

Improving Predictive Accuracy in Vincristine PopPK Models Through Model Averaging and Dataset Augmentation



Arsenii Zats^{1,3}, Andrew M Brandon², Tadao Tamura^{1,3}, Gareth J Veal², Shelby Barnett², Alwin D R Huitema⁴, Mirjam van de Velde⁵, Aniek Uittenboogaard⁵, Alaric Taylor^{1,3}, Jugal Suthar^{1,3}

¹ Vesynta Ltd, Innovation Gateway, The London Cancer Hub, Cotswold Road, Sutton, London, UK
² Newcastle University Centre for Cancer, Translational and Clinical Research Institute, Newcastle upon Tyne, UK
³ University College London, Dept of Chemical Engineering, London, UK
⁴ Department of Pharmacology, Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands
⁵ Emma Children's Hospital, Pediatric Oncology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Introduction

Background:

Vincristine is widely used in paediatric and adult cancer treatment but has a narrow therapeutic index and high inter-individual variability, making dosing challenging [1]. Existing population pharmacokinetic (popPK) models often perform inconsistently for a given patient group due to differences in model structure and source data [2, 3].

Modelling techniques such as averaging address model structure uncertainty by combining weighted predictions from multiple models. Alternatively, dataset augmentation improves model generalisability by integrating data from various sources [3,4]. This study investigates both methods to evaluate their impact on the accuracy of vincristine concentration predictions in clinical practice.

Aims:

- Explore model averaging and dataset augmentation for vincristine PopPK modelling to assess impact on predictive accuracy and generalisability.
- Compare their predictive performances against specialised single models on unexposed validation data splits.

Methods

Dataset Augmentation:

Vincristine pharmacokinetic data from four clinical trials [5-8] were combined to create an augmented dataset comprising 209 patients. Study inconsistencies were addressed by standardising units, harmonising covariates, and handling missing data and outliers. Eighteen unexposed patients were randomly selected and reserved for external performance evaluation. The remaining data were then stratified by key covariates and split into 10 training/validation sets: each comprised a training set (n=148) and an internal validation set (n=38), with representativeness confirmed using the Kolmogorov-Smirnov test.

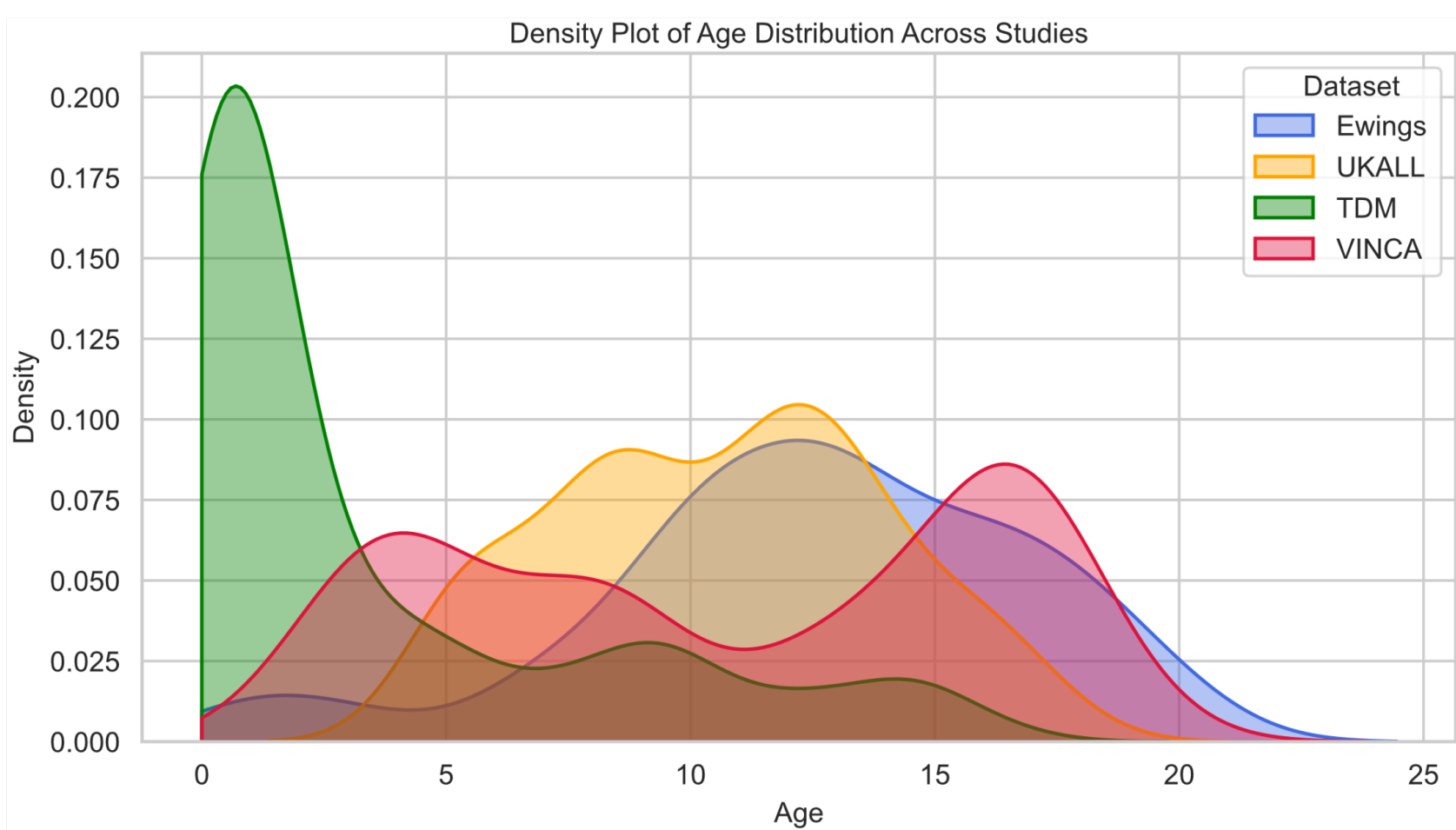


FIGURE 1. Density plot of age distribution across four clinical studies included in the augmented dataset.

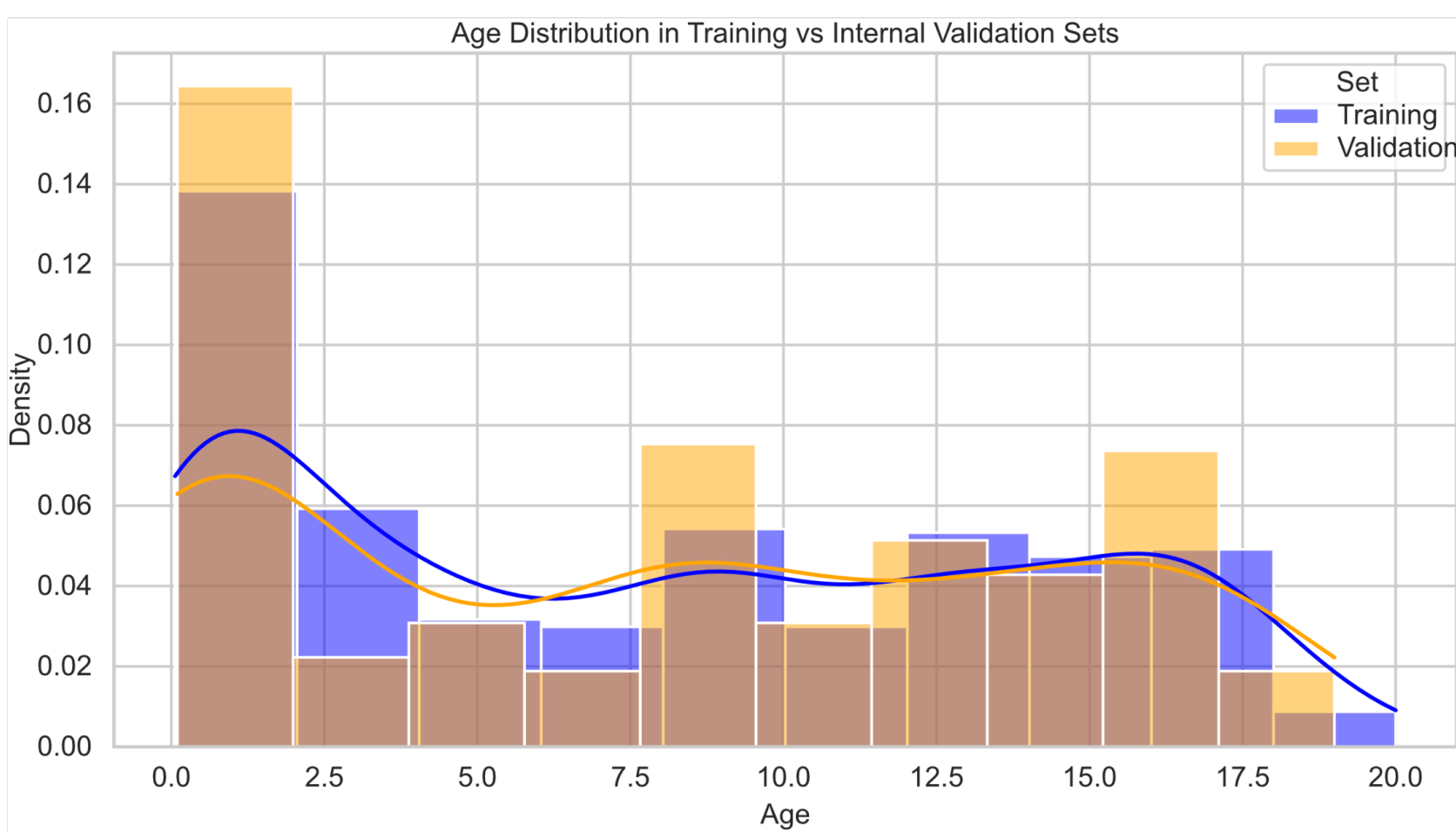


FIGURE 2. Age distribution to assess representativeness between training and validation sets.

A baseline model from Nijstad *et al.* was selected, featuring three-compartment structure with tubulin binding and multiple dosing occasions [1]. It was enhanced by adding inter-site variability on central and peripheral volumes and by applying the M3 method to handle concentrations below the limit of quantification. Parameters were re-estimated, and model performance was assessed using bootstrap, sampling importance resampling (SIR), and visual predictive checks (VPC).

Model Averaging:

Model averaging was implemented by combining predictions from six transcribed vincristine models [1, 8–12], with each model's contribution weighted inversely to its root mean square error (RMSE), following the methodology described by Uster *et al.* [2].

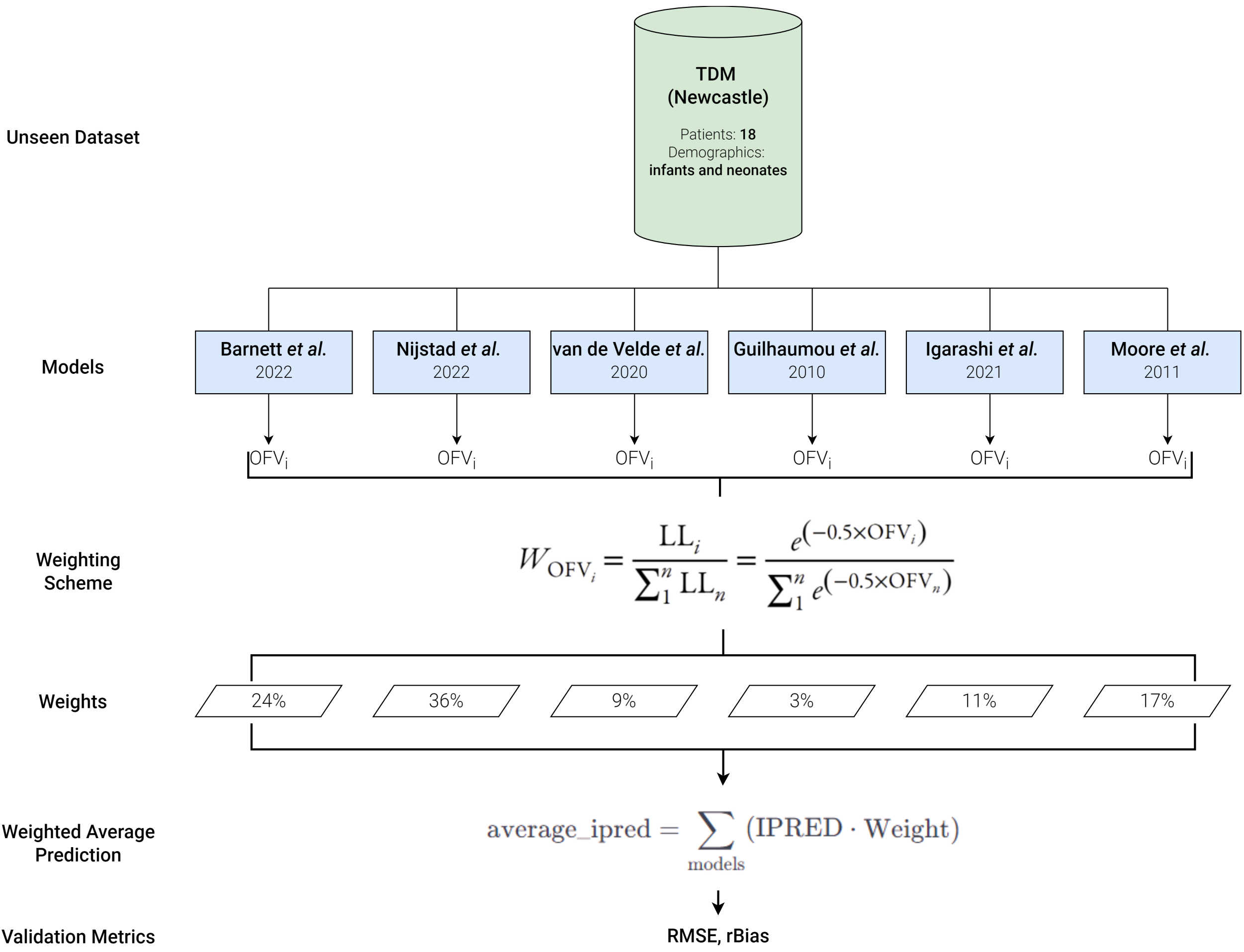


FIGURE 3. Workflow for model averaging using an unseen dataset.

Results

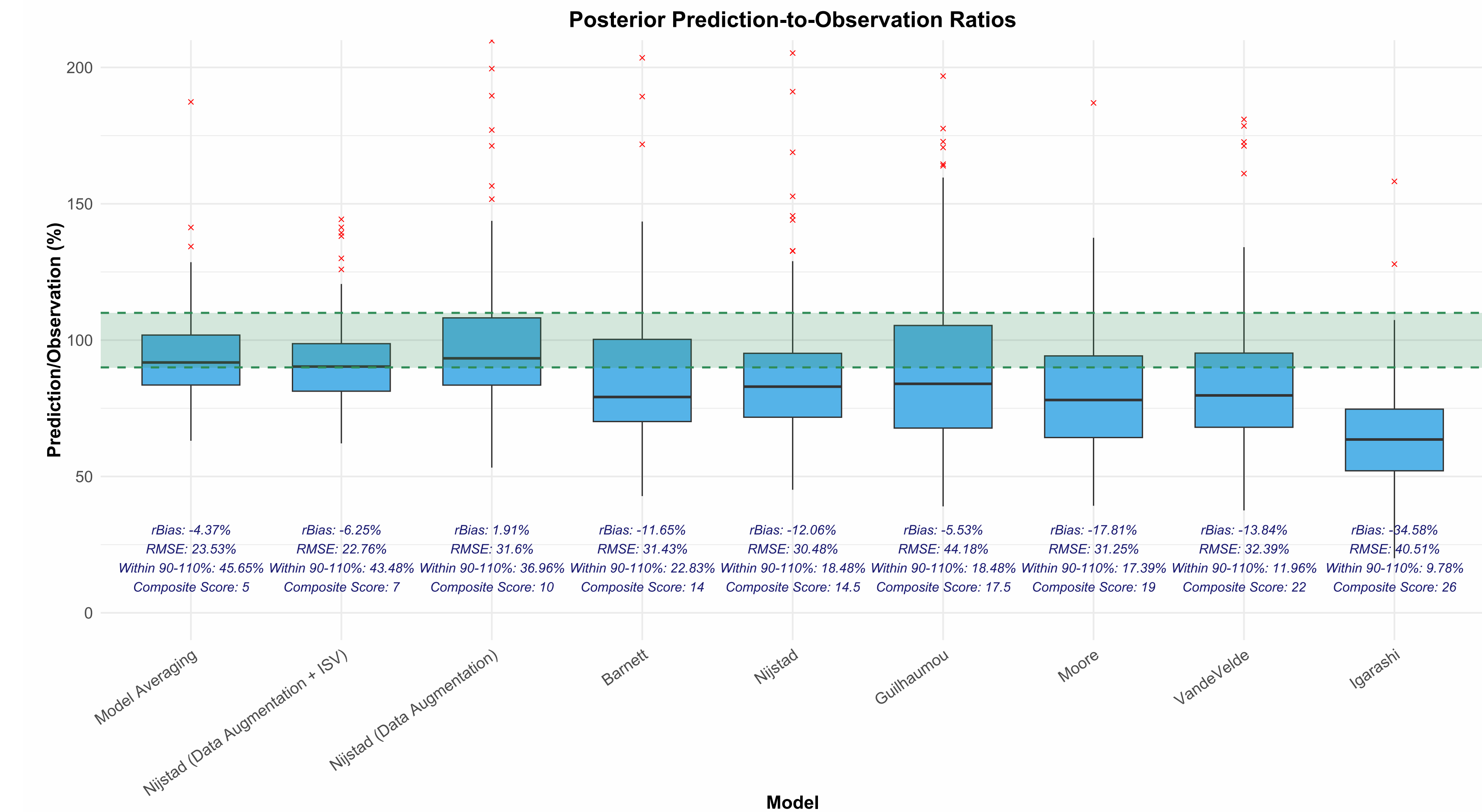


FIGURE 4. Comparison of posterior prediction-to-observation ratios across models.

- Refining the Nijstad *et al.* baseline model with inter-site variability and M3 handling for BLQ data significantly improved model fit (ΔOFV : -261).
- Across 10 data splits, model averaging outperformed all individual published models 8/10 times, while dataset augmentation outperformed individual models 7/10 times.
- Both approaches achieved lower prediction error (RMSE 23–31%) and bias (rBias between -6.3% and +1.9%), with 37–46% of predictions within the target range.
- Individual models showed greater variability and less accuracy, with RMSE values ranging from 30% to 44%, rBias from -34% to -6%, and only 10–23% of predictions within the target range.

Conclusion

- Model averaging and dataset augmentation both substantially improved predictive accuracy and consistency for vincristine concentrations compared to individual models.
- These approaches improve the reliability of popPK predictions, supporting safer and more precise model-informed dosing in clinical practice.

References:

[1] Nijstad AL *et al.* *Pharm Res.* 2022;39(10):2487-95.
[2] Uster DW *et al.* *Clin Pharmacol Ther.* 2021;109:175-83.
[3] Svensson E *et al.* *Br J Clin Pharmacol.* 2012;74(3):465-76.
[4] Agema BC *et al.* *Clin Pharmacokinet.* 2024;63:1449-61.
[5] ISRCTN Registry. ISRCTN36453794. 2014.

[6] ISRCTN10139334. <https://doi.org/10.1186/ISRCTN10139334>
[7] ISRCTN64515327. <https://doi.org/10.1186/ISRCTN64515327>
[8] van de Velde ME *et al.* *Cancers (Basel).* 2020;12(12):3745.
[9] Barnett S *et al.* *Eur J Cancer.* 2022;164:127-36.
[10] Moore AS *et al.* *J Paediatr Child Health.* 2011;47:875-82.
[11] Guilhaumou R *et al.* *Cancer Chemother Pharmacol.* 2011;68:1191-8.
[12] Igarashi T *et al.* *Cancer Chemother Pharmacol.* 2021;87(4):501-11.

Acknowledgements:

We would like to thank our collaborators from Newcastle Cancer Centre Pharmacology Group for providing access to the Ewings sarcoma study, TDM service and UKALL study data on Vincristine PK and Dr Mirjam Van de Velde for providing us with the data from the VINCA study.

The authors thank UK Research and Innovation for their financial support (Project ID 10111236).

Correspondence: arsenii.zats@vesynta.com

