ON-LINE APPENDIX

Transfer Learning Theory and Methodology

To briefly explain the mathematical details of our TL algorithm, assume N number of patients in a training dataset and a linear model between imaging features and cell density for patient k (ie, $y_k = X_k w_k + \varepsilon_k, k = 1, \ldots, N$). y_k consists of TCD measurements for n_k biopsy samples. X_k consists of MR imaging features for biopsy samples. w_k consists of model coefficients yet to be estimated. ε_k consists of random errors following a Gaussian distribution. To make the Gaussian distribution appropriate, we transformed the original TCD measurement (range, 0–1) using a sigmoid function. Furthermore, to couple models from different patients, we adopted a Bayesian framework[1] and assumed that the patient-specific model coefficients, $W = (w_1, \ldots w_K)$, share the same prior distribution, ie,

1)

$$p(W|\Omega, \Phi, b) \sim \Pi_{k=1}^{K} Laplace\,(w_k; b) \times MN(W; 0, \Omega, I).$$

Laplace $(w_k; b)$ is a Laplace distribution to facilitate sparsity in model estimation (ie, to produce a parsimonious model for better interpretability).[2] $MN(W; 0, \Omega, I)$ is a zero-mean matrix-variate normal distribution. Specifically, the covariance matrix, Ω, encodes the correlation among different patients.

Furthermore, given the prior distribution in Equation 1 and the likelihood based on the training data, $p(y_k|X_k W_k) \sim N(y_k; X_k W_k, \Sigma^2 I)$, we can obtain the posterior distribution of W as

2)

$$p(W|\{y_k, X_k\}_{k=1}^{K}, \Omega, \Phi, b) \sim p(W|\Omega\ \Phi, b)\Pi_{k=1}^{N} p(y_k|X_k, w_k).$$

Then, the maximum a posteriori estimator for W can be obtained by solving the following optimization problem:

3) $$\hat{W}, \hat{\Omega} = \underset{W,\Omega}{\operatorname{argmin}}\{\Sigma_{k=1}^{N}\|y_k - X_k w_k\|_2^2 + \lambda_1\|W\|_1 + \lambda_2(Q\log|\Omega| + tr(W\Omega^{-1}W^T))\}.$$

Here, $\|\cdot\|_2$ and $\|\cdot\|_1$ denote the L1 and L2 norms, respectively. $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are 2 regularization parameters to control the sparsity and the amount of knowledge transferred between the models of different patients, respectively. Equation 3 is a TL model in the sense that it allows a joint estimation of patient-specific model coefficients $w_k, k = 1, \ldots, N$. The most appealing part of the TL model in Equation 3 is that it does not require a prespecification on the correlation among patients, Ω, but can estimate it in a data-driven manner. To solve the optimization problem in Equation 3 (ie, to estimate W and Ω), we adopted an efficient alternating algorithm that estimates W and Ω iteratively.[1,3,4] That is, given Ω, the optimization problem with respect to W is convex and is solved using the accelerated gradient algorithm.[5] Given W, Ω can be solved analytically. This iterative algorithm is guaranteed to converge.

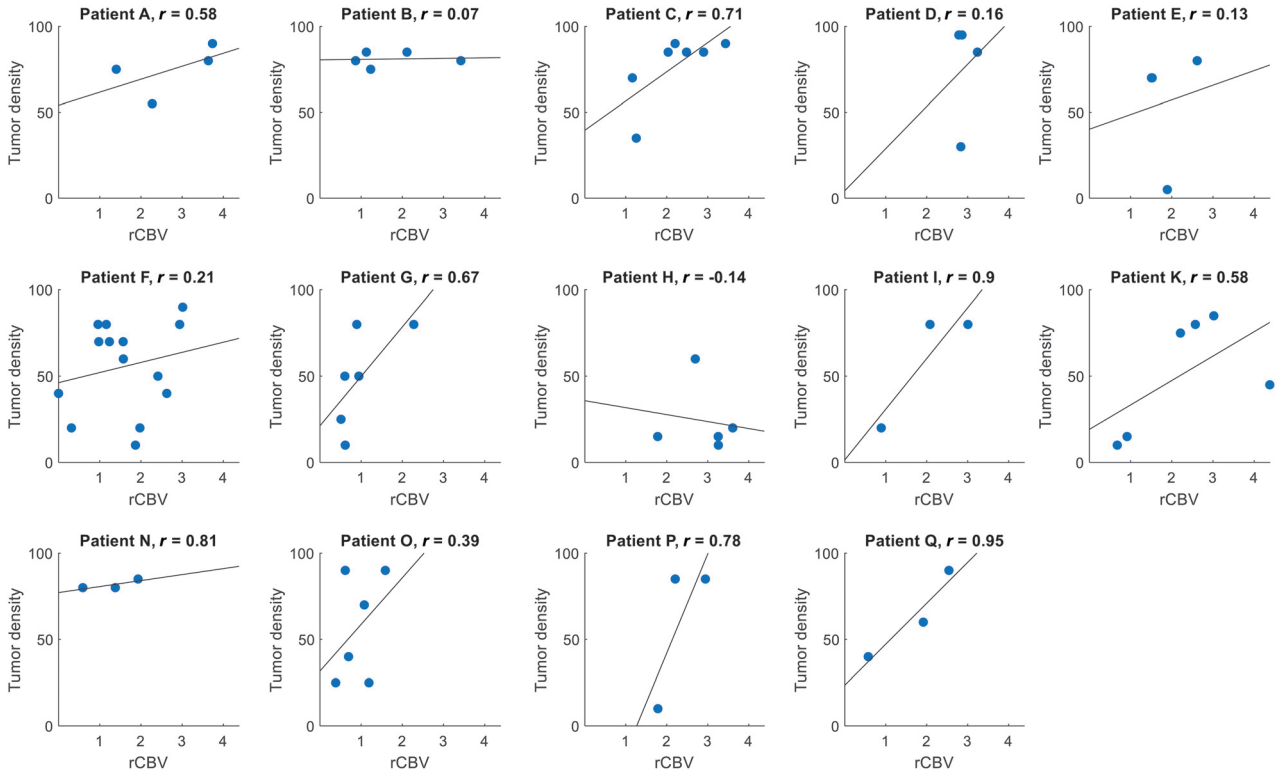
REFERENCES

1. Zou N, Baydogan M, Zhu Y, et al. **A transfer learning approach for predictive modeling of degenerate biological systems.** *Technometrics* 2015;57:362–73 CrossRef Medline
2. Tibshirani R. **Regression shrinkage and selection via the lasso**. *Journal of the Royal Statistical Society Series B (Methodological)* 1996;58: 267–88 CrossRef
3. Idier, J, ed. *Bayesian Approach to Inverse Problems.* New York: John Wiley & Sons; 2013
4. Zhang Y, Yeung DY. A Convex Formulation for Learning Task Relationships in Multi-Task Learning. *arXiv preprint* arXiv:1203.3536; 2012
5. Liu, J, Ji S, Ye J. **Multi-task feature learning via efficient L2, 1-norm minimization**. In: *Proceedings of the Twenty-Fifth Conference on Uncertain Artificial Intelligence,* Montreal, Quebec, Canada. June 18–21, 2009:339–48

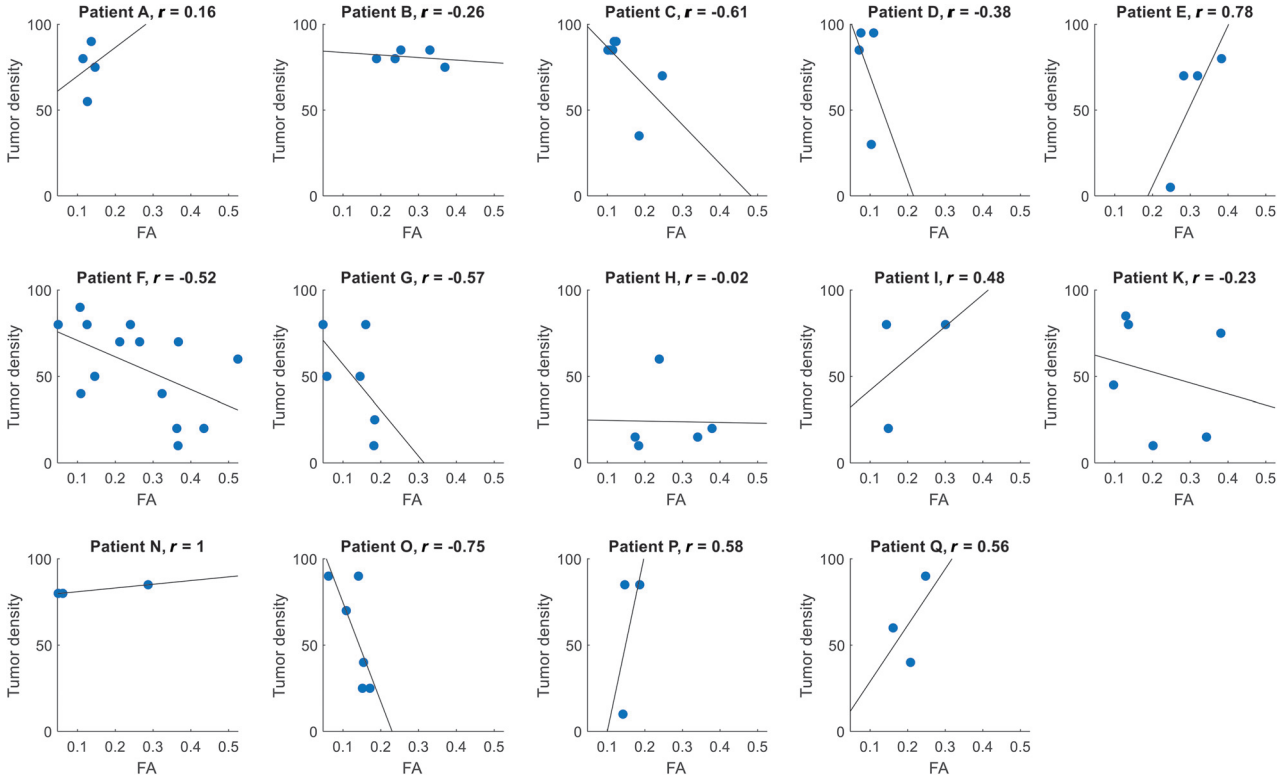
On-line Table: Pearson correlation coefficients from univariate analysis separately comparing the 6 MRI features with TCD for all samples and subgroups of only nonenhancing versus only enhancing biopsy samples[a]

	T1 + C	T2WI	rCBV	EPI + C	FA	MD
All samples (n = 82)	0.36 (<.001)	0.13 (.25)	0.33 (<.001)	−0.02 (.85)	−0.24 (.03)	0.03 (.79)
Enhancing only (n = 49)	0.18 (.22)	0.17 (.24)	0.26 (.07)	0.03 (.82)	−0.31 (.03)	0.04 (.79)
Nonenhancing only (n = 33)	−0.05 (.77)	0.00 (.99)	0.21 (.24)	0.03 (.87)	−0.18 (.32)	−0.01 (.97)

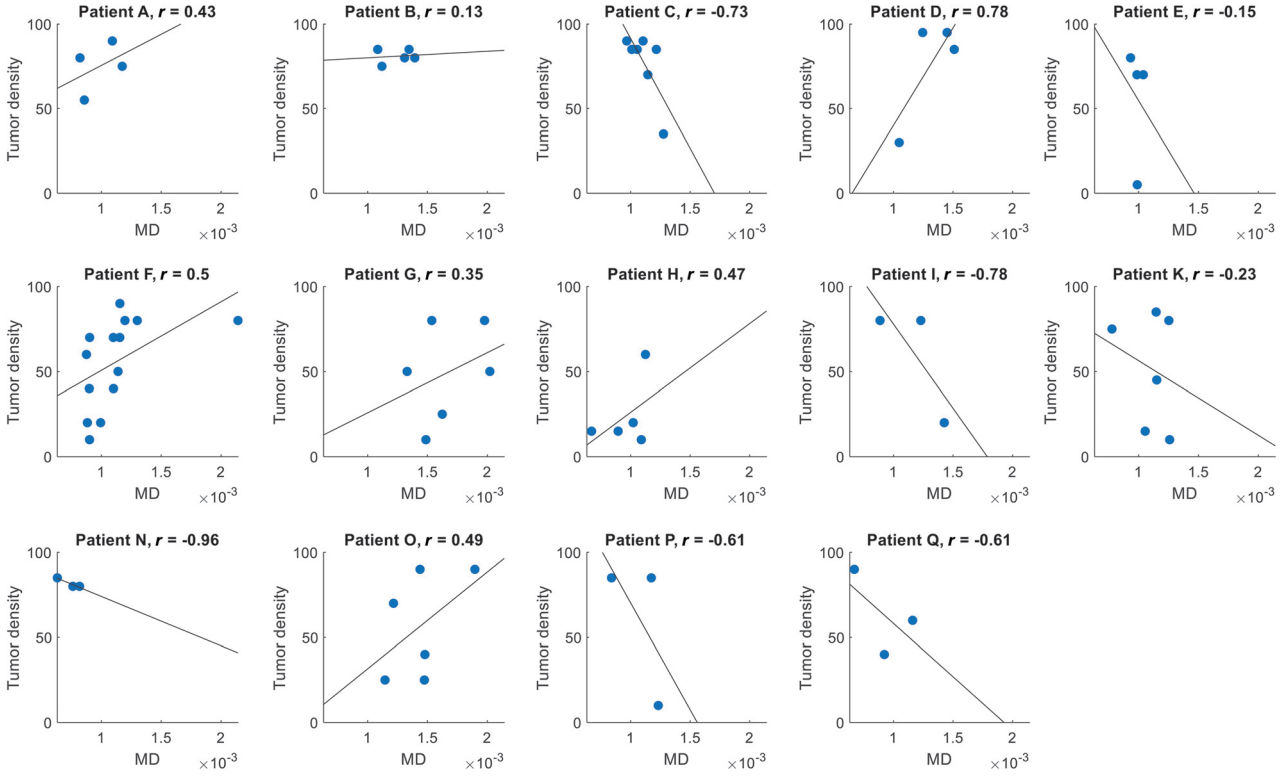
[a] *P* values are shown in parentheses.



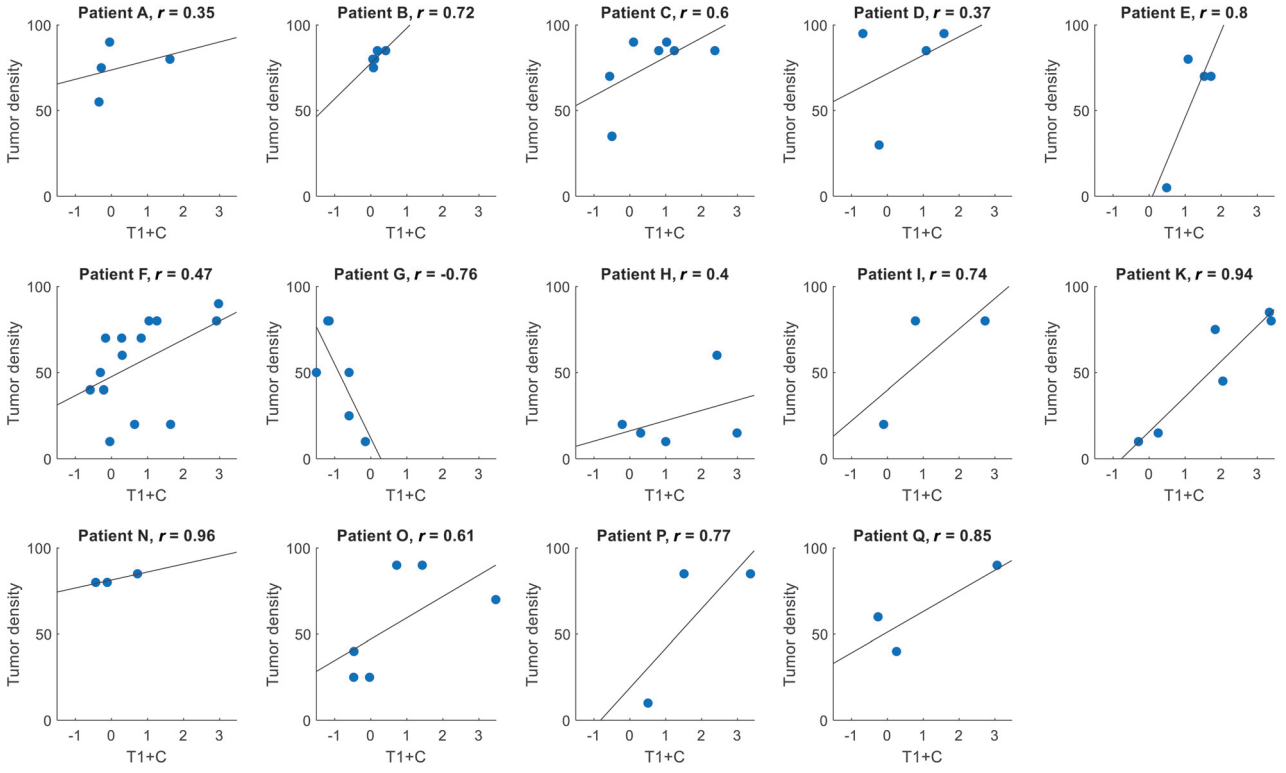
ON-LINE FIG 1. Individual patient (*A*–*Q*) plots of relative cerebral blood volume versus tumor cell density using spatially matched image-localized biopsies. We generated scatterplots for the 14 patients with primary GBM in our cohort who underwent at least 3 image-localized biopsies from their initial operation. There were 4 patients (not shown) who had only 2 image-localized biopsies each who were thus excluded from this analysis. The scatterplots for each patient consist of only that patient's histologic and MR imaging data and show the Pearson correlation coefficients for rCBV (x-axis) versus actual tumor cell density (y-axis) from corresponding spatially matched biopsies. Nearly all (13 of 14) patients demonstrated positive correlations so that rCBV increased with higher tumor cell density, though the strength of positive correlations varied from patient to patient.



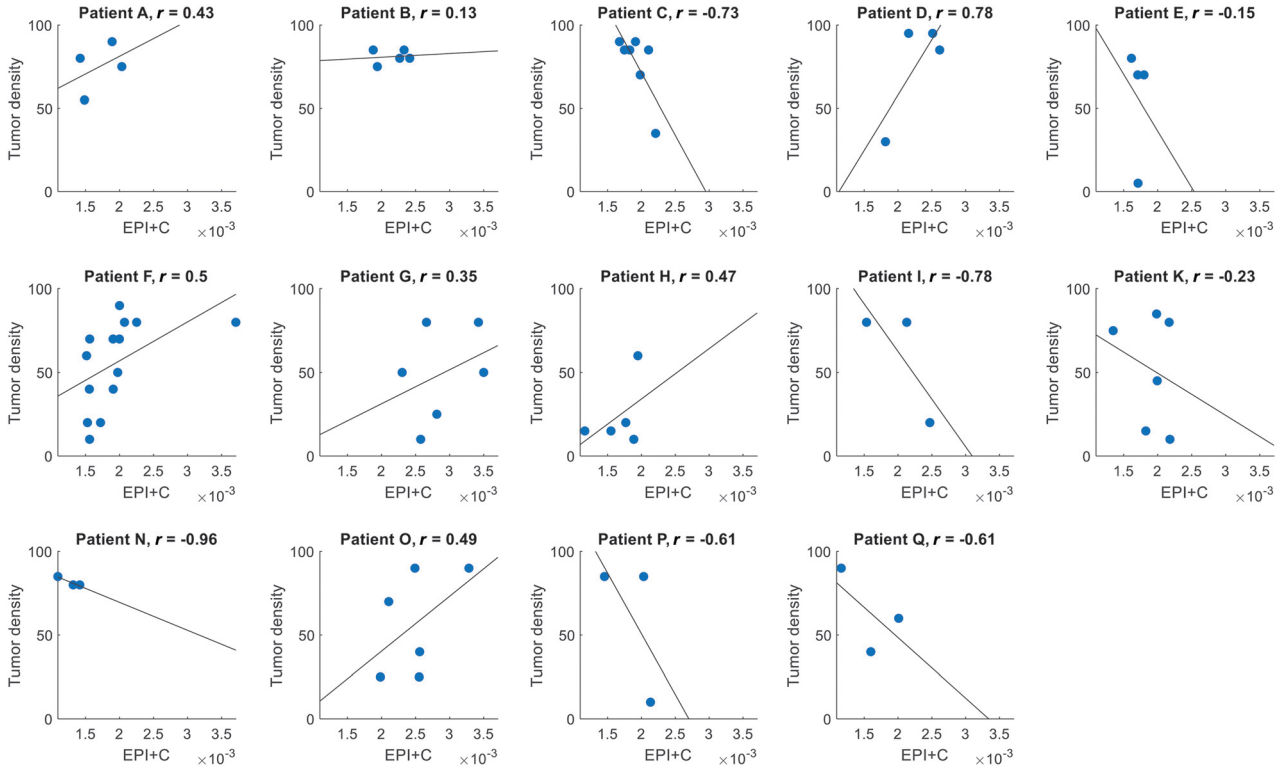
ON-LINE FIG 2. Individual patient (*A–Q*) plots of fractional anisotropy versus tumor cell density using spatially matched image-localized biopsies. We generated scatterplots for the 14 patients with primary GBM in our cohort who underwent at least 3 image-localized biopsies from their initial operation. There were 4 patients (not shown) who had only 2 image-localized biopsies each who were thus excluded from this analysis. The scatterplots for each patient consist of only that patient's histologic and MR imaging data and show the Pearson correlation coefficients for FA (x-axis) versus actual tumor cell density (y-axis) from corresponding spatially matched biopsies. Compared with rCBV plots in On-line Fig 1, the direction of FA correlations shows greater variability across patients, with only 57.1% (8/14) of patients having negative correlations with TCD (versus 42.9% with positive correlations).



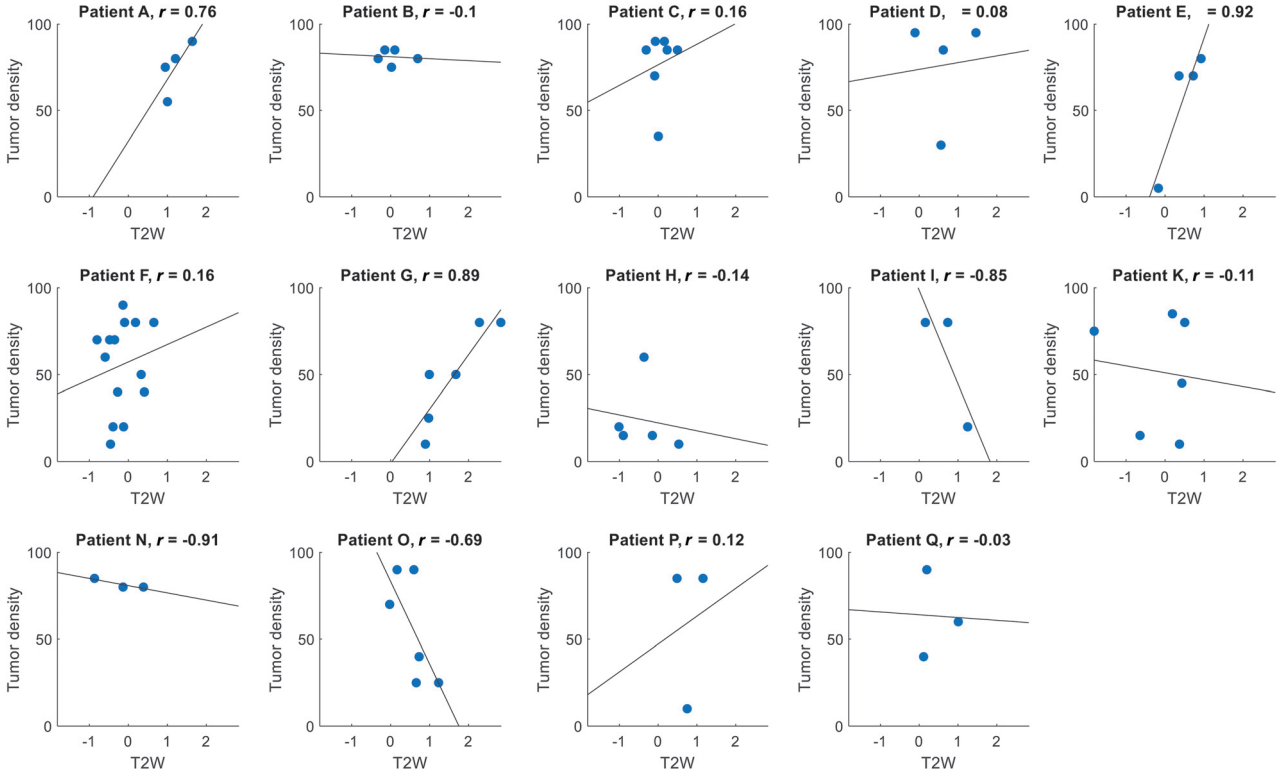
ON-LINE FIG 3. Individual patient (*A–Q*) plots of MD versus tumor cell density using spatially matched image-localized biopsies. We generated scatterplots for the 14 patients with primary GBM in our cohort who underwent at least 3 image-localized biopsies from their initial operation. There were 4 patients (not shown) who had only 2 image-localized biopsies each who were thus excluded from this analysis. The scatterplots for each patient consist of only that patient's histologic and MR imaging data and show the Pearson correlation coefficients for MD (x-axis) versus actual tumor cell density (y-axis) from corresponding spatially matched biopsies. Compared with rCBV plots in On-line Fig 1, the direction of MD correlations shows greater variability across patients, with 50% of patients split between negative and positive correlations.



ON-LINE FIG 4. Individual patient (*A*–*Q*) plots of T1-weighted postcontrast signal (T1 + C) versus tumor cell density using spatially matched image-localized biopsies. We generated scatterplots for the 14 patients with primary GBM in our cohort who underwent at least 3 image-localized biopsies from their initial operation. There were 4 patients (not shown) who had only 2 image-localized biopsies each who were thus excluded from this analysis. The scatterplots for each patient consist of only that patient's histologic and MR imaging data and show the Pearson correlation coefficients for T1 + C (x-axis) versus actual tumor cell density (y-axis) from corresponding spatially matched biopsies. Similar to rCBV plots in On-line Fig 1, the direction of T1 + C correlations shows high consistency across patients, with positive correlation in 13/14 patients.



ON-LINE FIG 5. Individual patient (*A–Q*) plots of postcontrast T2*WI signal (EPI+C) versus tumor cell density using spatially matched image-localized biopsies. We generated scatterplots for the 14 patients with primary GBM in our cohort who underwent at least 3 image-localized biopsies from their initial operation. There were 4 patients (not shown) who had only 2 image-localized biopsies each who were thus excluded from this analysis. The scatterplots for each patient consist of only that patient's histologic and MR imaging data and show the Pearson correlation coefficients for EPI+C (x-axis) versus actual tumor cell density (y-axis) from corresponding spatially matched biopsies.



ON-LINE FIG 6. Individual patient (*A–Q*) plots of T2-weighted signal versus tumor cell density using spatially matched image-localized biopsies. We generated scatterplots for the 14 patients with primary GBM in our cohort who underwent at least 3 image-localized biopsies from their initial operation. There were 4 patients (not shown) who had only 2 image-localized biopsies each and were thus excluded from this analysis. The scatterplots for each patient consist of only that patient's histologic and MR imaging data and show the Pearson correlation coefficients for T2-weighted (x-axis) versus actual tumor cell density (y-axis) from corresponding spatially matched biopsies.