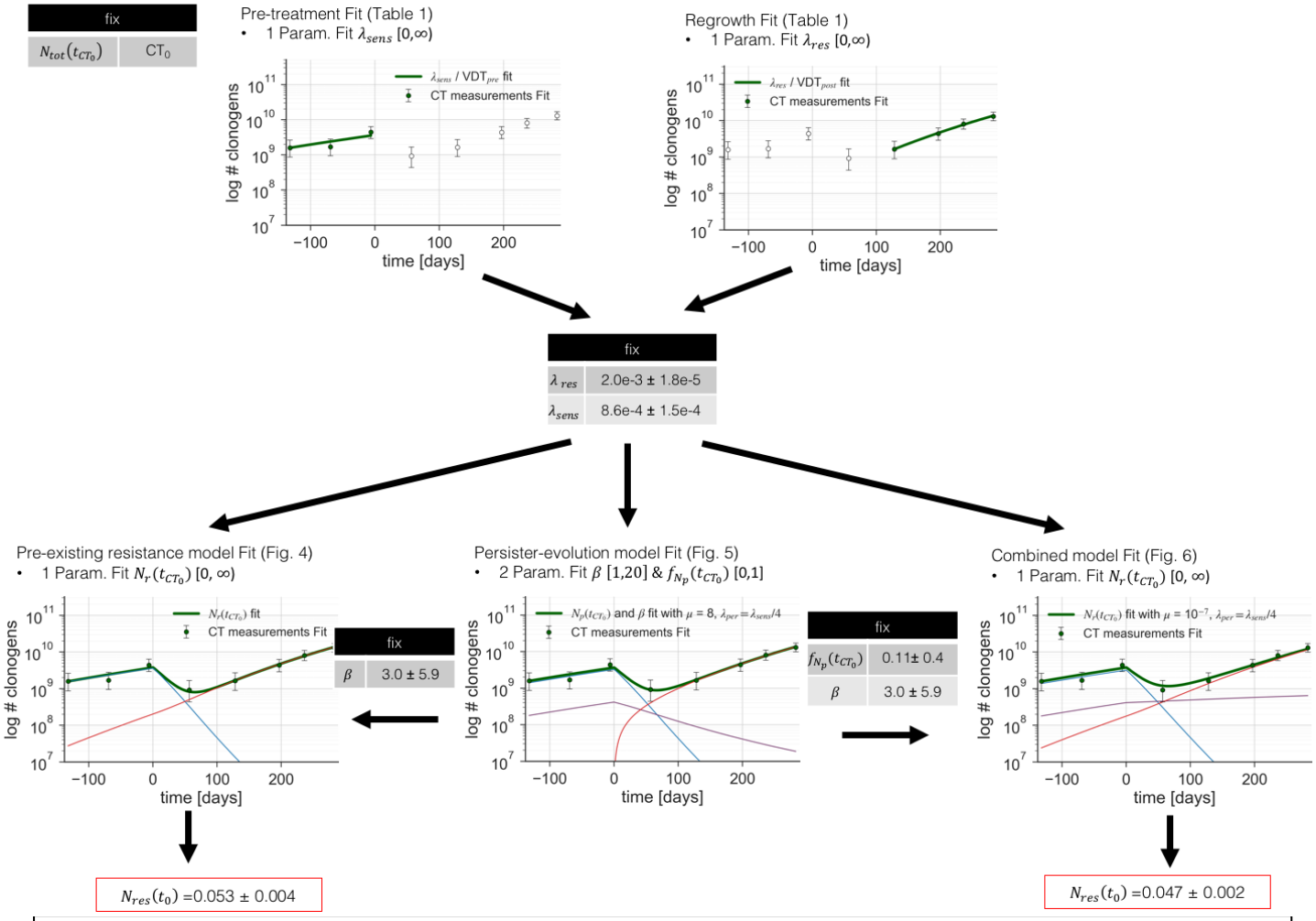
Supplementary Material: Patient-specific tumor growth trajectories determine persistent and resistant cancer cell populations during treatment with targeted therapies

Expanded Description of Fitting Procedure



Supplementary Figure 1. Flow diagram of the multi-step fitting procedure. A pictorial outline of the multi-step fitting procedure with example results from tumor #2.

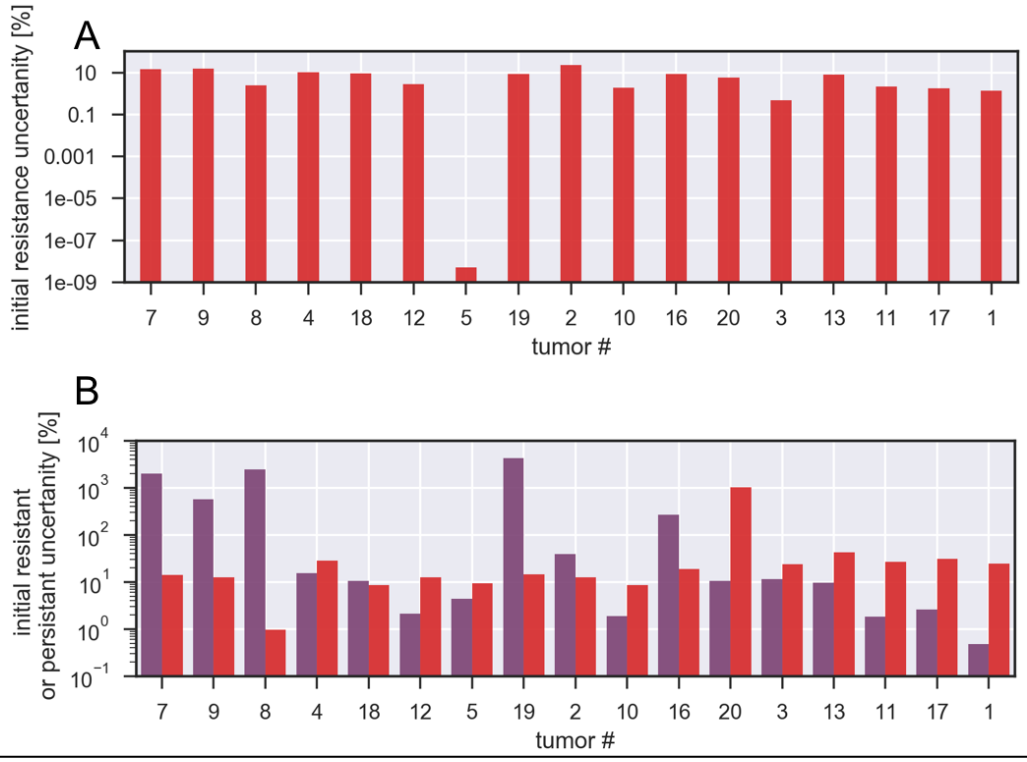
Supplementary Figure 1 demonstrates a step-by-step example of the sequential fitting procedure that was applied to all tumor growth trajectories. At the top level are plots of the CT data points with the pre- or post- treatment growth rate fits, which were only applied to the pre-treatment or post-recurrence data points respectively. Therefore, they are assumed to be representative of the sensitive and resistant growth rates respectively (the persister growth rate is assumed to be a fraction of the sensitive growth rate during TKI treatment). Note the green CT time points which are selectively used to fit the pre- or post- treatment growth rates independently. Similarly, the population of resistant cells is expected to constitute the majority of the tumor recurrence and beyond, and therefore the latter datapoints provide a good representation of size and growth rate of the resistant cell populations. The total clonogen population is actually never directly fit, but only assumed to be the sum of the underlying populations, e.g. only resister and sensitive cells in the resistance-only model (bottom left in supplementary figure 1).

The *sensitive* (from fit in the top left), *persistent* (assumed to be 0.25 of sensitive growth rate based on *in vitro* data, the robustness of this assumption was extensively investigated, see Figure 5 main manuscript), and *resistant* (from fit top right) growth rates are then fixed for the rest of the fitting procedure. In the table in the middle of the figure they are shown with estimation uncertainties reported from the LMFIT python package used in this analysis. On the bottom level, the resister only, persister-evolution only, and combined model fits are shown. The persister evolution model is performed first over a range of transition probabilities and persistent growth rates. The fitted cell-kill parameter β is fixed for both the resister only and combined models again to constrain the number of free fit parameters. Note that the value β is most senstivie to the first post-treatment CT data point, and is consequently bounded. While there may be significant uncertainity in β, there is minimal effect on the robustness of the fit over the entire time-course of CT measurements. This is due to the fact that the exact value chosen for β only affects one data point, the first CT after treatment initiation, and no other parts of the curve.

Additionally, the fitted initial persister fraction is fixed for the combined model. This is done to set a maximal value for the initial persister fraction, resulting in a minimal initial resister fraction, as the purpose of this study is to investigate which minimum fraction of cells has to exhibit pre-existing resistance to explain the observed recurrence trajectories. If both the initial persister and resister fractions were simultaneously fitted in the combined model, there would be a family of negatively correlated solutions with negligible initial persister fractions and unreasonably high initial resister fractions (>20%), similar to the resistance only model, which are not clinically observed (see discussion section). Finally, the independently fitted initial resister fraction is displayed for both the resister only and combined models with their estimated uncertainties.

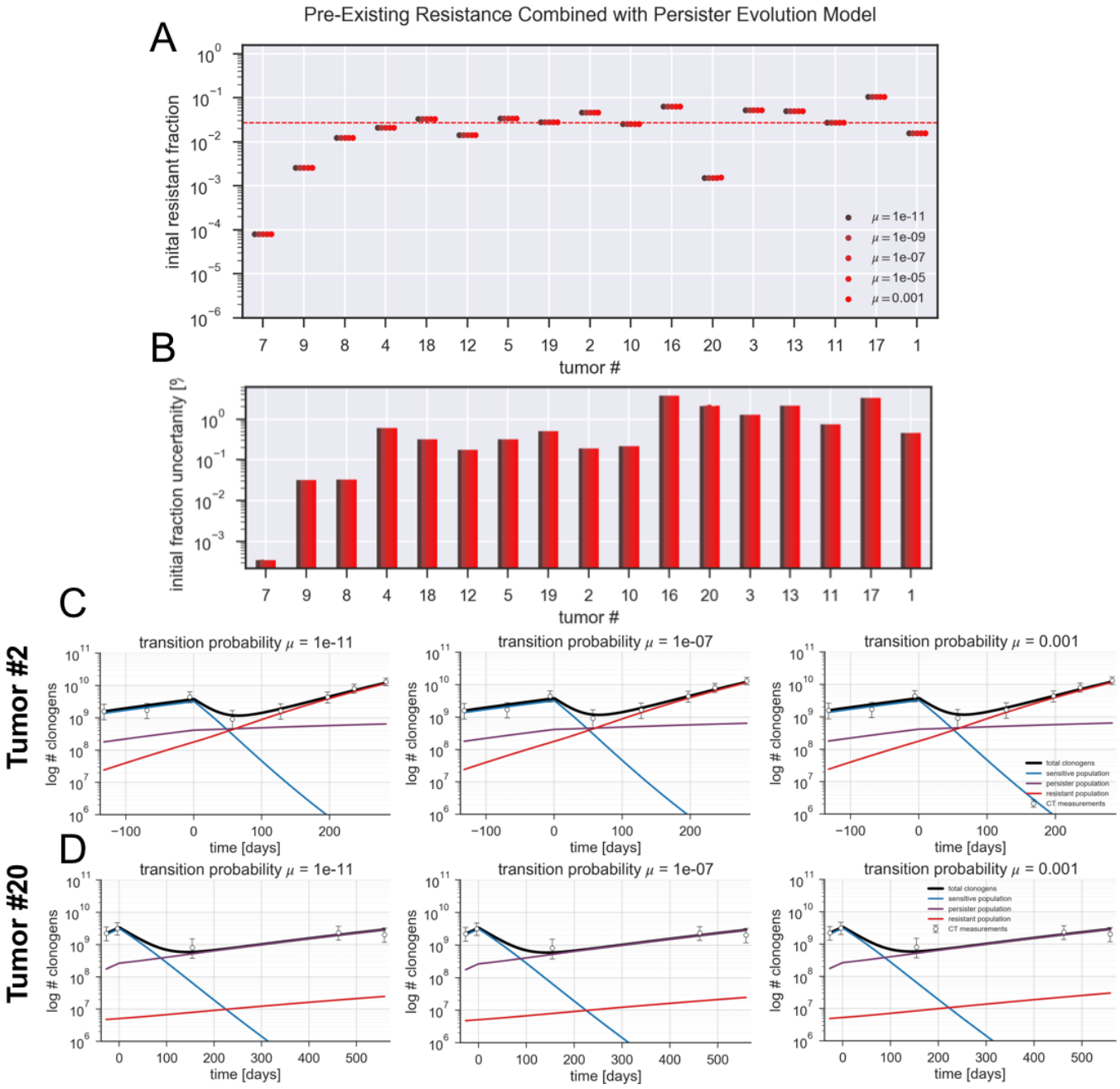
Note that for all fitting steps, the initial total clonogen population is fixed to the first CT data point to reduce the number of free fit parameters. Also note that for each step only one parameter is fitted, with the exception of the persister only model, where both the cell-kill parameter β and the initial persister fraction are fitted. The bounds over which each parameter was constrained for all tumors is displayed above each plot.

Uncertainty of Initial Population of Resistance and Persistence Estimation



Supplementary Figure 2. Uncertainties of the pre-existing resistant or persistent fraction. A: Uncertainties in the estimation of the initial population of resistance for the pre-existing only model. B: Uncertainties in the estimation of the initial population of resistance and persistence for the combined model. *All uncertainties are expressed as fraction of the population itself, which is different from the figures shown in the main manuscript.*

Sensitivity of Resistant Fraction Estimation to Mutation Probability



Supplementary Figure 3. Results of the combined model for different mutation induction probabilities. A: Fractions of pre-existing resistance in the combined model (similar to figure 6), for five values of mutation induction ranging from 10^{-11} to 10^{-3}. B: Uncertainties in the estimation of the pre-existing resistance fraction for each mutation induction. C & D: fits for two example tumors with mutation inductions of 10^{-11}, 10^{-7}, and 10^{-3}.