

Proposed Plate Reader Data Pipeline Application for the TERM Laboratory

Aubrey Schrameck¹

¹[Northern Arizona University, Department of Mathematics and Statistics]

Abstract

Microplate readers are central to quantitative analysis in cell and tissue engineering research, generating high-throughput absorbance, fluorescence, and luminescence data across multi-well experimental designs. This project proposes the development of an open-source, R-based data pipeline application to standardize, automate, and document the analysis of plate reader data generated in the Tissue Engineering and Regenerative Medicine (TERM) Laboratory.

Raw plate reader outputs and corresponding plate layout metadata will be supplied by the lab and ingested into a graphical user interface built using Shiny. If fully realized, the pipeline will normalize experimental readouts to user-defined controls and blanks, perform robust statistical analyses, including ANOVA, Tukey post-hoc comparisons, and bootstrapped variance testing, and generate publication-ready visualizations. In addition to analytical outputs, the application will produce a teaching-oriented PDF report that transparently documents each processing and statistical decision.

The overarching goal of this project is to replace manual, non-standardized analysis workflows with a reproducible, accessible system that improves analytical consistency, reduces training overhead for new lab members, and enhances statistical rigor. By integrating automation, visualization, and documentation into a single tool, this project aims to improve both research efficiency and data literacy within the laboratory.

Introduction

The Kellar Tissue Engineering and Regenerative Medicine (TERM) Laboratory conducts experimental research that relies heavily on quantitative cell-based assays, including measurements of viability, metabolic activity, and biomolecular response. These assays are typically performed using microplate readers capable of handling plate formats ranging from 6- to 96-well layouts and producing high-dimensional datasets with multiple experimental conditions and technical replicates.

Currently, plate reader outputs in the TERM lab are processed manually by importing raw files into GraphPad Prism. While Prism offers powerful statistical and visualization tools, this workflow presents several challenges. The process is time-intensive, requires access to proprietary software, and lacks standardization across users. New students must learn both the software interface and appropriate analytical practices, often resulting in inconsistent normalization strategies, statistical tests, and figure aesthetics. These inconsistencies complicate result interpretation, reduce reproducibility, and increase the cognitive burden of onboarding new lab members.

This project is motivated by the need for a unified, transparent, and reproducible analysis pipeline tailored specifically to the lab's experimental workflows. Rather than replacing Prism outright, the proposed application will complement existing practices by automating common analyses, enforcing consistent statistical standards, and producing outputs that remain compatible with Prism when further customization is desired. Several sample plate reader datasets with accompanying experimental metadata have been obtained and will be utilized to support the development, testing, and validation of the proposed analysis pipeline. While additional datasets may be employed, the central objective of this project is the creation of a robust, reproducible application rather than an exhaustive analysis of experimental outcomes.

	1	2	3	4	5	6	7	8	9	10	11	12	
A													PrestoBlue:530 /25.590/35
B		24345	24544	25730	26115	25199	27402		18989	19593	19430		PrestoBlue:530 /25.590/35
C		25405	25531	26007	27470	25978	26720		19365	19164	19347		PrestoBlue:530 /25.590/35
D		26852	25602	28100	27204	27326	29617		19265	19318	19295		PrestoBlue:530 /25.590/35
E		25743	649	26634	27072	28110	29412		19465	15599	19091		PrestoBlue:530 /25.590/35
F		27961	28419	29004	30302	31366	32250		19437	19299	19275		PrestoBlue:530 /25.590/35
G													PrestoBlue:530 /25.590/35
H													PrestoBlue:530 /25.590/35

Fig. 1: Sample plate reader output, purple and red indicate control and blank wells, respectively.

Methods

Data ingestion will begin with structured input of raw plate-reader output files alongside user-defined plate layouts that specify experimental groups, controls, and blanks. The upload interface will automatically detect and parse one or more plates contained within a single file, including Excel, CSV, or plain text formats, reducing the need for manual preprocessing. These inputs will be handled through a Shiny-based graphical user interface, enabling users with minimal programming experience to load, inspect, and process plate-level data in a consistent and reproducible manner.

The pipeline will implement standardized normalization techniques, including control-based scaling and optional background subtraction. Statistical analysis will focus on group-level inference using one-way or multi-factor ANOVA where appropriate, followed by Tukey's Honest Significant Difference tests for pairwise comparisons. Automated outlier detection will be implemented using standardized statistical thresholds to flag anomalous wells without silently discarding data.

To improve robustness beyond traditional parametric assumptions, bootstrapping methods will be used to estimate confidence intervals, assess variance stability, and evaluate effect size sensitivity. Optional F-tests with bootstrapped resampling will provide additional insight into variance differences across conditions, particularly for metabolic and viability assays known to exhibit heteroskedasticity.

Visualization outputs will emphasize clarity and transparency, including bar plots overlaid with jittered individual data points and standard curve representations where applicable. All figures will be generated with publication-ready formatting and consistent styling. The application will also export Prism-compatible data tables, enabling seamless downstream use in existing lab workflows.

The primary tools for this project include R, Shiny, tidyverse packages for data manipulation, and ggplot2 for visualization. Challenges may include accommodating heterogeneous plate formats and ensuring statistical methods remain appropriate across diverse assay types; these will be addressed through modular design, configurable analysis options, and the implementation of unit tests to validate core data parsing, normalization, and statistical functions. Establishing a testing framework will support future feature expansion, reduce the risk of regression errors, and promote long-term maintainability of the application.

Expected Outcomes

If successful, this project will deliver a streamlined graphical application that enables TERM laboratory researchers to move from raw plate reader files to interpretable results with minimal manual intervention. The application will load raw plate data and associated layouts, perform control-normalized analyses, and execute user-defined statistical tests, including ANOVA with Tukey post-hoc comparisons, through a simple, intuitive interface. Users will also be able to quickly generate publication-ready

bar-and-jitter visualizations, evaluate variance stability using bootstrapped F-tests, and synthesize metabolic indices for experiments involving multiple assay readouts.

A central outcome of the project is the automatic generation of a teaching-oriented PDF report that transparently documents each step of the analysis, including normalization decisions, statistical methods applied, and interpretation guidance. This report will function both as a permanent record of analytical choices and as a practical training resource for new lab members. Collectively, these features are expected to significantly reduce analysis time, lower the barrier to correct statistical practice, and promote consistent, reproducible data interpretation across the TERM laboratory.

Group Considerations

This project is conducted in collaboration with the TERM laboratory, with the work structured around a client–developer model. Christopher Groberg will serve as the client representative by providing representative raw plate reader data, describing typical analysis workflows, and offering feedback on usability and feature requirements. The primary developer, Aubrey Schrameck, will be responsible for all aspects of application design and implementation, including data ingestion, preprocessing, statistical analysis, visualization, testing, and documentation. While client input will guide requirements and validation, all core development, analysis, and reporting tasks will be completed independently.