AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®
732 North Washington Street, Alexandria, VA 22314-1943
(703) 684-1221 ▪ Fax: (703) 684-2037 ▪ Email: *asainfo@amstat.org*
Web: *www.amstat.org* *www.facebook.com/AmstatNews* @AmstatNews

October 20, 2015

Dr. Josephine Briggs
Interim Director, NIH Office of Precision Medicine
31 Center Drive, Room 2B11, MSC 2182
Bethesda, MD 20892-2182

Dear Director Briggs,

I write to you in your position as head of the NIH Precision Medicine Initiative to urge you to consider the essential role of statistical science in precision medicine when reviewing requests for applications (RFAs) for the Precision Medicine Initiative. Fundamentally, precision medicine relies on data and statistical methods to identify micro-groups of patients for tailored treatment.

Validation and quantification of the expected benefit to these selected individual patients is a statistical issue. The central role of statistical science is amplified by the need to seamlessly integrate electronic health records, mobile health data, and clinical data with high-dimensionality from technologies such as genomics, imaging, and related intensive measurement devices. Such integration leads to challenges in inference and inferring causality. In this setting, the role of statistical science is not simply to implement the technical execution of precision medicine, but, more importantly, to provide strategic guidance for prioritization and validation, grounded in robust evidence. Precision medicine teams funded by the NIH will generally yield better science and more reproducible results if they incorporate senior team members with formal statistical expertise.

As described in the American Statistical Association one-pager, "Statistics and Precision Medicine,"[1] there are many open statistical challenges in precision medicine. Addressing these challenges will be essential to its success. Precision medicine teams that address these challenges in an interdisciplinary way are far more likely to be successful. Examples include the development of methods to incorporate rigorous statistical analyses of large data sets into decision analyses tailored to the individual patient; methods to validate risk prediction models; statistical foundations for drawing rigorous inferences from electronic health records and other large data sets; data sharing approaches that facilitate learning from the small and highly dispersed set of cases similar to the patient for whom a decision is to be made; methods to

[1] http://www.amstat.org/policy/pdfs/PrecisionMedicineOnePager.pdf , also enclosed

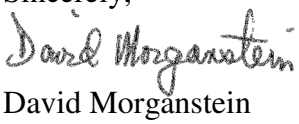
efficiently monitor and validate the precision of laboratory diagnostics and diagnostic imaging; and statistics curricula that teach the modeling and analysis procedures specifically addressing patient care and broader health outcome issues under increased emphasis on individual patient data.

In addition to addressing these scientific questions, statisticians can and should play critical roles within precision medicine teams. Statisticians can uniquely contribute to the development of study designs of precision medicine projects—to the development of rigorous approaches to optimally collect and use data to inform which treatment is best for a given individual at a given time under given circumstances. Statisticians, with their formal training in inference, are uniquely able to assess false positive and false negative rates, to design rigorous validation studies, and to develop methodologies to preserve data confidentiality. In addition, including formal statistical training on the study team will help with translating a medical question into a precise data-based question. This includes carefully describing data structures, the underlying system that generated the data, and how this system enhances/limits what the team aims to assess or predict; the development of data analytics for health scientists to test and further develop theories to yield more informative diagnostics and more effective interventions; and the enhancement of communication across disciplinary boundaries such as the computational, health, and behavioral disciplines.

To further bolster that statistics helps make the science better, I point out that *Nature's* list of 100 most-cited research articles includes nine articles by statisticians.[2] In another recent *Nature* article, the author discusses the role of statistical methods to alleviate some of the common human biases while making objective inferences.[3]

In closing, as NIH Precision Medicine RFAs are written, I urge consideration be given to the benefits of the engagement of statisticians as co-PIs and key personnel on project teams, to ensure NIH is funding the best possible science.

Sincerely,



David Morganstein
President, American Statistical Association

[2] http://community.amstat.org/browse/blogs/blogviewer?BlogKey=4b877e66-57e6-42f6-bee9-cb2058986fef&tab=recentcommunityblogsdashboard; http://www.nature.com/news/the-top-100-papers-1.16224
[3] http://www.nature.com/news/how-scientists-fool-themselves-and-how-they-can-stop-1.18517

Statistics and Precision Medicine

What is Statistics and What is Precision Medicine (PM)?

- Statistics is the science of learning from data, and accounting for relevant uncertainties. As such, it permeates the physical, natural, and social sciences, as well as public health, medicine, business, and policy.
- PM can be broadly framed as using a patient's specific characteristics in determining the best possible treatment for that individual at a given time under the given circumstances. PM complements traditional evidence-based medicine implemented by physicians by incorporating the increasing body of knowledge—including highly accurate laboratory and imaging diagnostics and epidemiological data sets—about individual forms of disease, individual responses to treatment, and new forms of intervention.

Precision Medicine Challenges of Statistical Nature

- Complex data challenges arise in a broad spectrum of PM activities, ranging from genomic medicine to comparative effectiveness research. As such, many of the statistical challenges of Big Data apply.
- When used for developing treatments and therapies, the analyses of large data sets should recognize that empirical correlations may not reflect cause and effect. The latter is investigated using a combination of substantive knowledge and statistical techniques called "causal inference".
- As each individual is unique, PM is inherently based on knowledge obtained from smaller pools of individuals, making inference much more difficult.
- It is essential to ensure the representative nature of a large cohort of subjects for precision medicine, in order for the results to generalize beyond the initial source of information.

Statistical Objectives in Precision Medicine

- Develop methods to incorporate rigorous statistical analyses of large datasets into decision analyses tailored to the individual patient.
- Develop better methods for validation of risk prediction models, and facilitate sharing of data so that investigators can validate their models.
- Develop solid foundations for drawing rigorous inferences from electronic health records and other large datasets.
- Develop data sharing approaches that facilitate learning from the small and highly dispersed set of cases similar to the patient for whom a decision is to be made.
- Develop methods to efficiently monitor and validate the precision of laboratory diagnostics and diagnostic imaging.
- Develop statistics curricula that teach the modeling and analysis procedures specifically addressing patient care and broader health outcome issues under increased emphasis on individual patient data.

Roles of Statisticians on Precision Medicine Teams

- Statisticians develop approaches to optimally collect and use data to inform which treatment is best for a given individual at a given time under the given circumstances.
- Statisticians translate a medical question into a precise data-based question. This includes carefully describing data structure, the underlying system that generated the data and what the team is trying to assess or predict.
- Statisticians develop data analytics for the health scientists to test and further develop theories in order to develop more informative diagnostics and more effective interventions.
- Statisticians enhance communication across disciplinary boundaries such as the computational, health, and behavioral disciplines.