October 14, 2019

Xujing Wang, PhD
National Institute of Diabetes, Digestive and Kidney Diseases
National Institutes of Health
Bethesda, MD 20892

Dear Dr. Wang,

The Endocrine Society appreciates the opportunity to comment on bioinformatics and computational needs for investigators engaged in diabetes, endocrinology and metabolic diseases research. Founded in 1916, the Endocrine Society is the world's oldest, largest, and most active organization devoted to research on hormones and the clinical practice of endocrinology. Our membership of over 18,000 includes clinical, translational, and basic scientists studying complex endocrine pathways, systems, and diseases. Given the growing use of large datasets in endocrine research, we have enormous needs for advanced bioinformatics tools and other resources to facilitate data analysis. We welcome NIDDK's willingness to identify and address the data-intensive needs of the biomedical research community. In our response, we identify specific challenges our members face and propose recommendations for NIDDK to consider as it addresses the bioinformatics needs of NIDDK-funded investigators.

We were pleased to see that NIDDK is interested in addressing challenges by connecting researchers and enhancing training. Our members report that one longstanding barrier to overcoming many of the issues highlighted in the RFI is difficulty in identifying and engaging with expert personnel to help with the analysis of voluminous bioinformatics data resulting from investigations of complex systems. To address this, NIDDK could support the development of detailed guidance on how bioinformaticians and statisticians can use different data sources e.g., on dknet.org, and current accepted expectations for analysis and data visualization. This would allow scientists to more efficiently identify and approach professionals earlier in the planning stages of a project based on their expertise. Additional training opportunities are importantly needed to build the bioinformatics workforce and ensure that the broader biomedical research community can generate, evaluate and reproduce research that require bioinformatics approaches. NIDDK could also invest in mechanisms to give investigators protected time (e.g., sabbaticals) at all career stages for training opportunities in bioinformatics or develop a Scholars program for masters-level bioinformatics training. As an example, NCATS programs that provide support for Masters degrees (e.g., TL1 Clinical Research Training Awards) could be a comparable initiative that provides an implementable framework should NIDDK pursue this concept.

We recognize the importance of good curation of data and the need to promote common data elements that can allow for integration of multiple data types. In the field of endocrinology, we sometimes lack gold-standard analytic approaches or a clear consensus for how best to define common data elements. For example, PCOS is a complex condition with a clear genetic component where best practices for analysis of whole-genome data would be extremely helpful. NIDDK could address these challenges by supporting the development of consensus approaches to data analysis and common data elements for diabetes and other endocrine diseases. These consensus approaches and best practices should be clearly communicated to grant reviewers to ensure that they can properly evaluate grants using these techniques.

Our members also report challenges related to the availability and ease of access of resources; bioinformatics core services are often overbooked due to both the frequency of requests, the amount of work required to address requests and the relative dearth of qualified bioinformaticians. For well-curated databases integrating multiple data types, access can take months if not years. Centralized or regional resources and easier access to such resources would be of immense benefit to researchers. We therefore recommend that NIDDK provide funding for programs that enable pooled financial resources to support service cores for detailed consultations and regional resources that would serve multiple institutions. Additionally, funding is needed for infrastructure necessary to handle the immense data needs along with new rules and best practices for handling computational demands.

In Conclusion

We appreciate NIDDK's recognition of the need for resources to address the analysis of complex data generated by high-dimensional and high-throughput technologies. To help address these needs, we recommend that NIDDK

- Focus on connecting scientists and experts in different types of data analysis through enhanced educational opportunities.
- Develop guidance for the community on consensus approaches and common data elements.
- Provide funding for service cores and regional resources, including computational infrastructure.

Thank you for considering the Endocrine Society's comments. If we can be of any further assistance, please contact Joe Laakso, Director of Science Policy at jlaakso@endocrine.org.

Sincerely,

E. Dale Abel MD PhD

President Endocrine Society