# Integrating Drug Structure and Target Binding Affinity for Improved Prediction of Survival in Cancer Patients Treated with Immune Checkpoint Inhibitors

ZEPHYR AI

**Sunil Kumar, Felicia Kuperwaser, Dillon Tracy, Jeff Sherman, Andrey Chursov, Emily Vucic and Maayan Baron**

Zephyr AI: 1800 Tysons Blvd Suite 901, McLean, VA, 22102 | https://www.zephyrai.bio

## ImmunoBERT: a deep learning framework to predict ICI patient response

Immune checkpoint inhibitors (ICIs) have transformed cancer therapy, yet their response rates remain modest, ranging from 20-40% across different cancer types [1]. There is a critical need for predictive tools to optimize treatments, avoid unnecessary side effects and identify patients most likely to respond to ICIs. Towards this goal, we developed a novel machine learning model for predicting overall survival (OS) in cancer patients undergoing treatment with ICIs, called ImmunoBERT, which takes as input clinical and molecular data currently available in real-world settings.

We curated a comprehensive clinicogenomics dataset of cancer patients treated with anti-PD1 and anti-CTLA4 checkpoint therapies (n=1700 patients), ICI drug structure embeddings, and binding affinity profiles of ICI drug targets. Using this dataset, we trained ImmunoBERT (Figure 1) which leverages large language models (LLMs) [2] and ProteinBERT (a deep learning model built upon the classic Transformer/BERT architecture) [3] to learn a generalization between ICI drugs, their protein targets, clinically available genomics data and patient outcome.

Correlations and higher-order interactions between 220 genes commonly sequenced on commercial NGS panels were also leveraged using ImmunoBERT architecture, to reconstruct features that improved ICI survival response prediction accuracy, including n=32 tumor microenvironment (TME) features, tumor mutational burden (TMB) and PDL1 expression.

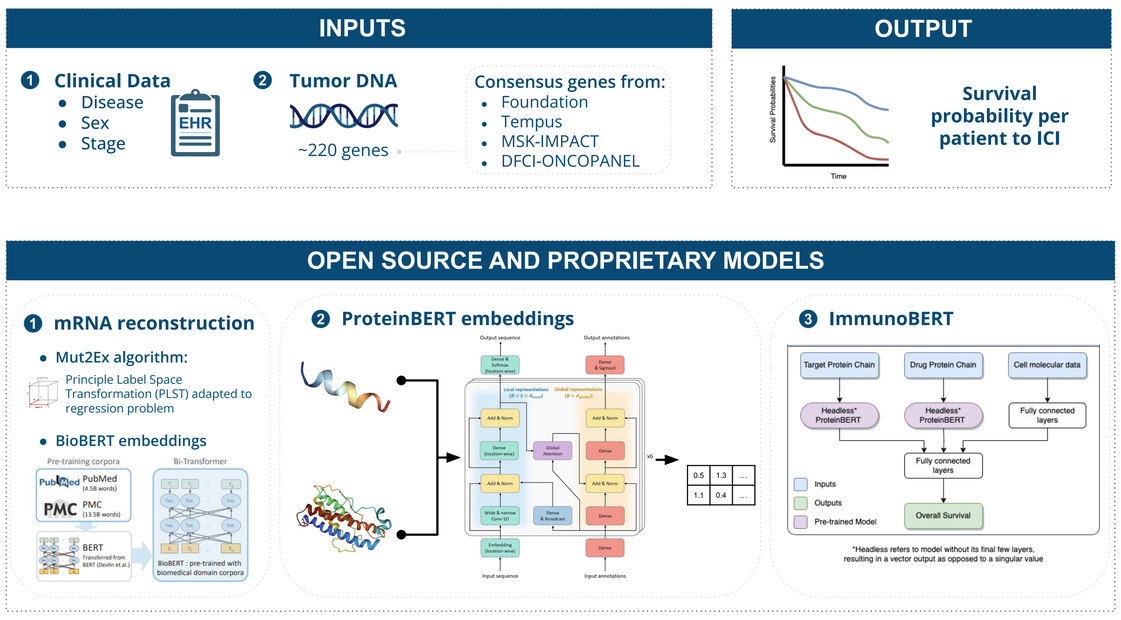


**Figure 1 | ImmunoBERT learns a generalization between ICI drugs, their protein targets, clinically available data, and patient outcome.** Clinical and NGS data available in real-world settings are used as inputs for ImmunoBERT (top, left). ImmunoBERT predicts a patient's survival probability to a specific ICI over an observation window (top, right). Several open source and proprietary models uncover signal in clinically available data that is useful for predicting outcome to immunotherapies. Real-world data is first used to reconstruct gene expression via two algorithms– Mut2Ex which leverages signal across cancer gene mutations and copy number alterations profiled on commercially available NGS panels, and BioBERT which leverages signal in patient clinical data (bottom, left). Drug and target protein chains are further used to encode protein binding information (bottom middle), which are fed into a deep neural network to estimate a patient's survival probability over an observation window (bottom right).

ImmunoBERT performance was benchmarked against top performing machine learning models from the Anti-PD1 Response Prediction DREAM Challenge [4]. The Anti-PD1 Response Prediction DREAM Challenge [5] specifically focuses on immunotherapies targeting the Anti-PD1 pathway in non-small cell lung cancer (NSCLC).

## Zephyr's models are purpose built to leverage RWD

Many of the top submissions to this challenge rely on gene expression data, which is often prohibitively expensive or otherwise impractical to obtain in clinical environments, outside of clinical trials or research settings.

Our approach leverages signal existing in molecular and clinical data commonly available in real-world settings to reconstruct a patient's tumor expression profile. This approach uses as input mutation and copy number data from widely used commercial next generation sequencing (NGS) multi-gene panels as well as clinically available annotations, such as disease type, patient gender, etc. Critically, we also supply the ICI drug structure embeddings and binding affinity profiles of ICI drug targets, so that the model can learn how the context of tumor's genetic profile affects the drug's efficacy.

We curated a comprehensive clinicogenomics dataset of cancer patients treated with anti-PD1 and anti-CTLA4 checkpoint therapies (n=1700 patients), ICI drug structure embeddings and binding affinity profiles of ICI drug targets.

Survival times recorded in these datasets were heavily right skewed (Fig 2A). In order to mitigate the impact on the model, patient survival times were truncated down to 24 months. Patients sustaining events past this point were re-labeled as censored.

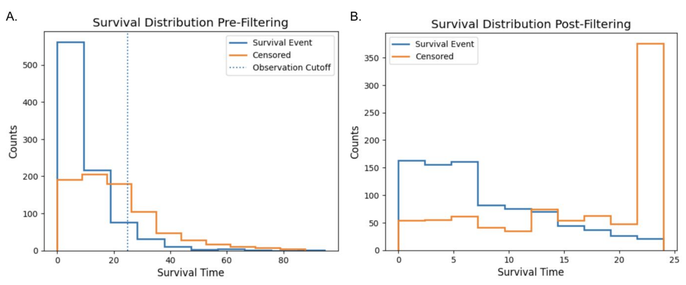


**Figure 2 | Event and censoring distributions are filtered to a tighter timeline.** Samples were processed to fit a shorter observation window of 24 months to balance the under- representation of survival events past this point.

This thresholding artificially spikes the number of patients censored at 24 months, but spreads the remaining patients more evenly (Fig 2B).

## ZephyrAI models outperform the state-of-the-art at predicting ICI survival

The C-index (or concordance index) statistic was used to evaluate and compare predictive accuracy of ImmunoBERT and different survival ML models, where higher C-index indicates better predictive accuracy of the model. Despite the limited specificity of our inputs, ImmunoBERT (C-index= 0.636) outperformed the top DREAM challenge models (C-index of 0.607 for top submission) in predicting patient response to ICI therapies.

| | Random Survival Forests | Fixed-Embedding Neural Net | ImmunoBERT | DREAM Challenge winner |
|---|---|---|---|---|
| **Concordance Index** | 0.621 | 0.630 | **0.636** | 0.607 |
| **Mean Dynamic AUC** | 0.335 | 0.323 | **0.437** | — |

**Figure 3 | Concordance Index and Dynamic AUC metrics comparing ZephyrAI models against the DREAM challenge.** All models that used target/drug embeddings outperformed the DREAM winner. Models that received fixed embeddings (Random Survival Forest, Fixed-Embedding Neural Net) achieved similar concordance scores as ImmunoBERT. Dynamic AUC was not reported by DREAM, but as shown in Fig.4A-B, it captures the performance difference between ImmunoBert and the other ZephyrAI models.

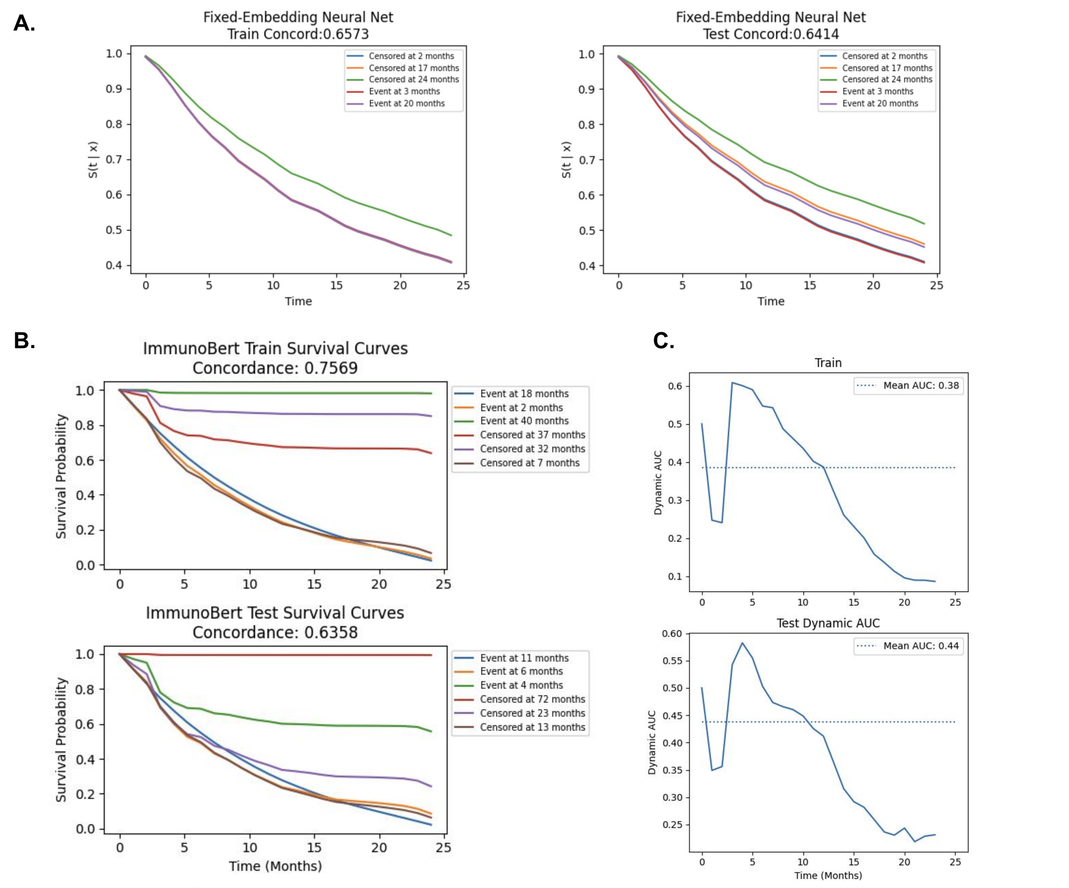


**Figure 4 | Learning ICI-specific drug/target embeddings improves the differentiation in predicted patient outcomes. A)** Survival Curves for a Fixed-Embedding Neural Net - Despite the variety of event and censoring times, there is very little differentiation between survival curves for each patient. **B)** Survival Curves for ImmunoBert - Although the concordance score is similar to that of the fixed embedding neural net, ImmunoBert is able to generate more specific and informative survival curves. **C)** Dynamic AUC Curve for ImmunoBert - AUC spikes to 0.58 when predicting survival 5 months into the future, with a steady drop off in AUC as the time of prediction reaches the end of the observation window.

## Incorporating target & drug protein structures improves ICI survival prediction

To test the efficacy of providing ICI drug and target embeddings to predict overall survival (OS), we first generated static embeddings for each drug/target combo using ProtBERT [6], a model pre-trained to reconstruct raw protein sequences. When fed to several out-of-the-box survival models, we immediately see concordance scores (on par with/outperforming) those from the top DREAM submissions (Fig 3). However, when plotting the survival curves, models were barely able to distinguish between samples with various survival times (Fig 4A).

ImmunoBert, on the other hand, utilizes transfer learning to specialize the generic protein embeddings to predict survival. While effects on concordance score were minimal, a clear differentiation in survival curve estimates for each sample was observed (Fig 4B).
When evaluating the accuracy and precision of ImmunoBert, we see its peak performance occurs when predicting survival 6 months out (Fig 4C), with a steady drop-off as survival time extends. The sharp drop in AUC from 2-3 months is likely due to the large proportion of samples that are right-censored within the first several months. We are accruing additional clinicogenomic ICI patient data to enhance overall model accuracy, particularly for extended time points.

## Conclusions

Our study demonstrates the value of integrating biologically relevant factors, such as drug structure, target binding affinity and genomic information, into machine learning models to improve accuracy of ICI response predictions. Moreover, by effectively leveraging real-world clinicogenomics data, including TME characteristics, we were able to reconstruct additional biologically relevant features, which further improved both the performance and interpretability of ImmunoBERT over current models. ImmunoBERT offers improved ICI prognostic capabilities, facilitating personalized treatment decisions to these promising drugs and enhancing patient care.

## References

1. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. Cell. 2017 Feb 9;168(4):707–23.
2. Lee J, Yoon W, Kim S, Kim D, Kim S, So CH, Kang J. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics. 2020 Feb 15;36(4):1234–40.
3. Brandes N, Ofer D, Peleg Y, Rappoport N, Linial M. ProteinBERT: a universal deep-learning model of protein sequence and function. Bioinformatics. 2022 Apr 12;38(8):2102–10.
4. Vincent BG, Szustakowski JD, Doshi P, Mason M, Guinney J, Carbone DP. Pursuing Better Biomarkers for Immunotherapy Response in Cancer Through a Crowdsourced Data Challenge. JCO Precis Oncol. 2021 Nov;5:51–4.
5. Mall, R., Ahmed, E., & Bedognetti, D. (2021, March 3). *Anti-PD1 Response DREAM Challenge I-MIRACLE*. Synapse. https://www.synapse.org/#!Synapse:syn23746812/wiki/608744
6. Elnaggar, A., Heinzinger, M., Dallago, C., Rehawi, G., Wang, Y., Jones, L., Gibbs, T., Feher, T., Angerer, C., Steinegger, M., Bhowmik, D., & Rost, B. (2020). ProtTrans: Towards cracking the language of life's code through self-supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1101/2020.07.12.199554

## Acknowledgements

The authors would like to acknowledge members of the Zephyr AI science, engineering and data teams. Special thanks to Candy Zhu for her valuable assistance in designing this poster.

Contact:
sunil@zephyrai.bio

**Abstract #1295**