Biology-driven, machine learning-based development of a biomarker to predict response to WEE1 inhibitor Debio 0123

Jeannette Fuchs¹, Luke Piggott¹, Kristian Urh², Matjaž Žganec², Eva Lavrenčič², Mark Uhlik², Carolina Haefliger¹ 1 Debiopharm International S.A., Lausanne, Switzerland | 2 Genialis Inc., Boston MA, United States

SUMMARY

Debio 0123 is an investigational, orally available, highly selective, and brain-penetrant adenosine triphosphate (ATP)-competitive inhibitor of the WEE1 tyrosine kinase, currently in phase one clinical trials. Inhibition of WEE1 presents an opportunity as a therapeutic target in cancer therapy, either in cells relying on cell cycle checkpoints regulated by WEE1 or to potentiate DNA damaging agents¹.

In this study, we present a first-generation digital biomarker leveraging the ResponderID[™] framework and Genialis[™] Supermodel to predict response to Debio 0123. Using an Extra-Trees machine learning algorithm trained on DNA damage response (DDR)-related pathways and validated on patient-derived organoid RNA-seq datasets, our predictor initially achieved an accuracy of 0.76, later optimized to 0.82. Notably, the optimized model successfully predicted independent patient-derived xenograft responses, distinguishing responders from non-responders, even in cases where ex vivo and *in vivo* outcomes differed. The biomarker model comprises biomodules previously connected to WEE1 inhibition as well as less characterized pathways.

These findings highlight the potential of a machine learning-driven approach to refine patient selection for WEE1 inhibitor therapies, providing a strong foundation for further clinical validation of Debio 0123.

BACKGROUND

Debio 0123 is a selective and orally available ATP-competitive inhibitor of WEE1 kinase

WEE1 acts at both the Intra-S-phase and G2/M checkpoints (Figure 1) making WEE1 inhibition amenable to combination with multiple chemotherapies with different mechanisms of actions.

The heterogeneity of cancer and the compensatory nature of DDR pathways pose significant challenges for therapies targeting DDR processes. Tumors often develop resistance to therapies through the engagement of alternative repair mechanisms. This limits the efficacy of current treatments, necessitating more refined approaches to predict and overcome therapeutic resistance². As such, patients are typically suboptimally stratified and experience widely varying treatment outcomes. There is a critical need for robust, clinically relevant biomarkers that can accurately identify patients who will benefit from DDR-targeted therapies. Such predictive biomarkers not only improve patient outcomes but also optimize clinical trial designs, informing therapeutic strategies that significantly de-risk development and increase trial success rates³.

To predict response to Debio 0123, we focused on a broad spectrum of DDR-related biomodules and applied a machine learning approach for modeling. The resulting predictor laid the groundwork for a versatile, biology-driven platform that can be further adapted and fine-tuned to clinical data or other relevant drugs in the portfolio.



▲ Figure 1.

In cancer cells, DDR pathways are often upregulated Inhibition of WEE1 reduces the phosphorylation of due to genomic instability. WEE1 is a key regulator CDC2 (CDK1) permitting cells to proceed through the of the Intra-S-phase and G2/M checkpoints where it cell cycle with an accumulation of DNA damage leading leads to cell cycle arrest allowing DNA damage repair. to mitotic catastrophe and ultimately cell death.



Genialis used 6,242 transcriptomic samples for unsupervised feature selection. The data were curated from public, partner and proprietary sources, spanning breast, colorectal, gastric, lung, pancreatic, and female reproductive cancers. Preclinical patient-derived xenograft organoid models (PDXO) of two cancer indications were used for machine learning training and testing. The dataset included indication 1 (n=30) and indication 2 (n=15) cancer samples, each with matched Debio 0123 IC50 values. Validation dataset consisted of three in vivo PDX models independently screened for response to Debio 0123. In

vivo PDX responders (n=1) were defined as tumor growth inhibition (TGI) >80% versus vehicle control, while nonresponders (n=2) were defined as TGI <10% versus vehicle control.



The two regression-based predictors (R0123-V1 and R0123-V2) each identified the same five DDR-related biomodules during supervised feature selection. The classification-based predictor (C0123-V1) identified four biomodules, all of which were a subset of the five found by the regression predictors (Table 1).

tive baseline (accuracy: 0.49 in both predictors) (Table 2). • Validation of the C0123-V1 predictor on an independent cohort accurately classified two out of three in vivo PDX models (Figure 3):

2 METHODS

Data

3 RESULTS

Predictive performance

• All predictors achieved superior performance compared to their respective baselines (Table 2).

• R0123-V1 achieved an accuracy of 0.76, while R0123-V2 improved to 0.82 by optimizing the decision threshold to 0.74 μM (*Table 2*, *Figure 2*).

• Both the regression-based predictor R0123-V2 (accuracy: 0.82) and the classification-based predictor C0123-V1 (accuracy: 0.80) significantly outperformed their respec-

- **1.** Correctly classified an *in vivo* PDX non-responder that showed <5% tumor growth inhibition
- 2. Correctly classified an in vivo PDX responder that showed >80% tumor growth inhibition
- **3.** Incorrectly predicted an *in vivo* PDX non-responder as a likely responder, with a low confidence score close to the decision boundary

▼ Table 1. Feature importance scores of individual **Genialis[™] Supermodel** biomodules that were used as input features in both regression- and classifi cation-based predictors.

Feature	Importance in regressor	Importance in classifer
Tumor suppressor pathway	0.29	0.23
Stress response pathway 1	0.21	/
Stress response pathway 2	0.19	0.30
Stress response pathway 3	0.16	0.34
Stress response pathway 4	0.15	0.13

Genialis™ Supermodel

Genialis[™] Supermodel is a large molecular model (LMM) which maps RNA-seq samples into cancer biology space and quantifies hundreds of biomodules that represent distinct biological processes and mechanisms.

A subset of 69 DDR-related biomodules was selected as an input feature set for machine learning, based on expert domain knowledge.

Feature selection

Unsupervised feature selection was used to identify tissueagnostic features by scoring biomodules based on similarity of distribution across tissue types.

Supervised feature selection was applied as a preprocessing step within cross-validation, using univariate feature selection and recursive feature elimination.



Figure 2. Scatter plot of predicted vs. measured IC50 values for indications 1 and 2 organoid samples using the R0123-V2 predictor (Extra Trees regressor, 0.74 µM optimized threshold). Predictions were generated in nested leave-one-out cross-validation. Blue dots (bottom left quadrant) are true responders, red dots (top right quadrant) are true non-responders, gray dots (bottom right quadrant) are false responders, and empty dots (upper left quadrant) are false non-responders.

▼ Table 2. Performance metrics of R0123-V1, R0123-V2, and C0123-V1 predictors computed in nested cross-validation and compared to their respective Dummy Baseline.

Performance metric	R0123-V1	Dummy Baseline	R0123-V2	Dummy Baseline	C0123-V1	Dummy Baseline
MSE [µM²]	0.15	0.41	0.15	0.41	/	/
Accuracy	0.76	0.60	0.82	0.49	0.80	0.49
Precision	0.78	0.73	0.87	0.55	0.78	0.48
Recall	0.94	0.73	0.80	0.55	0.88	0.48
F1-score	0.85	0.73	0.83	0.55	0.82	0.48



Modeling details

We evaluated three machine learning approaches to develop the first-generation digital biomarker for DDR: two regression-based and one classification-based. Performance was estimated using leave-one-out cross-validation and the predictors were benchmarked against a Dummy Baseline.

• R0123-V1:

An Extra Trees regressor trained to predict Debio 0123 IC50 values from indication 1 and indication 2 data. Samples were classified as responders or non-responders using a 1.00 μ M threshold that was informed by expert domain knowledge.

• R0123-V2:

An improved version of R0123-V1 with an optimized IC50 threshold for classification. The IC50 threshold was tuned to maximize the difference in classification accuracy between R0123-V2 and the Dummy Baseline. The optimal threshold was identified as $0.74 \mu M$.

• C0123-V1:

An Extra Trees classifier trained on indication 1 data to predict responder/non-responder status for Debio 0123. The target variable was defined using the IC50 threshold of 0.70 μ M optimized to maximize the difference in classification accuracy in comparison to the Dummy Baseline.

• Dummy Baseline: predictor.

Measured D0123 IC50 (ground truth) [µM]

The first-generation digital biomarker accurately predicts response to Debio 0123 in both patient derived organoid and in vivo xenograft models. The biology-driven, machine learning-based classifier outperformed the baseline classifier, supporting its potential clinical application.

Future work will focus on refining the predictor by incorporating additional DDR-related biomodules, validating performance on clinical datasets, and expanding its applicability to other DDR-targeting agents.

Sample rank

15

Figure 3. Rank plot of predicted response probabilities generated by the C0123-V1 predictor (Extra Trees classifier) in indication 1. Black circles represent measured responders, and white circles represent measured non-responders in the organoid setting. Blue indicates in vivo PDX responders, while red indicates in vivo non-responders. In blind validation using in vivo PDX data, the sample ranked 28 (red) was correctly predicted as a non-responder with <5% TGI. The sample ranked as 6 (blue) was correctly identified as an in vivo responder, despite being experimentally classified as a non-responder in the organoid setting. The sample ranked 12 (red), located near the decision boundary (dotted line), was incorrectly predicted and did not align with response in either the organoid or in vivo PDX model.



Debio 0123 is currently under phase I clinical investigation as a monotherapy (NCT05109975), in combination with carboplatin in patients with advanced solid tumors (NCT03968653), in combination with carboplatin and etoposide in patients with recurrent SCLC (NCT05815160), in combination with TMZ with or without RT in patients with GBM (NCT05765812) and in combination with the PKMYT1 inhibitor lunresertib in patients with advanced solid tumors (NCT04855656).

GGemialis

AACR 2025 | Abstract Number 3659

CONCLUSIONS

The biology-first machine learning predictors demonstrated solid performance in predicting response to Debio 0123 across preclinical models

Additional DDR-related biomodules will be integrated to refine predictions to clinical data

The predictors will be validated further on clinical data and other DDR targeting agents

These enhancements are expected to allow patient stratification in clinical trials, derisk development and allow patients access to precision medicine level treatment

ACKNOWLEDGEMENTS

We thank Debiopharm and the Genialis team, especially Marcel Levstek, Žan Kuralt, Roman Luštrik, Janez Kokošar, Miha Štajdohar, Luka Ausec, Eva Zupan Horaček and Rafael Rosengarten for their valuable contributions to this project.

CONTACT

Debiopharm International S.A., Lausanne, Switzerland. www.debiopharm.com | luke.piggott@debiopharm.com Genialis, Inc., 68 Harrison Ave #605, PMB 29417, Boston, MA 02111, USA, www.genialis.com

REFERENCES

- 1. Do K., et al. WEE1 kinase as a target for cancer therapy. Cell Cycle 2013;12:3159-64
- 2. O'Connor MJ. Targeting the DNA damage response in cancer. Molecular Cell 2015; 60:547-60
- 3. Parker JL., et al. Does biomarker use in oncology improve clinical trial failure risk? A large-scale analysis. Cancer Med 2021;10:1955-63

V SCAN FOR DOWNLOAD

- This poster is available via:
- www.debiopharm.com/medias/publications
- www.genialis.com/scientific-publications/



A baseline predictor that outputs IC50 values randomly drawn from the training dataset. For classification tasks, it applied the same decision threshold as the compared

