Tumor growth inhibition modeling in patients with second line biliary tract cancer and first line non-small cell lung cancer based on bintrafusp alfa trials

Ana-Marija Milenković-Grišić[1] | Nadia Terranova[2] | Diane R. Mould[3] | Yulia Vugmeyster[4] | Thomas Mrowiec[1] | Andreas Machl[4] | Pascal Girard[2] | Karthik Venkatakrishnan[4] | Akash Khandelwal[1]

[1]The healthcare business of Merck KGaA, Darmstadt, Germany

[2]Quantitative Pharmacology, Ares Trading S.A. (an affiliate of Merck KGaA, Darmstadt, Germany), Lausanne, Switzerland

[3]Projections Research Inc, Phoenixville, Pennsylvania, USA

[4]EMD Serono, Billerica, Massachusetts, USA

Correspondence

Ana-Marija Milenković-Grišić, The Healthcare Business of Merck KGaA, Frankfurter Str. 250, 64293, Darmstadt, Germany.
Email: ana-marija.milenkovic@emdgroup.com

Nadia Terranova, Merck Institute for Pharmacometrics (an affiliate of Merck KGaA, Darmstadt, Germany), Lausanne, Switzerland.
Email: nadia.terranova@emdgroup.com

Abstract

This analysis aimed to quantify tumor dynamics in patients receiving either bintrafusp alfa (BA) or pembrolizumab, by population pharmacokinetic (PK)-pharmacodynamic modeling, and investigate clinical and molecular covariates describing the variability in tumor dynamics by pharmacometric and machine-learning (ML) approaches. Data originated from two clinical trials in patients with biliary tract cancer (BTC; NCT03833661) receiving BA and non-small cell lung cancer (NSCLC; NCT03631706) receiving BA or pembrolizumab. Individual drug exposure was estimated from previously developed population PK models. Population tumor dynamics models were developed for each drug-indication combination, and covariate evaluations performed using nonlinear mixed-effects modeling (NLME) and ML (elastic net and random forest models) approaches. The three tumor dynamics' model structures all included linear tumor growth components and exponential tumor shrinkage. The final BTC model included the effect of drug exposure (area under the curve) and several covariates (demographics, disease-related, and genetic mutations). Drug exposure was not significant in either of the NSCLC models, which included two, disease-related, covariates in the BA arm, and none in the pembrolizumab arm. The covariates identified by univariable NLME and ML highly overlapped in BTC but showed less agreement in NSCLC analyses. Hyperprogression could be identified by higher tumor growth and lower tumor kill rates and could not be related to BA exposure. Tumor size over time was quantitatively characterized in two tumor types and under two treatments. Factors potentially related to tumor dynamics were assessed using NLME and ML approaches; however, their net impact on tumor size was considered as not clinically relevant.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Clinical trials of the investigational agent bintrafusp alfa in first-line non-small cell lung cancer (1L NSCLC) and second-line biliary tract cancer (2L BTC) did not meet the primary efficacy end points.

WHAT QUESTION DID THIS STUDY ADDRESS?

Are there patient-specific factors that explain the variability in antitumor activity in the respective clinical trials and can this be assessed using population tumor dynamic models with a comprehensive covariate assessment including machine-learning (ML) algorithms?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

Although an exposure effect on tumor shrinkage was demonstrated for BA in 2L BTC, no exposure-effect relationship could be discerned in 1L NSCLC and no clinically relevant covariate effects were identified in both populations.

HOW MIGHT THIS CHANGE DRUG DISCOVERY, DEVELOPMENT, AND/OR THERAPEUTICS?

Viewed from a broader perspective, these analyses illustrate a two-pronged approach incorporating both classical pharmacometrics and ML to covariate analyses, thereby contributing to the emerging area of pharmacometrics enhanced by advanced analytics, with opportunities to enable model-informed precision medicine.

INTRODUCTION

Bintrafusp alfa (BA; MSB0011359C or M7824) is a first-in-class bifunctional fusion protein composed of the extracellular domain of the transforming growth factor-β receptor II (a TGF-β "trap") fused to a human immunoglobulin G1 antibody blocking programmed death ligand 1 (PD-L1),[1] that thereby has potential to inhibit two key immunosuppression pathways in the tumor microenvironment simultaneously.[2] Two recent BA clinical trials, versus pembrolizumab in first-line non-small cell lung cancer (1L NSCLC; NCT03631706),[3] and in second-line biliary tract cancer (2L BTC; NCT03833661),[4] failed to replicate the encouraging results from earlier studies, warranting investigation to elucidate potential factors that might indicate sensitivity or resistance to BA or pembrolizumab treatment, as well as further elucidation of hyperprogression, defined as an unexpected rapid progression of disease in patients receiving immunotherapy,[5] also detected in these trials.

Previously, modeling and simulation approaches have been used to characterize the pharmacokinetics (PKs) of BA in various tumor types[6] and to support selection of the recommended phase II dose.[7] Furthermore, more recently population PK and exposure-safety analyses have been used to inform risk management of bleeding toxicities.[8] None of the analyses, however, investigated tumor dynamics under BA treatment so far. Tumor growth inhibition modeling enables estimation of drug exposure effects,

identification of significant covariates on tumor dynamics parameters, and quantification of their effects, thus potentially informing hypotheses around patient populations that might show higher or lower sensitivity to treatment. Moreover, machine-learning (ML) methods are increasingly being applied in different stages of drug discovery and development,[9,10] including more recent applications to guiding population model development,[11] and of particular interest in screening of covariates.[12]

The objectives of this analysis were to quantify tumor dynamics under BA and pembrolizumab treatment by tumor growth inhibition modeling, and to identify covariates that may explain the variability in tumor dynamics by using traditional pharmacometric covariate analysis and ML techniques.

METHODS

Clinical trial design and patients

This analysis (Figure 1) included data from two multicenter, international, open-label trials, in BTC and NSCLC. The BTC trial[13] was a phase II, single-arm trial designed to evaluate BA monotherapy in patients with locally advanced or metastatic BTC previously treated with 1L platinum-based chemotherapy. The 159 patients enrolled received BA at a dose of 1200 mg i.v. every 2 weeks
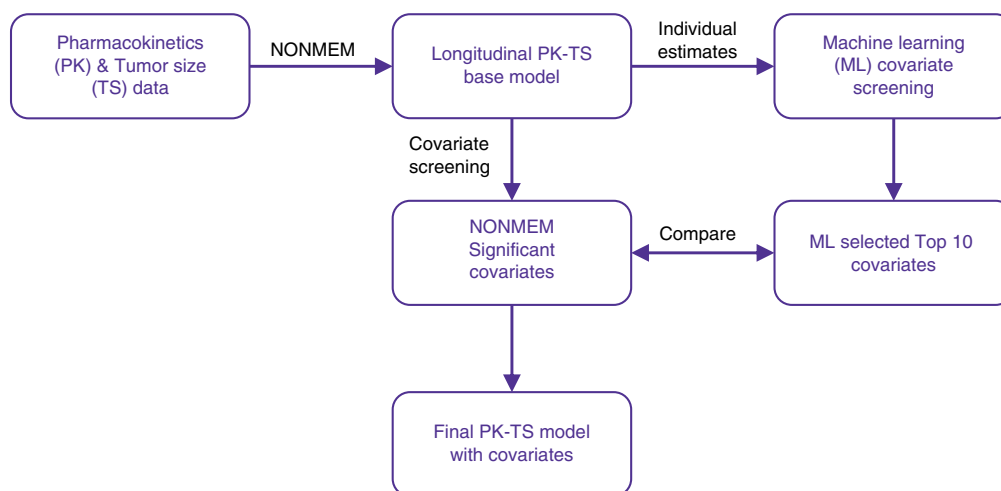
FIGURE 1 Overview of the analysis workflow.

(q2w). The NSCLC trial[14] was an adaptive phase III, randomized, controlled trial in patients with advanced NSCLC with high PD-L1 tumor expression (≥80% PD-L1 positive tumor cells). Patients ($n = 304$) who have not received previous treatment were enrolled and randomly assigned to receive either BA at a dose of 1200 mg i.v. q2w or pembrolizumab at a dose of 200 mg q3w in a 1:1 ratio. Schedule of tumor size (TS) assessments and PK sampling are given in Table S1.

Pharmacokinetics

The previously developed BA population PK model[8] was used to estimate the individual drug exposure (dynamic area under the concentration-time curve [AUC]) for BA for both trials, using the actual dosing information that accounts for treatment interruptions. In the absence of pembrolizumab concentration data in the current study, a previously developed population PK model was used to characterize pembrolizumab exposure.[15] Therefore, pembrolizumab exposure (AUC) reflects the value for a typical subject as defined by the covariate model alone instead of individual patient exposure.

Tumor size modeling

The TS measure used in this analysis was the sum of longest diameters, as per the Response Evaluation Criteria in Solid Tumors (RECIST).[16] In NSCLC the TS was measured mainly using computed tomography (CT) or X-ray; in BTC mainly CT and/or magnetic resonance imaging were used. In addition, patients' tumor growth was characterized as being hyperprogressive using the method

described in Matos et al.[5] that is, "based on RECIST [1.1] as progressive disease (PD) in the first 8 weeks after treatment initiation and minimum increase in the measurable lesions of 10 mm plus: (i) increase of ≥40% in sum of target lesions compared with baseline [which represents doubling in unidimensional target lesions compared with classic RECIST PD criterion (20%)]; and/or (ii) increase of ≥20% in sum of target lesions compared with baseline (the classic RECIST PD criterion) plus the appearance of new lesions in at least two different organs."[5] In NSCLC, all patients in the two arms had at least two tumor measurements. In BTC, there was one patient with only one tumor measurement, and this patient was retained in the analysis.

Several published models of tumor dynamics were evaluated, including: (1) the Claret et al.[17] model, which described an exponential decrease in the rate of tumor shrinkage to empirically account for resistance development, (2) a modified turnover model of Tham et al.,[18] (3) a simpler turnover model with first order growth and decay,[19,20] (4) a Gompertz function that described sensitive and resistant cell subpopulations,[21] and (5) a more complex model described by Chatterjee et al.[22] as a more physiologically based function that described a fraction of the tumor that is shrinking and a fraction that is not affected by the treatment.

Covariate evaluations

As one of the objectives of this work was to identify covariates that might explain variability in tumor dynamics, baseline values of preselected covariates were tested. Overall, the herein reported analyses evaluated covariates that were preselected for their potential association with

the mechanism of action of the drug and its efficacy, or with the disease, including biology of the drug targets and tumor. These included demographics (e.g., sex, age, and body size), disease-related characteristics (e.g., Eastern Cooperative Oncology Group status, histology, and necrotic tissue), disease-related laboratory values (e.g., albumin, C-reactive protein [CRP] and lymphocytes), therapy-related (e.g., prior therapy, anti-drug antibody status, and time since diagnosis to first dose), and genetic mutations (e.g., *IDH1*, *KDR* mutation, and minor allele frequency for all detected somatic variants in the sample). Detailed lists of covariates investigated for the BTC and NSCLC analyses are given in Supplementary Methods. Missing covariate information was imputed by the median value for continuous covariates and by the mode for categorical covariates. Covariate assessment comprised multiple steps, including (1) univariable covariate runs, (2) full covariate model followed by reduction, and (3) covariate evaluations using ML approaches.

Covariate evaluations via nonlinear mixed-effects approach

Single covariate selection was based on the log-likelihood criterion, goodness-of-fit plots, parameter precision, and scientific plausibility. All covariates resulting in a significant ($p < 0.01$) reduction in the objective function (i.e., χ^2(1 df) > 6.635) and a reduction in interindividual variability (IIV) for the modeled parameter compared to that for the base model were combined into a full model. Subsequently, a backward elimination step was conducted whereby the covariates were sequentially set to their null value and those that resulted in a statistical worsening of the model objective function ($p < 0.001$) or were poorly estimated (relative standard errors [RSEs] >50%) were eliminated from the model, resulting in the final models.

For all non-linear mixed-effects (NLME) modeling analyses NONMEM (version ≥7.4.4; ICON Development Solutions)[23,24] was used, and SAS version 9.4 (SAS Institute) was used for the exploratory analysis, model development, and post-processing of NONMEM output. In addition, Xpose version 4.4.0[25] PsN (psn.sourceforge.net),[26,27] and R[28] software packages were used.

Covariate evaluations via ML approach

In addition to the NLME modeling approach, ML was implemented to evaluate the baseline covariates of interest on empirical Bayes parameter estimates. Data assembly and ML methods used the Python Data Analysis Library, pandas 1.1.5.,[29] NumPy 1.21.6.[30] and scikit-learn library

1.0.2. Values were standardized and redundant features showing strong multicollinearity were dropped from the analysis dataset prior to modeling. Two different ML algorithms were used for the screening procedure: (1) a regularized linear regression and (2) tree-based model with recursive feature elimination. The methods for covariate selection are described below with further details provided in Supplementary Methods.

Within the elastic net model, L1, L2 penalties, and alpha weighting values for each covariate model were selected using grid search over a range of values. Indicators of the importance of each feature for the final elastic net covariate model were: (1) absolute values of the normalized co-efficient, (2) permutation importance, and (3) Shapley (SHAP) Feature Importance.

Within the random forest (RF) model, parameters for each random forest regressor covariate model were tuned via grid search over a range of different values. These parameters were: (1) number of estimators or trees, (2) maximum tree depth, and (3) minimum number of samples at which to split. Indicators of the importance of each feature in the final RF covariate model were: (1) impurity based feature importance (gini), (2) permutation importance, and (3) SHAP.

RESULTS

Structural tumor dynamic model across treatment and tumor types

The tumor trajectories by study and treatment are presented in Figure S1. Hyperprogression was identified in BA arms (43 and 11 patients with hyperprogression in 2L BTC and 1L NSCLC, respectively), but not in pembrolizumab arm. In all three evaluated cases, that is, under both treatments and for both tumor types, the model with log-normally distributed exponential tumor shrinkage component (describing the treatment-induced tumor shrinkage over time that would result in reducing the tumor size asymptomatically toward zero), linear tumor growth component,[31] and residual variability modeled using log-transform both sides approach with an additive error model described the data best:

$$TS_i(t) = BTS_i \bullet e^{-ks_i \bullet t} + kg_i \bullet t \qquad (1)$$

where $TS_i(t)$ is the tumor size at time t (weeks) for the i-th individual, BTS_i is the baseline tumor size, ks_i is the exponential tumor shrinkage rate constant, and kg_i is the linear tumor growth rate constant;

$$\ln(Y_{ij}) = \ln(TS_{ij}) + \epsilon_{ij} \qquad (2)$$

where Y_{ij} denotes the observed tumor size for the ith individual at time j, TS_{ij} denotes the corresponding predicted tumor size, and ε_{ij} denotes the residual random effect, which is assumed to have a normal distribution with a zero mean and variance σ.[2]

The model structure provided a good fit for all three herein investigated cases (i.e., both BA and pembrolizumab data in both BTC and NSCLC). The parameter estimates for k_s were 0.0159, 0.0368 and 0.0307 week^{-1} for BA in BTC, BA in NSCLC and pembrolizumab in NSCLC, respectively. Typical values for k_g were 0.624, 0.728, and 0.643 mm/week for BA in BTC, BA in NSCLC, and pembrolizumab in NSCLC, respectively.

Tumor dynamics and covariate assessment under BA treatment in 2L BTC

Of the 157 subjects assigned to treatment with BA, 143 (91.1%) had evaluable PK and tumor size data. In this subpopulation, the median age was 65 years (range: 39 to 83 years), median body weight 65 kg (range: 38 to 100 kg), and 85 patients (59.4%) were men.

The tumor dynamics model parameters were well-estimated (Table 1). Although some of the identified covariate parameters had relatively high standard errors (70.1%), IIV for baseline TS, k_g and k_s (50, 72.7, and 103.4% coefficient of variation [CV], respectively) was moderate to high, and estimated with acceptable accuracy (RSE < 20% CV). Shrinkage for the three tumor dynamic model parameters was low (<20%), allowing for covariates assessments (Table S2). The final model included effect of AUC on tumor shrinkage portion of the tumor size equation, with increasing rate of tumor shrinkage with increasing AUC, and the following covariates: minor allele frequency (+), time since diagnosis (−), aspartate aminotransferase (AST; +), CRP; (+) and Asian race on baseline TS (−), *IDH1* mutation on k_s (+), and minor allele frequency (+), *IDH1* mutation (+), CRP (+), and blood tumor mutational burden (+) on k_g (Table 1).

Of the investigated covariates, the univariable NLME analyses identified several covariates as significant on baseline TS (10 covariates), k_g (10 covariates), and k_s (5 covariates). Likewise, ML identified numerous covariates for baseline TS, k_s and k_g. Goodness of fit metrics for the ML models is provided in Table S3. A summary of the comparison of the covariates identified by ML and univariable NLME analyses is provided in Table S4. SHAP plots have been provided as Figure S3.

TABLE 1 Parameters of the tumor dynamics model for BTC under 2 L bintrafusp alfa treatment.

Parameter (units)	Typical value	%RSE
Baseline TS (mm)	51.3	11.4
k_s (week^{-1})	0.0159	26.7
k_g (mm/week)	0.624	20
AUC effect	0.339	15.9
Proportional residual unexplained variability (% CV)	15.6	6.2
Minor allele frequency on baseline TS	0.106	26.2
Time since diagnosis on baseline TS	−0.151	35.9
AST on baseline TS	0.189	54
CRP on baseline TS	0.0777	43.8
Asian race on baseline TS	−0.2	35.2
IDH1 mutation on k_s	2.13	55.4
Minor allele frequency on k_g	0.262	33.2
IDH1 mutation on k_g	1.17	51.5
CRP on k_g	0.152	55.1
Blood tumor mutational burden on k_g	0.629	70.1
IIV baseline TS (% CV)	50	6.5
IIV k_s (% CV)	103.4	17
IIV k_g (% CV)	72.7	12.3
Corr(baseline TS, k_s)	−0.362	NE
Corr(baseline TS, k_g)	0.465	NE
Corr(k_s, k_g)	−0.845	NE

Abbreviations: AST, aspartate aminotransferase; AUC, area under the curve; BTC, biliary tract cancer; CRP, C-reactive protein; CV, coefficient of variation; IIV, interindividual variability, k_g, tumor growth rate constant; k_s, tumor shrinkage rate constant; NE, not evaluable; RSE, relative standard error; TS, tumor size.

Tumor dynamics and covariate assessment under BA treatment in 1L NSCLC

Of the 152 subjects assigned to treatment with BA, 139 (91.4%) had evaluable PK data as well as TS data. The median age in this subpopulation was 67 years (range: 42 to 90), median body weight 66.5 kg (range: 35.0 to 104.3), and 111 patients (79.9%) were men.

The tumor dynamics model parameters were generally well-estimated (Table 2). IIV for baseline TS, k_g and k_s (55.4, 122.9, and 62.8% CV, respectively) were moderate to high, estimated with high accuracy (RSE < 30% CV). Shrinkage for the three tumor dynamic model parameters was acceptable (≤20%), allowing for covariate assessments (Table S2).

Covariates (list available in Supplementary Methods) were evaluated on all three tumor dynamic model parameters (baseline TS, k_g, and k_s). As AUC was not found to be a significant explanatory factor, herein presented are the covariate assessments performed without AUC in the model. However, as a sensitivity analysis, the covariate assessment was performed both without and with AUC in the model, with the results being virtually identical as expected with the inclusion/exclusion of a nonsignificant covariate.

With the NLME modeling approach the covariates were independently, univariably assessed for significance on baseline TS, k_g, and k_s. The univariable analyses identified the following covariates as significant: γ-glutamyl transferase and CRP on baseline TS; lymphocytes as a percentage of total leukocytes on k_g; and no covariates had significant effect on k_s. Of these, CRP on baseline TS (+) and lymphocytes as a percentage of total leukocytes on k_g (−) met the criteria for retainment in the final model after backward elimination step (Table 2).

Similarly, ML did not identify any strong covariates for k_g and k_s; the normalized coefficient and feature importance as well as permutation importance indicated similar likelihood of all covariates evaluated. For baseline TS, on the other hand, the ML algorithm identified the following covariates as being potentially important: Asian race, creatinine, γ-glutamyl transferase, CRP, monocyte, and lymphocyte levels (with goodness-of-fit metrics in Table S3). SHAP plots have been provided as Figure S4.

Tumor dynamics and covariate assessment under pembrolizumab treatment in 1L NSCLC

Of the 152 subjects assigned to treatment with pembrolizumab, 143 (94.1%) had evaluable PK and TS data. The median age of this subpopulation was 67 years (range: 36 to 88 years), median body weight 66.9 kg (range: 40.0 to 141.5 kg), and 111 patients (77.6%) were men.

The parameters of tumor dynamics model under pembrolizumab treatment were well-estimated (Table 2). IIV in baseline TS, k_g, and k_s (55.5, 92.7, and 70.1% CV, respectively) were moderate to high, estimated with high accuracy (RSE < 25% CV), and acceptable shrinkage (<20%; Table S2).

The analysis of pembrolizumab was performed in the same manner as described above for BA and the same set of covariates was evaluated. Similar to BA, as AUC was not significant, the results below stem from covariate analysis without AUC in the model. Univariable analyses identified the proportion of necrotic tissue in the whole tumor region and fibroblast density in the whole tumor region as significant covariates on baseline TS; no covariates were identified as significant with respect to k_g and k_s. None of the covariates was retained in the final model after the backward elimination step.

Similarly, and using the same criteria as for BA (i.e., similar feature and permutation importance with all evaluated covariates) no covariates for k_g or k_s were identified using ML. For baseline TS, the following covariates were identified as important by ML: lymphocytes as a percentage of total leukocytes, minor allele frequency, AST, and lymphocyte level (with goodness-of-fit metrics in Table S3). SHAP plots have been provided as Figure S5.

DISCUSSION

In this work, we report tumor dynamics' models for patients with BTC and NSCLC treated with BA or pembrolizumab which, together with ML methods, made it possible to identify factors that describe part of the variability in tumor dynamics. Although artificial intelligence (AI)/ML approaches are valuable to assist with model selection,[11] the scope of application of ML algorithms in this analysis was limited to covariate evaluations and the model building step relied on conventional pharmacometric approaches that are established for modeling tumor growth dynamics.[32]

The developed tumor growth models for patients with BTC receiving 2L BA treatment and patients with NSCLC receiving 1L BA or 1L pembrolizumab treatment all performed well (as assessed by standard diagnostics, e.g., goodness-of-fit plots and visual predictive checks). Regardless of the tumor type (BTC vs. NSCLC) or treatment (BA vs. pembrolizumab), tumor growth rates were of similar and overlapping extent (Figure 2). However, tumor shrinkage rates differed, being 59% lower in the BTC population than the NSCLC population, whereas they were comparable between BA and pembrolizumab arms of the NSCLC trial (Figure 2). The effect of BA exposure (AUC) on tumor shrinkage was significant in the

TABLE 2 Parameters of the tumor dynamics model for NSCLC under 1 L bintrafusp alfa and pembrolizumab treatment.

Parameter (units)	Bintrafusp alfa		Pembrolizumab	
	Typical value	%RSE	Typical value	%RSE
Baseline TS (mm)	81.1	5.75	80.4	4.91
k_s (week^{-1})	0.0368	12.8	0.0307	9.06
k_g (mm/week)	0.728	17.9	0.643	12.9
Proportional residual unexplained variability (% CV)	12.8	2.29	11.3	2.42
Effect of CRP on baseline TS	0.125	23.8	–	–
Effect of lymphocyte as percentage of total leukocytes on k_g	−0.623	43.7	–	–
IIV baseline TS (% CV)	55.4	14.0	55.5	13.02
IIV k_s (% CV)	62.8	26.3	70.1	24.03
IIV k_g (% CV)	122.9	22.1	92.7	20.7
Corr(baseline TS, k_s)	−0.0411	NE	−0.0228	NE
Corr(baseline TS, k_g)	0.423	NE	0.288	NE
Corr(k_s, k_g)	−0.0639	NE	−0.0834	NE

Abbreviations: CRP, C-reactive protein; CV, coefficient of variation; IIV, interindividual variability; k_g, tumor growth rate constant; k_s, tumor shrinkage rate constant; NE, not evaluable; NSCLC, non-small cell lung cancer; RSE, relative standard error; TS, tumor size.

BTC model, whereas neither BA nor pembrolizumab exposure was significant in the corresponding NSCLC models. The final models also included significant covariates on the tumor dynamics parameters (baseline TS, k_g, and/or k_s). Under the BTC setting (Table 1), several covariates were found significant, including patient-related (race), disease-related (AST and CRP), and genomic factors (minor allele frequency, *IDH1* mutation, and tumor mutational burden). Although the covariates did explain some of the variability in tumor growth and shrinkage parameters, the net impact on tumor size over time (illustrated in Figure 3a–h) was of very limited extent for all covariates, indicating them clinically unimportant as contributors to overall variability in tumor dynamics. The only covariate identified to be related to tumor shrinkage was *IDH1* mutation, in accordance with the previous reports of better prognosis in patients with *IDH1* mutation,[33] however, IDH1 was at the same time directly related to tumor growth, resulting in a more complex relationship with TS change over time (Figure 3f). In contrast to BTC, the NSCLC analysis identified only two covariates (Figure 3i,j) as significant in the BA arm, and none in the pembrolizumab arm. In the BA arm, baseline TS was found to be related to CRP concentration, which was expected as CRP is related to the disease severity, and lymphocyte percentage was inversely related to tumor growth rate, which is in accordance with the previous reports of low lymphocyte percentage as an indicator of poor prognosis in lung cancer.[34] Interestingly, none of the covariates was identified as related to the tumor shrinkage rate.

A second part of this analysis utilized ML and univariable NLME approaches to identify trends in relationships between covariates and tumor dynamics, at the same time enabling comparison of these approaches. In BTC analysis, there was an overall agreement among covariates identified as significant by univariable NLME analyses and top features identified by ML (Table S4). Overall, the identified covariates fit into expected categories, including genomic (minor allele frequency, tumor mutational burden, and *IDH1* mutation), disease-related (CRP, lymphocytes, and alkaline phosphatase), and patient-related factors (Asian race). Of note, the race was found only to be related to baseline TS and the overall magnitude of the effect was low (Figure 3e). In NSCLC analyses, there was a limited agreement between the two approaches, which might be partly attributable to the overall lack of significance of covariates in these analyses. Furthermore, whereas in the NLME setting there are clear statistical criteria that define significance, with ML assignment of significance relies on human intervention and decision in setting a predefined parameter (threshold or a predefined number of the top or most important covariates), thus influencing the comparisons of the two methods. In the BA arm analysis, covariates identified by NLME analysis were also captured by ML, but ML identified further covariates not captured by NLME. No covariates were associated with k_s with either of the methods; and the one covariate on k_g identified by NLME was not captured by ML. In the pembrolizumab arm analysis, neither of the methods linked any covariates to k_g nor k_s, whereas the
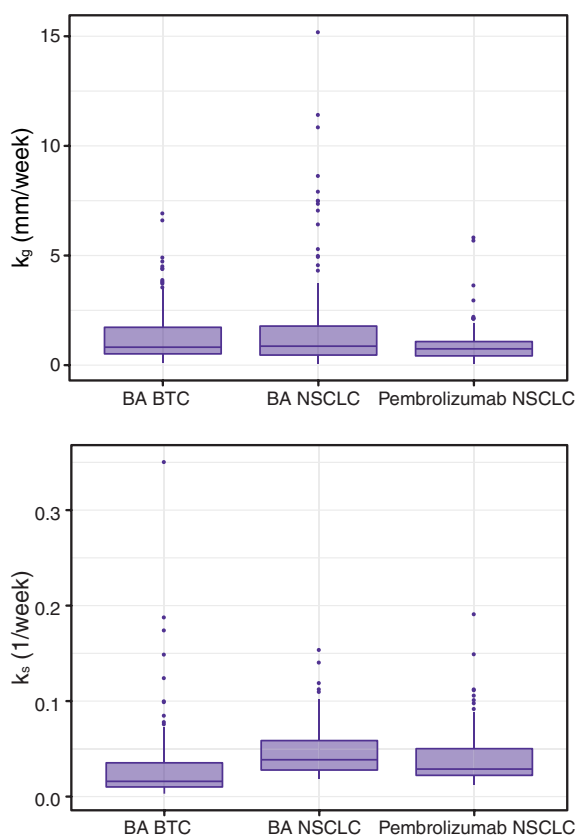
FIGURE 2 Distribution of tumor growth (upper panel) and tumor shrinkage (lower panel) rate constants among the developed tumor dynamics models. AST, aminotransferase; AUC, area under the curve; MAF, minor allele frequency; TS, tumor size.

covariates identified by the two methods for baseline TS differed without overlap. In summary, there was lower agreement between the two methods for NSCLC than for the BTC trial, and for BA than for pembrolizumab. As a phase I analysis in esophageal carcinoma[35] suggested that immune-excluded phenotype might be related to response to BA, our covariate analyses included the immune phenotype, these were, however, not found to be significant in any of our analyses. Similarly, the proportion of necrotic tissue in the tumor area[36] and inflammatory microenvironment (macrophage, lymphocyte, and fibroblast density)[37] were considered of interest, but none of them were found to be significant in our analyses.

As a primary purpose of the present research was to identify sources of variability (i.e., covariates) that may explain the observed heterogeneity in tumor growth dynamics in the evaluated datasets, it was important to qualify the developed population tumor dynamics models as having adequately low shrinkage (<20%) in the parameters where IIV could be described. In ML-enabled covariate analyses, making the evaluation of shrinkage in model parameters is important. When shrinkage is high, the individual parameter estimates are not reflective of the actual individual patient making covariate identification difficult. Of note, the parameters for all developed models had little shrinkage in the present evaluation, supporting the results of the covariate analyses. It is important to acknowledge that the deployed ML methods in the analyses also do not take into account correlation between model parameters. However, covariate evaluations using ML run more quickly than with traditional modeling, and there is no need to create multiple model files to test each, making ML useful, particularly when model run times are long and many covariates need to be evaluated. Although AI/ML has shown great potential in assisting model selection,[11] this model building step relied on conventional approaches.

Finally, because tumor hyperprogression was observed in certain patients, a subanalysis in the BA treatment groups compared patients who had experienced hyperprogression to patients who had not. With respect to tumor dynamics, the patients who experienced hyperprogression could be identified by higher tumor growth rate and lower tumor kill rate in both BTC and NSCLC (Figure S2). Interestingly, no meaningful difference in the BA exposure (first cycle AUC) was observed between these two patient subgroups (Figure S2d,h), suggesting that BA exposure was not a driver for hyperprogression, that is, that factors beyond BA treatment warrant further investigation with respect to hyperprogression. Due to the low number of hyperprogressing patients in our data, these findings warrant further research.

In conclusion, the herein reported analyses quantitatively characterized TS change over time in two different tumor types and under different two treatments. These studies did not meet the primary end point, thus, the main focus of this work was to assess the potential existence of subpopulations who might be sensitive to the effects of BA. To this purpose, two approaches, NLME and ML, were used to identify factors potentially related to the tumor dynamics. Given their demonstrated potential to mine and assess covariates in the high-dimensional space,[12] ML approaches were run to complement conventional methods in a comprehensive investigation. Results were generally consistent across methods, and several covariates were suggested to have an effect on population variability in model parameters characterizing tumor dynamics. However, their net impact on TS was not considered clinically relevant. Viewed from a broader perspective, these analyses illustrate a two-pronged approach incorporating both classical pharmacometrics and ML to covariate analyses, thereby contributing to the emerging area of pharmacometrics enhanced by advanced analytics, with opportunities to enable model-informed precision medicine.[10,38,39]
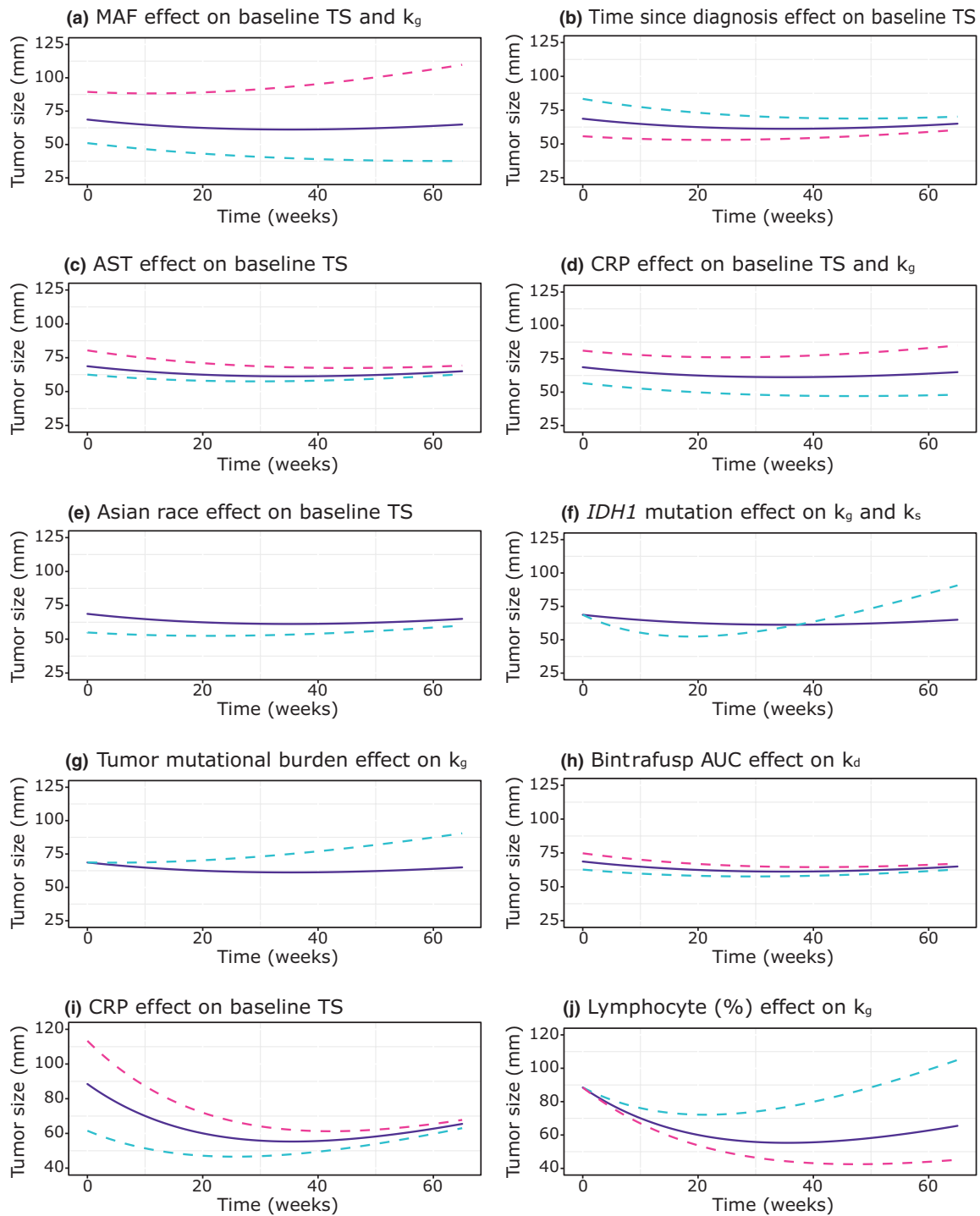
FIGURE 3 Effects of covariates identified as significant in the analysis of BTC trial (a–h) and NSCLC trial (i, j) data on tumor size dynamics. For continuous covariates purple full line denotes median, blue dashed line 5th percentile, and red dashed line 95th percentile; for categorical covariates purple full line denotes the reference category (non-Asian race, no *IDH1* mutation, and blood tumor mutational burden [TMB] <50th percentile) and blue dashed line the significant category (Asian race, *IDH1* mutation, and blood TMB ≥50th percentile). BTC, biliary tract cancer; NSCLC, non-small cell lung cancer.

AUTHOR CONTRIBUTIONS
A.-M.M.-G., N.T., D.R.M., Y.V., P.G., K.V., and A.K. wrote the manuscript. A.-M.M.-G., N.T., D.R.M., Y.V., P.G., K.V., and A.K. designed the research. All authors performed the research and analyzed the data. All authors read and approved the final manuscript.

CONFLICT OF INTEREST STATEMENT

A.-M.M.-G., T.M., and A.K. are employees of the healthcare business of Merck KGaA, Darmstadt, Germany. D.M. reports personal fees from the healthcare business of Merck KGaA, Darmstadt, Germany during the conduct of the study. Y.V., A.M., and K.V. are employees of EMD Serono. N.T. and P.G. are employees of Quantitative Pharmacology, Ares Trading S.A., Lausanne, Switzerland, an affiliate of Merck KGaA, Darmstadt, Germany.

ORCID

Ana-Marija Milenković-Grišić https://orcid.org/0000-0002-1093-6879
Nadia Terranova https://orcid.org/0000-0002-0033-3695
Akash Khandelwal https://orcid.org/0000-0002-0472-7173

REFERENCES

1. Lind H, Gameiro SR, Jochems C, et al. Dual targeting of TGF-beta and PD-L1 via a bifunctional anti-PD-L1/TGF-betaRII agent: status of preclinical and clinical advances. *J Immunother Cancer*. 2020;8:e000433.
2. Wrzesinski SH, Wan YY, Flavell RA. Transforming growth factor-beta and the immune response: implications for anticancer therapy. *Clin Cancer Res*. 2007;13:5262-5270.
3. Cho BC, Lee JS, Wu Y-L, et al. Bintrafusp alfa versus pembrolizumab in patients with treatment-naive, PD-L1–high advanced non-small cell lung cancer: a randomized, open-label, phase 3 trial. *J Thorac Oncol*. Forthcoming 2023.
4. Yoo C, Oh D-Y, Choi HJ, et al. Phase I study of bintrafusp alfa, a bifunctional fusion protein targeting TGF-β and PD-L1, in patients with pretreated biliary tract cancer. *J Immunother Cancer*. 2020;8:e000564.
5. Matos I, Martin-Liberal J, García-Ruiz A, et al. Capturing hyperprogressive disease with immune-checkpoint inhibitors using RECIST 1.1 criteria. *Clin Cancer Res*. 2020;26:1846-1855.
6. Wilkins JJ, Vugmeyster Y, Dussault I, Girard P, Khandelwal A. Population pharmacokinetic analysis of bintrafusp alfa in different cancer types. *Adv Ther*. 2019;36(9):2414-2433.
7. Vugmeyster Y, Wilkins J, Koenig A, et al. Selection of the recommended phase 2 dose for bintrafusp alfa, a bifunctional fusion protein targeting TGF-β and PD-L1. *Clin Pharmacol Ther*. 2020;108:566-574.
8. Vugmeyster Y, Grisic A-M, Wilkins JJ, et al. Model-informed approach for risk management of bleeding toxicities for bintrafusp alfa, a bifunctional fusion protein targeting TGF-β and PD-L1. *Cancer Chemother Pharmacol*. 2022;90:369-379.
9. Terranova N, French J, Dai H, et al. Pharmacometric modeling and machine learning analyses of prognostic and predictive factors in the JAVELIN gastric 100 phase III trial of avelumab. *CPT Pharmacometrics Syst Pharmacol*. 2022;11:333-347.
10. Terranova N, Venkatakrishnan K, Benincosa LJ. Application of machine learning in translational medicine: current status and future opportunities. *AAPS J*. 2021;23:74.
11. Sibieude E, Khandelwal A, Girard P, Hesthaven JS, Terranova N. Population pharmacokinetic model selection assisted by machine learning. *J Pharmacokinet Pharmacodyn*. 2022;49:257-270.
12. Sibieude E, Khandelwal A, Hesthaven JS, Girard P, Terranova N. Fast screening of covariates in population models empowered by machine learning. *J Pharmacokinet Pharmacodyn*. 2021;48:597-609.
13. Yoo C, Javle MM, Verdaguer Mata H, et al. Phase 2 trial of bintrafusp alfa as second-line therapy for patients with locally advanced/metastatic biliary tract cancers. *Hepatology*. 2023;78:758-770.
14. Ahn M-J, Barlesi F, Garon E, et al. Randomized, open-label study of bintrafusp alfa vs. pembrolizumab as first-line (1L) treatment in patients with PD-L1-expressing advanced non-small cell lung cancer (NSCLC). *J Thorac Oncol*. 2021;16:S27-S28.
15. Ahamadi M, Freshwater T, Prohn M, et al. Model-based characterization of the pharmacokinetics of pembrolizumab: a humanized anti-PD-1 monoclonal antibody in advanced solid tumors. *CPT Pharmacometrics Syst Pharmacol*. 2017;6:49-57.
16. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst*. 2000;92:205-216.
17. Claret L, Girard P, Hoff PM, et al. Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. *J Clin Oncol*. 2009;27:4103-4108.
18. Tham L-S, Wang L, Soo RA, et al. A pharmacodynamic model for the time course of tumor shrinkage by gemcitabine + carboplatin in non-small cell lung cancer patients. *Clin Cancer Res*. 2008;14:4213-4218.
19. Jusko WJ. Pharmacodynamics of chemotherapeutic effects: dose-time-response relationships for phase-nonspecific agents. *J Pharm Sci*. 1971;60:892-895.
20. Zhi J, Nightingale C, Quintiliani R. A pharmacodynamic model for the activity of antibiotics against microorganisms under non saturable conditions. *J Pharm Sci*. 1986;25:1063-1067.

21. Yano Y, Oguma T, Nagata H, Sasaki S. Application of a logistic growth model to pharmacodynamic analysis of in vitro bacteriocidal kinetics. *J Pharm Sci*. 1998;87:1177-1183.

22. Chatterjee MS, Elassaiss-Schaap J, Lindauer A, et al. Population pharmacokinetic/pharmacodynamic modeling of tumor size dynamics in pembrolizumab-treated advanced melanoma. *CPT Pharmacometrics Syst Pharmacol*. 2017;6:29-39.

23. Jonsson EN, Karlsson MO. Automated covariate model building within NONMEM. *Pharm Res*. 1998;15:1463-1468.

24. Beal SL, Sheiner LB. *NONMEM Users Guide: Part V*. University of California – San Francisco; 1992.

25. Jonsson EN, Karlsson MO. Xpose—an S-PLUS based population pharmacokinetic/pharmacodynamic model building aid for NONMEM. *Comput Methods Programs Biomed*. 1999;58:51-64.

26. Lindbom L, Pihlgren P, Jonsson EN. PsN-toolkit—a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput Methods Programs Biomed*. 2005;79:241-257.

27. Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN)—a Perl module for NONMEM related programming. *Comput Methods Programs Biomed*. 2004;75:85-94.

28. European Environment Agency. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2020.

29. McKinney W. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. 2nd ed. O'Reilly Media; 2017.

30. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. *Nature*. 2020;585:357-362.

31. Wang Y, Sung C, Dartois C, et al. Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development. *Clin Pharmacol Ther*. 2009;86:167-174.

32. Bruno R, Bottino D, de Alwis DP, et al. Progress and opportunities to advance clinical cancer therapeutics using tumor dynamic models. *Clin Cancer Res*. 2020;26:1787-1795.

33. Pirozzi CJ, Yan H. The implications of IDH mutations for cancer development and therapy. *Nat Rev Clin Oncol*. 2021;18:645-661.

34. Huang H, Li L, Luo W, et al. Lymphocyte percentage as a valuable predictor of prognosis in lung cancer. *J Cell Mol Med*. 2022;26:1918-1931.

35. Tan B, Khattak A, Felip E, et al. Bintrafusp alfa, a bifunctional fusion protein targeting TGF-β and PD-L1, in patients with esophageal adenocarcinoma: results from a phase 1 cohort. *Target Oncol*. 2021;16:435-446.

36. Richards CH, Mohammed Z, Qayyum T, Horgan PG, McMillan DC. The prognostic value of histological tumor necrosis in solid organ malignant disease: a systematic review. *Future Oncol*. 2011;7:1223-1235.

37. De Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system during cancer development. *Nat Rev Cancer*. 2006;6:24-37.

38. Venkatakrishnan K, Benincosa LJ. Diversity and inclusion in drug development: rethinking intrinsic and extrinsic factors with patient centricity. *Clin Pharmacol Ther*. 2022;112:204-207.

39. Hutchinson L, Steiert B, Soubret A, et al. Models and machines: how deep learning will take clinical pharmacology to the next level. *CPT Pharmacometrics Syst Pharmacol*. 2019;8:131-134.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.