Prediction intervals based on historical controls for the micronucleus test

Jonathan Rathjens[1)], Max Menssen[2)] and Hannes-Friedrich Ulbrich[3)]

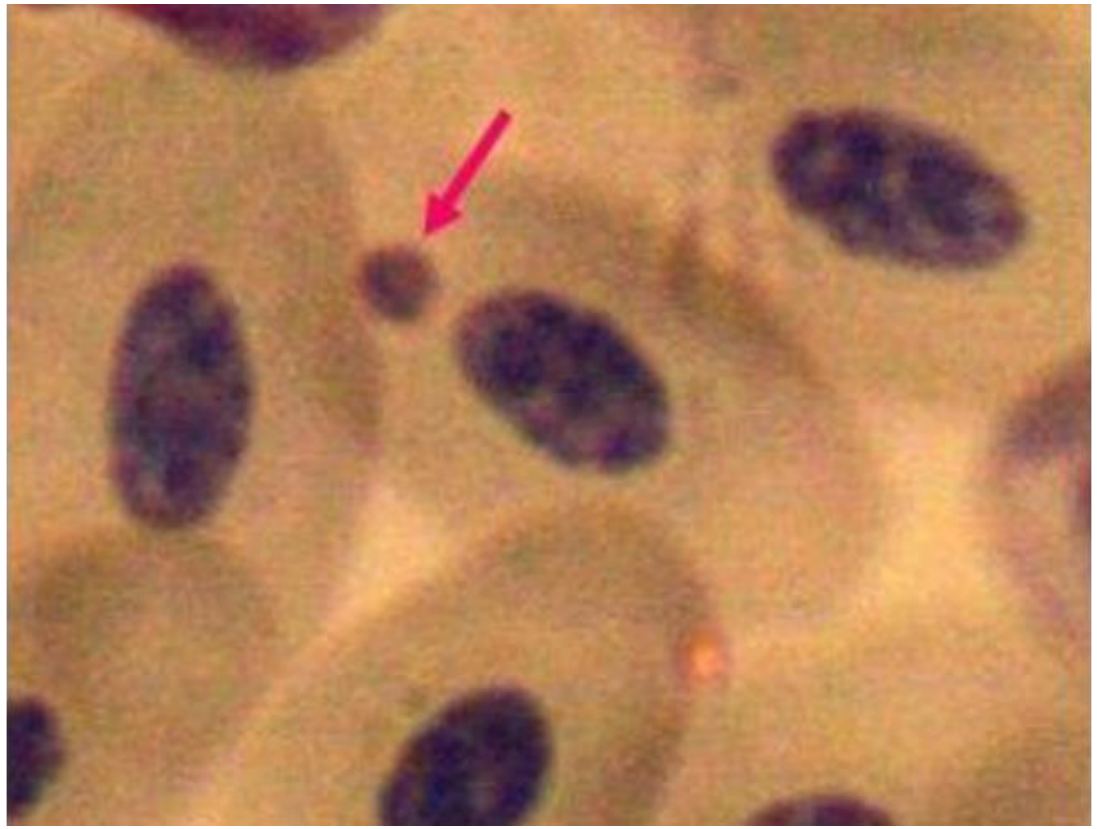
[1)] Early Development Statistics, Chrestos GmbH, Essen

[2)] Biostatistics, Institute of Cell Biology and Biophysics, Leibniz University Hannover

[3)] Scientific Insight Solutions, Bayer AG, Berlin

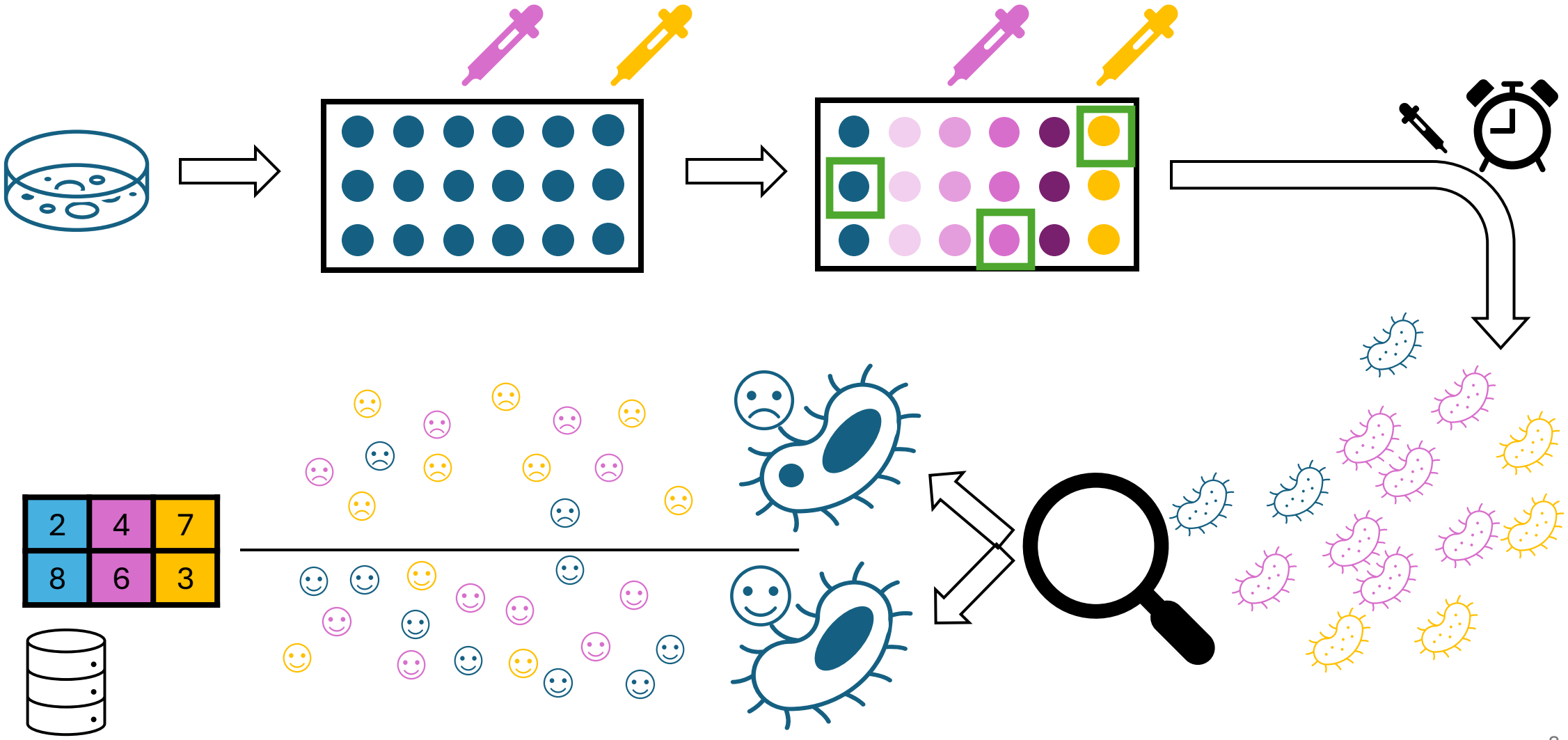
Workshop of the IBS-DR Working Groups 'Non-Clinical Statistics' and 'Bayes Methods', Göttingen, 5/6 December 2024

Micronucleus



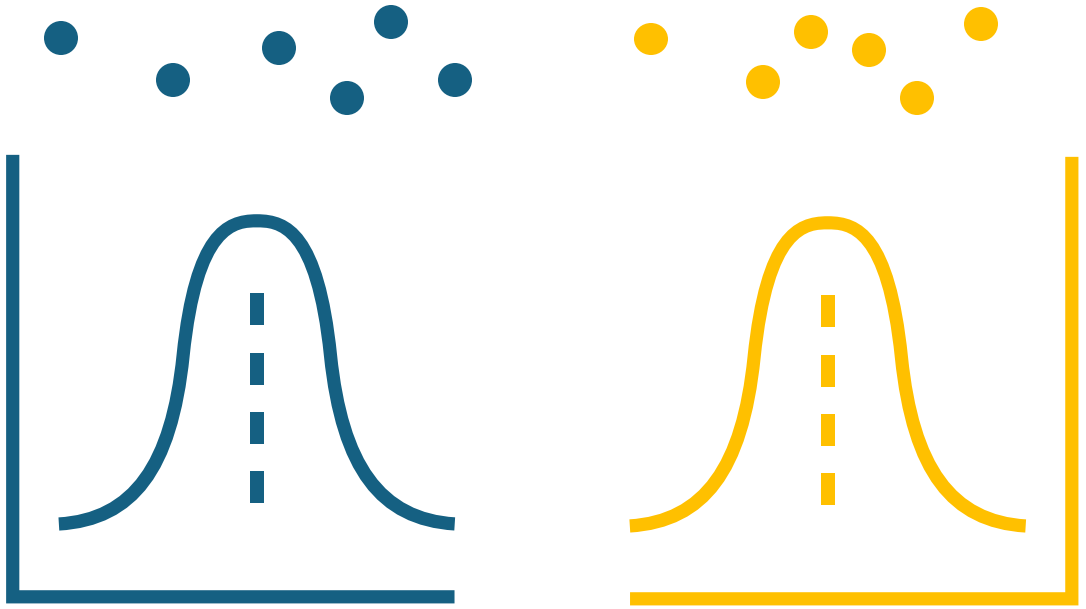
- adverse event during cell division
- small fragment of chromosomes outside new nuclei
- hint to genotoxic conditions

Micronucleus test (MNT) in-vitro



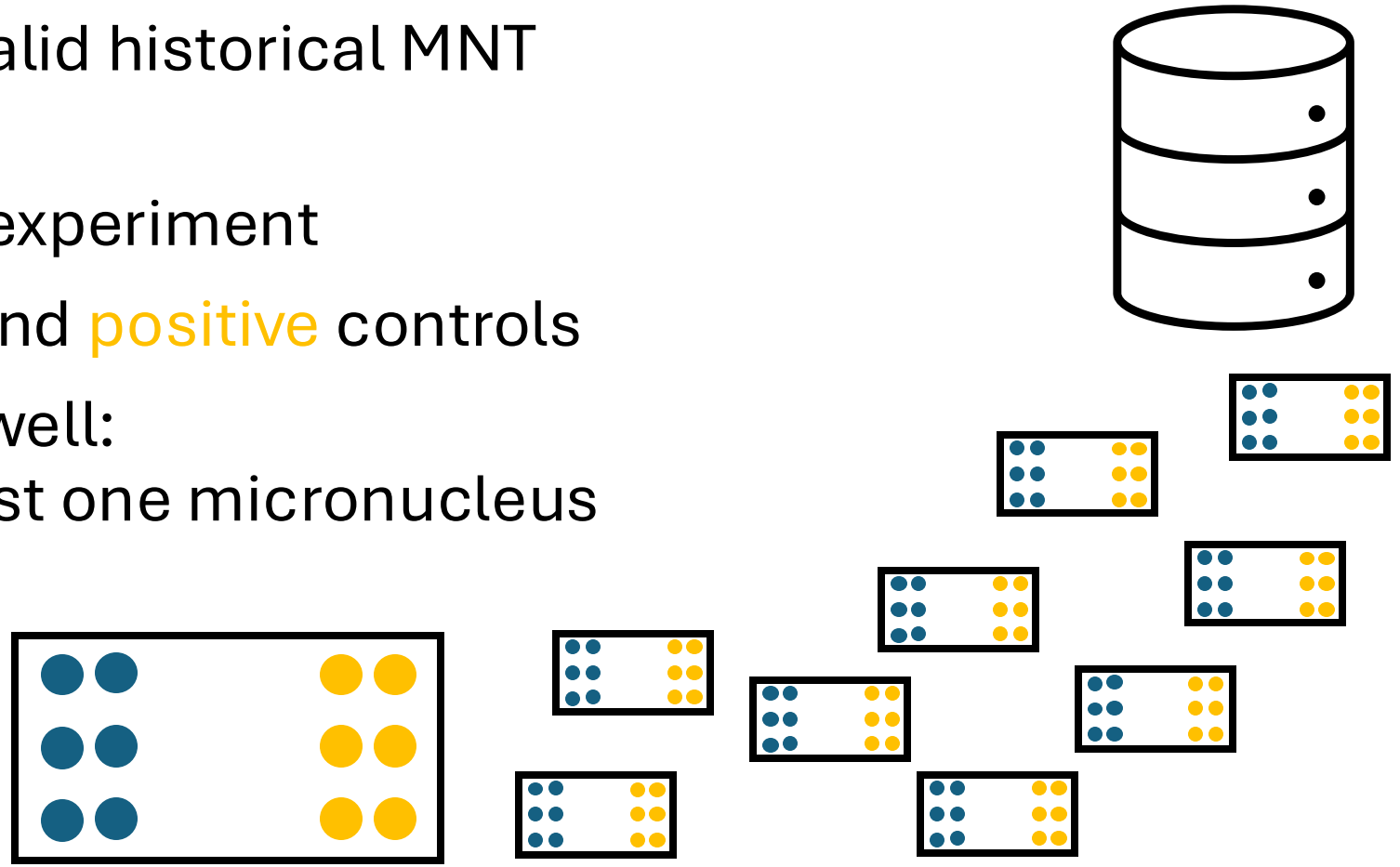
Test assay validation

- ensure validity of results
- OECD Guideline for the Testing of Chemicals 487
- considering negative and positive control groups
- current experiment's wells' counts compared to historical distributions



Historical control group data

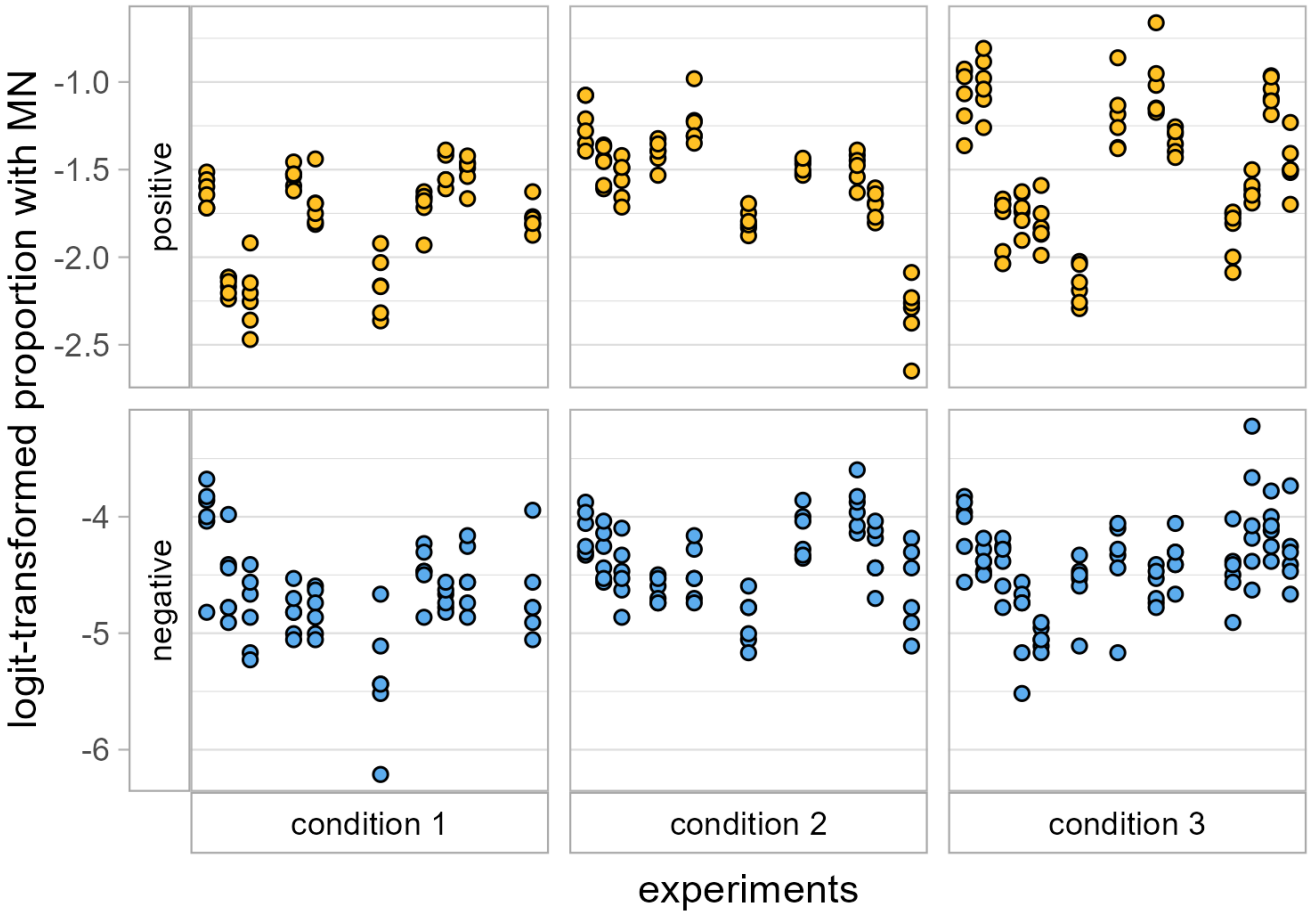
- about $H = 30$ valid historical MNT experiments
- $I = 6$ wells per experiment
- … for negative and positive controls
- count data per well:
  cells with at least one micronucleus



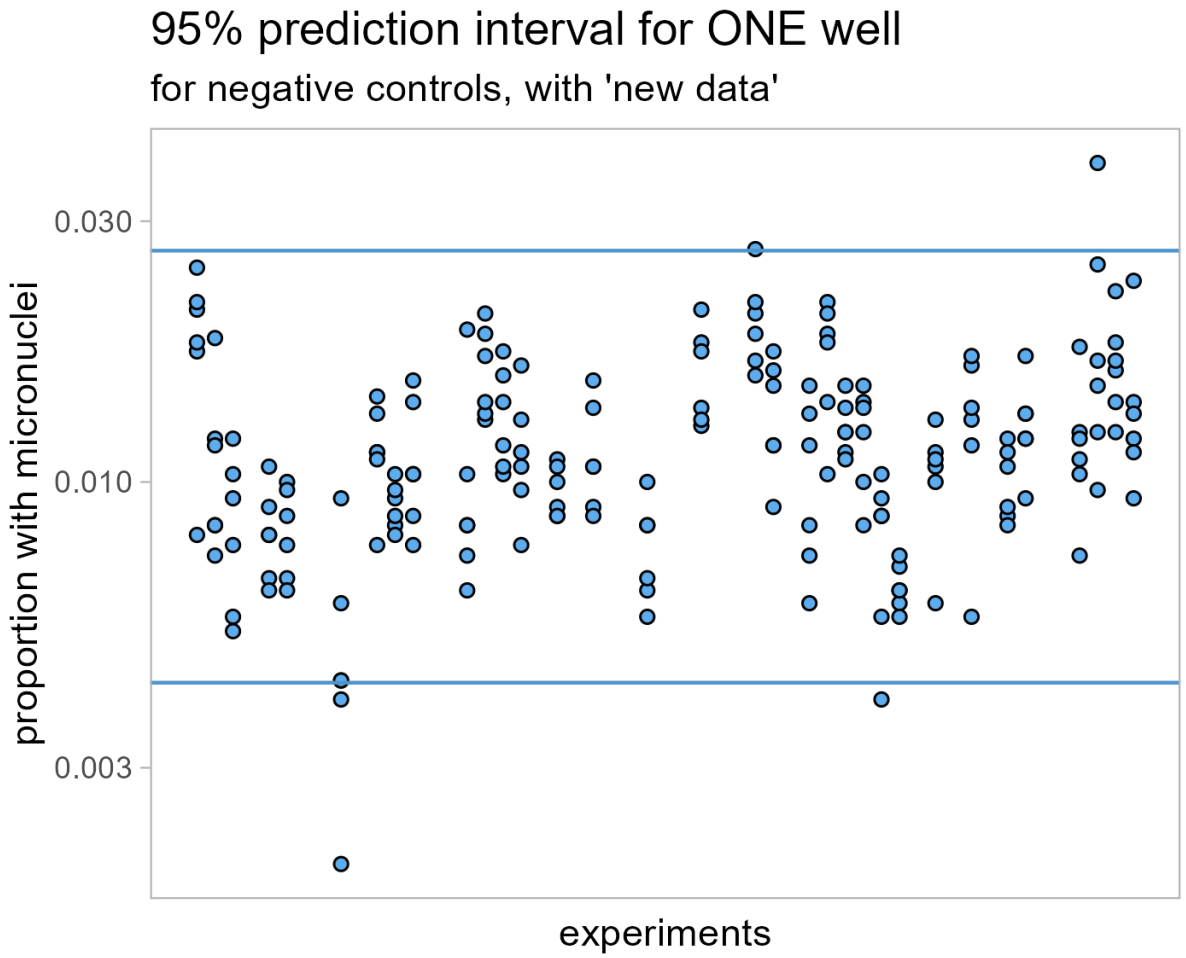
Endpoint information per well

experiment $h = 1, \dots, H$;
well $i = 1, \dots, I$;
$n = 3000$ cells as target

- **raw**: x_{hi} cells out of n have at least one micronucleus
- **proportional**: $p_{hi} = x_{hi}/n \in (0, 1)$
- **transformed**: $z_{hi} = \log(p_{hi}/(1 - p_{hi})) \in \mathbb{R}$



Prediction intervals



- for micronucleus proportion $p_{(H+1)1}$ of ONE well in a future experiment
- directly interpretable on data level

- several model candidates
- evaluation by cross validation, leaving out one historical experiment per run

Conventional models

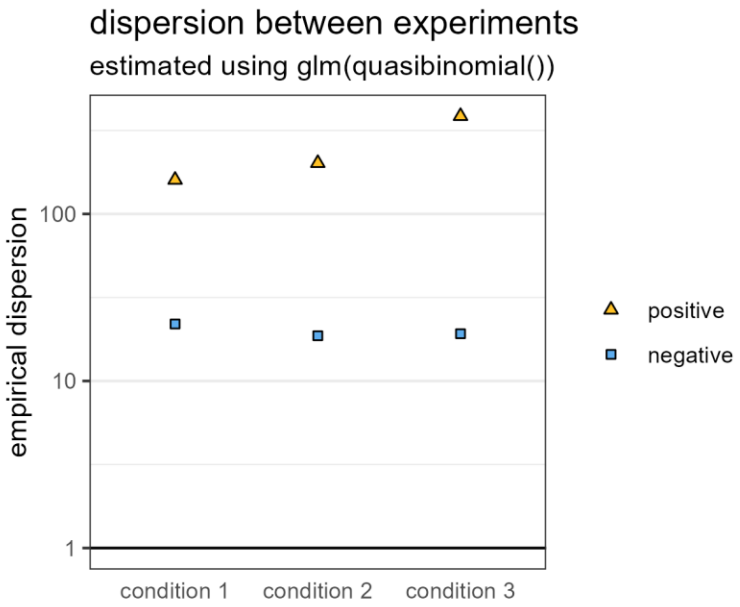
$x_{hi} \sim \mathrm{Bin}(n, \pi)$ or

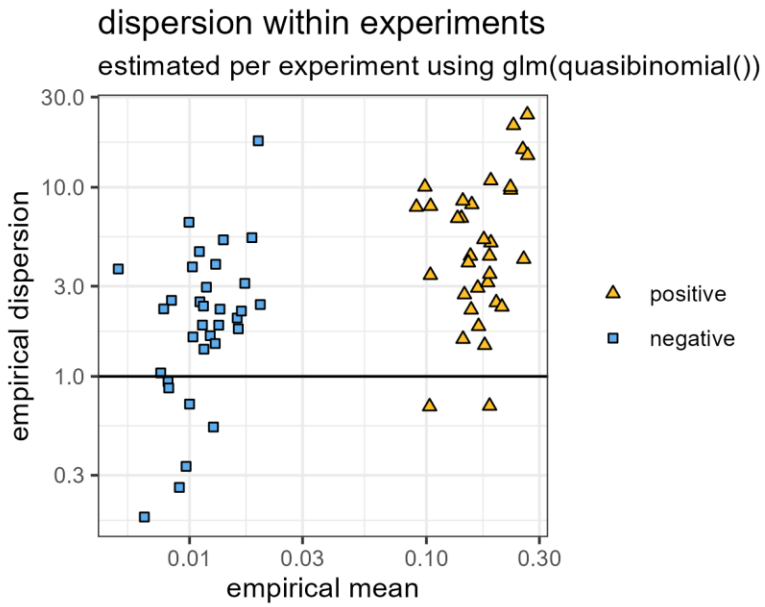
$z_{hi} \sim N(\mu, \sigma^2)$

using all data without experiment information

not accounting for

- overdispersion between experiments (random effects of experiments)
- overdispersion within experiment (random effects of wells)





Bootstrap-calibrated prediction intervals

- `R` package `predint`
  (Menssen and Schaarschmidt 2022, *Stat Neerl* 76(3);
  Menssen et al. 2024, *Pharm Stat* online-first)
- accounting for between-experiment-overdispersion
  - using aggregated count data $X_h = \sum x_{hi}$ on experiment level
  - not accounting for within-experiment-overdispersion
  - predicting total count per experiment
- **beta-binomial**:

$$\pi_h \sim \mathrm{Beta}(a, b)$$
$$X_h \sim Bin(n \cdot I, \pi_h)$$

- or equivalently (n and I constant) **quasibinomial**:

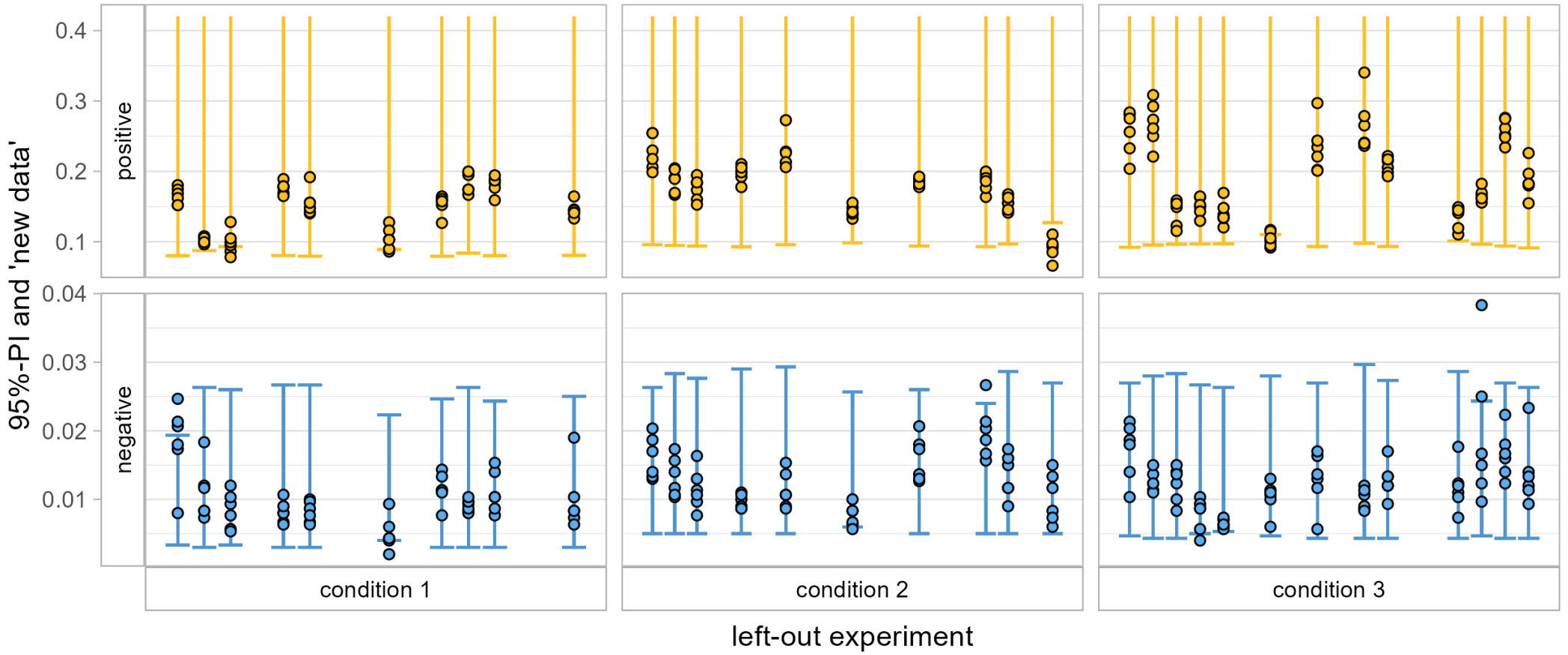
$$\mathrm{E}(X_h) = \pi \cdot n \cdot I$$
$$\mathrm{Var}(X_h) = \phi \cdot \pi \cdot (1 - \pi) \cdot n \cdot I$$

Bayesian (generalized) random effect models

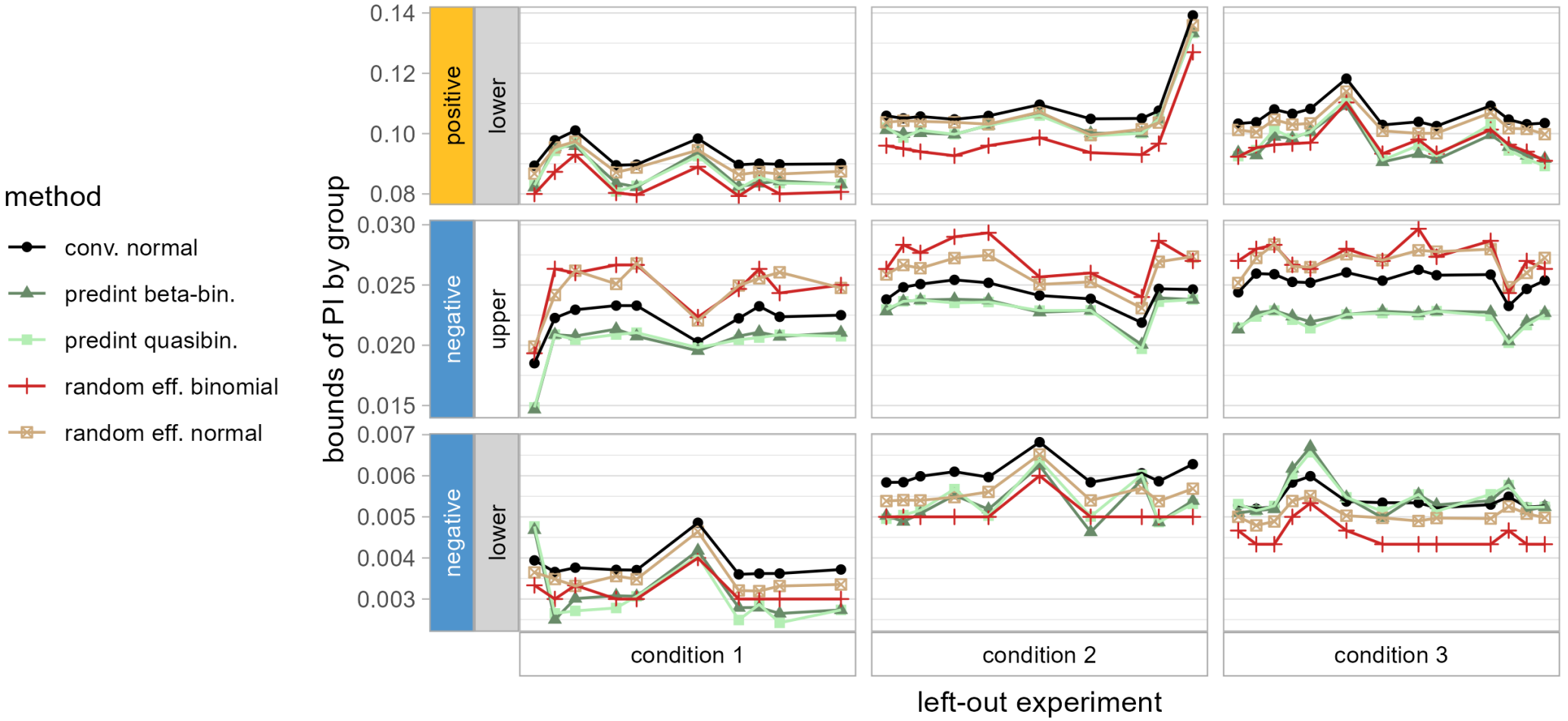
	overdispersed binomial	**approximative normal**
data	$x_{hi} \sim \mathrm{Bin}(n, \pi_{hi})$	$z_{hi} \sim N(\mu_h, \sigma^2)$
transformation	$\pi_{hi} = \mathrm{logit}^{-1}(\mu_{hi})$ $\mu_{hi} = \alpha + \beta_h + \gamma_{hi}$	$\mu_h = \alpha + \beta_h$
random effects	$\beta_h \sim N(0, \tau^2)$ $\gamma_{hi} \sim N(0, \sigma^2)$	$\beta_h \sim N(0, \tau^2)$
model-parameters	$\alpha \sim (-\infty, +\infty)$ $\tau^2 \sim (0, +\infty)$ $\sigma^2 \sim (0, +\infty)$	$\alpha \sim (-\infty, +\infty)$ $\tau^2 \sim (0, +\infty)$ $\sigma^2 \sim (0, +\infty)$

Cross-validated prediction intervals

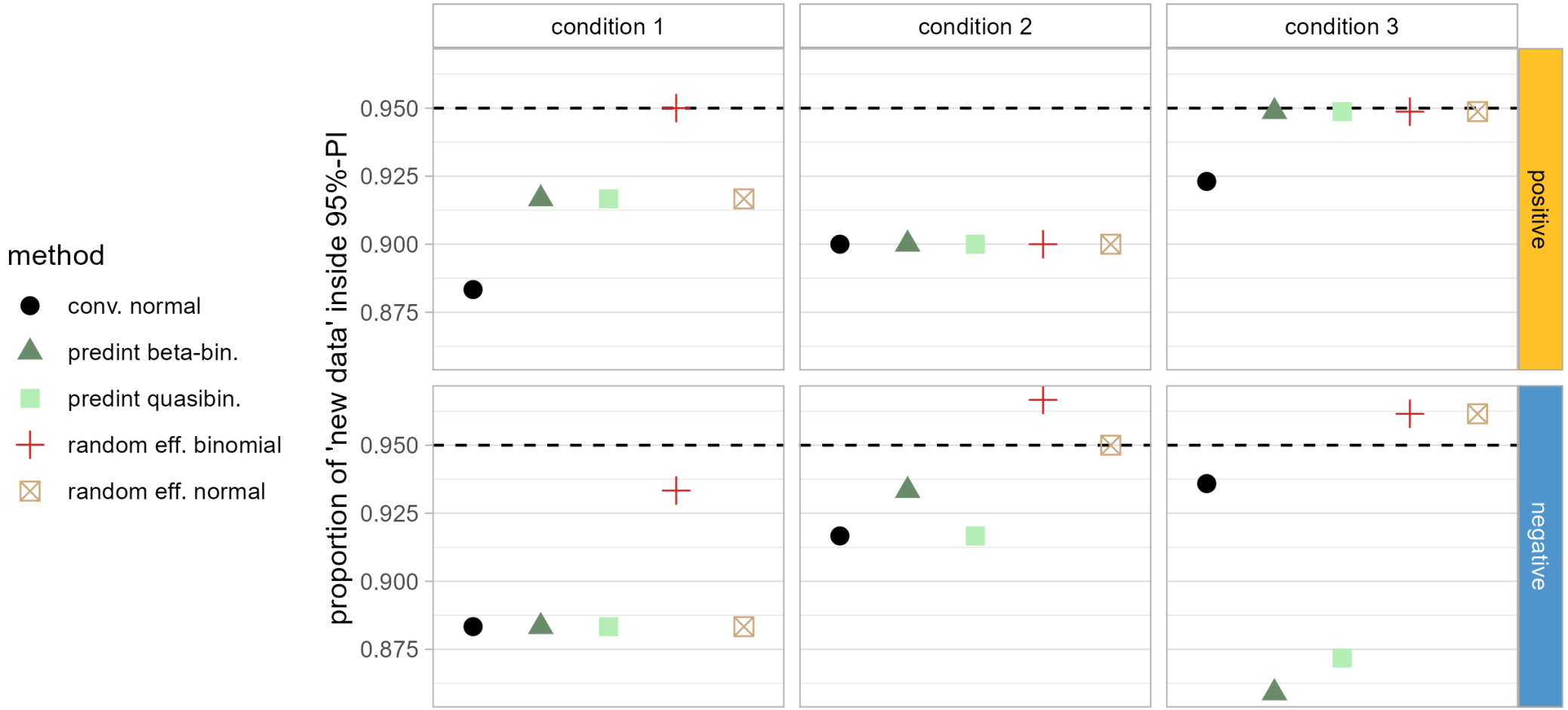
Example: Bayesian overdispersed binomial



Comparison of prediction interval bounds



New data inside interval – overall proportions



Results‘ summary

- small differences between methods
- Bayesian models using individual data

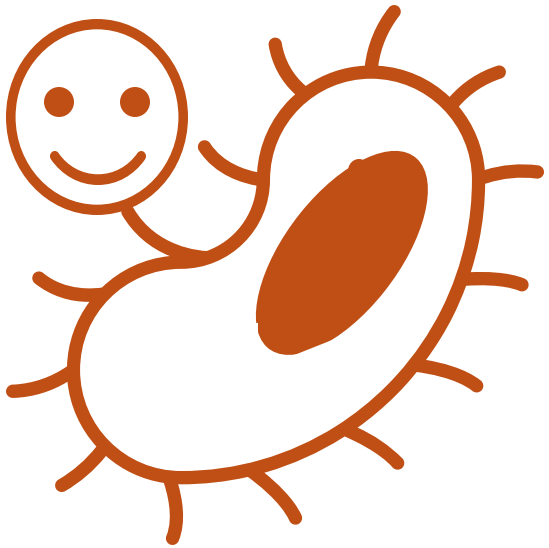
… return wider prediction intervals

… have proportion of ‘new’ well data within 95%-prediction-interval closer to 95%

Perspectives

- joint analyses of negative and positive controls
  - model all data jointly or
  - consider derived data representing the ratio of positives and negatives
- temporal correlation structure between experiments
- more stable results using larger historical database
- simultaneous prediction for all wells of one future experiment

Conclusions



- prediction intervals from historical controls suitable for validating new experiment
- consider both negative and positive controls
- model individual values
  - with experiment random effects
  - with overdispersion

Thank you!