PERSPECTIVE

Pharmacometrics Golems: Exposure-Response Models in Oncology

Akash Khandelwal[1,*], Ana-Marija Grisic[1], Jonathan French[2] and Karthik Venkatakrishnan[3]

Pharmacometric models are prone to different types of bias, which can confound the analysis and challenge the credibility of causal inference. This holds particularly true for exposure-response analysis with time-to-event end points. With ever wider use of pharmacometric models and increased recognition of confounding factors, advanced methods addressing these biases are being developed and increasingly utilized. Herein, we provide a perspective highlighting the limitations introduced by the biases and considerations for future applications.

Golems—creatures made from inanimate matter by humans—are not only a folklore motif, but can also be made by pharmacometricians *in silico*.[1] In the context of pharmacometrics, these are population pharmacokinetic/pharmacodynamic and exposure-response models used in drug discovery and development for predictions, mechanistic insights, and decision making. They too are typically not designed to have intent nor wisdom of their own and without our guidance, control, and prudence can lead to unintended consequences, such as false inferences from confounded analyses. Herein, we share our perspectives on one of the Golems of pharmacometrics, exposure-response models in oncology, with special emphasis on the potential of immortal time bias to lead to incorrect decision making.

Immortal time bias (ITB), also known as guarantee-time bias, arises when a time-to-event end point, such as disease-free survival, progression-free survival (PFS), or overall survival (OS), is related to predictors or compared between groups defined by a classifying variable that is measured post-baseline.[2] Although ITB is well-recognized in various fields, such as economics, military, and pharmacoepidemiology, its presence and impact in exposure-response analyses has only been appreciated in recent years. In the context of exposure-response analysis, the classifying variable susceptible to ITB could include drug exposure or post-baseline covariates. Some of the common examples in time-to-event models are objective response to treatment, maximal reduction of a biomarker, duration of treatment, onset of an adverse event (AE), post-baseline detection of anti-drug antibodies, and measurable drug exposure. In these cases, the observation of the classifying variable may be conditional on the main outcome not occurring—a patient cannot contribute observations after they died—thus introducing bias (**Figure 1**). In other words, a patient from whom the observation (e.g., an AE) was made is from the analysis perspective "immortal" until that time point. Of note, in this paper, we distinguish ITB from so-called survivor(ship) bias, which occurs in analyses where entering the study is conditional on survival (e.g., patients need to be alive long enough to receive any treatment), associated primarily with observational studies, and closely related to ITB.

Immortal time bias was first recognized in the 1980s, in a report by Anderson *et al.*,[2] wherein they investigated OS between responders and nonresponders to cancer treatment. In that case, ITB originates from the fact that responders need to survive long enough to respond to the therapy, and patients who die before the first response assessment are assigned to the nonresponse group. Thereby, the response group is given an "advantage," resulting in overestimation of survival in this group. More recently, Montomoli *et al.*[3] investigated the association of KRAS mutations with 1-, 2-, and 5-year survival after diagnosis in patients with metastatic colorectal cancer using the Kaplan–Meier technique. As the authors recognize in the paper, because only patients who were still alive at the time of KRAS testing could be included in the study, this analysis was confounded by ITB, likely resulting in an overestimation of survival compared to the general metastatic colorectal cancer population. Another example is use of duration

[1]The Healthcare Business of Merck KGaA, Darmstadt, Germany; [2]Metrum Research Group, Tariffville, Connecticut, USA; [3]EMD Serono, Billerica, Massachusetts, USA. *Correspondence: Akash Khandelwal (akash.khandelwal@emdgroup.com)

Received December 21, 2021; accepted February 17, 2022. doi:10.1002/cpt.2564

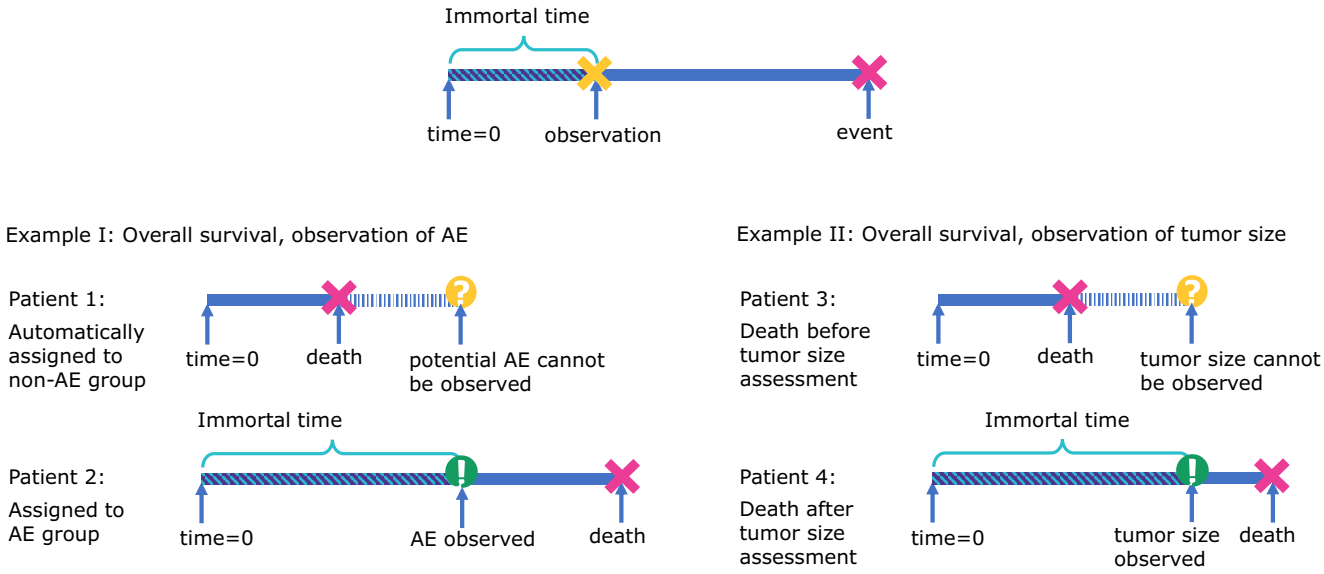

Figure 1 Illustration of immortal time bias (ITB) on examples of adverse event (AE) and tumor size (TS) observations.

of treatment as a predictor, as demonstrated in the study by Souhami *et al.*[4] That analysis investigated OS, disease-free survival, cause-specific mortality, local failure, and distant metastasis in patients with locally advanced prostate cancer who received hormonal therapy for up to 1 year, 1–5 years, and more than 5 years, concluding that therapy duration of > 5 years was associated with improved outcomes. This analysis is, however, confounded by ITB, as the patients who received therapy for > 5 years per definition had to survive for 5 years, resulting with 100% survival until year 5 in this subgroup. Furthermore, a study in patients with advanced non-small cell lung cancer compared PFS in patients who experienced immune-related AE and those that did not.[5] Although Kaplan–Meier analysis found a significant difference in PFS between the two groups, subsequent landmark analysis, which addresses ITB, found the difference not to be significant, pointing to the potential confounding impact of ITB in the Kaplan–Meier analysis. These are just some of the examples which are relevant in the context of covariate analysis in exposure-response models, where treating these variables as static covariates in exposure-response analysis would introduce ITB.

Certain sources of bias are inherent to the data and cannot be addressed by methodological approaches, such as selection bias in nonrandomized trials or ITB by design (more common in an observational study than a randomized control trial). However, with increasing incorporation of real-world data in pharmacometrics analysis, care must be taken at the analysis planning stage, as ITB can be introduced if time of eligibility, start of follow-up, and treatment assignment are not identical. In a typical randomized control trial, these times are identical and, hence, ITB by design is unlikely. ITB can also emerge in the context of therapeutic drug monitoring/precision dosing studies when treatment is determined based on a time-varying biomarker that is potentially itself a consequence of prior treatment, such that methods, such as semiparametric or parametric time-to-event models, with a time-varying biomarker may not reliably estimate treatment effect.

ITB introduced by naïve analysis (**Table S1**) can be addressed by several methodological approaches (**Table 1**). Although the use of such advanced methods in pharmacometric modeling has so far been scarce due to their complex implementation, they offer a means of mitigating ITB and thus warrant further recognition and broader use. Landmarking approaches (**Table 1**, approaches 1 and 2) address ITB by including only subjects who are still at risk (i.e., alive and did not drop out) by the prespecified landmark time(s). In these analyses, follow-up time is measured from the landmark time and all covariates and intercurrent events measured prior to the landmark time are eligible predictors of the landmarked time to event. The landmark analysis is, however, not without limitations, as predictors that occur after the landmark time are not considered, introducing a different bias in the context of the question being addressed by the analysis. For instance, use of cycle 1 systemic exposure (landmark time) in a time-to-event exposure-response analysis would not consider the changes in exposure over time (e.g., from dose modifications to manage AEs that occur during the course of treatment).

In the cloning and inverse probability of censoring weighting within the emulated trial approach (**Table 1**, approach 3), all subjects that experienced the observation and/or the event contribute their survival times to all categories of the classifying variable, thus controlling for ITB. The last three covered approaches (**Table 1**, approaches 4–6) address ITB by considering only data available by the current time point in the analysis (e.g., time-varying covariate information) in a dynamic analysis of longitudinal data, also making them akin to pharmacodynamic analysis of biological systems and therefore appealing from a clinical pharmacology perspective. Such dynamic models of longitudinal data, when coupled with dropout models where relevant, should importantly be well-suited to clinical trial

Table 1 Advanced exposure-response models in oncology

Approach[a]	Brief description	Question addressed	Limitations
(1) Landmark analysis	"Landmark time," (i.e., a time point after the trial start), is chosen and only subjects who survived/did not dropout by that time are included in the analysis	What is the treatment effect estimate in patients who lived long enough to reach landmark time (not the whole population)? What is the probability of success of phase III (GO/NOGO) given limited OS data from phase II in cases where the probability of early event is low?	Still selection bias, lack of randomization. Selection of landmark time arbitrary, results can change if different time used. Omission of events before landmark time. Based on data subset → affects statistical power.
(2) Landmark analysis with dynamic prediction	A landmarking approach, which dynamically adjusts predictive models for survival data during follow-up, by directly fitting models for patients still at risk at the landmark point (i.e., survival model is fitted from a series of landmarks as a function of predictors measured up to the landmark time)	What is the treatment effect estimate in the investigated patient population? What is the probability of success of phase III (GO/NOGO) given limited OS data from phase II?	Although straightforward implementation and computational simplicity, it uses less information from the longitudinal covariates than joint modeling. Not fitted using a full likelihood. Inferior to joint modeling, although less sensitive to misspecification of longitudinal process.
(3) Cloning and IPCW within the emulated trial approach	This method is only applicable to the analysis of observational data in the emulated trial framework wherein immortal-time bias is introduced when the start of follow-up and treatment initiation do not coincide. In this method, exact copies of each individual are created ("cloned") and each subject is simultaneously assigned to all classifier categories. Once the category deviates from the observed category for the subject, the subject is at that time censored in that category. Thus, until the time when the observation of interest is made, the subject contributes to all categories. Subsequently, potential selection bias introduced by censoring is adjusted for through "inverse probability of censoring weighting," whereby uncensored subjects are up-weighted to represent censored subjects.	What is the treatment effect estimate in the investigated patient population?	Complex approach. All confounders should be measured, which is difficult to assess or achieve in practice.
(4a) Extended time-dependent Cox regression (with model-derived covariates)	Cox regression modeling including observed or (better) model-derived time-varying covariates	What is the treatment effect estimate in the investigated population? What is the impact of time-varying covariates on clinical end points (i.e., progression-free survival and OS)? What is the probability of success of phase III (GO/NOGO) given limited OS data from phase II?	Compared to standard Cox less intuitive interpretation. Potentially biased and imprecise estimate of treatment effect compared with joint model. If modeling time-varying covariates, complex implementation, functional representation of time-varying covariates needs to be known. If using observed covariates, limited ability to simulate outcomes due to lack of model for time-varying covariate(s).

(Continued)

Table 1 (Continued)

Approach[a]	Brief description	Question addressed	Limitations
(4b) Sequential parametric model (with model-derived covariates)	Parametric time to event model including observed or (better) model-derived time-varying covariates	What is the treatment effect estimate in the investigated population? What is the impact of modeled time-varying covariates on clinical end points (i.e., progression-free survival and OS)? What is the probability of success of phase III (GO/NOGO) given limited OS data from phase II?	Potentially biased and imprecise estimate of treatment effect compared to joint model. Complex implementation; Functional representation of time-varying covariates needs to be known. If using observed covariates, limited ability to simulate outcomes due to lack of model for time-varying covariate(s).
(5) Joint model of longitudinal covariates and TTE model	For all covariates of interest, longitudinal models are developed simultaneously with TTE model	What is the estimate of treatment effect in the investigated population? What is the impact of modeled time-varying covariates on clinical end points (i.e., progression-free survival and OS)? What is the probability of success of phase III (GO/NOGO) given limited OS data from phase II?	Complex implementation, incl. potential computation difficulties (e.g., overfitting due to model complexity, challenging estimation. Only covariates for which longitudinal model is developed can be investigated. Concerns about robustness. In case of OS, cannot determine effect of time-varying covariates on hazard post-progression.
(6) Multi-state model	Transitions between predefined states (e.g., response, progression, death) are modeled by using information up until the TTE	What are relationships between time to response, the duration of response, progression-free survival and the OS in a unified framework? What is the impact of time-varying covariates on clinical end points (i.e., response, progression and death)? What is the probability of success of phase III (GO/NOGO) given response and limited OS data from phase II?	Complex implementation, incl. potential computation difficulties (e.g., overfitting due to high number of covariates). For certain cases, prediction can difficult, dependent on the data.

IPCW, inverse probability of censoring weighting; OS, overall survival; TTE, time of transition.

[a]The reader is referred to the Supplementary data for in-depth literature on details and mathematical formulations for the advanced exposure-response models.

simulations to evaluate the performance of alternate trial designs, biomarker selection algorithms, or dosing regimens, thereby enabling model-informed precision medicine development.

Although all advanced approaches are complex in terms of implementation, data requirements and, in some cases, interpretation, the advantages they offer in terms of achieving unbiased estimation of exposure-response relationships and enabling causal inferences are tantalizing to consider their broader evaluation to enable adoption. Within the oncology clinical pharmacology and pharmacometrics community, ITB is increasingly recognized, and alternative approaches are being utilized to partially circumvent ITB. Nishino *et al.*[6] performed 3-month conditional landmark analyses (**Table 1**, approach 1) as well as extended Cox models with time-dependent covariates (**Table 1**, approach 4) to assess the tumor size-OS relationship of pembrolizumab in patients with melanoma. The OS model developed by Zheng *et al.*[7] for durvalumab in the urothelial cancer population investigated model-predicted tumor response as a time-varying covariate (**Table 1**, approach 4). Netterberg *et al.*[8] linked the time course of atezolizumab systemic exposure (area under the curve) to relative change of IL-18 from baseline, which, in turn, was linked to tumor dynamics and ultimately used as a predictor of OS in patients with non-small cell lung cancer (**Table 1**, approach 4). Krishnan *et al.*[9] developed a multistate model (**Table 1**, approach 6) that described the transitions between states capturing jointly the overall event and survival data in patients with HER2-negative breast cancer. The longitudinal tumor size model derived metrices were used as covariates on the transition hazard in a prospective manner to alleviate ITB.

The benefit of advanced analysis to mitigate ITB compared to naïve analysis is more noticeable when the number of early outcome events is high or the classifying variable occurs later in time. For example, when relating early changes in tumor size to OS, a naïve analysis is more prone to ITB in populations in which the hazard of death during the early follow-up is high (i.e., higher number of early events) than in populations in which the hazard of death during the early follow-up is low (i.e., lower number of early events).[10] There are instances when the presence of ITB can be easily recognized (e.g., if a Kaplan–Meier curve of one of the compared groups is flat at the beginning), implying 100% early survival for this group. This is exemplified in the above-mentioned analysis[4] where the Kaplan–Meier plot clearly shows 100% survival rate for first 5 years in the group treated for > 5 years.

Comparison of the results from multiple approaches (e.g., ref. 6) is beneficial for evaluation of the impact of ITB on the results, however, in practice, this may have limited feasibility and lead to substantial prolongation of analysis timelines. Therefore, in our opinion, the context of use of the model and the potential for ITB must be kept in mind before defining the analysis strategy, including selection of any of the advanced approaches to mitigate impact of ITB. Furthermore, we posit that continued refinement of approaches to mitigate biases (including but not limited to ITB) will be important to enhance fidelity of model-informed drug development frameworks to ultimately increase the probability of success in oncology drug development.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

FUNDING

No funding was received for this work.

CONFLICT OF INTEREST

The authors declared no competing interests for this work. As an Associate Editor for *Clinical Pharmacology & Therapeutics*, Karthik Venkatakrishnan was not involved in the review or decision process for this paper.

1. McElreath, R. Statistical rethinking: A Bayesian course with examples in R and Stan, 2nd edn. CRC Press, Boca Raton, FL and Abingdon, Oxon (2020).
2. Anderson, J.R., Cain, K.C. & Gelber, R.D. *J. Clin. Oncol.* **1**, 710–719 (1983).
3. Montomoli, J., Hamilton-Dutoit, S.J., Frøslev, T., Taylor, A. & Erichsen, R. Retrospective analysis of KRAS status in metastatic colorectal cancer patients: a single-center feasibility study. *Clin. Exp. Gastroenterol.* **5**, 167–171 (2012).
4. Souhami, L., Bae, K., Pilepich, M. & Sandler, H. Impact of the duration of adjuvant hormonal therapy in patients with locally advanced prostate cancer treated with radiotherapy: a secondary analysis of RTOG 85–31. *J. Clin. Oncol.* **27**, 2137–2143 (2009).
5. Sato, K. *et al*. Correlation between immune-related adverse events and efficacy in non-small cell lung cancer treated with nivolumab. *Lung Cancer* **115**, 71–74 (2018).
6. Nishino, M. *et al*. Immune-related tumor response dynamics in melanoma patients treated with pembrolizumab: identifying markers for clinical outcome and treatment decisions. *Clin. Cancer Res.* **23**, 4671–4679 (2017).
7. Zheng, Y. *et al*. Population modeling of tumor kinetics and overall survival to identify prognostic and predictive biomarkers of efficacy for durvalumab in patients with urothelial carcinoma. *Clin. Pharmacol. Ther.* **103**, 643–652 (2018).
8. Netterberg, I. *et al*. Tumor time-course predicts overall survival in non-small cell lung cancer patients treated with atezolizumab: dependency on follow-up time. *CPT Pharmacometrics Syst. Pharmacol.* **9**, 115–123 (2020).
9. Krishnan, S.M., Friberg, L.E., Bruno, R., Beyer, U., Jin, J.Y. & Karlsson, M.O. Multistate model for pharmacometric analysis of overall survival om HER2-negative breast cancer patients treated with docetaxel. *CPT Pharmacometrics Syst. Pharmacol.* **10**, 1255-1266 (2021).
10. Liu, P.Y., Voelkel, J.O., Crowley, J. & Wolf, M. Sufficient conditions for treatment responders to have longer survival than non-responders. *Stat. Probab. Lett.* **18**, 205–208 (1993).