Appendix 1

List of investigators

Makoto Nishio, Shinji Atagi, Koichi Goto, Yukio Hosomi, Takashi Seto, Toyoaki Hida, Kazuhiko Nakagawa, Hiroshige Yoshioka, Naoyuki Nogami, Makoto Maemondo, Seisuke Nagase, Isamu Okamoto, Noboru Yamamoto, Masahiro Fukuoka, Nobuyuki Yamamoto, Kazuto Nishio.

Methods

Statistical analysis for gene polymorphisms

SNPs are coded by two dummy variables in the analysis, focusing on the minor allele (the less common allele). The first dummy variable takes a value of 1 if the patient has a SNP genotype with at least one copy of the minor allele (dominant effect of the minor allele) and the second takes value 1 if the subject has two copies of the minor allele (recessive effect of the minor allele). This coding allows for consideration of both dominant and recessive genetic effects in the model. If both dummy variables show significance, it means there is an additive effect.

MFPI approach

The MFPI analysis was performed using the following method by Royston and Sauerbrei (27):

Step 1. Let Z denote a continuous variable of the biomarker, and Z was transformed into Z^{p1} for the fractional polynomials-1 (FP1) model. The powers p_1 were chosen from a set, S = {p_1 =–2, –1, –0.5, 0, 0.5, 1, 2, 3}, where Z^{p1} denotes ln(Z) if p_1=0. In order to choose the best FP1 model, the best fitted p_1 was selected, while minimizing the model fit statistics for the likelihood ratio test based on the Cox proportional hazard model including the following covariates: treatment arm (0 or 1), Z^{p1} and the interaction between treatment arm and Z^{p1}.

Step 2. Z was transformed into Z^{p1} and Z^{p2} for the fractional polynomials-2 (FP2) model. The powers p_1 and p_2 were also chosen from a set, S = {–2, –1, –0.5, 0, 0.5, 1, 2, 3}. If p_i was 0, Z^{pi} denotes ln(Z), where i =1 or 2. Z^{p2} denotes Z^{p1} ln(Z) if p_1 = p_2. In order to choose the best FP2 model, the best combination of p_1 and p_2 was selected, while minimizing the model fit statistics for the likelihood ratio test based on the Cox proportional hazard model including the following covariates: treatment arm (0 or 1), Z^{p1}, Z^{p2} and two interaction terms between treatment arm and Z^{p1} or Z^{p2}.

Step 3. Determine which model is better between the best FP1 in Step 1 and the best FP2 in step 2 based on the comparison of the model fit statistics for the likelihood ratio test between the two models, using a χ^2 test with 3 degrees of freedom (df).

Step 4. The interaction P value was estimated based on the difference of the model fit statistics for the likelihood ratio test between the models with and without the interaction term(s) in Cox proportional hazard models, using a χ^2 test with 1 df if FP1 was selected, or a χ^2 test with 2 df if FP2 was selected.

Step 5. We also performed MFPI analysis to estimate the interaction P value when adjusting for the stratification factors (gender, disease stage, smoking history, and type of *EGFR* mutation). These factors were incorporated in the Cox proportional hazards model as categorical covariates along with the continuous biomarker covariates, and MFPI was conducted from step 1 to step 4 described above.

STEPP

STEPP methodology was used to visualize the interaction between bevacizumab plus erlotinib treatment and a continuous valuable biomarker (26). Two types of STEPP pattern have been proposed, named as sliding window STEPP (SW-STEPP) and TO-STEPP by Bonetti and Gelber (26). We selected TO-STEPP in the current study since it has been reported to be more stable than SW-STEPP (3). TO-STEPP was performed using the following method by Bonetti and Gelber (26):

Step 1. Let the subpopulations be defined with respect to a continuous biomarker value Z*, and let Z_i* be the value of such covariate for patient i. Considering a set of increasing values of Z* {z_1, z_2, ..., z_g}, with the exception of the duplicated values, we constructed an increasing collection of subpopulations P_l, l =1, 2, ..., g by including in P_l the patients for whom $Z_i^* \geq z_l$. Similarly, we constructed the subpopulations P_l, $l = g + 1$, ..., $2g–1$ by including in Pl the patients for whom $Z_i^* > z_{l-g}$. Let p

https://dx.doi.org/10.21037/tlcr-22-632

denote the number of subpopulations after excluding the subpopulations involving less than 30 patients.

Step 2. Let $\hat{\beta}_l$ and $\hat{\sigma}_l$ denote the estimated logarithm of the HR and standard error in the subpopulation P_l, (l=1, ..., p), respectively. We defined the 95% confidence band as $\left\{\beta_l* \in \hat{\beta}_l* \pm \gamma 1.96\sigma_l, l = 1,..., p\right\}$, where β_l* was the component of $\hat{\beta}_l$. The value of γ was estimated to meet the following equation by Monte Carlo simulation; $\mathrm{P}\left[\cap l = {}^{p} 1\left\{\beta_l* \in \hat{\beta}_l* \pm \gamma 1.96\sigma_l\right\}\right] = 0.95$.

Step 3. The treatment effects in each subpopulation and confidence band were plotted, where the horizontal axis showed the biomarker values, and the vertical axis showed the treatment effect (logarithm of HR). A lower value in the vertical axis denoted the better bevacizumab treatment effect.

Cutpoint for dichotomizing a continuous biomarker

To dichotomize a continuous biomarker, we estimated the cutpoint value for potential biomarkers referring to the methods by Jiang *et al.* as follows (29):

Step 1. Let c be any cutoff value of the biomarker. The likelihood ratio test statistic S(c) for treatment effect (HR for PFS in patients treated with bevacizumab and erlotinib combination therapy compared with E) was calculated in the subpopulation with biomarker value below or above c for all potential cutpoints by using a Cox proportional hazards model. The optimal cutpoint $\hat{c}$ was estimated as the one corresponding to the maximum S(c).

Step 2. Confidence interval of $\hat{c}$ was estimated by the bootstrap method. Let A_j (j =1, ..., 1,000) be each bootstrap sample from the observed data. As per the same methods mentioned above, cutpoint $\hat{c}_j$ was estimated in each A_j. The 95% confidence intervals of $\hat{c}$ were estimated based on the empirical distribution of $\hat{c}_j$.

Logistic regression analysis

Correlations between follistatin levels dichotomized at the cutpoint and baseline serum concentrations of angiogenesis-related proteins were evaluated by univariate logistic regression analysis. A Wald test was used to evaluate the statistical significance of coefficients in the model. The FDR was estimated using Benjamini-Hochberg methods (30).

References

53. Sauerbrei W, Royston P, Zapien K. Detecting an interaction between treatment and a continuous covariate: A comparison of two approaches. Comput Stat Data Anal 2007;51:4054-63.

https://dx.doi.org/10.21037/tlcr-22-632

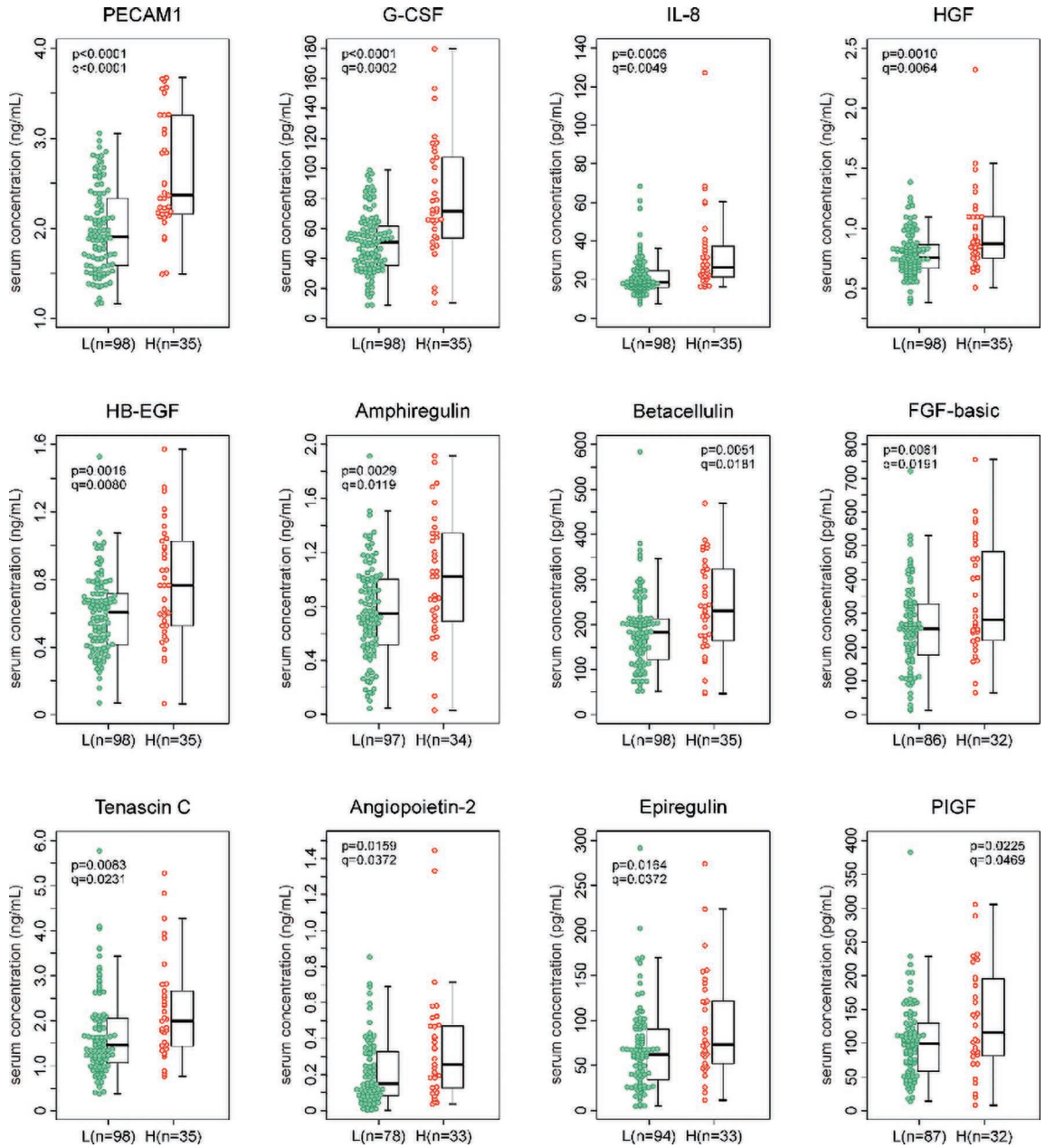

Figure S1 Distributions of serum angiogenesis-related proteins in subgroups dichotomized at the cutpoint of serum follistatin. L and H represent the patients with follistatin level below and above (≥) the cutpoint, respectively. The cutpoint was estimated at 490.5 pg/mL based on the interaction of follistatin with the PFS prolongation effects of EB treatment. Univariate logistic regression analysis was performed. All test results were shown as Wald test P value and q value (FDR) by Benjamini-Hochberg methods in each figure. There were 12 proteins significantly associated with H-follistatin ($q<0.05$) among 25 angiogenesis-related proteins analyzed. The box plot shows the summary statistics of serum concentrations of the proteins in each subgroup. The bottom and top of the box are the 1Q and 3Q. The horizontal bar within each box represents the median. The upper whisker extends from the 3Q to the highest value within 1.5× the IQR (the distance between 3Q and 1Q). The lower whisker extends from the 1Q to the lowest value within 1.5× IQR. PECAM1, platelet/endothelial cell adhesion molecule 1; G-CSF, granulocyte colony-stimulating factor; IL-8, interleukin-8; HGF, hepatocyte growth factor; HB-EGF, heparin-binding epidermal growth factor-like growth factor; FGF, fibroblast growth factor; PlGF, placental growth factor; PFS, progression-free survival; EB, combination therapy of erlotinib 150 mg/day and bevacizumab 15 mg/kg every 3 weeks; FDR, false discovery rate; 1Q, 25th percentile; 3Q, 75th percentile; IQR, interquartile range.

https://dx.doi.org/10.21037/tlcr-22-632

Table S1 Baseline demographics and clinical characteristics in each analysis

Characteristics	Whole analysis population (n=152)		pVEGFA (n=101)		Serum (n=133)		Tissue mRNA (n=24)		SNPs/VNTR (n=135)		NRP1 IHC (n=28)	
	EB (n=75)	E (n=77)	EB (n=48)	E (n=53)	EB (n=62)	E (n=71)	EB (n=11)	E (n=13)	EB (n=63)	E (n=72)	EB (n=14)	E (n=14)
Age (years)												
Median	67.0	67.0	70.5	67.0	68.0	68.0	71.0	70.0	68.0	67.5	69.5	69.0
<75 years	63 (84.0)	62 (80.5)	38 (79.2)	41 (77.4)	52 (83.9)	56 (78.9)	9 (81.8)	8 (61.5)	53 (84.1)	57 (79.2)	12 (85.7)	9 (64.3)
≥75 years	12 (16.0)	15 (19.5)	10 (20.8)	12 (22.6)	10 (16.1)	15 (21.1)	2 (18.2)	5 (38.5)	10 (15.9)	15 (20.8)	2 (14.3)	5 (35.7)
Sex												
Male	30 (40.0)	26 (33.8)	21 (43.8)	20 (37.7)	26 (41.9)	23 (32.4)	5 (45.5)	7 (53.8)	26 (41.3)	24 (33.3)	6 (42.9)	7 (50.0)
Female	45 (60.0)	51 (66.2)	27 (56.3)	33 (62.3)	36 (58.1)	48 (67.6)	6 (54.5)	6 (46.2)	37 (58.7)	48 (66.7)	8 (57.1)	7 (50.0)
Smoking status												
Never/former light	51 (68.0)	51 (66.2)	32 (66.7)	34 (64.2)	40 (64.5)	47 (66.2)	7 (63.6)	7 (53.8)	41 (65.1)	47 (65.3)	10 (71.4)	8 (57.1)
Other	24 (32.0)	26 (33.8)	16 (33.3)	19 (35.8)	22 (35.5)	24 (33.8)	4 (36.4)	6 (46.2)	22 (34.9)	25 (34.7)	4 (28.6)	6 (42.9)
Clinical stage												
IIB	1 (1.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
IV	60 (80.0)	62 (80.5)	39 (81.3)	40 (75.5)	50 (80.6)	56 (78.9)	1 (9.1)	2 (15.4)	51 (81.0)	57 (79.2)	2 (14.3)	2 (14.3)
Recurrence	14 (18.7)	15 (19.5)	9 (18.8)	13 (24.5)	12 (19.4)	15 (21.1)	10 (90.9)	11 (84.6)	12 (19.0)	15 (20.8)	12 (85.7)	12 (85.7)
EGFR mutation type												
Exon 19d	40 (53.3)	40 (51.9)	27 (56.3)	31 (58.5)	34 (54.8)	38 (53.5)	6 (54.5)	6 (46.2)	34 (54.0)	38 (52.8)	7 (50.0)	7 (50.0)
L858R	35 (46.7)	37 (48.1)	21 (43.8)	22 (41.5)	28 (45.2)	33 (46.5)	5 (45.5)	7 (53.8)	29 (46.0)	34 (47.2)	7 (50.0)	7 (50.0)

Data are for n (%), unless otherwise specified. pVEGFA, plasma vascular endothelial growth factor-A; mRNA, messenger RNA; SNP, single nucleotide polymorphism; VNTR, variable number of tandem repeats; NRP1, neuropilin 1; IHC, immunohistochemistry; EB, combination therapy of erlotinib 150 mg/day and bevacizumab 15 mg/kg every 3 weeks; E, 150 mg/day erlotinib monotherapy; EGFR, epidermal growth factor receptor.

https://dx.doi.org/10.21037/tlcr-22-632