

# Combining parametric models and Gaussian processes for longitudinal data

BAYES 2023 conference, Utrecht

Juho Timonen

Generable Inc. & Aalto University

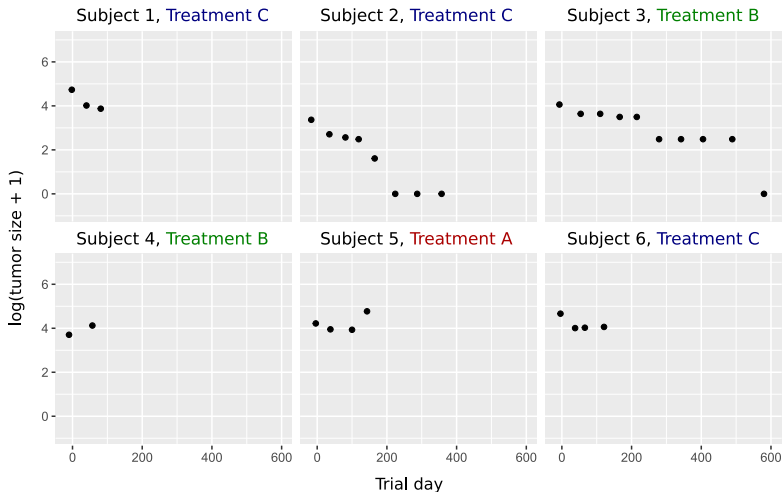
26th October 2023

*Generable*



# Longitudinal data

- ▶ Repeated measurements of same patients over time
- ▶ Example: tumor sizes of patients in a drug trial

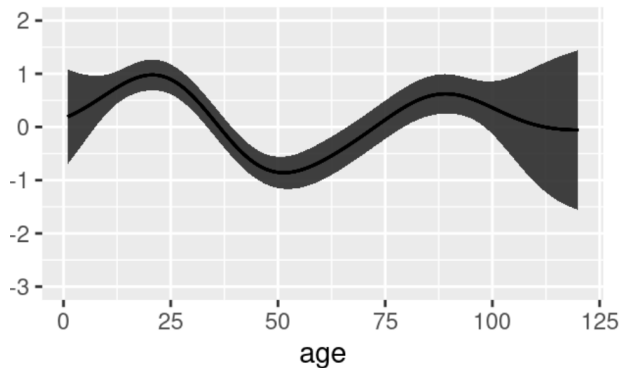


# Modeling motivation

- ▶ The data is scarce and irregularly sampled
- ▶ Model = description of the data generating process
- ▶ Helps us understand effects of different predictors on the modeled variable
- ▶ We want to use domain knowledge to fix parts of the model, and infer the rest using Bayesian inference

# Gaussian processes

- ▶ Functional form of an effect can be an unknown part



# Additive models

- ▶ Gaussian processes can be used additively with a parametric term whose functional form is based on domain knowledge
- ▶ Stein-Fojo+GP = exp. decay + exp. growth + GP

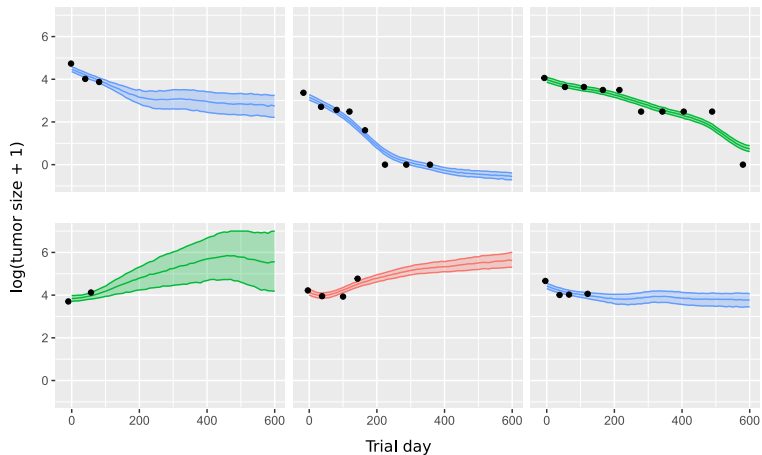
# Stein-Fojo+GP example

Posterior predictive  
Median and 50% Credible Interval

• Treatment A

• Treatment B

• Treatment C



# Stan and R

- ▶ At Generable, we have developed an R package which allows specifying models with an R formula such as

```
tumor_size ~ gp(day) + gp(day, arm) + sf(day,  
subject | arm)
```

- ▶ Under the hood, Stan is used for Bayesian inference

# Interpretable additive Gaussian processes

- ▶ The `lgpr` R-package fits additive GPs for longitudinal data using Stan
- ▶ Can model rapid effects during disease initiation

