

Parameter Estimation and Model Selection for the Quantitative Analysis of Oncolytic Virus Therapy in Zebrafish^{*}

Yuhong Liu^{*} Dilan Pathirana^{*} Jan Hasenauer^{*}

^{} Life and Medical Sciences (LIMES) Institute & Bonn Center for
Mathematical Life Sciences, University of Bonn, 53115 Bonn,
Germany (e-mail: yuhong.liu@uni-bonn.de,
dilan.pathirana@uni-bonn.de, jan.hasenauer@uni-bonn.de).*

Abstract: Oncolytic virus therapy (OVT) is emerging as a potent alternative to conventional cancer treatments by employing engineered viruses that selectively infect and lyse tumor cells while sparing normal tissues. Although mathematical models have been developed to elucidate the dynamics of OVT and inform personalized therapies, they are often specific to certain organisms. Mathematical models tailored to more recently developed animal models of OVT, such as zebrafish, are not yet available. Here, we introduce the first mathematical model of OVT trained on zebrafish data from published studies to bridge the gap. We explore a variety of mathematical model structures and perform parameter estimation and model selection. The selected model effectively captures the observed tumor dynamics, i.e. delayed tumor shrinkage, and provides valuable insights into the underlying mechanisms of OVT in zebrafish. Our work establishes the groundwork for advancing experimental studies in zebrafish, contributing to the design of more effective cancer treatment strategies in the future.

Keywords: Model predictive control, Systems biology, Synthetic biology, Metabolic flux modeling, Dynamic modeling and simulation for control and operation

1. INTRODUCTION

Oncolytic virus therapy (OVT) is a promising alternative to conventional cancer treatments and has demonstrated success in recent human clinical trials (Ling et al., 2023). OVT employs oncolytic viruses (OVs) to eradicate malignant cancer cells through two main mechanisms. Firstly, these viruses are typically engineered strains that selectively infect, replicate within and promote lysis in tumor cells while remaining safe for normal cells (Fig. 1a). Secondly, the lysis of tumor cells by OVs exposes tumor antigens, thereby stimulating the immune system to recognize and attack cancer cells. The latter process helps overcome immunosuppression, a significant challenge since tumor cells often appear non-threatening to the immune system. Antonio Chiocca (2002) and Tian et al. (2022) have comprehensively summarized different types of OVs, how their selectivity is engineered, and how they enhance the effects of antitumor immunity.

The experimental assessment of (i) the interaction of OVs and tumor cells, and (ii) the contribution of the two mechanisms described above is non-trivial. Accordingly, math-

ematical models which capture the underlying processes have been developed. These models provide insights into emergent dynamics, guide future experiments, and build the basis for personalized OVT. Various mathematical models have been used to study various process characteristics: Mahasa et al. (2017) studied how infection of normal cells with OVs in the vicinity of tumor cells enhances OVT; Eftimie et al. (2011) explored biological conditions that could lead to the permanent elimination of tumor cells and examined the multistability and instability of the nonlinear dynamics; and Kumar et al. (2021) approached the design of more effective OVs using a fractional-order delay differential equation model that incorporates the viral lytic cycle. Mathematical models, and especially their parameters, are typically tailored to specific animal models. However, there are animal models that are important for cancer research, such as zebrafish that enable rapid screening of OVT candidates (Fazio et al., 2020) and are used for OVT studies (Mealiea et al., 2021), for which no mathematical models are available.

In this manuscript, we propose the first mathematical model for OVT in zebrafish, to our knowledge. To this end, we consider a range of model topologies and perform parameter estimation and model selection using published experimental data from Mealiea et al. (2021). Our selected model can quantitatively capture the measured dynamics and provide insights into underlying mechanisms contributing to, for example, delayed tumor shrinkage. Subsequently, we refine the model to account for the heterogeneity between individual zebrafish through additional pa-

^{*} The authors acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2047—390685813 (Hausdorff Center for Mathematics), EXC 2151—390873048 (ImmunoSensation2)), and under project ID 432325352 – SFB 1454 (Metaflammation), by the European Union via ERC grant INTEGRATE, grant agreement number 101126146, and by the University of Bonn (via the Schlegel Professorship of Prof. Dr. Jan Hasenauer)

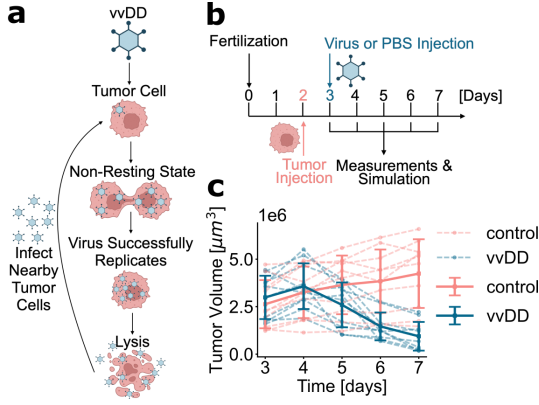


Fig. 1. **OVT mechanism and zebrafish experiment.** (a) vvDD replication requires cell division. (b) Experimental setup for zebrafish study. Fertilization occurs on day zero. (c) Measured tumor volumes for individual zebrafish (dashed) and population mean (solid, error bars are ± 1 std. dev.). *Sub-figures (a) and (b) were created using Biorender.com.*

parameters for each individual. We show that the individual numbers of initially homed tumor cells can substantially explain the population variability in tumor response. Finally, we discuss important extensions of our models in the direction of mixed-effects and stochastic models for extinction analysis. Our work represents a foundational first step in the mathematical modeling of OVT on zebrafish data. This enables a platform for future research to advance experimental studies in this animal model that can inform the design of effective cancer treatment strategies.

2. DATA AND METHODS

2.1 Experimental Setup and Measurement

In this manuscript, we analyze and model the OVT data collected by Mealiea et al. (2021) for zebrafish. In their experimental study (Fig. 1b), 20 zebrafish were each injected with 400 fluorescently-labeled tumor cells on day 2 (after fertilization on day 0). The zebrafish were then split into two groups. The treatment group was injected with 10^9 plaque-forming units of double-deleted vaccinia virus (vvDD) on day 3, while the control group remained untreated and thus injected with only phosphate-buffered saline (PBS). Both groups were observed until day 7.

vvDD is a double-stranded DNA virus with deletions that increase tumor specificity and safety for other cells (McCart et al., 2001), and was reported to only replicate efficiently in tumor cells. In the zebrafish study (Mealiea et al., 2021), the virus is also fluorescently-labeled.

The tumor volume was recorded daily on days 3 to 7 (Fig. 1c). In the vvDD-treated group, the tumor volume declined from day 4 onwards, while it continued to grow in the control group. Overall, both groups show substantial inter-individual heterogeneity in tumor volume.

2.2 Mathematical Modeling

We model the interaction between OVVs and tumor cells using ordinary differential equation (ODE) models. These ODE systems are of the form:

$$\frac{dx}{dt} = f(x, \theta, u), \quad x(0) = x_0(\theta, u) \quad (1)$$

with state vector $x(t) \in \mathbb{R}^{n_x}$ at time t , parameter vector $\theta \in \mathbb{R}^{n_\theta}$, and input vector $u \in \mathbb{R}^{n_u}$ that encodes the experimental conditions. The vector field $f: \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_u} \mapsto \mathbb{R}^{n_x}$ encodes the dynamics of the process and is constructed to be Lipschitz continuous, ensuring that the ODE solution exists and is unique. The function $x_0: \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_u} \mapsto \mathbb{R}^{n_x}$ defines the initial condition. The inputs $u = (u_1, u_2)^T$ are used to encode the experimental condition, namely the injected amount of tumor cells (u_1) and virus particles (u_2). The inputs in the vvDD-treated and control conditions are $u^{(v)} = (400, 10^9)^T$ and $u^{(c)} = (400, 0)^T$, respectively.

The parameters θ , e.g., the growth rate of tumor cells, are inferred from the measured data using maximum likelihood estimation. The observable is the tumor volume,

$$y = h(x, \theta),$$

with $h: \mathbb{R}^{n_x} \times \mathbb{R}^{n_\theta} \mapsto \mathbb{R}$ denoting the observation mapping. The zebrafish measurements are denoted by

$$\mathcal{D}_i^{(k)} = \{(t_j = j, y_{i,j,k}^m)\}_{j=3}^7,$$

with condition $i \in \{v, c\}$, time point index $j \in \{3 \dots 7\}$ in days, and zebrafish individual index $k \in \{1 \dots 10\}$. The measurement noise is assumed to be additive and normally distributed, $y^m = y + \epsilon$, with a noise variance depending on the tumor volume, $\epsilon \sim \mathcal{N}(0, \sigma_a^2 + \sigma_b^2 y^2)$, with $\sigma_a > 0$ and $\sigma_b > 0$. Accordingly, the conditional probability of observing y^m given y , σ_a and σ_b is given by

$$p(y^m | y, \sigma_a, \sigma_b) = \frac{1}{\sqrt{2\pi(\sigma_a^2 + \sigma_b^2 y^2)}} \exp\left(-\frac{1}{2} \frac{(y^m - y)^2}{\sigma_a^2 + \sigma_b^2 y^2}\right).$$

Assuming that the zebrafish are identical, the objective function with negative log-likelihood is

$$\begin{aligned} J(\theta, \sigma_a, \sigma_b) &= - \sum_{i \in \{v, c\}} \sum_{j=3}^7 \sum_{k=1}^{10} \log p(y_{i,j,k}^m | y(t_j, \theta, u^{(i)}), \sigma_a, \sigma_b) \end{aligned} \quad (2)$$

with $y(t, \theta, u) = h(x(t, \theta, u), \theta)$ and $x(t, \theta, u)$ denoting the solution of the ODE model (1) for parameters θ and input u evaluated at time t . To account for differences between zebrafish, we introduce parameters $\xi_{i,k}$ which are specific to the individual zebrafish, yielding the objective function

$$\begin{aligned} J(\theta, \sigma_a, \sigma_b, \xi) &= - \sum_{i \in \{v, c\}} \sum_{j=3}^7 \sum_{k=1}^{10} \log p(y_{i,j,k}^m | y(t_j, \theta, u^{(i)}, \xi_{i,k}), \sigma_a, \sigma_b). \end{aligned} \quad (3)$$

Details on how ξ influences the solution are in Section 3.4.

The maximum likelihood estimates are obtained by minimizing the objective functions $J(\theta, \sigma_a, \sigma_b)$ or $J(\theta, \sigma_a, \sigma_b, \xi)$, i.e., $\{\hat{\theta}, \hat{\sigma}_a, \hat{\sigma}_b, \hat{\xi}\} = \arg \min_{\theta, \sigma_a, \sigma_b, \xi} J(\theta, \sigma_a, \sigma_b, \xi)$ for the latter, subject to parameter bounds, with each parameter bound in \mathbb{R}_+ .

2.3 Software

We use PETA (Schmiester et al., 2021) to define the parameter estimation problems and SBML (Hucka et al.,

2003) for standardized representations of the mathematical models. Parameter optimization and uncertainty analysis are implemented using pyPESTO (Schälte et al., 2023), wherein we used AMICI (Fröhlich et al., 2021) for simulation and Fides (Fröhlich et al., 2022) for optimization. We used STRIKE-GOLDD to perform structural identifiability analysis (Díaz-Seoane et al., 2023).

3. RESULTS

3.1 Model formulation and analysis

In the **baseline model**, we distinguish uninfected tumor cells (U), infected tumor cells (I), and extracellular virus (V). The cells are denoted by $[U]$, $[I]$ and $[V]$, yielding the state vector $x = ([U], [I], [V])^T$. This state vector changes over time due to six processes (Fig. 2a):

(P1) The quantity of uninfected and infected tumor cells increases due to tumor cell division,

$$\begin{aligned} U &\rightarrow 2 \times U, & \text{rate} &= \rho[U] \left(1 - \frac{[U] + [I]}{\kappa}\right), \\ I &\rightarrow 2 \times I, & \text{rate} &= \rho[I] \left(1 - \frac{[U] + [I]}{\kappa}\right). \end{aligned}$$

We assume a logistic growth model with maximal division rate constant ρ and carrying capacity κ .

(P2) The quantity of uninfected tumor cells decreases and the quantity of infected tumor cells increases due to infection with the virus,

$$U + V \rightarrow I, \quad \text{rate} = \psi[V][U],$$

with infection rate constant ψ .

(P3) The quantity of infected tumor cells decreases due to cell death, resulting in the release of viruses,

$$I \rightarrow \beta \times V, \quad \text{rate} = \alpha[I],$$

with age-of-infection-independent death rate constant α , and the quantity of released virus particles β .

(P4) The quantity of extracellular viruses decreases due to virus removal processes,

$$V \rightarrow \emptyset, \quad \text{rate} = \delta[V],$$

with removal rate constant δ .

(P5, P6) Tumor cell injection on day 2 and virus injection on day 3 are processes happening at discrete time points and marking singular events. At these injection points, the respective abundances become nonzero:

$$\begin{aligned} [U](t_{i,U}) &= u_1, \\ [V](t_{i,V}) &= u_2. \end{aligned}$$

The injection time points are denoted by $t_{i,U} = 2$ and $t_{i,V} = 3$ (Fig. 1b), while the quantity of injected uninfected tumor cells and viruses are denoted by u_1 and u_2 .

Given the processes P1 to P6, the dynamics can be separated into three intervals:

Interval 1 is the time interval before the injection of tumor cells. In this interval, all state variables are zero:

$$\forall t \in [0, t_{i,U}) : [U](t) = [I](t) = [V](t) = 0.$$

Interval 2 is the time interval between the injections of tumor cells and viruses. In this interval, the quantity of

infected tumor cells and viruses is zero, and the quantity of uninfected tumor cells grows (P1):

$$\begin{aligned} \forall t \in [t_{i,U}, t_{i,V}) : [I](t) &= [V](t) = 0, \text{ and} \\ \frac{d[U]}{dt} &= \rho[U] \left(1 - \frac{[U]}{\kappa}\right), \quad [U](t_{i,U}) = u_1. \end{aligned}$$

The time-dependent state $[U](t)$ can be computed analytically using the separation of variables, yielding

$$\forall t \in [t_{i,U}, t_{i,V}) : [U](t) = \frac{\kappa}{1 + \frac{\kappa - u_1}{u_1} e^{\rho(t - t_{i,U})}}.$$

Interval 3 is the time interval after the injection of the virus. In this interval, all state variables can change dynamically:

$\forall t \in [t_{i,V}, \infty) :$

$$\begin{aligned} \frac{d[U]}{dt} &= \rho[U] \left(1 - \frac{[U] + [I]}{\kappa}\right) - \psi[V][U], \\ \frac{d[I]}{dt} &= \rho[I] \left(1 - \frac{[U] + [I]}{\kappa}\right) + \psi[V][U] - \alpha[I], \\ \frac{d[V]}{dt} &= \alpha\beta[I] - \psi[V][U] - \delta[V], \end{aligned}$$

with initial conditions $[U](t_{i,V}) = \kappa / (1 + \frac{\kappa - u_1}{u_1} e^{\rho(t - t_{i,U})})$ (the final value of the previous interval), $[I](t_{i,V}) = 0$, and $[V](t_{i,V}) = u_2$. For this interval, no analytical solution is known, and the ODE solution is approximated with AMICI.

The theoretical analysis of the model reveals that it is Lipschitz continuous (proof provided in the Supplementary Material). Furthermore, for non-negative values of the parameters and inputs, the non-negative quadrant, i.e. $0 \leq [U]$, $0 \leq [I]$, $0 \leq [V]$, is invariant. This means that, with an initial condition in this quadrant, the states remain non-negative, which is biologically reasonable.

The data provide information about the volume of the tumor, which is assumed to be proportional to the quantity of tumor cells. Thus, the observable mapping is

$$y = s([U] + [I]),$$

where the scaling factor s is the mean volume of a tumor cell and y is the tumor volume. Structural identifiability analysis with STRIKE-GOLDD reveals that this parameter is non-identifiable (see Supplementary Material). To resolve this, we searched the literature to approximate s . The tumor cells used in the experiment are from the MC38 murine colorectal adenocarcinoma cell line, which has an average doubling time of about 24 hours (Morimoto-Tomita et al., 2005). Accordingly, one would expect 800 tumor cells at day 3. As the mean tumor volume at day 3 is $2808614 \mu\text{m}^3$, we estimate $s \approx \frac{2808614}{800} \mu\text{m}^3/\text{cell} \approx 3511 \mu\text{m}^3/\text{cell}$. With this estimate of s , the baseline model achieved ‘‘Full Input-State-Parameter Observability’’ (FISPO) (Díaz-Seoane et al., 2023), suggesting that with sufficient measurement data of y , all parameters are identifiable. A limitation of this approximation is that the tumor doubling time during the first day is assumed to be comparable between zebrafish and mice.

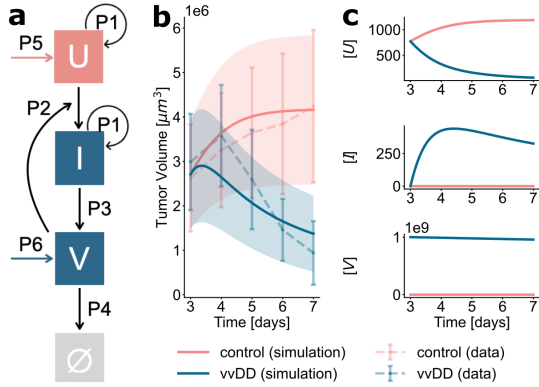


Fig. 2. **The baseline model.** (a) Process diagram. (b) Optimized model simulation. Data are means with ± 1 empirical std. dev. error bars. Simulations have ± 1 estimated std. dev. error bands. (c) Optimized state variable trajectories.

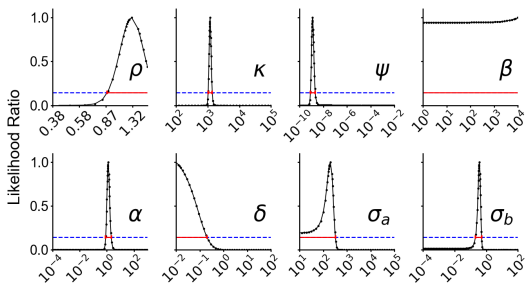


Fig. 3. **Baseline model profile likelihoods.** Domains are the parameter bounds (Table 1). Blue lines are the likelihood ratio threshold that was used to compute the 95% confidence intervals (red lines).

3.2 Parametrization of Baseline Model

To assess whether the baseline model can describe the data, we obtain the maximum likelihood estimate by minimizing (2). Individual measurements are used for the optimization but the heterogeneity between individual zebrafish is not accounted for in the baseline model. Parameter bounds are provided in Table 1.

The analysis of the fitting results reveals that the optimization converges robustly, suggesting that globally optimal parameters are found (see Supplementary Material). The model provides a good description of the measurement data (Fig. 2b). However, it fails to capture the quantitative dynamics of the tumor volume after virus injection on day 3. Although the measurements show that mean tumor volume increases until day 4, the best model fit exhibits a decrease. Hence, the baseline model cannot capture the observed delay in tumor shrinkage.

To assess parameter uncertainties, we compute 95% confidence intervals with the profile likelihood method (Raue et al., 2009). This reveals that the parameters α , κ , ψ , and σ_b are practically identifiable, as their confidence intervals are finite (Fig. 3). The parameters δ , ρ , σ_a , and β cannot be bounded from both above and below, and hence are practically non-identifiable.

3.3 Modeling delayed tumor shrinkage

As the baseline model fails to describe key properties of the measurement data, we consider an **age-of-infection model**, which accounts for the experimental observation that cell lysis does not occur directly after infection but instead after a large number of virus particles are produced (Dominguez et al., 2015). This could be captured using a discrete delay (and a delay differential equation), yet, here we use the linear chain trick (Hurtado et al., 2020). Therefore, we split the state variable for infected cells I into L state variables, I_l , with $l \in \{1 \dots L\}$. Tumor cells in state I_1 are infected but carry a low viral load, while tumor cells in state I_L carry a high viral load. Cell division occurs for all tumor cells at the same rate (P1), but the infection of uninfected tumor cells only yields I_1 (P2). Infected cells in state I_l (except $l = L$) can transition into state I_{l+1} according to process P2':

$$I_l \rightarrow I_{l+1}, \quad \text{rate} = \phi[I_l],$$

which captures different stages of the infected cell due to an increase in the number of virus particles, with the transition rate constant ϕ . Here we assume one constant transition parameter for simplicity. Cells that reach the state I_L can undergo cell lysis, which results in virus release (P3). This yields the governing ODE system for interval 3, $\forall t \in [t_{i,V}, \infty)$:

$$\frac{d[U]}{dt} = \rho[U] \left(1 - \frac{[C]}{\kappa}\right) - \psi[V][U],$$

$$\frac{d[I_1]}{dt} = \rho[I_1] \left(1 - \frac{[C]}{\kappa}\right) + \psi[V][U] - \phi[I_1],$$

for $l \in \{2, \dots, L-1\}$:

$$\frac{d[I_l]}{dt} = \rho[I_l] \left(1 - \frac{[C]}{\kappa}\right) + \phi([I_{l-1}] - [I_l]),$$

$$\frac{d[I_L]}{dt} = \rho[I_L] \left(1 - \frac{[C]}{\kappa}\right) + \phi[I_{L-1}] - \alpha[I_L],$$

$$\frac{d[V]}{dt} = \beta\alpha[I_L] - \psi[V][U] - \delta[V],$$

with $[I_l](0) = [I_l](t_{i,U}) = [I_l](t_{i,V}) = 0, \forall l \in \{1 \dots L\}$, the same initial conditions as in the baseline model for $[U]$ and $[V]$, and total tumor cell number $[C] = [U] + \sum_{l=1}^L [I_l]$. In this study, we use $L = 5$.

We estimate the parameters of the age-of-infection model using the same approach as for the baseline model, with the corresponding observable mapping $y = s[C]$. The optimization also achieved good convergence, suggesting that the global optimum is found (see Supplementary Material). The assessment of the fit reveals a good agreement with the measurement data (Fig. 4b). Indeed, the model simulation exhibits a delay in tumor shrinkage. The obtained state variables' trajectories show that virus accumulation (Fig. 4c) causes a delay.

We used the same literature-based estimate of s as in the baseline model. The structural identifiability analysis with STRIKE-GOLDD found that the model is FISPO. Assessment of the parameter uncertainties revealed that ρ , ψ , β , and δ are practically non-identifiable (Fig. 5).

Table 1. **Informed parameter bounds of all models.** The lower and upper bounds for the parameters and the reasoning are provided. For parameters that are not defined, a lower bound of $1e-8$ and an upper bound of $1e8$ is used.

Parameter	Effect	Bounds	Information
ρ	Cell division	[0.42, 1.66]	Tumor doubling time is ≈ 1 day (Morimoto-Tomita et al., 2005)
κ	Carrying capacity	[$1e2$, $1e5$]	Tumor growth slows at day 7 with < 2000 tumor cells (scaled by s) (Fig. 1)
ψ	Infection	[$1e-10$, $1e-2$]	Arbitrary interval based on estimate from Iwami et al. (2012)
β	Virus burst size	[$1e0$, $1e4$]	Arbitrary interval based on estimate from Iwami et al. (2012)
α	Cell death	[$1e-4$, $1e3$]	Arbitrary interval based on estimate from Iwami et al. (2012)
δ	Virus degradation	[$1e-2$, $1e2$]	Arbitrary interval based on estimate from Neumann et al. (1998)
$\xi_{i,k}$	Individual homing	[$1e-7$, $1e0$]	Scaling factor for effective tumor inoculation amount, necessarily < 1
σ_a	Additive noise	[$1e1$, $1e5$]	Similar upper bound as κ
σ_b	Proportional noise	[$1e-4$, $1e2$]	Pairwise relative differences in the data were always < 20 (Fig. 1)

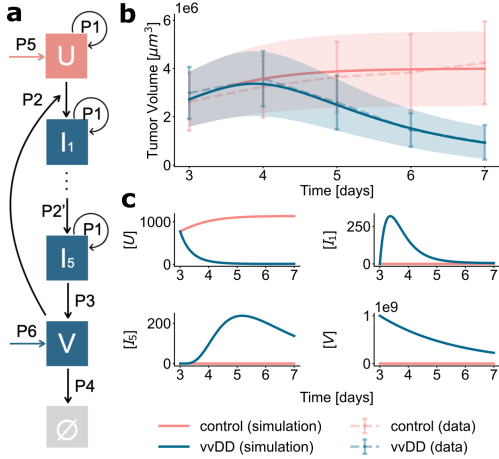


Fig. 4. **The age-of-infection model.** (a) Process diagram. (b) Optimized model simulation. Data are means with ± 1 empirical std. dev. error bars. Simulations have ± 1 estimated std. dev. error bands. (c) Optimized state variable trajectories.

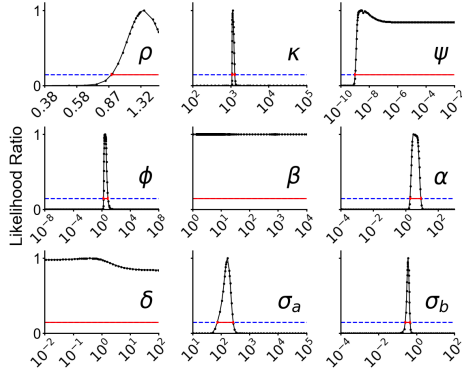


Fig. 5. **Age-of-infection model profile likelihoods.** Domains are the parameter bounds (Table 1). Blue lines are the likelihood ratio threshold that was used to compute the 95% confidence intervals (red lines).

3.4 Modeling Inter-Individual Variability

The age-of-infection model provides a good description of the mean tumor growth and shrinkage dynamics amongst zebrafish. Yet, it does not explain the observed inter-individual variability. We thus consider an **individual-based** age-of-infection model; specifically, we allow for differences in the initial amount of tumor cells, e.g., due to different numbers of injected cells or different homing

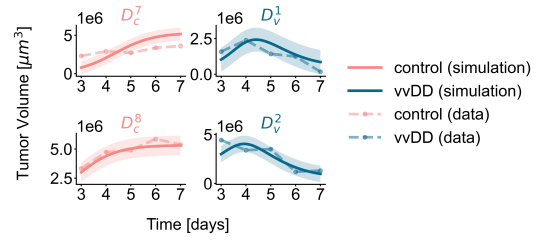


Fig. 6. **Optimized fits with the individual-based model.** Simulations have ± 1 estimated std. dev. error bands. The fits for the remaining 16 zebrafish are provided in the Supplementary Material.

Table 2. **AIC and AICc values for different model candidates.**

Model	AIC	AICc
Baseline	1475.8	1477.3
Age-of-infection	1464.7	1466.7
Individual-based	1437.0	1461.8

rates. Accordingly, the initial condition in time interval 2 can differ between individual zebrafish,

$$[U](t_{i,U}) = \xi u_1,$$

with individual-specific homing parameter $0 < \xi \leq 1$.

We estimate the shared and individual parameters of the model by minimizing objective function (3). In contrast to the previous estimation problem for the age-of-infection model, we have 20 additional parameters: $\xi_{v,1}$ to $\xi_{c,10}$. Despite this increase, we observe that the optimization also converges (see Supplementary Material).

The assessment of the fitting results reveals relatively good agreement of data and model simulation for individual zebrafish (Fig. 6). Indeed, individual-specific homing parameters describe individual dynamics better for the most part. However, not all variance is explained, as shown in some individuals. Model selection using the standard (AIC) and the corrected (AICc) Akaike Information Criterion shows that the individual-based age-of-infection model performs best (Tab. 2). Accordingly, the individual-based model provides a better balance between model fit and model complexity. The uncertainty analysis using the profile likelihood reveals that some parameters are practically non-identifiable (see Supplementary Material).

4. DISCUSSION

OVT is a promising alternative to established anti-tumor therapies. Yet, the understanding of the underlying pro-

cesses is limited. Here, we presented three mathematical models for OVT in zebrafish, an experimental animal model that enables rapid screening of OVT candidates. Our analysis reveals that a baseline model cannot capture the delay in tumor shrinkage. In contrast, the observed delay in tumor volume reduction can be accurately described using the age-of-infection model with multiple infection states representing different intracellular virus concentrations. The formulation of the individual-based age-of-infection model, which accounts for differences in the number of homed tumor cells, provides a better description of the observed variability.

Yet, our analysis also reveals that the homing rates of tumor cells do not fully capture the observed variability. As a next step, we consider two model extensions that may improve this. Firstly, the joint fitting of the individuals could be improved by with nonlinear mixed-effects models. Such models allow for improved integration and dissection of inter-individual variability and uncertainty. This can be particularly beneficial in the case of non-identifiability. Besides ξ , additional inter-individual differences, e.g., individual-specific carrying capacity κ , could also be considered. A structured approach should then be used to select the most relevant sources of inter-individual variability. Secondly, tumor clearance is inherently stochastic due to heterogeneity at low tumor cell numbers. An immediate extension of our study is to perform extinction analysis with a stochastic implementation of the described biological processes. This would enable identification of OVT parameters that can optimally drive tumors to extinction.

By making our code and models publicly available, we aim to facilitate reuse of our work and support further research in this area. The proposed models provide a foundation for future studies to advance OVT strategies.

ACKNOWLEDGEMENTS

We thank the members of the research group of Prof. Dr. Jan Hasenauer for valuable feedback on drafts. We acknowledge the Marvin and Unicorn clusters hosted by the University of Bonn.

SUPPLEMENTARY MATERIAL

All additional figures and results, together with all SBML and PETA files, and the code and instructions to perform all analyses, is available at <https://doi.org/10.5281/zenodo.14802838>.

REFERENCES

Antonio Chiocca, E. (2002). Oncolytic viruses. *Nature Reviews Cancer*, 2(12), 938–950.

Dominguez, T. et al. (2015). *The early phase of the vaccinia virus replication induces cytokine gene expression in macrophages*. Ph.D. thesis, Ludwig-Maximilians-Universität.

Díaz-Seoane, S. et al. (2023). STRIKE-GOLDD 4.0: user-friendly, efficient analysis of structural identifiability and observability. *Bioinformatics*, 39(1), btac748.

Eftimie, R. et al. (2011). Multi-stability and multi-instability phenomena in a mathematical model of

tumor-immune-virus interactions. *Bulletin of Mathematical Biology*, 73, 2932–2961.

Fazio, M. et al. (2020). Zebrafish patient avatars in cancer biology and precision cancer therapy. *Nature Reviews Cancer*, 20(5), 263–273.

Fröhlich, F. et al. (2021). AMICI: high-performance sensitivity analysis for large ordinary differential equation models. *Bioinformatics*, 37(20), 3676–3677.

Fröhlich, F. et al. (2022). Fides: Reliable trust-region optimization for parameter estimation of ordinary differential equation models. *PLoS Computational Biology*, 18(7), e1010322.

Hucka, M. et al. (2003). The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4), 524–531.

Hurtado, P. et al. (2020). A procedure for deriving new ode models: Using the generalized linear chain trick to incorporate phase-type distributed delay and dwell time assumptions. *Mathematics in Applied Sciences and Engineering*, 1(4), 412–424.

Iwami, S. et al. (2012). Identifying viral parameters from in vitro cell cultures. *Frontiers in Microbiology*, 3, 319.

Kumar, P. et al. (2021). Fractional time-delay mathematical modeling of oncolytic virotherapy. *Chaos, Solitons & Fractals*, 150, 111123.

Ling, A.L. et al. (2023). Clinical trial links oncolytic immunooactivation to survival in glioblastoma. *Nature*, 623(7985), 157–166.

Mahasa, K.J. et al. (2017). Oncolytic potency and reduced virus tumor-specificity in oncolytic virotherapy: a mathematical modelling approach. *PLOS ONE*, 12(9), e0184347.

McCart, J.A. et al. (2001). Systemic cancer therapy with a tumor-selective vaccinia virus mutant lacking thymidine kinase and vaccinia growth factor genes. *Cancer Research*, 61, 8751–8757.

Mealiea, D. et al. (2021). Modeling oncolytic virus dynamics in the tumor microenvironment using zebrafish. *Cancer Gene Therapy*, 28(7), 769–784.

Morimoto-Tomita, M. et al. (2005). Mouse colon carcinoma cells established for high incidence of experimental hepatic metastasis exhibit accelerated and anchorage-independent growth. *Clinical & Experimental Metastasis*, 22, 513–521.

Neumann, A.U. et al. (1998). Hepatitis c viral dynamics in vivo and the antiviral efficacy of interferon- α therapy. *Science*, 282(5386), 103–107.

Raue, A. et al. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15), 1923–1929.

Schälte, Y. et al. (2023). pyPESTO: a modular and scalable tool for parameter estimation for dynamic models. *Bioinformatics*, 39(11), btad711.

Schmiester, L. et al. (2021). PETA—interoperable specification of parameter estimation problems in systems biology. *PLoS Computational Biology*, 17(1), 1–10.

Tian, Y. et al. (2022). Engineering strategies to enhance oncolytic viruses in cancer immunotherapy. *Signal Transduction and Targeted Therapy*, 7(1), 117.