## **Munich Health Foundation Model Symposium**



Contribution ID: 26

Type: Talk

## **Mixed Models with Multiple Instance Learning**

Predicting patient features from single-cell data can help identify cellular states implicated in health and disease. Linear models and average cell type expressions are typically favored for this task for their efficiency and robustness, but they overlook the rich cell heterogeneity inherent in single-cell data. To address this gap, we introduce MixMIL, a framework integrating Generalized Linear Mixed Models (GLMM) and Multiple Instance Learning (MIL), upholding the advantages of linear models while modeling cell state heterogeneity. By leveraging predefined cell embeddings, MixMIL enhances computational efficiency and aligns with recent advancements in single-cell representation learning and by extension foundation models. Our empirical results reveal that MixMIL outperforms existing MIL models in single-cell datasets, uncovering new associations and elucidating biological mechanisms across different domains.

**Primary authors:** ENGELMANN, Jan (Helmholtz Munich); CASALE, Francesco Paolo **Session Classification:** Break + Posters session

Track Classification: Poster