of the data life cycle and corresponding data stewardship practices in the context of documentation and training, organization, and governance and ownership.
   - Explain how data security regulations may affect research quality and reproducibility.
   - Identify best practices for management of small datasets in spreadsheets.
   - Discuss special considerations and common pitfalls in managing image data.
   - Recognize best practices for the sharing and archiving of data.

RCR topics shown in red; R&R topics shown in blue; Topics that integrate both elements shown in green.

- Identify funder requirements for formal data management plans and data sharing.

3. Data and Resource Sharing

Learning outcomes

- Articulate the importance of sharing and critically evaluating data and data-related resources.
- Identify stakeholder requirements and regulations for data and resource sharing.
- Identify the pros and cons of data sharing.
- Explain the importance of methodological transparency in the effective reuse of data.
- Describe the ecosystem of resources for post-publication data sharing.
- Discuss current advocacy for including incentives for data sharing.

RCR topics shown in red; R&R topics shown in blue; Topics that integrate both elements shown in green.