

Pharmacokinetic-Pharmacodynamic informed labelling for Machine Learning classification tasks: predicting severe neutropenia using Real World Data



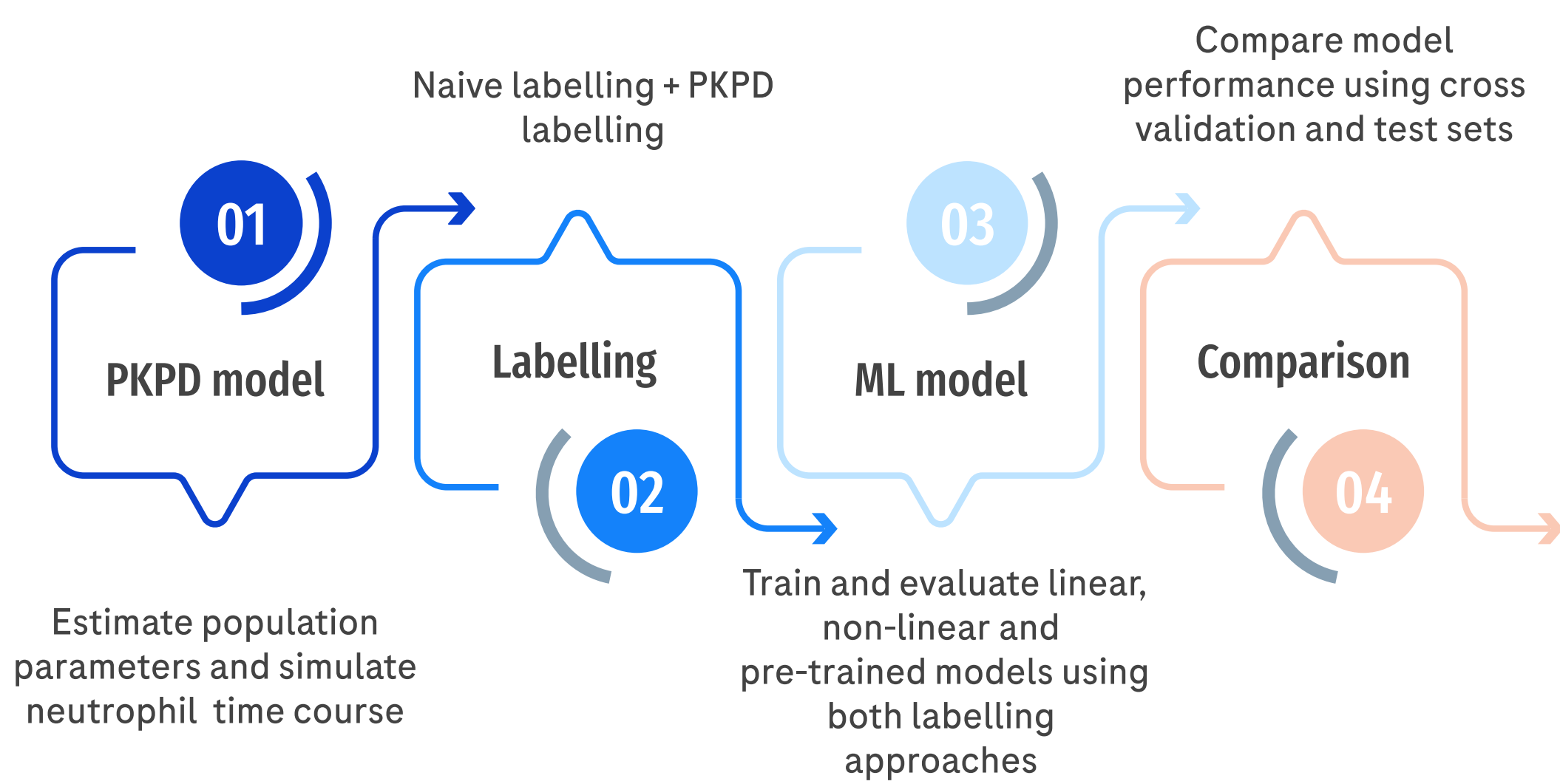
Conor O’Hanlon¹, Jonas Denck¹, Elif Ozkirimli¹, Stefanie Bendels¹, Candice Jamois², Oscar Brück^{3,4,5}, Eric Fey⁶, Kimmo Porkka^{3,4}, Ken Wang²

(1) Roche Informatics, F. Hoffmann-La Roche AG, Kaiseraugst, Switzerland, (2) Roche Pharmaceutical Research and Early Development, Roche Innovation Center, Basel, Switzerland, (3) Comprehensive Cancer Center, Helsinki University Hospital, Helsinki, Finland, (4) Department of Oncology, University of Helsinki, Helsinki, Finland, (5) Department of Clinical Chemistry, HUS Diagnostic Center, Helsinki University Hospital, Helsinki, Finland, (6) iCAN Digital Precision Cancer Medicine Flagship, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Context	Challenge	Opportunity	Use-case	Objective
Real-world data (RWD) from electronic health records provide a rich source of information about diverse patient trajectories and outcomes.	Leveraging RWD for Machine Learning (ML) modelling can be challenging due to sparse, imbalanced, and missing data, and difficulty in defining clinical outcomes.	Pharmacokinetic-Pharmacodynamic (PKPD) modelling can offer a mechanistic basis to a prediction task and could be used to inform label definition for training datasets, potentially enhancing ML classification and data utilisation.	Predicting severe neutropenia (neutrophil count < 0.1 cells x 10 ⁹ /L) at baseline following docetaxel administration was selected as an example to investigate the impact of PKPD informed labelling of ML training set data.	Develop a methodology to incorporate PKPD information in the training dataset through labelling and compare model performance between a ‘naïve’ approach (using some of the available observations) with labels determined using PKPD model predictions.

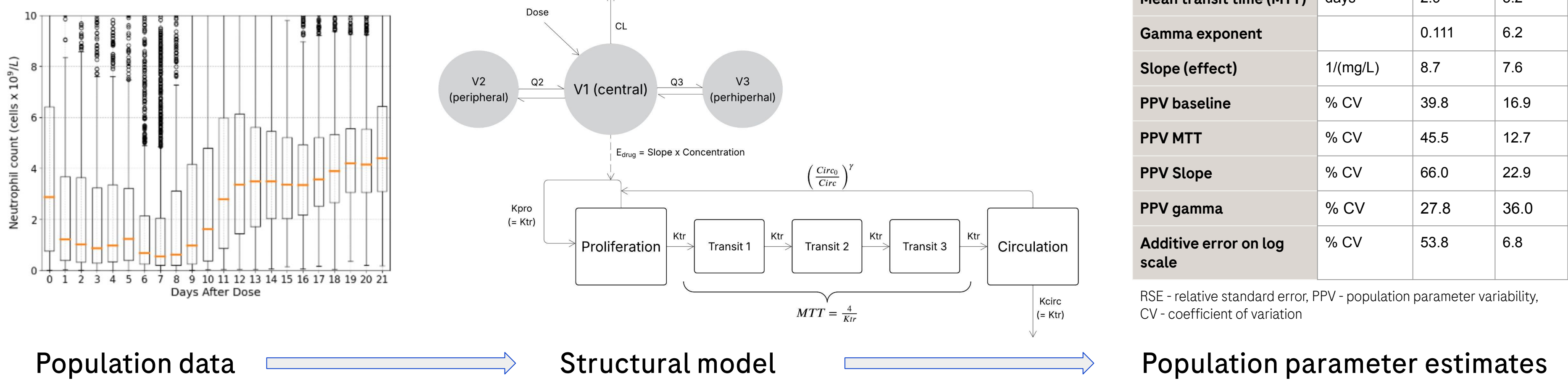
Data and workflow overview

- Data source: Helsinki University Hospital Finland electronic health records
- 14,707 patients administered docetaxel
- Data cleaning and exclusion criteria (e.g. no co-administered anticancer or G-CSF drugs) resulted in **4508 patients and 16 153 instances** (dose + associated neutrophil observations) available for analysis



1. Pharmacokinetic-pharmacodynamic model

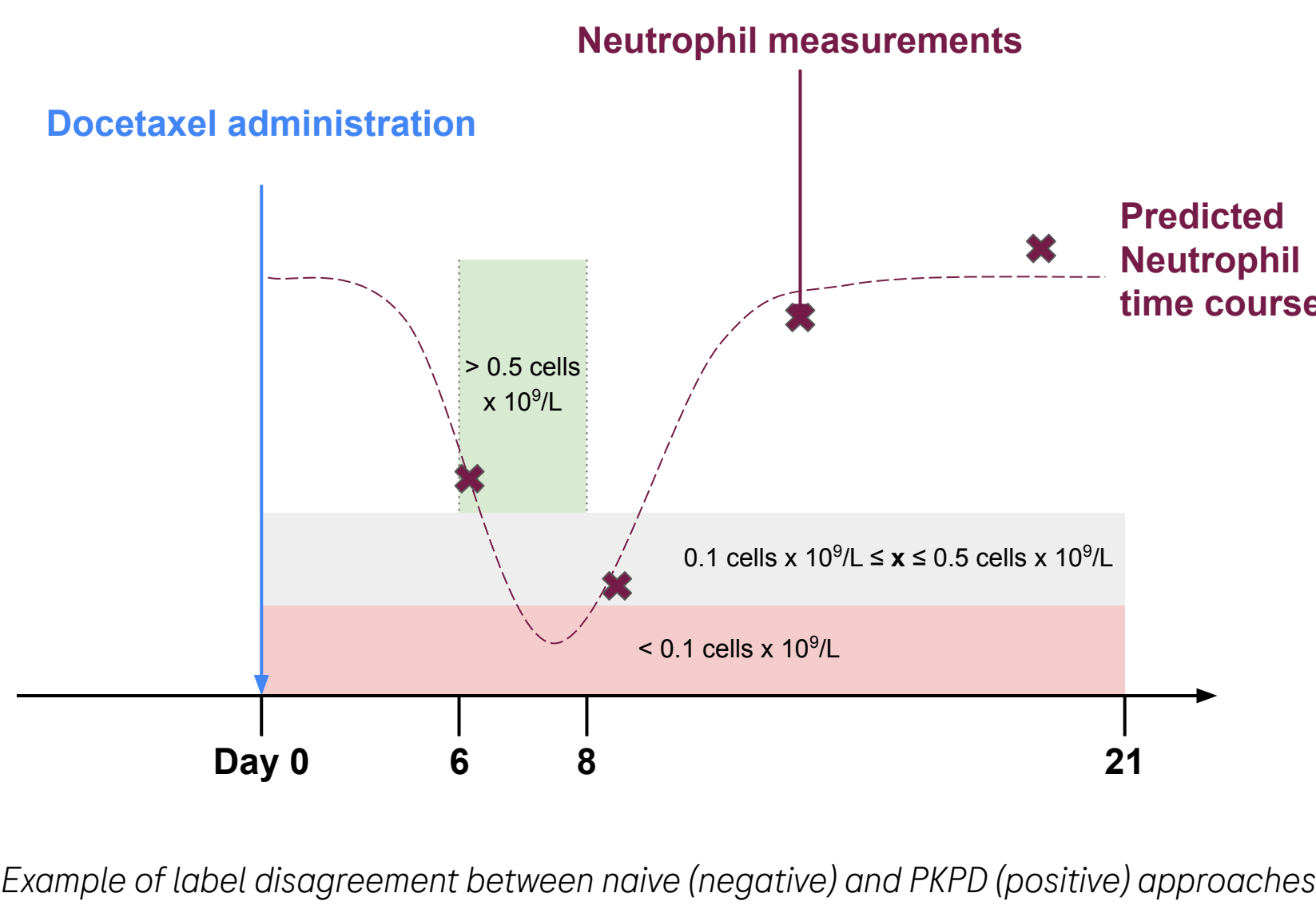
Individual docetaxel exposure was predicted using a published PK model (Fukae et al. 2016). Docetaxel PD model based on the Friberg model for drug-induced neutropenia (Friberg et al. 2002).



2. Labelling

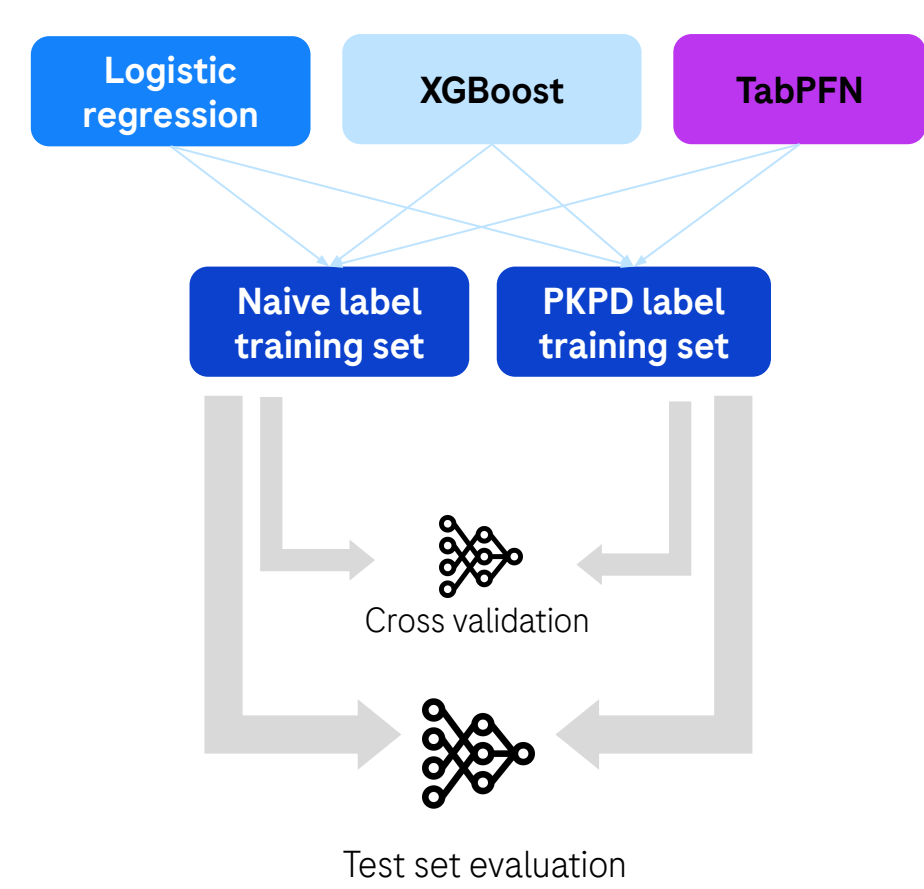
	Naive	PKPD
Positive	Any observation < 0.1 cells x 10 ⁹ /L	Lowest individual prediction < 0.1 cells x 10 ⁹ /L
Negative	Must have an observation between 6 - 8 days after dose, and it must be > 0.5 cells x 10 ⁹ /L	Lowest individual prediction > 0.5 cells x 10 ⁹ /L
Label	2,477 instances, 702 positive, 1,775 negative	4,633 instances, 2,639 positive, 1,994 negative

The 0.1 - 0.5 cells x 10⁹/L exclusion criterion was applied to reduce the risk of false negatives in the naive labelling approach, but at the expense of label sensitivity. This threshold was not required in the PKPD training set, but was applied to enable fair comparison



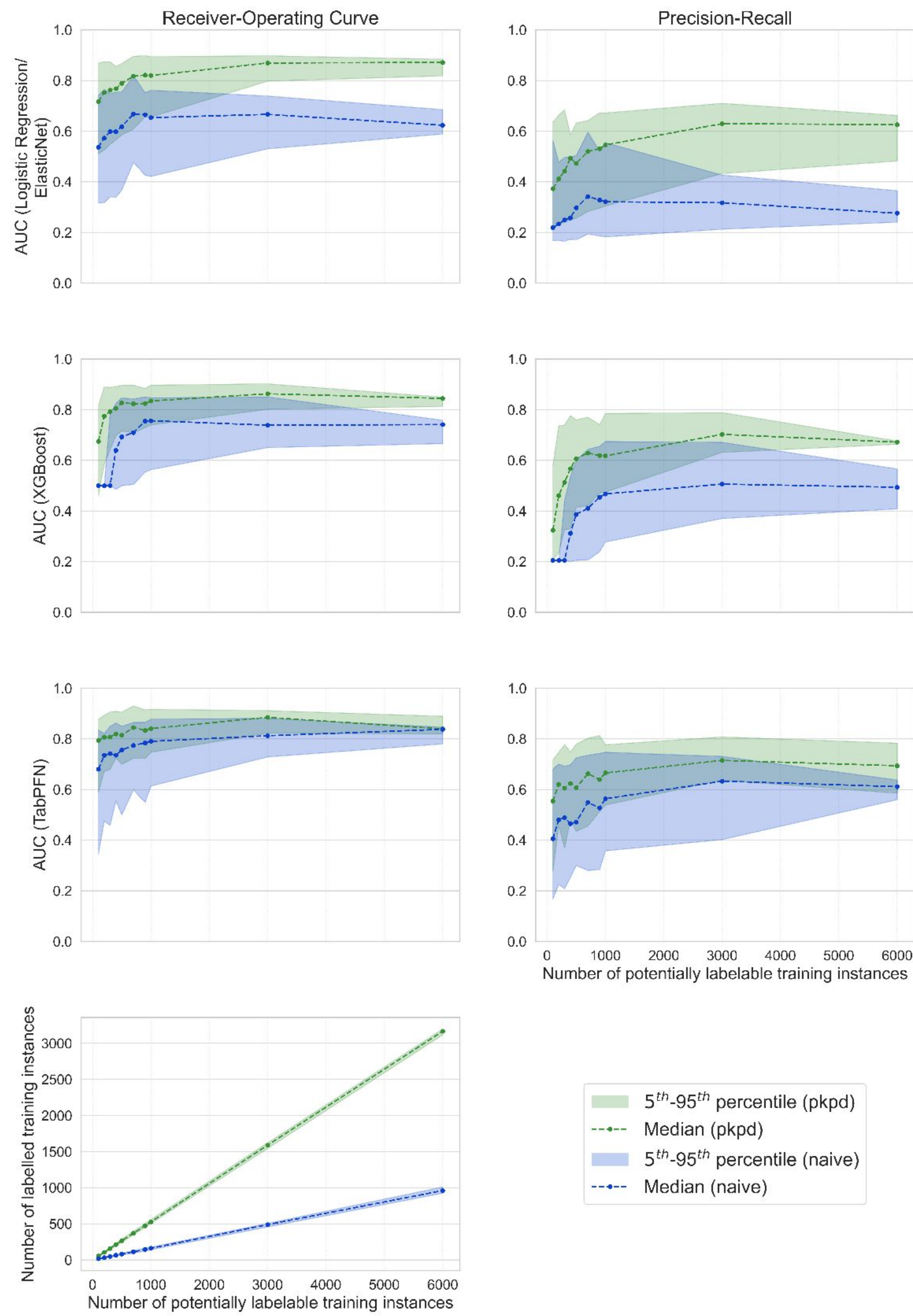
3. Machine learning model

Models algorithms and features: ML algorithms (Logistic Regression, XGBoost, TabPFN) used baseline lab tests (up to 90 days prior) as input features
Iterative Training: Trained the models using datasets from each labelling approach
- Incrementally increased training set size, sampled hyperparameters and conducted a cross validation procedure to assess the impact of these factors on overall model performance
Test set evaluation: Reported results from evaluation using a strictly held-out test subset of data where 'naive' and 'PKPD-informed' labels were in agreement, ensuring an unbiased assessment.

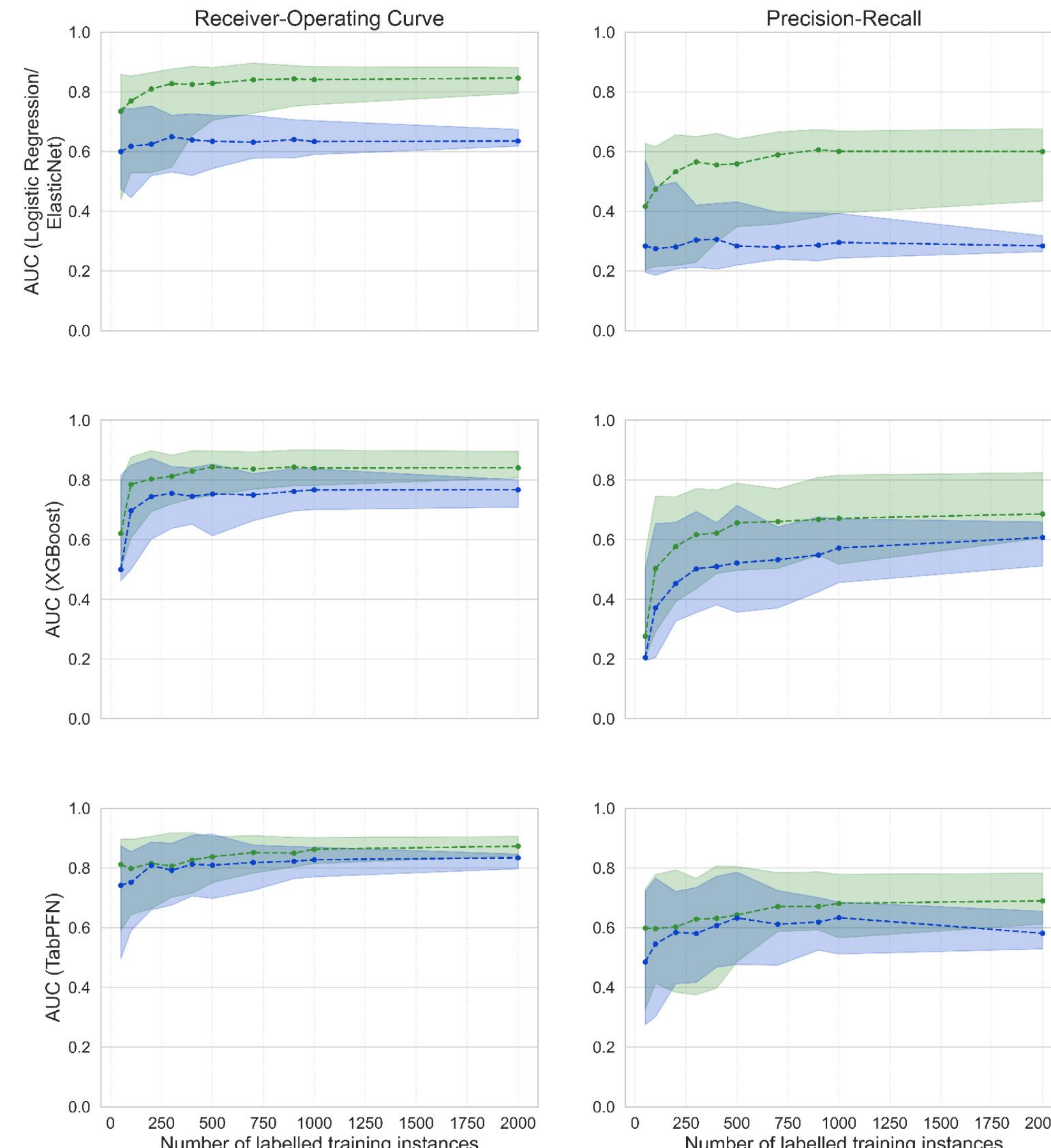


4. Comparison

Performance metrics using both labelling approaches
Difference due to number of instances that can be labelled using naive (blue) and PKPD (green), and ‘quality’ of label.



Matched training instances between naive and PKPD
Difference due to ‘quality’ of label only.



Conclusion

- ML models trained on **PKPD-informed labels outperformed models trained on naive labels** across various training dataset sizes and model types
- PKPD approach to label a significantly larger training cohort**, but even when the number of training instances was matched, the PKPD-informed labelling strategy increased performance across all algorithms (almost all data points were statistically significantly different using Mann-Whitney U test with Bonferroni Correction for multiple comparisons - results not shown)
- This sustained increase, observed with the same number of training instances, suggests that **PKPD-informed labels provide a qualitatively superior and more informative signal** compared to naive labels. This improvement may arise because PKPD-informed labels incorporate quantitative physiopathological constraints, resulting in more plausible neutrophil trajectories
- We provide a **methodology** to integrate biological/mechanistic information (constraints) into ML models at baseline (no on treatment-data is required once the model is deployed)