**Investigator:** Dominic D. LaRoche

**Proposal Title:** Novel methodology for evaluation of next-generation sequencing measurements.

**1 Significance**

The accuracy and precision of many RNA measuring systems is largely unknown, despite the importance of

RNA quantification in modern biological research. We define an RNA measurement system as any sequence of procedures designed to quantify RNA from a biological sample and provide data for analysis. Our proposed research will provide a methodology for studying measurement error in NGS-based RNA measurement systems. Specifically, we will provide a method to directly compare the precision of any two RNA measurement systems using only technical replicates from each system. Even though RNA is widely studied, imprecise measurements may limit the quality of inference and slow the rate of scientific progress. Continued progress in fields that rely on NGS-based RNA measurement systems will depend on understanding how, and why, errors occur.

There is currently no robust methodology for comparing measurement precision between NGS-based RNA

measurement systems. The current approach for comparing measurement systems is to indirectly compare each system’s reproducibility and repeatability, an approach with limited utility (Lovell et al. 2015). Alternatively, others have used the final outcome measure, i.e. differential expression of genes, as a measure of performance. However, the outcome of an analysis is only a crude indicator of measurement precision at best, even when using simulated data where the actual precision is known. Our method will allow direct comparison of the precision of any two measurement systems. This will allow for: 1) evaluating measurement systems for diagnostic procedures; 2) improving a single measurement system through experimentation to determine where errors arise; 3) evaluating the relative impact on precision from the numerous data normalization methods currently in widespread use; and 4) selecting the best measurement system and normalization method for any given experiment. Finally, we will provide all of our methodology as a freely accessible web application.

**2 Innovation**

We will apply a little known mathematical framework, Relative Sensitivity, originally developed in 1985 by John Mandel for evaluating measurement precision in analytical chemistry. This method does not require knowledge of the true amount of a sequence being measured, a quantity that is often difficult or impossible to know when measuring RNA. Instead, relative sensitivity provides the framework for making comparisons among normally distributed measurements. Our approach is conceptually innovative because we will need to derive the formulas necessary to apply relative sensitivity to less well-understood count data, often with multiplicative errors, arising naturally from NGS-based measurement systems.

A new methodology is only useful if it is accessible to a wide range of researchers and scientists. However,

many advances in statistical methodology take considerable time for adoption by the general scientific community because of the barriers to implementation. Our approach is technically innovative because we will provide a software suite for implementing our methods through the recently developed web platform for R programs, shinyapps.io. This platform allows users to interact with statistical analyses pre-programmed in R through a web-based graphical user interface (GUI). The platform greatly reduces the barriers to method implementation because it does not require anything to be installed on user computers and can provide elaborate point-and-click documentation to guide the user through the analyses. We will provide our suite of methods as a dynamic GUI available for free to researchers who wish to compare measurement or normalization systems on their own data. This platform has never been used for widespread implementation of a new method but holds great potential as a translational statistics tool.