

## Supplementary Materials

### Supplementary Table 1

We conducted a sensitivity analysis including only the subset of studies with a high degree of compliance. To determine compliance, for each trial, we first calculated the percentage of its sample size included in our original analysis (see the CONSORT Diagram in Figure 1 of our manuscript). The average percentage across all studies was 23%. We then defined a trial as “compliant” if its percentage was at least 30%. Six trials were identified as compliant based on this threshold. The predictive ability measures – specifically the c-indices and H-L statistics – are similar to those we obtained in the original analysis using all 13 trials, summarized in the tables below.

All Trials						
	Breast		NSCLC		Colon	
	Slope (N=189)	% Change (N=182)	Slope (N=746)	% Change (N=734)	Slope (N=322)	% Change (N=320)
c-index (95% CI)*	0.58 (0.53-0.65)	0.55 (0.52-0.63)	0.58 (0.55-0.61)	0.57 (0.55-0.61)	0.62 (0.58-0.68)	0.64 (0.59-0.69)
RECIST	0.52		0.57		0.60	
H-L	0.05	0.04	0.03	0.01	0.12	0.05
RECIST	0.05		0.01		0.04	

Compliant Trials						
	Breast		NSCLC		Colon	
	Slope (N=107)	% Change (N=101)	Slope (N=715)	% Change (N=703)	Slope (N=296)	% Change (N=294)
c-index (95% CI)*	0.61 (0.53-0.69)	0.57 (0.49-0.65)	0.57 (0.54-0.60)	0.57 (0.54-0.60)	0.63 (0.57-0.69)	0.64 (0.58-0.70)
RECIST	0.54		0.59		0.60	
H-L	0.04	0.03	0.03	0.02	0.13	0.04
RECIST	0.12		0.01		0.06	

\* 95% CI for c-indices are bootstrap confidence intervals.

**Supplementary Figure 1.** Scatterplots of 6-12 week vs. 0-6 week slope and % changes. Size of plotting symbol proportional to the average baseline sum.

