Longitudinal Analysis of In Vivo mRNA Expression Data Using Bayesian P-Splines

Authors: Hakem Ben Addi[‡], Steve Lenhard[†], Emiliano Chiarot[#], Laura Lessen[§], Sonia Budroni[‡‡] & Cedric Taverne[‡]

Affiliation: [‡]GSK Development Biostatistics (Rixensart, Belgium), [†] GSK Research In Vivo/In Vitro Translation (Upper Providence, USA), [#]GSK Vaccines R&D (Siena, Italy), [§]GSK Vaccines R&D (Rockville, USA), [‡‡]GSK Development Biostatistics (Siena, Italy)

The advent of mRNA-based drugs has necessitated human-translatable methods to compare the dynamic expression profiles of the encoded proteins. Current research employs longitudinal *in vivo* imaging on live mice and rats, using bioluminescence emission as readout. Existing parametric methods, such as linear and non-linear mixed models can be used for such analyses. However, these parametric approaches often face limitations due to the lack of mathematical equations adequately describing the expression profiles.

The current study sought to address these limitations by applying a Bayesian p-spline model to data assumed lognormally distributed and considering the autocorrelation of residuals. The flexible brms R package was employed to fit the model. The resultant posterior distribution facilitated the comparison of the overall expression dynamic and exploration of specific time range of interest. The analyses demonstrated how varying mRNA drug substances influenced the *in vivo* expression profile at different time points.

Results underscore the advantages of Bayesian p-spline approaches for modelling longitudinal *in vivo* expression data. This methodology offers versatility and can be generalized to any type of longitudinal data, including bacterial load measurements. In addition, Bayesian p-splines demonstrated reliable performance in settings with small sample sizes, between 5 and 10 animals per group, which is common in preclinical research setups. Unevenly distributed time points, another prevalent issue in longitudinal studies, had a minimal impact on the model's effectiveness, providing further evidence of the Bayesian p-spline model's robustness. This research underscores the potential of this approach for the analysis of longitudinal expression data in the development and evaluation of mRNA-based therapeutics.