

Letter to the editor: Quality criteria for computational models predicting individual outcomes in CAR-T cell therapy

Anna M Mc Laughlin ,¹ Cassian Yee ^{2,3}

To cite: Mc Laughlin AM, Yee C. Letter to the editor: Quality criteria for computational models predicting individual outcomes in CAR-T cell therapy. *Journal for ImmunoTherapy of Cancer* 2023;**11**:e006990. doi:10.1136/jitc-2023-006990

Accepted 06 April 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Pharmetheus AB, Uppsala, Sweden

²Department of Melanoma Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

³Department of Immunology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Correspondence to

Dr Anna M Mc Laughlin;
anna.mclaughlin@pharmetheus.com

Chimeric antigen receptor (CAR)-T cells have revolutionized the treatment of relapsed/refractory hematological malignancies. Using CAR-T cell therapy, a high percentage of patients respond initially, yet up to 60% of patients eventually relapse. Predictive and prognostic factors allowing to predict long-term response are not yet fully established.

Liu *et al* recently introduced a computational model of CAR-T cell immunotherapy with the aim to predict late responses from early stage clinical data in patients with leukemia.¹ By separately fitting a set of ordinary differential equations to different groups of patient data, for example, patients with complete response (CR), no response (NR), CD19⁺ relapse, and CD19⁻ relapse, the authors estimated distinct model parameter sets for each outcome.

Computational models for the prediction of long-term response in CAR-T cell therapy can be helpful and various methodologies may be applied for their development. However, regardless of the methodology chosen, critical quality criteria should be met. We have several concerns regarding the approach presented by Liu *et al*, with the issues falling into the following areas: the data and methodology applied for the model's development as well as communication of the model's abilities and limitations.

The authors highlight the development of a computational model for the prediction of individual responses to CAR-T cell therapy as their goal. As with any model (computational or not), it is a key and intuitive rule that the data used to develop the model should be representative of the data it will be later used to predict. In their manuscript, the authors state that data from 209 patients were used for the model's development. It is later specified that individual-level data were available for

only 8.6% (18/209) of these patients and that for the remaining patients, the 'individual data had been preprocessed and only statistical values such as medians were provided'. The authors justify this approach by claiming that statistical values could be regarded as representative individuals. Considering the high interindividual variability observed in CAR-T cell therapy, this assumption can clearly not be made. As the authors emphasize themselves, cell kinetics are highly variable between patients and, most importantly, often predictive of response. Therefore, using statistical summary parameters instead of individual-level data will dilute signals, create artificial, never observed data, and be overoptimistic in the prediction of individual trajectories. Methodologies allowing to simultaneously fit both individual and aggregate-level data in an unbiased way exist.² However, such methods do not seem to have been applied here.

Liu *et al* state that they performed non-linear mixed-effects (NLME) modeling to estimate both population-level and individual-level data. While they describe estimating random effects to quantify the interindividual variability in the population parameters, no such variability estimates are provided. It therefore is unknown on which parameters interindividual variability was implemented and how high the estimated variabilities were. Moreover, relative standard errors or confidence intervals, which quantify the uncertainty in the parameter estimates, are not provided.

In general, while the authors applied a NLME model algorithm, their model development strategy does not follow established NLME modeling practices. The advantage of a NLME approach is the simultaneous quantification of general trends (using fixed-effects population parameters) and unexplained

variability (using random-effects parameters) observed within the cohort. Once quantified, patient or treatment characteristics can be added to the model as covariates to explain parts of the unexplained interindividual variability. Once the impact of covariates on one or more model parameters has been estimated, these estimates can give new insights into the underlying reasons for the interindividual differences in the observed cell kinetics. This approach has been used in several excellent computational CAR-T cell models, whose review would be beyond the scope of this letter. The interested reader is referred to a recent review paper summarizing computational modeling approaches to support CAR-T cell clinical pharmacology strategies.³

In contrast to the established approach of estimating typical values, interindividual variability, and covariates for all patients simultaneously, Liu *et al* chose a different approach. Principally, they predicted early CAR T-cell response based on long-term outcome (responder status) rather than the other way around. By separately estimating a parameter set each for patients with CR, NR, CD19⁺ relapse, and CD19⁻ relapse, the authors missed the opportunity to quantify the interindividual cell kinetic differences and then investigate covariates to explain the observed differences in one joint model. In addition, by dividing their dataset into four subsets, the authors decreased the size of the development dataset for each outcome even further. Concretely, individual patient numbers remaining for the parameter estimation were five in the CR dataset, four in the NR dataset, six in the CD19⁺ relapse dataset, and three in the CD19⁻ relapse dataset. The additional statistical summary data increased the cohort sizes of the respective datasets, however, the value of statistical summary data for providing insight into individual cell kinetic and treatment outcome is low at best. Considering the large number of estimated fixed-effects (ie, typical) parameters (n=17) and an additional unreported number of random-effects parameters, it is unlikely that the model parameters were estimated with reliable precision for each of the four datasets. Finally, with their parameter estimates, the authors could successfully capture previously reported kinetic differences in patients with different outcomes (eg, patients with beneficial outcome usually have higher CAR-T cell expansion^{4,5}), but no new insights into patient or treatment characteristics associated with each outcome were generated. Thus, while the model correctly captures kinetic differences in patients with different outcomes, it is currently of a largely descriptive nature.

To test the performance of the model on a 'larger scale with higher reliability', the authors generated virtual patient cohorts by sampling values from an assumed Gaussian distribution of the estimated population-level parameters. The predicted peak values and areas under the concentration-time curve in the first 28 days (AUC28s) for the virtual patients were then compared with the peak values and AUC28s in the respective model development datasets. Because the virtual patient parameters were

generated from the model-estimated parameters, the similarity of the virtual patient cohorts with the observed data must be expected and cannot be viewed as a proof of the model's generalizability. Generally, the simulation of concentration-time profiles in a virtual population capturing the interindividual variability is an often-used tool to better understand the expected exposure ranges in a real-life population. The model's generalizability, however, would have been better addressed using external evaluation, that is, comparing predictions with observations of a new 'real' dataset, which would have not been used for the model's development.

A crucial quality criterion for a developed model is its translatability into providing potentially helpful clinical information. Liu *et al* identified different parameter sets for patients with different outcomes and proposed several model-derived secondary parameters as factors for use in early prediction of late-stage response. The practical usefulness of most of the proposed factors, however, is low as no link is made between the prediction factors and patient or treatment characteristics. Therefore, should one want to use the developed computational model to make an early response prediction for a new patient by calculating the CAR-T cell function or negative relapse factor, this would require fitting the model to the patient's individual cell kinetic data. However, for a computational model to potentially provide value for predicting clinical efficacy and inform researchers, its use should not be restricted to fellow modelers. Indeed, computational models for cancer immunotherapy are nowadays increasingly developed with a focus on clinical usability by providing simpler outputs in clinical language. An excellent overview on such mathematical models for personalized clinical translation is provided in a recent review article.⁶

The 'response prediction factor' is the authors' only proposed model-derived prediction factor that could be calculated without fitting the model, and it is defined as the log₂ of the product of the CAR-T cell peak value and the CAR-T cell area under the concentration-time curve in the first week after infusion. While certainly being accurate, the novelty of this factor is low, as the positive correlation between CAR-T cell exposure and beneficial outcome has been well established previously. Moreover, supported by an independent clinical analysis,⁵ we have identified a clinical composite score of CAR-T cell peak expansion normalized to the baseline tumor burden to be a better predictor for outcome than expansion alone.⁴

In general, while we are pleased that more computational research is being applied to the field of cell therapy, further development of any computational model (including our own) will need to include rigorous (re-) examination of modeling strategies, proper validation, as well as judicious and critical discussion of findings and their limitations before any of the developed models or model-derived prediction factors can be considered for use beyond research purposes.

Novel computational models for early response prediction in CAR-T cell treatment can be helpful, yet cell

kinetic data for their development are sparse. Appropriate methodology is available to jointly analyze individual and statistical summary data in an unbiased way and can be used together with a non-linear mixed-effects or another appropriate modeling approach to generate new insights. To potentially have true value in cell therapy design, the output or a derived simplified output of the model should be presented in a format and language well recognized by researchers and clinicians.

Twitter Anna M Mc Laughlin @anna_mclaughl and Cassian Yee @tcellsrus

Acknowledgements The authors thank Martin Bergstrand, PhD, E Niclas Jonsson, PhD, and Elodie Plan, PhD, of Pharmetheus AB (Uppsala, Sweden), and Professor Mats Karlsson, PhD, of Pharmetheus AB and Uppsala University (Uppsala, Sweden), for valuable input and discussions.

Contributors AMMcL and CY contributed equally.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests CY is on scientific advisory boards in the field of cellular therapy but none that is directly involved in the use of computational models.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Anna M Mc Laughlin <http://orcid.org/0000-0002-5936-1877>

Cassian Yee <http://orcid.org/0000-0002-4429-7307>

REFERENCES

- 1 Liu L, Ma C, Zhang Z, *et al.* Computational model of car T-cell immunotherapy dissects and predicts leukemia patient responses at remission, resistance, and relapse. *J Immunother Cancer* 2022;10:e005360.
- 2 Väitalo PAJ. Pharmacometric estimation methods for aggregate data, including data simulated from other pharmacometric models. *J Pharmacokinet Pharmacodyn* 2021;48:623–38.
- 3 Huang W, Li J, Liao MZ, *et al.* Clinical pharmacology perspectives for adoptive cell therapies in oncology. *Clin Pharmacol Ther* 2022;112:968–81.
- 4 Mueller-Schoell A, Puebla-Osorio N, Michelet R, *et al.* Early survival prediction framework in CD19-specific CAR-T cell immunotherapy using a quantitative systems pharmacology model. *Cancers (Basel)* 2021;13:2782.
- 5 Locke FL, Rossi JM, Neelapu SS, *et al.* Tumor burden, inflammation, and product attributes determine outcomes of axicabtagene ciloleucel in large B-cell lymphoma. *Blood Adv* 2020;4:4898–911.
- 6 Butner JD, Dogra P, Chung C, *et al.* Mathematical modeling of cancer immunotherapy for personalized clinical translation. *Nat Comput Sci* 2022;2:785–96.