Use of Statistical Analyses for Evaluating Large Historical Datasets of AAV Productions to Assess Process Parameters – A Case Study

John Ketz, Byron Carper, Sarah Bergman, Samir Acharya.

INTRODUCTION

Andelyn Biosciences, a CDMO launched out of Nationwide Children's Hospital in 2020, has developed Curator® Platforms for adherent and suspension AAV productions. Using the Curator® adherent platform we have manufactured more than 450 Phase I and II clinical batches spanning over 17 years and covering a variety of serotypes and indications. To assess the robustness of the adherent platform and identify potential critical process parameters, historical data generated at each process step of GMP productions involving 74 sublots of Serotype A and 44 sublots of Serotype B were compiled and analyzed.

METHODS

Exploratory Data Analysis and Visualization: In-process parameters such as volumes, cell confluency, and pressures were extracted from batch records from over 150 different AAV production lots and sublots supporting Phase I and II clinical trials. In-process and final results, such as vector genome titer, full:empty ratio, and vector purity, were compiled from the same batches to investigate possible critical process parameters that may predict key outcomes. Two AAV serotypes were manufactured more than 40 times each, allowing for a visual examination of productivity, recovery, and purity results from productions using various clinically-relevant transgenes.

Correlation Analysis: To screen for relationships between variables, a Pearson correlation matrix was generated with all in-process measurements and final results for each serotype. The Pearson r values and two-tailed p values were calculated to determine the direction, strength and statistical significance of all pairwise linear correlations. Where appropriate, a 1% ROUT method was applied to detect and remove outliers. Correlations were individually plotted and investigated if the r value had an absolute value greater than 0.7 or if the p value was less than 0.05.

FINDINGS

Though harvest titers and overall vector genome recoveries vary by AAV serotype and transgene (Figures 1 and 2), the trends in step recoveries were similar between Serotype A and Serotype B (Figure 3). The purity of the final product was high (>95%) regardless of serotype or transgene (Figure 4).

After excluding trivial correlations – e.g., concentrate volume and concentration factor – a majority of pairwise correlations were weak or not statistically significant (Figure 5), indicating that most in-process measurements do not correlate to each other or to final results. Figure 6 shows examples of final release results not depending on in-process parameters, indicating that the Curator® adherent platform is robust to changes in transgene and typical batch-to-batch variations. Figures 7 shows the weak but statistically significant correlation between two pairs of in-process parameters. These warrant further experimentation to demonstrate whether they constitute critical process parameters and if stricter controls over these process steps should be implemented

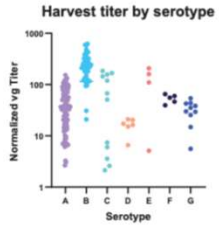

Figure 1. Vector genome titer at harvest of 154Phase I and II productions in the adherent Curator® platform. Seven AAV serotypes were produced four or more times with various transgenes. Titers were normalized and plotted on a logarithmic scale. Serotypes A and B were both produced more than 40 times with various transgenes, and those results will be investigated further.

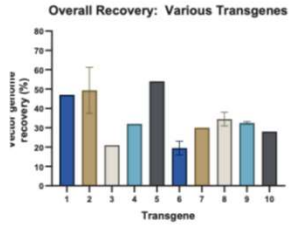

Figure 2. Overall vector genome recovery from harvest to final product for 18 productions of Serotype A, using 10 different clinical transgenes. Values were calculated by dividing the yield in the final drug product by the yield in the harvest of the respective lot or sublots.

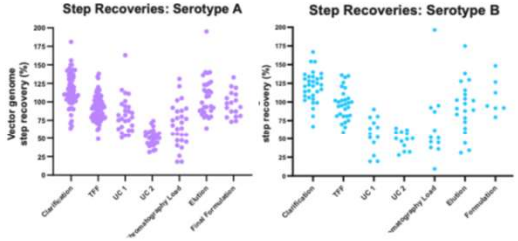

Figure 3. Vector genome recoveries for all process steps for two serotypes. The vector genome yield was determined for the individual processing steps: clarification, tangential flow filtration (TFF), ultracentrifugation spin 1 (UC1), ultracentrifugation spin 2 (UC2), chromatography column load, chromatography elution, and final formulation. The yield was divided by the yield in the previous step to determine step recovery. N decreases at later process steps, as production lots or sublots were pooled.

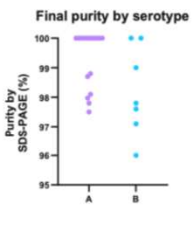

Figure 4. Purity of the final drug product from 24 productions of Serotype A and Serotype B. Purity is determined by SDS-PAGE analysis.

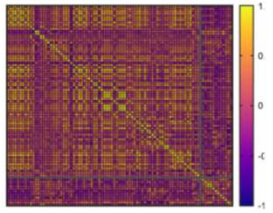

Figure 5. Pearson correlation matrix for pairwise correlations between in-process measurements extracted from batch records and in-process and final testing results. Yellow shading indicates a strong positive association, and dark purple shading indicates a strong negative association. Also calculated but not shown were the two-tailed p-values for each correlation.

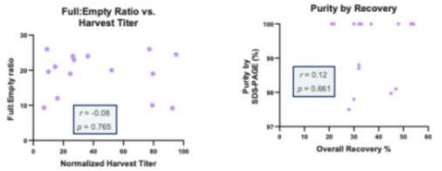

Figure 6. Statistically insignificant linear correlations. Left: Full:empty ratio of the final drug product by AUC versus normalized harvest titer. Results from 15 sublots. **Right:** Purity of the final drug product by SDS-PAGE versus overall vector genome recovery from harvest to final fill. Results from 16 sublots.

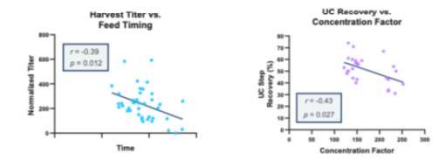

Figure 7. Statistically significant linear correlations. Left: Normalized harvest titer versus the timing of a feed step. Results from 14 sublots; no outliers were detected using a 1% ROUT method. **Right:** Vector genome step recovery from an ultracentrifugation step versus concentration factor. Results from 27 pooled sublots; two outliers were detected using a 1% ROUT method.