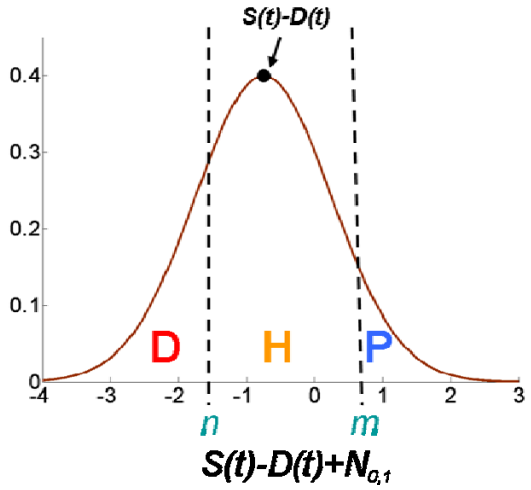
Supplementary Material

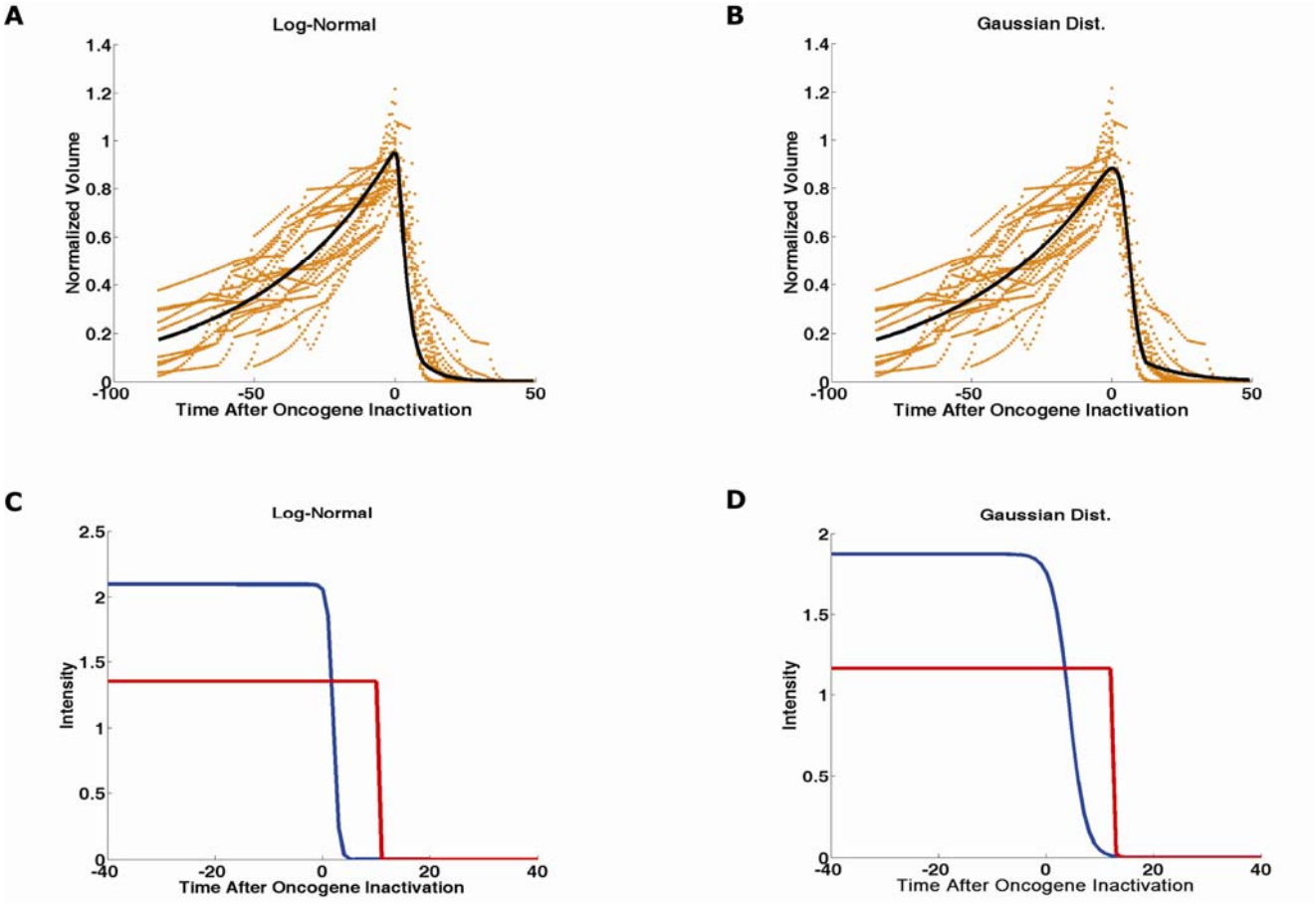
Materials and Methods

Mathematical Modeling of Signal Behavior. Our mathematical model represents the temporal changes in tumor volumes before and after oncogene inactivation as a balance of two aggregate signals, a survival (*S(t)*) and a death (*D(t)*) signal (*S1*). At any given time cells may react to the balance of these signals through one of three states, proliferation (P), homeostasis (H) or apoptosis (A). We have defined the homeostatic population of cells as non-cycling cells and thus may contain cells in G0, differentiated cells or dormant tumor stem cells:



Cells within the same tumor undergo different programmatic decisions as is observed empirically with simultaneous proliferation and apoptosis within a given tumor. The stochastic difference in cell behavior to input signals at the microenvironmental level is modeled using a normal distribution CDF $\Phi(\mu,\sigma^2)$ on the difference between survival and death signals, where μ is the difference between the input signals, and σ is 1. To minimize the number of unknowns and to eliminate a redundant degree of freedom, we fixed the

value σ to 1 and thus, the scale of the signaling intensities is in arbitrary units. Along with the normal distribution sampling over time (Fig. 4B), the percentage of cells in one of the three states is also determined by two different thresholds, *m* and *n* (Fig. 4B). We assumed that the large numbers of cells in the tumor are independent random variables. Based on the central limit theorem, the re-averaged sum of the large number of random variables will be approximately distributed normally with finite mean and variance. Therefore, we assumed normal (or Gaussian) distribution to represent the stochastic difference in cell behaviors. In addition, the primary determinants of the fate of cells in this mathematical model are the thresholds *m* and *n*. Therefore, the specific form of distributions used in the mathematical model should not greatly change the overall prediction.



Comparison of the use of normal distribution *versus* log-normal distribution reveals that both functions produce very similar model fit curves for aggregate survival and death signals as shown in the figure above. Regardless of which functions we use to model the behavior of single tumor cells, survival and death signals have similar shapes and decay rates. The model fit well to both of the original tumor volumes using (A) log-normal distribution and (B) normal distribution. The survival signals were short-lived following oncogene inactivation compared to the death signals in both (C) log-normal distribution and (D) normal distribution. The figure above shows that the results generated using normal distribution and log-normal distribution are almost identical. For simplicity, we used normal distribution in the rest of the paper and for setting up the mathematical model.

Therefore, the fraction of cells in each state in response to the balance of the input signals can be summarized as follows,

fraction of cells proliferating = $1-\Phi(m-S(t)+D(t))$

fraction of cells in apoptosis = $\Phi(n-S(t)+D(t))$ (1)

fraction of cells in homeostasis = 1 – (fraction of cells in proliferation) – (fraction of cells in apoptosis)

The rate of volume change over time observed in microCT images is therefore determined by the rates of cell proliferation and cell apoptosis, which are then determined by the balance of input signals. These rates can be determined by dividing the number of cells in each state by the amounts of time required for cell proliferation (T_p) and cell apoptosis (T_a). Therefore, our model to explain the temporal changes over time based on

the balance of the *S(t)* and *D(t)* signals can be mathematically summarized by an ordinary differential equation,

$$\frac{dV}{dt} = \frac{1}{T_p} \cdot (1 - \Phi(m - S(t) + D(t))) \cdot V - \frac{1}{T_a} \cdot (\Phi(n - S(t) + D(t))) \cdot V \qquad (2)$$

The exponential-like individual tumor volume curves were linearly interpolated in semi-log space to minimize interpolation error. Tumors that were too small to be identified on microCT were treated as a single voxel to avoid log(0). In addition, we modeled the behaviors of the survival *S(t)* and death *D(t)* signals as sigmoidal curves because the signals were found to be in a steady state before perturbing the system (inactivation of the oncogene) and we assumed they would reach another steady state some time after perturbation. Therefore, we used the simplest sigmoid function, i.e. the logistic function, commonly used in biological models to represent these intracellular signals which can be mathematically summarized as follows,

$$S(t) = \frac{a}{1 + e^{b(t-c)}}, \quad D(t) = \frac{d}{1 + e^{f(t-g)}} \qquad (3)$$

Parameters *b* and f were the rates of signal decay, *c* and *g* were the amounts of time it takes for the signals to begin dropping off, and *a* and *d* were the starting intensities of the signals. These parameters were estimated using the Levenberg-Marquardt optimization technique (*S2*) to obtain optimal values based on the actual volumetric measurement obtained from microCT images and the IHC data described in the following section.

Equations (4) and (5) also summarizes the temporal rate changes in the IHC data, in particular the measurements of cell apoptosis (represented by cleaved caspase 3 and TUNEL staining) and proliferation (represented by Ki-67). These measurements provide a

measure of cells in a state of proliferation and apoptosis, quantified as an instantaneous percentage. However, in order to quantify the rate as events per unit time, one must also consider the duration for which cells express the markers of proliferation and apoptosis, in this case, t_p for the duration for which casp-3 is expressed and t_a for the duration for which Ki-67 is expressed. For example, a fast rate of events with a shorter duration of detectability could have the same instantaneous percentage as a slower rate of events with a longer duration of detectability (*S3*). Because of a lack of estimates of these durations for these particular tumors in the published literature, we have treated these variables as unknown parameters that are then estimated along with other model parameters. Based on the same assumption for the signaling model, we reasoned that temporal changes in the proliferation and apoptosis rates were also determined by the differences between the survival *S(t)* and death *D(t)* signals. Therefore, the relationship between the signals and the IHC measurements can be mathematically summarized by the following equations:

$$\text{PI} = \textit{(fraction of cells in proliferation)} \cdot \frac{t_p}{T_p} = (1 - \Phi(m - S(t) + D(t))) \cdot \frac{t_p}{T_p} \qquad (4)$$

$$\text{AI} = \textit{(fraction of cells in apoptosis)} \cdot \frac{t_a}{T_a} = (\Phi(n - S(t) + D(t))) \cdot \frac{t_a}{T_a} \qquad (5)$$

where PI is the proliferation index from Ki-67 IHC data, AI is the apoptosis index from cleaved caspase 3 and TUNEL staining IHC data, t_p and t_a were the durations that proliferation and apoptosis, respectively, could be detected by IHC. A couple of assumptions were made in the modeling, 1) PI and AI were at a stable rate before oncogene inactivation, and 2) the asymptotic behavior of PI continued beyond 10 days of oncogene inactivation for *K-ras*^{G12D}-induced lung tumors (Fig. 2L). Parameters in

equations (2), (4), and (5) were estimated using the Levenberg-Marquardt (LM) algorithm on combined data of volumetric measurements of microCT images and the Ki-67 and cleaved caspase 3 measurements. The task of the LM algorithm can be stated as follows: given three sets of data points, V_i (normalized volume), P_i (PI) and A_i (AI), determine all parameters $\boldsymbol{\beta}=\{a,b,c,d,f,g,t_a,t_p,T_a,T_p,n,m\}$ of the model curve $f(x,\boldsymbol{\beta})$ to minimize the error function $E(\boldsymbol{\beta})$ in equation 6. $\hat{V}_i(\beta)$, $\hat{P}_i(\beta)$ and $\hat{A}_i(\beta)$ are the fitted values of normalized volume, PI and AI for a given set of parameter values $\boldsymbol{\beta}$. To ensure equal weighting between volume, PI and AI datasets, all three were normalized to their maximum value and mean values were used instead of sums.

$$E(\beta)=\sqrt{\frac{1}{n_v}\sum_{i=1}^{n_v}(\hat{V}_i(\beta)-V_i)^2}+\frac{1}{\max P_i}\sqrt{\frac{1}{n_p}\sum_{i=1}^{n_p}(\hat{P}_i(\beta)-P_i)^2}+\frac{1}{\max A_i}\sqrt{\frac{1}{n_a}\sum_{i=1}^{n_a}(\hat{A}_i(\beta)-A_i)^2} \quad (6)$$

Assessment of the Robustness of the Model. To validate our proposed mathematical model, we first performed a hold-out like validation analysis on these datasets, *V*, *PI* and *AI* in order to ensure that there was no overlap between training and test data. We assumed the duration of cell proliferation and apoptosis, T_a, T_p, detectability of cell proliferation and apoptosis by IHC markers, t_a, t_p, and thresholds that determine cell stochasticity, n, m, remain the same for all oncogene-addicted tumors. We therefore used the values of these parameters from all 20 oncogene-addicted tumors estimated using LM optimization

algorithm and kept them fixed throughout this part of the experiment. Different combinations of two of the three datasets, *{(V,PI), (V,AI), (AI,PI)}*, were used for parameter optimization to assess whether the estimated signal curves would be alike. Each permutation of the hold-out-like validation experiments also resulted in a delayed decay of the aggregate death signal relative to the survival signal (fig. S3A-C).

In addition, we utilized bootstrap statistical method to estimate the sampling distribution and the bias and variance of the optimization algorithm by re-sampling with replacement from the original sample. This is done to validate the robustness of our mathematical model to the data itself using random subsets of the data and to determine the stability of the estimated parameters.

The bootstrap algorithm can be formally described as follows.

For n = 1 to 10,000,

1) For i = 1 to 20,
    a. Randomly select one tumor in the dataset with replacement using uniform distribution.
2) Combine the volume data from the selected tumors with the IHC data to obtain ***H***. Repeats in the tumors selected in step 1 are allowed.
3) Estimate all parameters of equation (2) by curve fitting the model with ***H*** using the LM optimization technique and the objective function described in equation (6).

Finally, to obtain the distribution of each parameter and compute the coefficient of variation (CV) from the bootstrap experiments.

In order to assess the robustness of the proposed method to the values of the estimated parameters from the optimization method, we further performed a sensitivity analysis of each variable. We assessed the average RMSE of the model fit to the actual volume, PI and AI data within the ± 5% of the estimated value. The set of our estimated values for all parameters is considered stable if all values are at the local minimum of the assessed range. Root mean square error (RMSE) for the model fitting of *K-ras*^{G12D} data was 5% error for normalized volume fit, 8.5% error for PI fit (Ki-67), and 5% error for AI fit (cleaved caspase 3) or 6.2% average overall. *MYC* model fit: RMSE was 8% for volume fit, 18% error for PI fit (Ki-67), and 67% for AI (cleaved caspase3) fit or 31% overall.

The fig. S3D-I show the statistics estimated from bootstrap resampling technique and the sensitivity analysis. Both data showed that our mathematical model is robust to variation in the data and to the values estimated from the optimization. Typically, distributions with CV < 1 are considered low-variance, while those with CV > 1 are considered high variance. fig. S3D showed that all parameters are low variance except for the variables that represent the decay rate of aggregate death signal (variable f) and the time when the signal decay is initiated (variable g). In addition, the sensitivity analysis (fig. S3E-I) also demonstrated similar instability in variable g. These observations demonstrated the uncertainty of the duration of the presence of death signal after oncogene inactivation. One possible explanation is that, as tumor volumes become progressively smaller, the limits of the microCT based volumetric measurements become more apparent such that the validity of the absolute time of signal decay is uncertain.

Histology and Immunohistochemistry.

Tissues were fixed in 10% buffered formalin for 24 h and then transferred to 70% ethanol until embedded in paraffin. Tissue sections 5 μm thick were cut from paraffin embedded blocks, placed on glass slides and hematoxylin and eosin (H&E) staining was performed using standard procedures (Stanford Histology Core). We measured Ki67 and TUNEL-staining as described previously (*S4*). Antibodies used in our study: c-Myc (C19) (Santa Cruz Biotech.), cleaved caspase 3 (Cell Signaling Tech.), phospho-AKT-S497 (Cell Signaling Tech.), phospho-Erk1/2-T202/Y204 (Cell Signaling Tech.), phospho-Stat3-Y705 (Cell Signaling Tech.), phospho-p38-T180/Y182 (Cell Signaling Tech.) and phospho-Stat5-Y694 (Cell Signaling Tech.). Samples were dewaxed in xylene and rehydrated in a graded series of ethanols. Antigen retrieval for c-Myc, cleaved caspase 3 and phospho-AKT were performed by 14 min microwave irradiation in citrate-based Antigen Unmasking Solution (Vector Laboratories, Burlingame, CA, USA). Antigen retrieval for phospho-Stat3 and -Stat5 were performed by 14 min microwave irradiation in EDTA, pH 8.0, and antigen retrieval for phospho-Erk1/2 and phospho-p38 was performed by10 min incubation in Pronase (Roche, Basel, Switzerland). Endogenous peroxidases were blocked in either 3% hydrogen peroxide in deionized water (phospho-AKT, -pErk, -p38 and -pStat3/5) or 0.3% hydrogen peroxide in methanol (c-Myc and cleaved caspase 3) for 10-20 minutes. Non-specific binding was blocked with 5 -10% goat serum for 60 minutes. Primary antibodies were used at appropriate dilutions (c-Myc, phospho-AKT, -pErk and -p38 at 1:100; cleaved caspase 3 at 1:150; phospho-Stat5 at 1:200; and phospho-Stat3 at 1:50) and sections

incubated overnight at 4 degrees Celsius. Detection was conducted using the Vector Elite ABC detection kit (Vector Laboratories) with 3,3'-diaminobenzidine tetrahydrochloride as the chromogen. Sections were counterstained with Gill's hematoxylin (Vector Laboratories).

Tumor Segmentation.

A semi-automated open source image analysis application, ITK-Snap (*S5*), was employed for segmentation of the lung tumors from the microCT images. The tool uses a level set algorithm to semi-automatically delineate tumor from background by expanding a level set from a user-defined seed point. We viewed 2D images in sagittal, coronal and transverse views to detect tumors and track them over time. The segmentation algorithm was seeded manually and stopped once the tumor boundary was detected. The post-processing of the segmented data provided the voxel counts, the volume (in cubic millimeters) and displayed the shape of the segmented structure. Matching tumors across time points was performed manually by simultaneous viewing of the serial data. This volume information was then used for analysis of the temporal changes in lung tumor nodules. MicroCT imaging has been shown to correlate with the number and volume of murine lung tumors found on necropsy (*S6*).

Normalization.

Peak tumor volumes at the time of oncogene inactivation had significant variability due to the subjective determination of when the oncogene was inactivated and also due to the variability in tumor sizes in the same mouse. However, simple normalization by peak

tumor volume was artificially biased toward a single time point due to zero variance at time zero (peak) and exaggerated variance at large negative and positive times. Instead, we weighted each time point equally in the normalization process. Due to the exponential-like growth and regression pattern observed in volume data, two straight lines were fitted by least squares to tumor volume in semi-log space, one for the growth phase and one for the regression phase. The value at the intersection of the two straight lines was then used to normalize the volume for each tumor time series (fig. S2A). Normalized tumor volumes at each scan time were averaged across all tumors to obtain a general trend of the growth and regression patterns leading to much more consistent variance in tumor volume measurements across time.

MYC-induced Lymphoma Mouse System.

MYC-induced lymphoma cells expressing luciferase were transduced using retrovirus carrying either pBabe-puro, pMx-Stat3d358L-puro, or pBabe-MyrAktHA-puro. Infected cells were then selected with 1 μg/ml puromycin. Ten million *MYC*-induced lymphoma cells were injected subcutaneously and then were scanned every 2-5 days prior to oncogene inactivation and 20 days after oncogene inactivation using BLI imaging (*S7*). Average radiance (p/s/cm2/sr) of each scan was collected and treated similarly as volumetric measurements for the *K-ras*^{G12D}–induced lung tumors and was used to study tumor kinetics. A similar normalization method was used on the BLI measurements. We fitted the BLI data to equation (2) to obtain the necessary parameters for estimation of *S(t)* and *D(t)* behaviors.

Support Vector Machine (SVM) Classification.

v-SVM and leave-one-out techniques were used for the classification and prediction of tumor genotype based on the first few serial weekly scans after oncogene inactivation. A recent paper by Basavanhally *et. al.* (S8) has shown that for various medical image classification tasks, SVM has best performance (compared to kNN and C4.5) for n>13 training set size. Our data set was n~20 for the SVM. Although the datasets and tasks differ, we believe this indicates that SVM is a reasonable choice of classifier. Two classes of data points were constructed: (1) 20 *K-ras*^{G12D} oncogene-induced tumors, and (2) 8 non-oncogene addicted *MYC*-induced mouse tumors also following simulated *MYC* oncogene targeted treatment. *v*-SVM algorithm can be summarized as follows. Given a dataset *D* of *n* samples, $(\boldsymbol{S}_1,L_1),\cdots,(\boldsymbol{S}_n,L_n)$, where $\boldsymbol{S}_i = <s_1,s_2,\cdots,s_m>$ was a set of feature vectors which in our case were the tumor volumes from weekly scan, and $L_i \in \{-1,+1\}$ represented the binary class membership, SVMs mapped the input data into a high-dimensional feature space, and constructed a maximum margin hyperplane to separate the two classes in order to predict the label of a new sample. In this work, we used the Gaussian kernel $k(S,S') = \exp(-d(S,S')^2 / \sigma^2)$ to map the input data to the high-dimensional feature space, where ***S, S'*** were two feature vectors, *d* was the Euclidean distance, and σ the width parameter of the kernel. We selected the width σ of the kernel by minimizing the leave-one-out error estimate of the classification error between the true label and the predicted value of the *v*-SVM classification function for each σ. In our experiments simulating a more heterogeneous population, we included 20 oncogene addicted *K-ras*^{G12D} tumors, 8

non-oncogene addicted *MYC* tumors and 6 double mutant, *MYC*/ *K-ras*^{G12D}, tumors for training and classification. Training data was derived from normalized tumors volumes.

Mouse and Human Lung Tumor Regression Analysis.

Human lung tumor measurements from CT scan data were performed according to the World Health Organization (WHO) standards. Bi-dimensional tumor measurements were performed before and 4 weeks following erlotinib therapy by obtaining the longest diameter of the tumor in the axial plane and the greatest perpendicular to it. Tumor volume was calculated as a sphere and percent change at follow up was calculated as a percentage of baseline value. The median doubling time of human lung tumors is 144 days (S9-S11) while the conditional transgenic mouse lung tumors in our study double every 37.04 days. To correct for different growth kinetics between species, mouse tumor kinetics were expanded by a factor of 3.89 (human tumor doubling time divided by mouse tumor doubling time). Using this analysis, at one month after oncogene inactivation, the average change in volume of oncogene addicted mouse lung tumors (*K-ras*^{G12D}) was 58.5% while patients with *EGFR* mutations had average change in tumor volume of 60.9%. Similarly, the average change in volume of non-oncogene addicted mouse lung tumors (non-*K-ras*^{G12D}) was 12.5% while patients with wildtype *EGFR* had an average change of 5.7%.

Supplementary Figures

Fig. S1

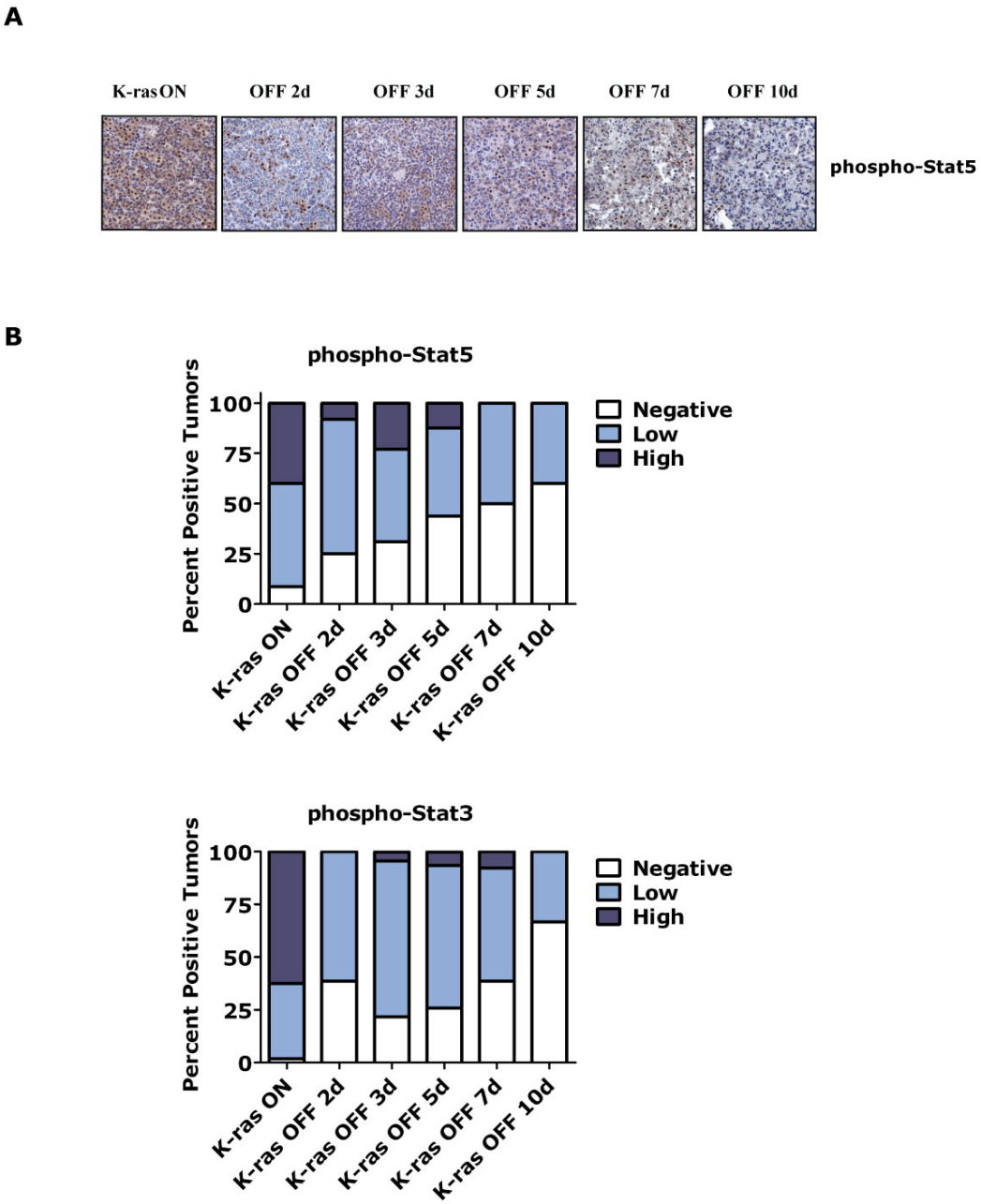
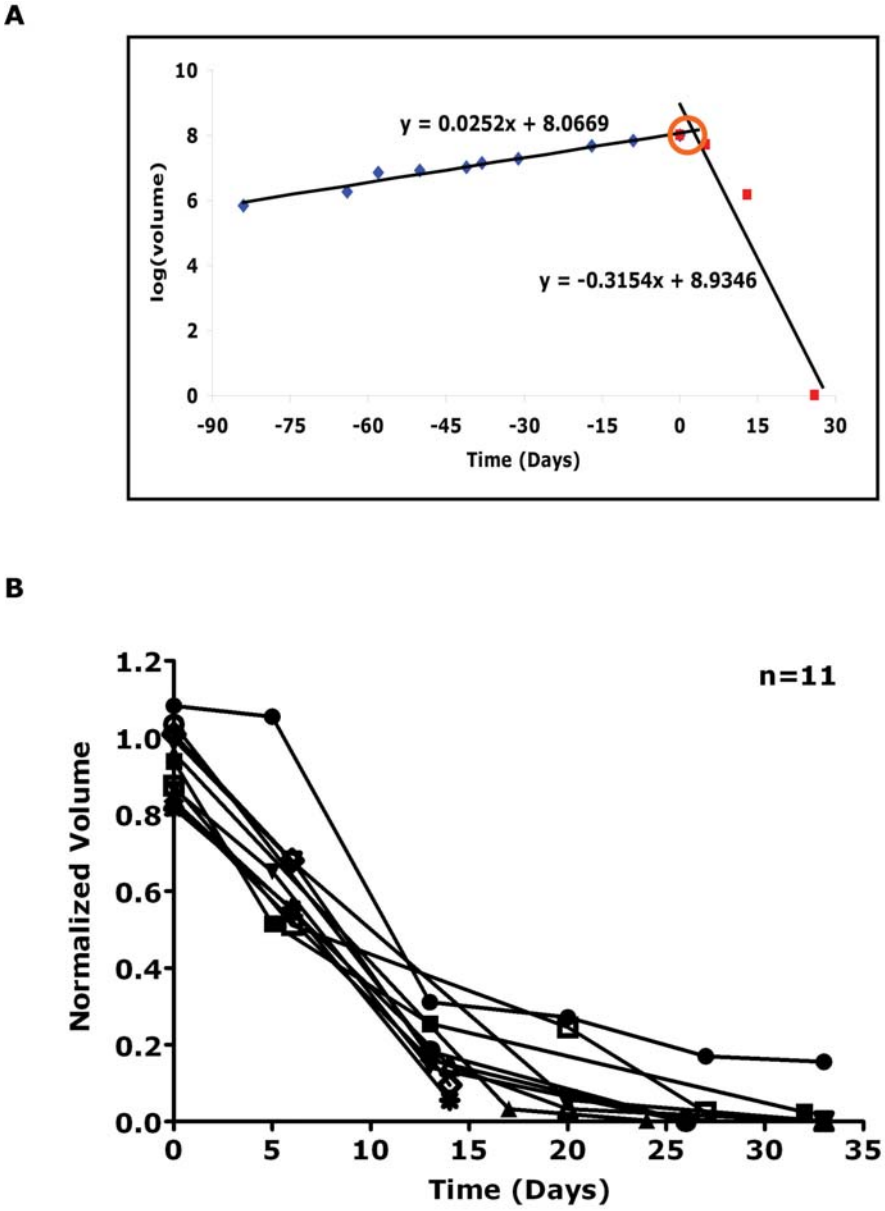


Fig. S1. Attenuation of pro-survival and pro-death signaling pathway mediators, during regression of *K-ras*^{G12D} –induced lung tumors *in vivo*. (A) *K-ras*^{G12D} induced lung tumors with

K-ras^{G12D} oncogene activated (or "ON") and at time points following oncogene inactivation (or "OFF") were evaluated for phosphorylation of signaling pathway mediators by IHC. Representative examples show pro-survival pathways mediated by STAT5 are phosphorylated in lung tumors when *K-ras*^{G12D} is "ON," but are de-phosphorylated after turning *K-ras*^{G12D} "OFF" (day 2-3). (B) Phospho-Stat5 and –Stat3 IHC was scored using similar system as in Fig. 3B.

Fig. S2


A
y = 0.0252x + 8.0669
y = -0.3154x + 8.9346
log(volume)
Time (Days)
B
n=11
Normalized Volume
Time (Days)

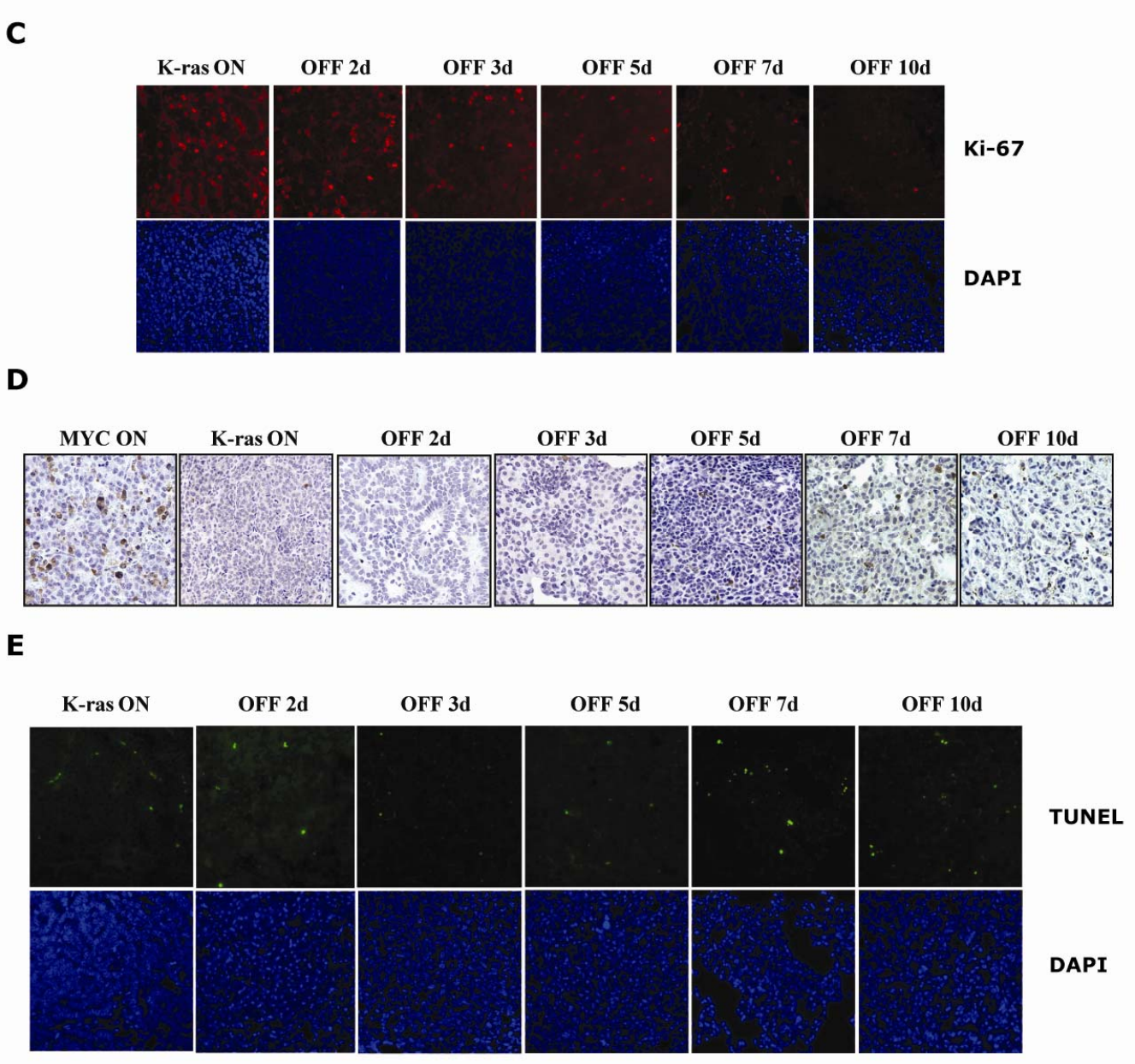
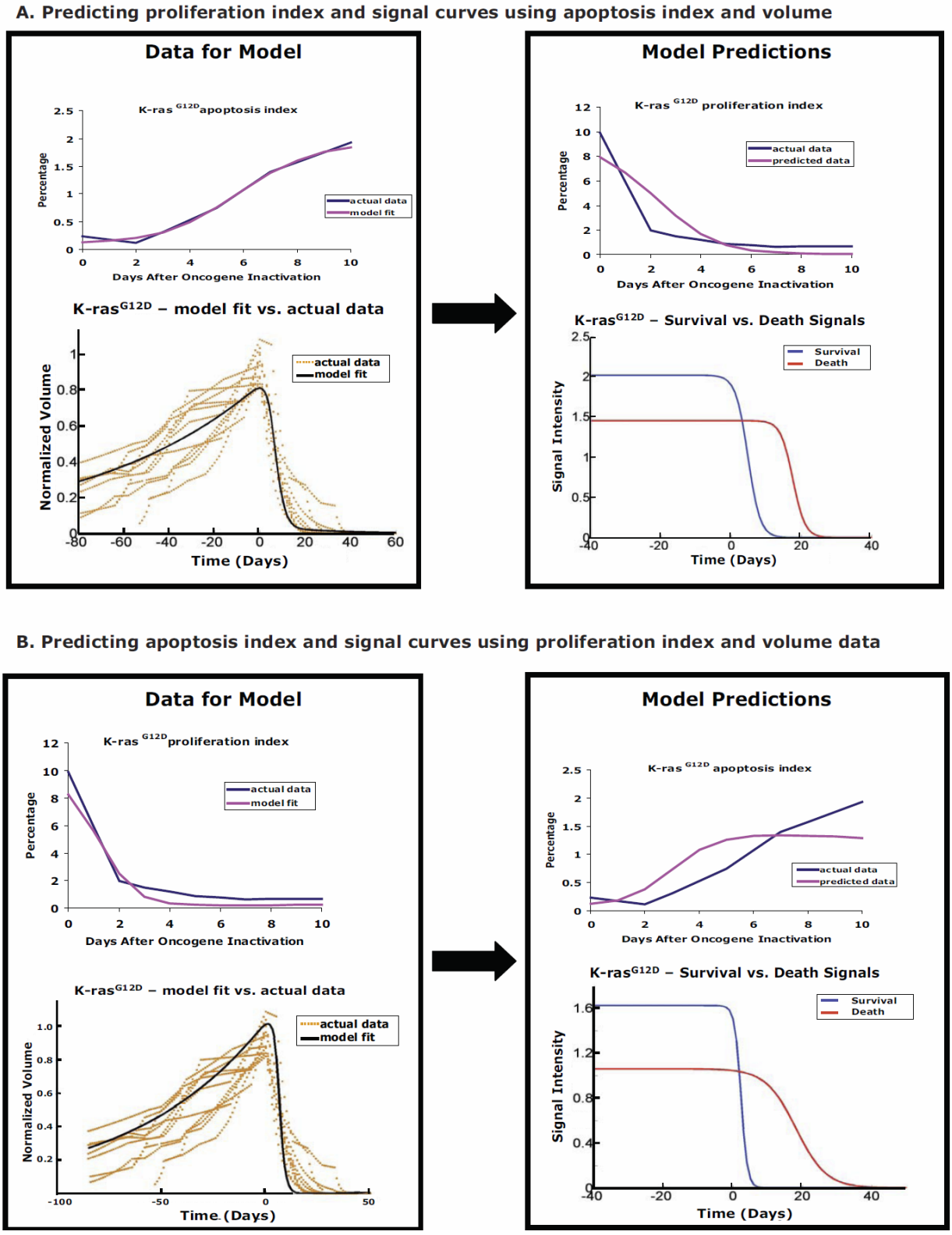
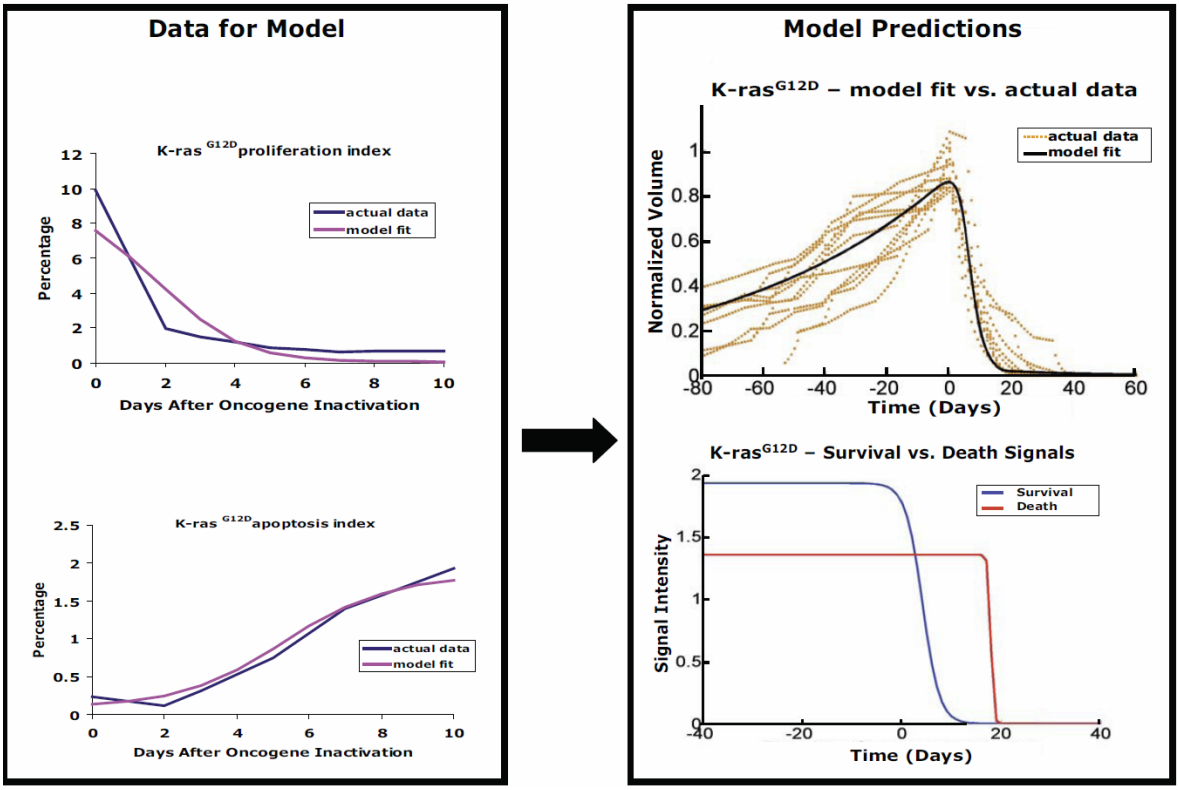

C
K-ras ON
OFF 2d
OFF 3d
OFF 5d
OFF 7d
OFF 10d
Ki-67
DAPI
D
MYC ON
K-ras ON
OFF 2d
OFF 3d
OFF 5d
OFF 7d
OFF 10d
E
K-ras ON
OFF 2d
OFF 3d
OFF 5d
OFF 7d
OFF 10d
TUNEL
DAPI

Fig. S2. *K-ras*^{G12D} induced murine lung tumors are oncogene-dependent. (A) Volumetric quantization of *K-ras*^{G12D} lung tumors using serial microCT. Representative tumor volume plotted as two straight lines fitted by least squares in semi-log space, one for the growth phase and one for the regression phase. (B) Oncogene inactivation in *K-ras*^{G12D} lung mice show individual lung tumors quantified and normalized over time as described in Supporting Materials and Methods show consistent complete regression of CR lung tumors (n=11). (C) Sections of *K-ras*^{G12D}–induced lung tumors following doxycycline withdrawal time course show decreasing proliferation with time as demonstrated by immunofluorescence (IF) against Ki-67. (D & E) Increased frequency of apoptotic cells with time after switching *K-ras*^{G12D} "OFF." (D) Similar CR lung tumor sections were stained for cleaved caspase 3 and showed increased apoptosis after switching *K-ras*^{G12D} "OFF." (E) Lung tumor samples were assayed for apoptotic cells using the TUNEL procedure and imaged using IF at the indicated time points relative to switching *K-ras*^{G12D} "OFF". IF sections were counterstained with DAPI to indicate cell number.

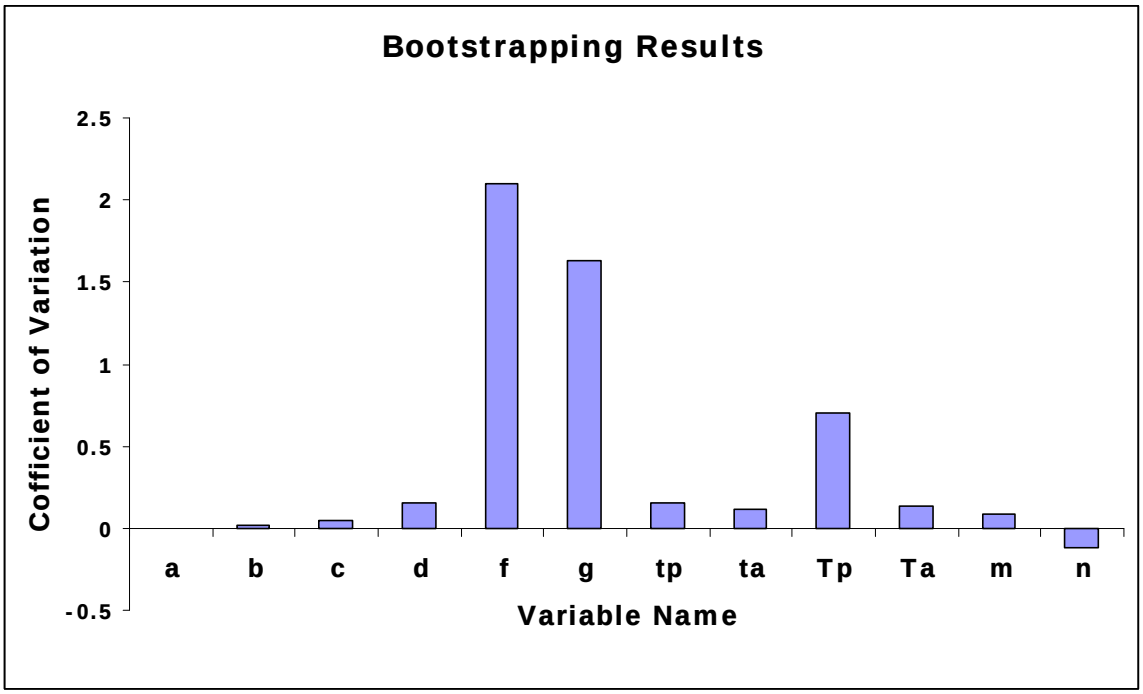
Fig. S3


A. Predicting proliferation index and signal curves using apoptosis index and volume
Data for Model
K-ras G12D apoptosis index
Percentage
Days After Oncogene Inactivation
actual data
model fit
K-rasG12D – model fit vs. actual data
Normalized Volume
Time (Days)
Model Predictions
K-ras G12D proliferation index
actual data
predicted data
K-rasG12D – Survival vs. Death Signals
Survival
Death
Signal Intensity
Time (Days)
B. Predicting apoptosis index and signal curves using proliferation index and volume data
Data for Model
K-ras G12D proliferation index
Days After Oncogene Inactivation
K-rasG12D – model fit vs. actual data
Time (Days)
Model Predictions
K-ras G12D apoptosis index
K-rasG12D – Survival vs. Death Signals

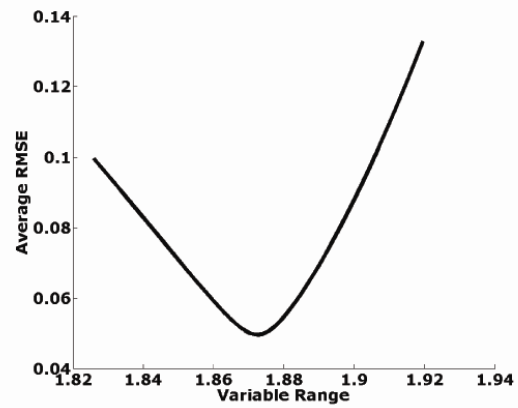
C. Predicting volume data and signal curves using proliferation and apoptosis indices

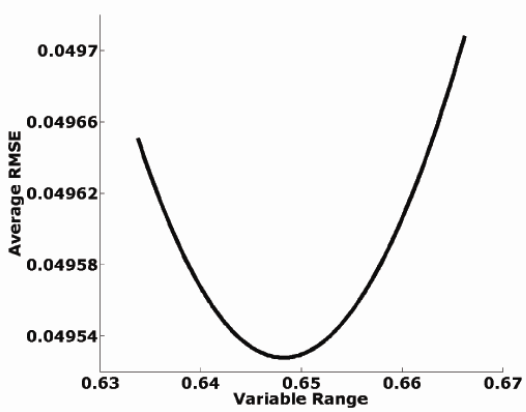


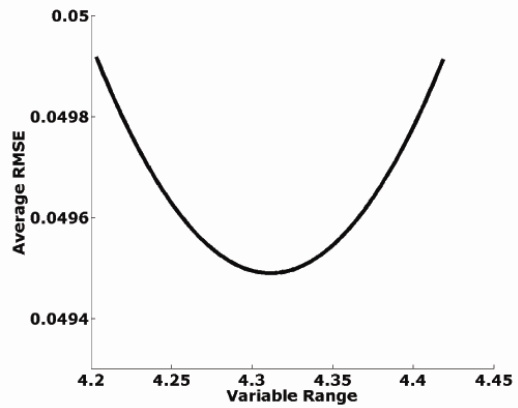
D



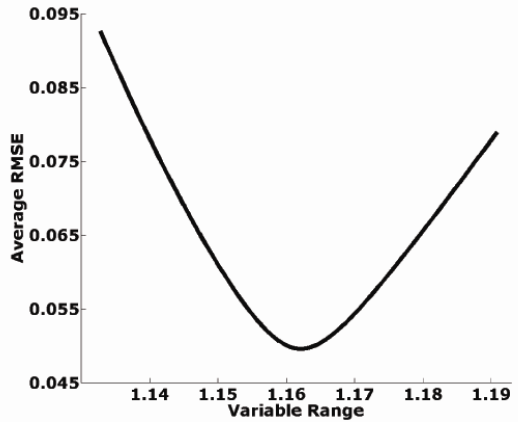
E

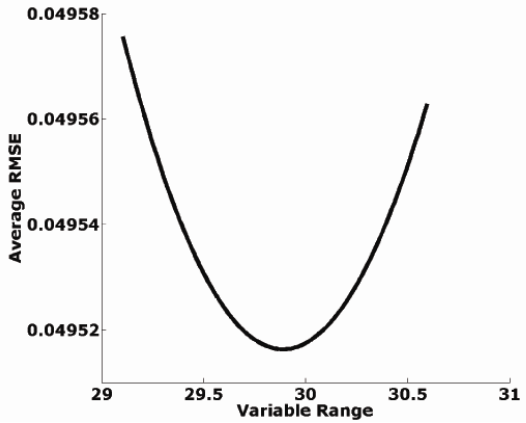

Average RMSE
0.14
0.12
0.1
0.08
0.06
0.04
1.82
1.84
1.86
1.88
1.9
1.92
1.94
Variable Range

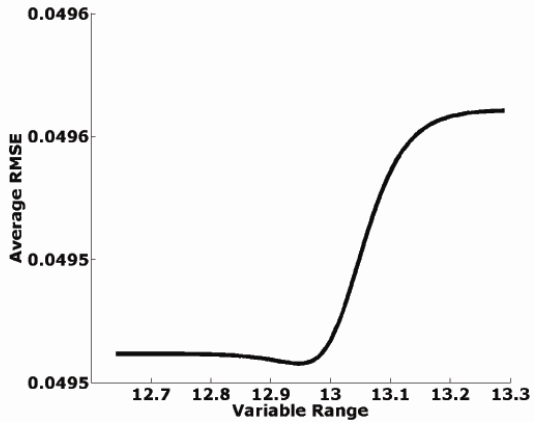

Average RMSE
0.0497
0.04966
0.04962
0.04958
0.04954
0.63
0.64
0.65
0.66
0.67
Variable Range


Average RMSE
0.05
0.0498
0.0496
0.0494
4.2
4.25
4.3
4.35
4.4
4.45
Variable Range

F


Average RMSE
0.095
0.085
0.075
0.065
0.055
0.045
1.14
1.15
1.16
1.17
1.18
1.19
Variable Range


Average RMSE
0.04958
0.04956
0.04954
0.04952
29
29.5
30
30.5
31
Variable Range


Average RMSE
0.0496
0.0496
0.0495
0.0495
12.7
12.8
12.9
13
13.1
13.2
13.3
Variable Range

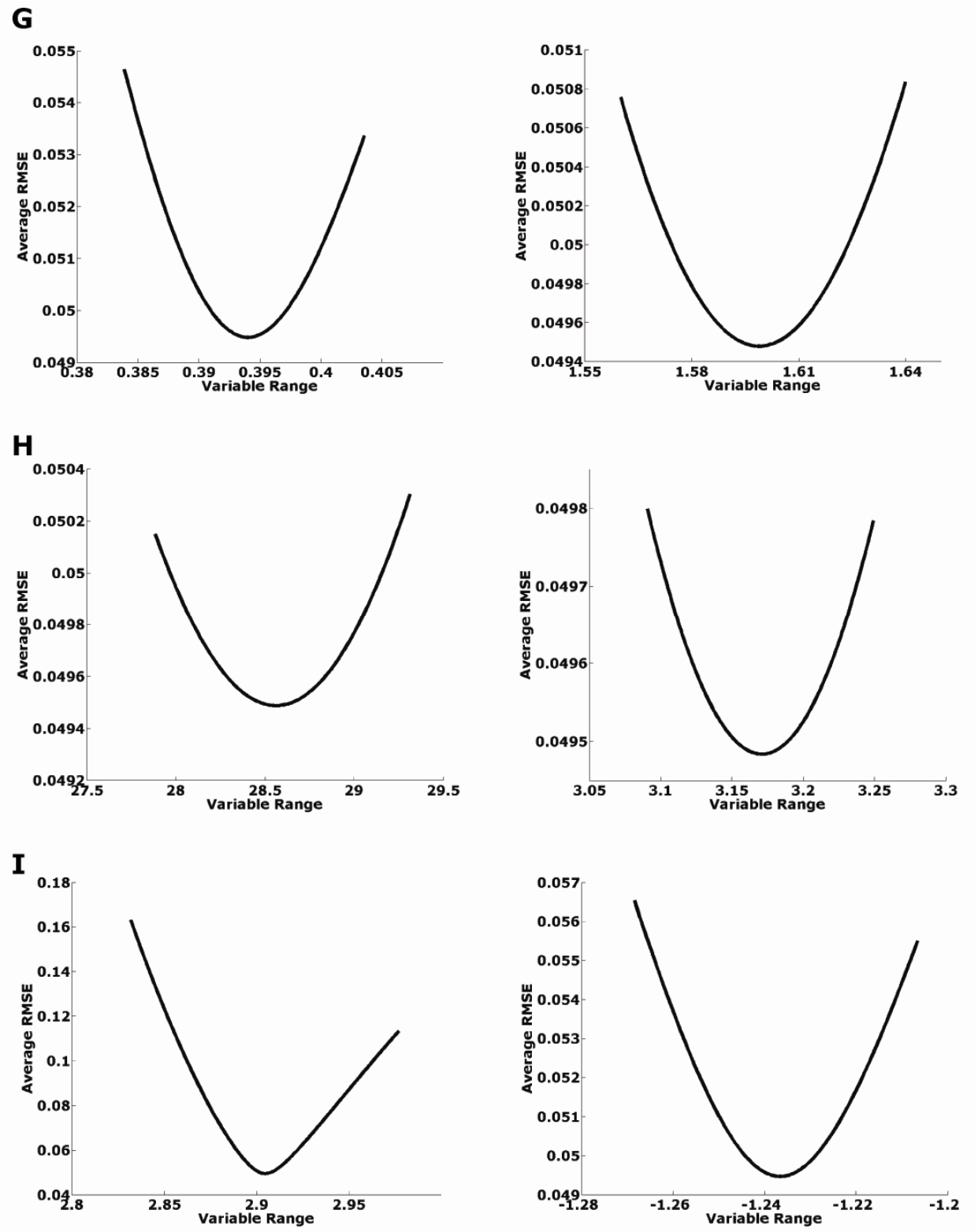
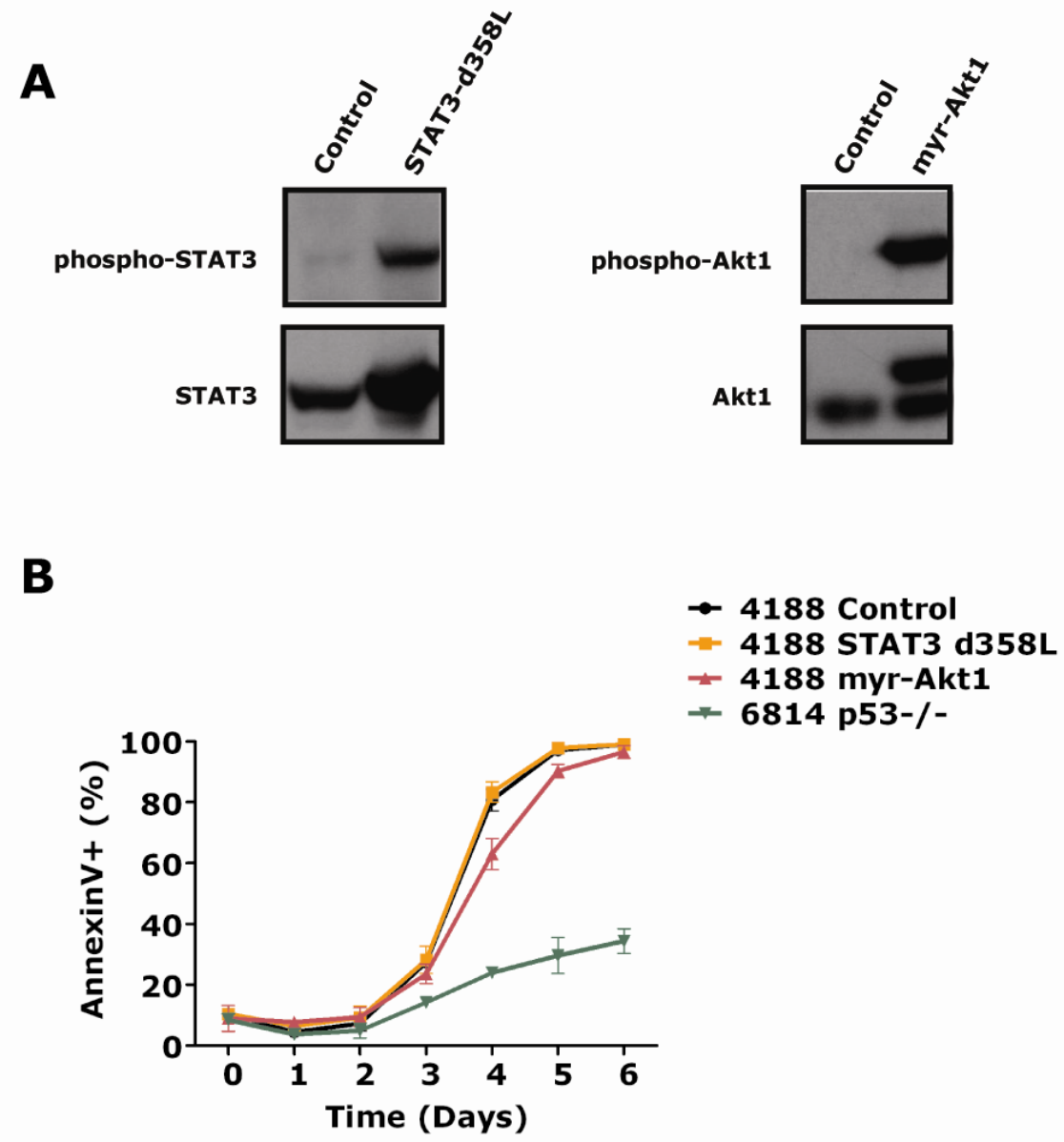

G
Average RMSE
Variable Range
0.055
0.054
0.053
0.052
0.051
0.05
0.049
0.38
0.385
0.39
0.395
0.4
0.405
0.051
0.0508
0.0506
0.0504
0.0502
0.05
0.0498
0.0496
0.0494
1.55
1.58
1.61
1.64
H
0.0504
0.0502
0.05
0.0498
0.0496
0.0494
0.0492
27.5
28
28.5
29
29.5
0.0498
0.0497
0.0496
0.0495
3.05
3.1
3.15
3.2
3.25
3.3
I
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
2.8
2.85
2.9
2.95
0.057
0.056
0.055
0.054
0.053
0.052
0.051
0.05
0.049
-1.28
-1.26
-1.24
-1.22
-1.2

Fig. S3. Validation of the ODE mathematical model for oncogene addicted tumor behavior. Hold-out validation consistently demonstrates differential attenuation of aggregate survival and death signal curves. (A-C) We fixed all parameters, except for a, b, c, d, f, and g, obtained from previous model fitting using all 3 datasets [volumetric measurements (*V*), Ki-67 (*PI*) and cleaved capase3 (*AI*) IHC data). Using different combinations of two of the three datasets, we were able to predict what the signal curves would look like. (A) Apoptosis index (*AI*, cleaved caspase3) and volumetric measurements (*V*) were model fitted to predict the signal curves. (B) Proliferation index (*PI*, Ki-67) and volume data (*V*) were used to predict the signals curves. (C) *PI* and *AI* were used to predict the signal curves. Signal curves predicted from all three combinations are qualitatively concordant and show differential attenuation of the two signals, indicating that our mathematical model is robust. (D) Boostrap resampling of variables used in the mathematical model confirm reproducibility. The bootstrap observational unit is the individual tumor and we used 10,000 bootstrap samples. The results of bootstrapping analysis demonstrated that the proposed variables of our mathematical model are robust to the variation in the data itself and that all parameters are reproducible (within 5% error). Variable names: *a*-starting intensity of the aggregate survival signal; *b*-rate of survival signal decay; *c*-time when survival signals begin dropping off; d-starting intensity of the death signal; *f*-rate of death signal decay; *g*-time when death signals begin dropping off; T_p-duration of intracellular proliferation; T_a-duration of intracellular apoptosis; t_p-duration of proliferation detectable by IHC; t_a-duration of apoptosis detectable by IHC; and *m,n*-thresholds that determine cellular program. (E-I) Sensitivity analysis of the estimated parameters from the Levenberg-Marquardt optimization method demonstrate the precision

of the estimated values. The results showed that all values in our estimated set are at the minimum of the assessed range. (E) All parameters (*a*-starting intensity of the aggregate survival signal; *b*-rate of survival signal decay; *c*-time when survival signals begin dropping off) for survival signals, and (F) similar parameters for aggregate death signals (*d*, *f* & *g*). (G) Durations for intracellular proliferation (T_p, top panel) and apoptosis (T_a, bottom panel). (H) Durations for which proliferation (t_p, top panel) and apoptosis (t_a, bottom panel) can be detected by IHC, and (I) thresholds *m* (top) and *n* (bottom) that determine cellular states.

Fig. S4


A
Control
STAT3-d358L
phospho-STAT3
STAT3
Control
myr-Akt1
phospho-Akt1
Akt1
B
4188 Control
4188 STAT3 d358L
4188 myr-Akt1
6814 p53-/-
AnnexinV+ (%)
100
80
60
40
20
0
0
1
2
3
4
5
6
Time (Days)

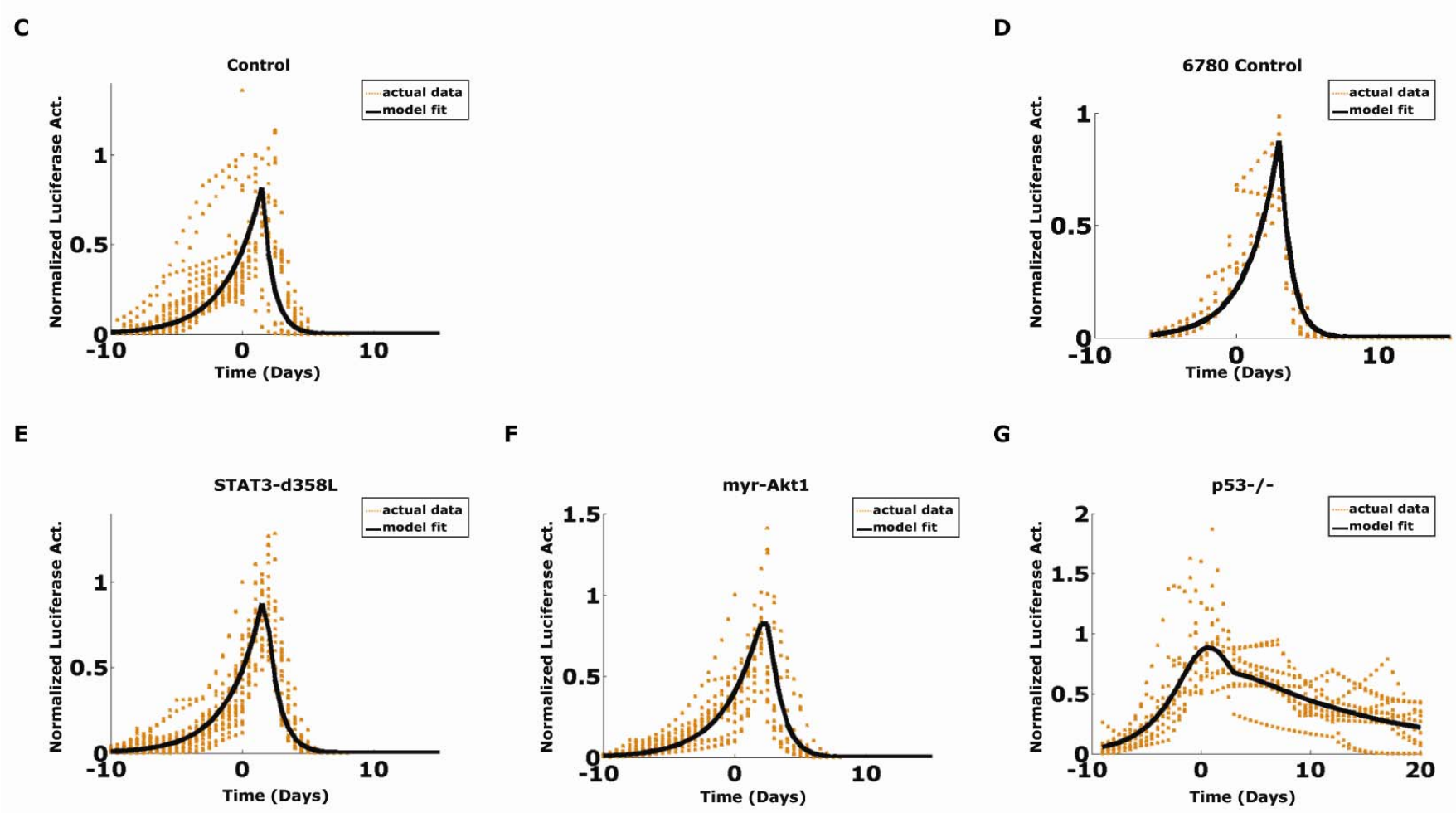

C
Control
actual data
model fit
Normalized Luciferase Act.
Time (Days)
D
6780 Control
actual data
model fit
Normalized Luciferase Act.
Time (Days)
E
STAT3-d358L
actual data
model fit
Normalized Luciferase Act.
Time (Days)
F
myr-Akt1
actual data
model fit
Normalized Luciferase Act.
Time (Days)
G
p53-/-
actual data
model fit
Normalized Luciferase Act.
Time (Days)

Fig. S4. Genetic perturbation of pro-survival and pro-death pathways in *MYC*-induced lymphomas can impede tumor regression following *MYC* inactivation *in vivo*. (A) *MYC*-induced lymphomas were transduced with constructs to express mutant active forms of Stat3-d358L and myr-Akt1. The resultant lines were checked for expression of Stat3-d358L and myr-Akt1 *via* phospho-western. (B) A FACS time course using Annexin V staining was performed on the cell lines from (A) and a *MYC*-induced lymphoma derived from a p53-/- background (6814) following *MYC* inactivation *in vitro*. The myr-Akt1 and p53-/- *MYC*-induced lymphomas demonstrate decreased apoptosis beginning on day 3 and at later time points following *MYC* inactivation in comparison to the control and Stat3-d358L lymphomas. (C-G) Mathematical modeling of *MYC*-induced lymphomas with bioluminescence imaging (BLI) data demonstrated that genetic perturbations of pro-survival and pro-death pathways can impede tumor regression following *MYC* inactivation *in vivo*. The kinetics of tumor regression following oncogene inactivation was examined in *MYC*-induced lymphomas in the different mutant backgrounds as shown (E-G). Tumor cell lines labeled with luciferase were transduced with various constructs and then inoculated into syngeneic hosts generating tumors (numbers of tumors in parentheses). *MYC*-induced tumors were transduced with (C) an empty Control construct (21 tumors); (E) a Stat3-d358L construct (24 tumors); or (F) myr-Akt1 construct (16 tumors). Also, (D) a non-transduced but luciferase labeled tumor line (5 tumors) was used as a control for (G) a p53-/- *MYC*-induced lymphoma that was labeled with luciferase (10 tumors). Ten million cells of each variant (E-G) were subcutaneously injected into FVBN mice. Tumor growth and regression were measured by BLI before and after *MYC* inactivation. Our mathematical model for *K-ras*^{G12D} -induced lung

tumors adapted for BLI (see Supporting Materials and Methods) was able to fit the imaging data: (C) RMSE 12.6%; (D) RMSE 7.0%; (E) RMSE 5.9%; (F) RMSE 5.6%; and (G) RMSE 10.4%.

Fig. S5

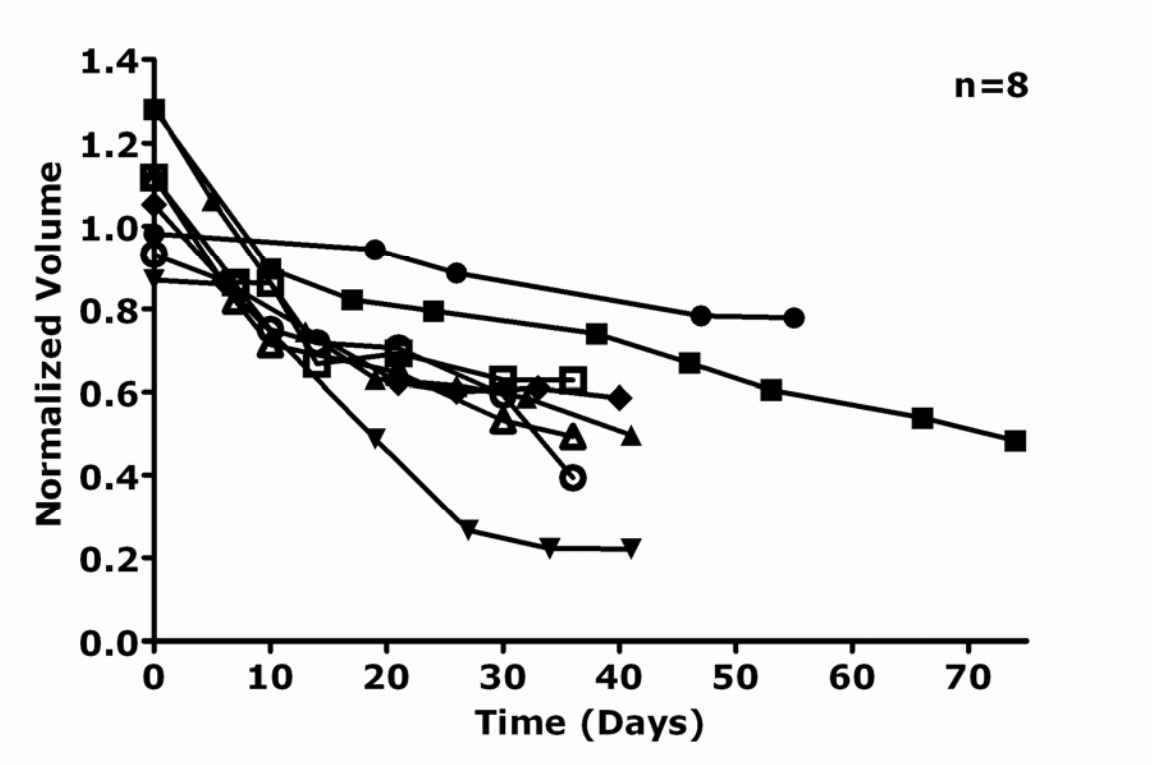
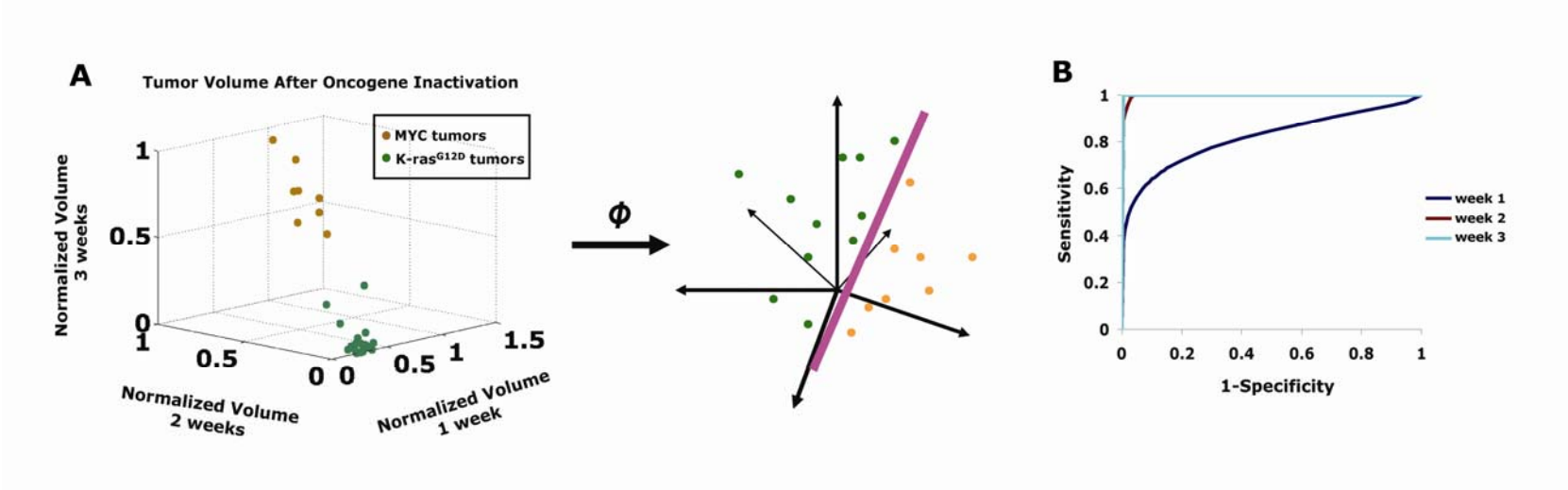


Fig. S5. *MYC*-induced lung tumors are not oncogene addicted. Volumetric quantization of *MYC*-induced lung tumors following oncogene inactivation using serial microCT with normalization as performed for fig. S2 (described in Supporting Materials and Methods). *MYC*-induced murine lung tumors are oncogene-independent, despite > 6-10 weeks of *MYC* inactivation CM mice still harbor gross tumors (n=8).

Fig. S6


A
Tumor Volume After Oncogene Inactivation
MYC tumors
K-ras^{G12D} tumors
Normalized Volume 3 weeks
Normalized Volume 2 weeks
Normalized Volume 1 week
ϕ
B
Sensitivity
1-Specificity
week 1
week 2
week 3

C

After 1 week of inactivation

Predicted	K-ras (%)	MYC (%)
K-ras	19 (0.950)	2 (0.250)
MYC	1 (0.050)	6 (0.750)

After 2 weeks of inactivation

Predicted	K-ras (%)	MYC (%)
K-ras	20 (1.00)	1 (0.125)
MYC	0 (0.00)	7 (0.875)

After 3 weeks of inactivation

Predicted	K-ras (%)	MYC (%)
K-ras	20 (1.00)	0 (0.00)
MYC	0 (0.00)	8 (1.00)

D

After 1 week of inactivation

Predicted	K-ras (%)	Non-K-ras (%)
K-ras	19 (0.950)	6 (0.429)
Non-K-ras	1 (0.050)	8 (0.571)

After 2 weeks of inactivation

Predicted	K-ras (%)	Non-K-ras (%)
K-ras	19 (0.950)	2 (0.143)
Non-K-ras	1 (0.050)	12 (0.857)

After 3 weeks of inactivation

Predicted	K-ras (%)	Non-K-ras (%)
K-ras	20 (1.00)	2 (0.143)
Non-K-ras	0 (0.00)	12 (0.857)

After 4 weeks of inactivation

Predicted	K-ras (%)	Non-K-ras (%)
K-ras	20 (1.00)	1 (0.071)
Non-K-ras	0 (0.00)	13 (0.929)

Fig. S6. Support vector machine (SVM) trained with quantified imaging data can be used to distinguish *K-ras*^{G12D} and *MYC*-induced lung tumors. (A) An illustration of the SVM shows that the technique maps the original dataset in 3-D to a higher dimensional space where a maximal separating hyperplane is constructed that best separates the data points between two different genotypes, *K-ras*^{G12D} and *MYC*, for prediction. (B) Receiver operating characteristics (ROC) curves were used to show the accuracy of the SVM technique in predicting the oncogene dependent genotypes based on tumor volumes obtained from different length of time after oncogene targeted therapy. (C-D) Sensitivity and specificity tables for predicting genotype using support vector machine (SVM) trained with serial imaging data from *K-ras*^{G12D} and non-*K-ras*^{G12D} tumors following oncogene addiction. Confusion matrices show the sensitivity and specificity of the SVM technique in predicting the oncogene-dependent genotypes based on the number of tumor volumes measured after oncogene inactivation. The matrices are trained using tumors following simulated targeted therapy from (C) *K-ras*^{G12D} - and *MYC*-induced lung tumors or from (D) *K-ras*^{G12D} - and non-*K-ras*^{G12D} (*MYC* and *MYC*/*K-ras*^{G12D} double animals)-induced lung tumors.

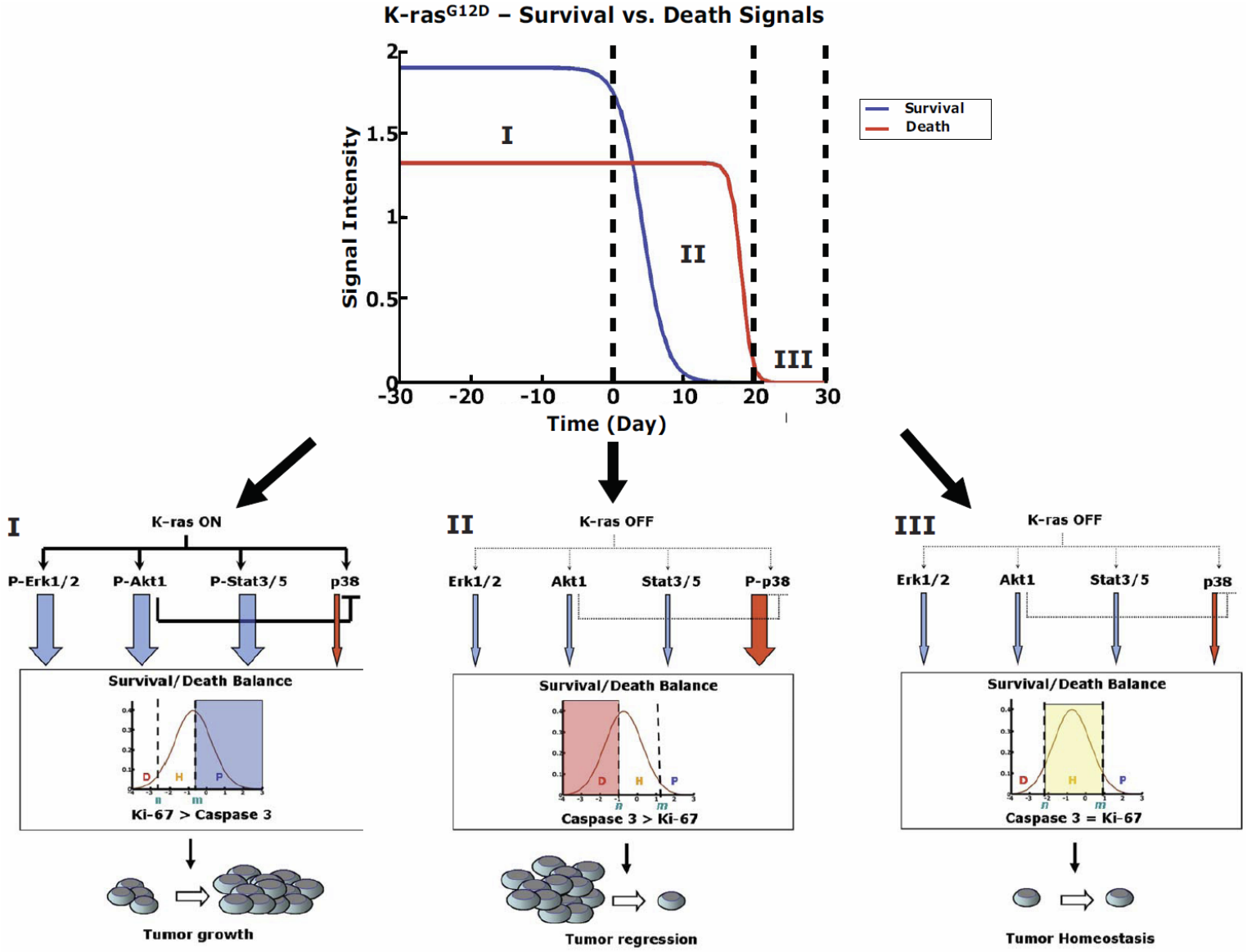

K-ras^{G12D} – Survival vs. Death Signals
Signal Intensity
Time (Day)
Survival
Death
I
II
III
K-ras ON
P-Erk1/2
P-Akt1
P-Stat3/5
p38
Survival/Death Balance
Ki-67 > Caspase 3
Tumor growth
K-ras OFF
Erk1/2
Akt1
Stat3/5
P-p38
Survival/Death Balance
Caspase 3 > Ki-67
Tumor regression
K-ras OFF
Erk1/2
Akt1
Stat3/5
p38
Survival/Death Balance
Caspase 3 = Ki-67
Tumor Homeostasis

Fig. S7. Schematic integrating the temporal phospho-IHC in relation to *K-ras*^{G12D} activation state. The balance of aggregate survival and death signaling determines whether *K-ras*^{G12D}–induced lung tumors grow, regress or become static. (I) *K-ras*^{G12D} signals to Erk1/2, Akt1 and Stat3/5 keeping them phosphorylated denoted by "P-" and width of downstream blue arrows. The pro-death factor p38 is also activated by a signal from *K-ras*^{G12D}, but negative counter-regulation by phosphos-Akt1 de-phosphorylates p38 (as denoted by a thin downstream red arrow). The integration of aggregate survival and death signals results in the shift of thresholds "m" and "n" in our mathematical model (Fig. 4) resulting in a net preponderance of cells proliferating and therefore net tumor growth. (II) When *K-ras*^{G12D} is inactivated Erk1/2, Akt1 and Stat3/5 are quickly de-phosphorylated and negative counter regulation of p38 is released, resulting in transient phosphos-p38 (denoted by a wider downstream red arrow). Similarly, integration of these aggregate survival and death signals results in the shift of thresholds "m" and "n" in our mathematical model (Fig. 4) resulting in a net preponderance of cells dying and therefore net tumor regression. This model is consistent with the delayed attenuation of pro-death factors such as phospho-p38 [larger red arrow as compared to red arrow in panel (I)]. (III) Following a longer period of *K-ras*^{G12D} inactivation the aggregate pro-death effectors have decreased as exemplified by p38 (denoted by a thin downstream red arrow) and the net balance is in favor of homeostatic states such as quiescence or differentiation. Width of the blue and red downstream arrows below intracellular mediators is representative of the degree of phospho-IHC signaling observed in Fig. 3A.

Supplementary References

S1. S. W. Lowe, E. Cepero, G. Evan, *Nature* **432**, 307 (Nov 18, 2004).

S2. W.H. Press *et al.,* Numerical Recipes in C: The Art of Scientific Computing (*2nd ed*). Cambridge University Press (2002).

S3. R.G. Renehan, C. Booth, C. S. Potten, *BMJ* **322**, 1536 (2001).

S4. S. Beer *et al.*, *PLoS biology* **2**, e332 (Nov, 2004).

S5. P. A. Yushkevich *et al.*, *Neuroimage* **31**, 1116 (Jul 1, 2006).

S6. D. D. Cody *et al.*, *Invest Radiol* **40**, 263 (May, 2005).

S7. S. Giuriato *et al.*, *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16266 (Oct 31, 2006).

S8. A.N. Basavanhally *et. al. Proceedings of IEEE ISBI* p229 (2010).

S9. D.F.Yankelevitz *et al*., Cancer **97,** 1271 (2003).

S10. R.M. Lindell, *et al*., Radiology **242,** 555 (2007).

S11. S.Y. El Sharouni, et. al. Br J Cancer **89,** 2184 (2003).

Supplementary Videos: (A) Regression of *K-ras*^{G12D} -induced lung tumors imaged serially using microCT. Movie of CR lung tumors before inactivation of *K-ras*^{G12D} by withdrawal of doxycycline is representative of 11 CR lung tumor bearing mice serially scanned by microCT as described in the Supporting Materials and Methods (AVI; 1.7 MB). (B) Regression of *K-ras*^{G12D} -induced lung tumors imaged serially using microCT. Movie of CR lung tumors after inactivation of *K-ras*^{G12D} by withdrawal of doxycycline is representative of 11 CR lung tumor bearing mice serially scanned by microCT as described in the Supporting Materials and Methods (AVI; 1.3 MB).