Online supplemental material

Lieb et al. - Vascular endothelial growth factor, its soluble receptor and hepatocyte growth factor: Clinical and Genetic Correlates and Association with Vascular Function

Study sample

The Framingham Heart Study was established in 1948 with the enrolment of 5,209 men and women from Framingham in a prospective epidemiological study (original cohort).[1] The Framingham Offspring cohort was initiated in 1971, including 5,124 offspring of the original cohort and their spouses.[2] Starting in 2002, 4,095 participants with at least one parent in the Offspring cohort were enrolled in the Generation 3 cohort.[3] Examination 1 was attended by 4,095 Generation 3 participants. A total of 341 individuals were excluded from the present investigation for the following reasons: prevalent cardiovascular disease (n=66), serum creatinine >2 mg/dl (n=1), missing data on one or more biomarkers (n=149), missing one or more covariates (n=125). After exclusions, 3,754 participants remained eligible for the present analysis.

During their first examination at the Framingham Heart Study (2002 to 2005), Generation 3 cohort participants underwent a targeted medical history, physical examination, anthropometry, and laboratory assessment of traditional cardiovascular risk factors. Blood pressure (BP) was measured twice by a physician after the participant had been sitting in a chair for about 5 and 10 minutes, respectively. The arithmetic mean of both measurements is considered the examination BP. Body mass index (BMI) was calculated as the weight in kilograms divided by the square of height in meters. The glomerular filtration rate was estimated with the MDRD formula.[4] All participants provided written informed consent and the study protocol was approved by the Institutional Review Board at the Boston University Medical Center.

Laboratory Measurements of VEGF, sFlt-1 and HGF

Blood was drawn after an overnight fast, immediately centrifuged and stored at -80 ^{o}C until biomarkers were assayed. Serum VEGF, sFlt-1 and HGF were measured with commercial assays (R&D Inc.). These three biomarkers were chosen from a broad spectrum of endothelium-derived biomarkers because preliminary data from smaller clinical and epidemiological studies suggest that these biomarkers play an important role in cardiovascular disease, as detailed in the introduction. Opinion is divided regarding the choice of serum versus plasma for VEGF measurements.[5, 6] We, therefore, measured VEGF in a subset (n=18) of matched plasma and serum samples. Plasma VEGF was lower (42±28 pg/ml, mean±SD) than serum VEGF (361±223 pg/ml); 7/18 (39%) of plasma samples were below the lowest point of the standard curve (31.25pg/ml). There was a strong correlation between plasma and serum VEGF (r=0.76, p=0.0002). Based on the strong correlation between plasma and serum VEGF and the large number of plasma samples with very low VEGF, we chose to measure serum VEGF in our sample. The average inter-assay coefficients of variation were 2.1% for VEGF, 6.4% for sFlt-1, and 1.6% for HGF, respectively.

Flow Mediated Dilation and Reactive Hyperemia Measurements

After obtaining baseline brachial diameter and baseline flow velocity using a Toshiba SSH-140A ultrasound system,[7] a cuff was inflated on the right forearm to at least 50 mm Hg above the participant's systolic blood pressure (BP) to interrupt arterial blood flow for 5 minutes. Flow mediated dilation (FMD) was defined as percent change in brachial diameter 60 sec after deflation as compared with baseline brachial diameter. The coefficients of variation for baseline and hyperemic diameters were 0.5% and 0.7%, respectively.[7]

Assessments of Doppler flow at baseline and during reactive hyperemia were performed using a 3.75 MHz probe, correcting for the insonation angle. Semiautomated signal

averaging of the Doppler flow spectrum was used to analyze digitized audio data to yield measurements of mean baseline and hyperemic flow velocities.

Statistical analyses - addendum

Distribution of baseline variables

Descriptive statistics on the baseline variables suggested that the distributions were approximately normal with the exception of alcohol consumption which was standardized. In addition, the association of each variable with the outcome variables was examined individually as well as the homogeneity of errors across the range of the predictor variables was evaluated via graphical analyses. All variables had graded and monotonic associations with the outcome variables.

Assessment of predictive accuracy of the models describing clinical correlates of each biomarker

Candidate variables were chosen based on biological plausibility and previous smaller clinical studies. Using the stepwise procedure, we identified the significant predictors among all candidate variables for each biomarker and we formed our final models. To assess predictive accuracy, we then compared our final models (for each biomarker) to models of similar size to confirm that the selected set of variables provided the highest R^2 value among all possible combinations. To further validate our models, we performed a bootstrap bagging procedure with 200 samples, and for each sample, we ran our stepwise selection procedure. **Online supplementary Table 2** summarizes the variables selected, **online supplementary Table 3** the regression coefficients and their standard errors.

Highly significant variables in the final stepwise model (with $p < 0.001$, reported herein) were selected in at least 96% of the 200 replications; those variables with marginal p values were selected less often, but still more frequently than correlates that were not part of the final

model. Furthermore, regression coefficients (betas) in the final model in the manuscript were very consistent with mean betas from the bootstrapping analyses (**online supplementary Table 3).**

As an alternative way of model validation, we calculated the shrinkage factor[8] for the final models created from the stepwise selection procedure (**online supplementary Table 4**). For all three biomarker models the shrinkage factors were close to 1 indicating that the models were not overfitted.

Heritability estimates

Heritability was estimated from variance-component models using the software Sequential Oligogenic Linkage Analysis Routines (SOLAR)[9] to find maximum likelihood estimates of log-biomarker heritabilities (ratio of trait variance due to additive polygenetic effects to total trait variance) in complex pedigrees. For each growth factor, we provided two heritability estimates: first accounting for sex and age and then accounting for sex, age and all other covariates that were significantly associated with the respective biomarker in our prior analyses (clinical correlates).

Batch effect

Because some family members attended Heart Study examinations on or near the same date, we performed additional analyses to assess whether the heritability estimates might have been inflated due to an assay 'batch effect' in the laboratory. To assess the influence of batch, we used linear mixed-effect models to estimate sibling correlations for each marker (not accounting for covariates) with and without a random batch effect incorporated in the model. Results are shown in **Online Supplementary Table 5**. Batch effects were small for VEGF and HGF and did not materially alter the sibling correlations; for sFlt-1 a very modest batch effect had slightly inflated the sibling correlation (0.07 versus 0.05).

References

1. Dawber TR, Kannel WB, Revotskie N, Stokes J, III, Kagan A, Gordon T. Some factors associated with the development of coronary heart disease: six years' follow-up experience in the Framingham study. *Am J Public Health Nations Health* 1959;**49**:1349-1356.

2. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* 1979;**110**:281-290.

3. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ, D'Agostino RB, Sr., Fox CS, Larson MG, Murabito JM, O'Donnell CJ, Vasan RS, Wolf PA, Levy D. The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 2007;**165**:1328-1335.

4. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. *Ann Intern Med* 1999;**130**:461-470.

5. Dittadi R, Meo S, Fabris F, Gasparini G, Contri D, Medici M, Gion M. Validation of blood collection procedures for the determination of circulating vascular endothelial growth factor (VEGF) in different blood compartments. *Int J Biol Markers* 2001;**16**:87-96.

6. Lee JK, Hong YJ, Han CJ, Hwang DY, Hong SI. Clinical usefulness of serum and plasma vascular endothelial growth factor in cancer patients: which is the optimal specimen? *Int J Oncol* 2000;**17**:149-152.

7. Kathiresan S, Gona P, Larson MG, Vita JA, Mitchell GF, Tofler GH, Levy D, Newton-Cheh C, Wang TJ, Benjamin EJ, Vasan RS. Cross-sectional relations of multiple biomarkers from distinct biological pathways to brachial artery endothelial function. *Circulation* 2006;**113**:938-945.

8. Copas JB. Using regression models for prediction: shrinkage and regression to the mean. *Stat Methods Med Res* 1997;**6**:167-183.

9. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998;**62**:1198-1211.

Online Supplementary Table 1. Number of siblings per family in the study sample

Number of siblings per participant	n	Subtotal n
0	708	708
1	562	1124
2	301	903
3	131	524
4	45	225
5	23	138
6	14	98
7	2	16
8	2	18
Total n		3754

Online supplementary Table 2. Clinical correlates of VEGF, sFlt-1 and HGF. Variables selected using a bootstrap procedure for 200 samples.

Dependent: VEGF			
Candidate Variables	**% of times each variable is selected**	**Variables in final Model**	**P-value (final model)**
Age	74.00%	*	0.10
Male Sex	97.00%	*	<0.001
Smoking	98.00%	*	<0.001
Systolic BP	46.50%	*	0.014
Triglycerides	62.00%	*	0.054
Body mass index	100.00%	*	<0.001
Diastolic BP	55.00%		-
Antihypertensive medication	29.50%		-
Diabetes	18.00%		-
Total cholesterol	11.50%		-
HDL cholesterol	33.00%		-
Alcohol consumption	8.00%		-
eGFR	41.50%		-

Dependent: sFlt-1			
Candidate Variables	**% of times each variable is selected**	**Variables in final Model**	**P-value (final model)**
Age	67.00%	*	0.041
Male sex	100.00%	*	<0.001
Smoking	100.00%	*	<0.001
eGFR	83.00%	*	0.017
Systolic BP	12.50%		-
Diastolic BP	14.50%		-
Antihypertensive medication	12.00%		-
Diabetes	13.50%		-
Total cholesterol	38.50%		-
HDL cholesterol	50.50%		-
Triglycerides	42.50%		-
Body mass index	11.00%		-
Alcohol consumption	19.50%		-

Dependent: HGF			
Candidate Variables	**% of times each variable is selected**	**Variables in final Model**	**P-value (final model)**
Age	100.00%	*	<0.001
Male sex	100.00%	*	<0.001
Diastolic BP	71.00%	*	0.004
Antihypertensive medication	73.00%	*	0.005
Diabetes	96.00%	*	<0.001
HDL cholesterol	69.00%	*	0.039
Triglycerides	58.00%	*	0.039
Smoking	100.00%	*	<0.001
Body mass index	100.00%	*	<0.001
Systolic BP	36.50%		-
Total cholesterol	15.00%		-
Alcohol consumption	10.50%		-
eGFR	47.00%		-

Abbreviations: BP, blood pressure; eGFR, estimated glomerular filtration rate; HDL, high-density lipoprotein; HGF, hepatocyte growth factor; sFlt-1, soluble fms-like tyrosine kinase-1; VEGF, vascular endothelial growth factor.

Online supplementary table 3. Mean Beta Estimates, SD of Betas, and Standard Errors of Betas from boostrapping analyses.

	Dependent: VEGF			
		From Bootstrap analysis		
	Beta (Manuscript)	**Mean Beta**	**SD of Beta**	**Mean SE**
Age	0.023	0.0384	0.0107	0.0136
Male sex	-0.120	-0.1138	0.0300	0.0274
Systolic BP	0.033	0.0449	0.0121	0.0152
Diastolic BP	-	0.0456	0.0126	0.0147
Antihypertensive medication	-	-0.0944	0.0617	0.0492
Diabetes	-	-0.0906	0.1542	0.0823
Total cholesterol	-	-0.0195	0.0296	0.0144
HDL cholesterol	-	0.0373	0.0087	0.0156
Triglycerides	0.026	0.0409	0.0125	0.0142
Smoking	0.127	0.1390	0.0339	0.0355
Body mass index	0.074	0.0766	0.0134	0.0142
Alcohol consumption	-	0.0026	0.0303	0.0137
eGFR	-	0.0331	0.0081	0.0134

	Dependent: sFlt-1			
		From Bootstrap analysis		
	Beta (Manuscript)	**Mean Beta**	**SD of Beta**	**Mean SE**
Age	0.0220	0.0280	0.0078	0.0098
Male sex	0.1100	0.1007	0.0212	0.0199
Systolic BP	-	0.0277	0.0147	0.0114
Diastolic BP	-	-0.0220	0.0212	0.0113
Antihypertensive medication	-	0.0541	0.0480	0.0347
Diabetes	-	0.0179	0.1281	0.0598
Total cholesterol	-	0.0281	0.0073	0.0103
HDL cholesterol	-	-0.0307	0.0090	0.0111
Triglycerides	-	-0.0366	0.0141	0.0109
Smoking	-0.1030	-0.1114	0.0280	0.0260
Body mass index	-	-0.0055	0.0217	0.0100
Alcohol consumption	-	-0.0210	0.0086	0.0099
eGFR	-0.0270	-0.0306	0.0096	0.0096

	Dependent: HGF			
		From Bootstrap analysis		
	Beta (Manuscript)	**Mean Beta**	**SD of Beta**	**Mean SE**
Age	0.027	0.0286	0.0050	0.0044
Male sex	-0.041	-0.0432	0.0102	0.0090
Systolic BP	-	0.0152	0.0037	0.0050
Diastolic BP	0.013	0.0166	0.0041	0.0047
Antihypertensive medication	0.044	0.0435	0.0123	0.0157
Diabetes	0.09	0.0935	0.0239	0.0262
Total cholesterol	-	0.0094	0.0053	0.0044
HDL cholesterol	-0.01	-0.0135	0.0037	0.0049
Triglycerides	0.01	0.0119	0.0035	0.0045
Smoking	0.115	0.1149	0.0120	0.0113
Body mass index	0.053	0.0516	0.0049	0.0046
Alcohol consumption	-	0.0037	0.0093	0.0043
GFR	-	0.0102	0.0026	0.0043

Abbreviations: BP, blood pressure; eGFR, estimated glomerular filtration rate; HDL, high-density lipoprotein; HGF, hepatocyte growth factor; SD, Standard deviation; SE, standard error; sFlt-1, soluble fms-like tyrosine kinase-1; VEGF, vascular endothelial growth factor.

Online supplementary Table 4. Shrinkage factors for the final models created from the stepwise selection procedure

	VEGF	**sFlt-1**	**HGF**
LogLikelihood (full model)	-4350.8158	-3163.9456	-44.026
LogLikelihood (null model)	-4401.6855	-3199.9035	-296.9775
minus 2Loglikelihood (full)	8701.6316	6327.8912	88.052
minus 2Loglikelihood (null)	8803.371	6399.807	593.955
Diff between minus 2LogL	101.7394	71.9158	505.903
Degrees of Freedom	6	4	9
Shrinkage Factor	**0.950854831**	**0.958284549**	**0.984186692**

Abbreviations HGF, hepatocyte growth factor; sFlt-1, soluble fms-like tyrosine kinase-1; VEGF, vascular endothelial growth factor.

Online Supplementary Table 5. Sibling correlation with and without considering a random batch effect.

	Sibling correlation		
Batch effect	**LogVEGF**	**Log HGF**	**Log sFlt-1**
Ignored	0.3872	0.1944	0.0697
Included	0.4102	0.1971	0.0505