



Published in final edited form as:

*J Clin Pharmacol.* 2014 June ; 54(6): 682–687. doi:10.1002/jcph.254.

## Technical Considerations in the Development of Circulating Peptides as Pharmacodynamic Biomarkers for Angiogenesis Inhibitors

Vasiliki Thomeas, MS<sup>1</sup>, Selina Chow, MD<sup>1</sup>, Jose O. Gutierrez, BS<sup>1,†</sup>, Sanja Karovic, BS<sup>1</sup>, Kristen Wroblewski, MS<sup>2</sup>, Emily Kistner-Griffin, PhD<sup>2,‡</sup>, Theodore G. Karrison, PhD<sup>2,3</sup>, and Michael L. Maitland, MD, PhD<sup>1,3,4</sup>

<sup>1</sup>Section of Hematology/Oncology, Department of Medicine, University of Chicago, Chicago, IL, USA

<sup>2</sup>Department of Health Studies, University of Chicago, Chicago, IL, USA

<sup>3</sup>Comprehensive Cancer Center, University of Chicago, Chicago, IL, USA

<sup>4</sup>Committee on Clinical Pharmacology and Pharmacogenomics, University of Chicago, Chicago, IL, USA

### Abstract

To determine the biological reproducibility and estimate relevant covariates for candidate circulating biomarkers of angiogenesis, we conducted 3 sub-studies with 15 subjects each. In study 1, 6 healthy subjects provided 13 blood samples across 14–24 days. In study 2, 15 advanced solid tumor patients provided single blood samples before, and approximately 8 and 40 days after sorafenib treatment. In study 3, 4 healthy subjects provided blood samples on 3 occasions over 14 days, processed simultaneously in 2 different laboratories at a single institution. Vascular endothelial growth factor (VEGFA), soluble VEGF receptor-2 (sVEGFR2), and angiopoietin-2 (Ang2) concentrations in plasma and serum were determined by standard immunoassays. Ang2 and sVEGFR2 demonstrated low variance within and high variance across individuals reflected by the high intraclass correlation coefficient (for Ang2: 0.86 for plasma, 0.89 for serum; for sVEGFR2: 0.91 for plasma, 0.87 for serum). Repeated measures linear modeling from 15 patients demonstrated increased Ang2 ( $P = 0.05$ ) and decreased sVEGFR2 ( $P = 0.05$ ) after exposure to sorafenib. VEGFA had high intraindividual variance, and study 3 demonstrated the laboratory to have significant effects on plasma measurements ( $P = 0.05$ ). The biological reproducibility of sVEGFR2 and Ang2 support further use of these markers in studies of vasculature-targeted therapeutics.

### Keywords

biological marker; angiogenic proteins; angiogenesis inhibitor; VEGFR2; angiopoietin 2

---

Corresponding Author: Michael L. Maitland, MD, PhD, Section of Hematology/Oncology, Department of Medicine, University of Chicago, 5841 S. Maryland Avenue, Chicago, IL 60637, MC 2115, USA, mmaitlan@medicine.bsd.uchicago.edu.

<sup>†</sup>Present address of J. O. Gutierrez, Harvard School of Medicine, Cambridge, MA.

<sup>‡</sup>Present address of E. Kistner-Griffin, Medical University of South Carolina, Charleston, SC

Biomarkers, characteristics that are objectively measured and evaluated as indicators of normal biological processes, pathogenic processes, or pharmacologic responses to therapeutic interventions<sup>1,2</sup> could more rapidly advance development and safe, effective use of therapeutics. Basic scientific technological progress has accelerated discovery of candidate biomarkers for disease and therapy and rapid development of commercially available assays. Stepwise evaluation of biomarker candidates as fit-for-purpose in the clinical development setting should conserve resources<sup>3</sup> and avoid unnecessarily wasteful expenditures in early biomarker and therapeutics development programs.

Antiangiogenic agents were rapidly developed for the treatment of cancer without successful concurrent biomarker development. Nine drugs that inhibit signaling through the vascular endothelial growth factor (VEGF) signaling pathway (VSP) have now been approved by regulatory agencies for treatment of human diseases: aflibercept, axitinib, bevacizumab, cabozantinib, pazopanib, regorafenib, sorafenib, sunitinib, and vandetanib. Others remain in late phases of development or have already failed to prove safe or effective for intended indications. Despite millions of dollars spent on laboratory and clinical investigation of imaging, tumor tissue, circulating cellular and circulating peptide biomarkers, there remains no qualified predictive biomarker for selecting patients likely to benefit from these agents and no pharmacodynamic biomarker with which to rationally improve the therapeutic index of these drugs.

A basic principle of diagnostics development: is to assess biological reproducibility of novel biomarkers. This could be especially helpful in the early development of novel pharmacodynamic biomarkers to estimate whether and how they could be established as “fit-for-purpose.”<sup>4,5</sup> Here we present 3 small studies that were designed to estimate the sources and magnitude of variance for candidate pharmacodynamic biomarkers of VEGF signaling inhibition. The markers and assays in this study for angiopoietin-2 (Ang2), vascular endothelial growth factor A (VEGFA) and soluble VEGF receptor-2 (sVEGFR2) have been analyzed in numerous disease and therapeutic intervention settings, but not in this specific stepwise approach or with these particular serial sampling, comparative strategies for estimation and validation. These approaches might be applied to future efforts to discover and develop novel circulating peptide biomarkers in early therapeutics development.

## Methods

### Study Subjects

**Study 1**—Six healthy adults (3 men, 3 women), ages 26–38, were recruited by the study investigator for University of Chicago Study 13176, *Analysis of variation in measurement of circulating levels of biological mediators in healthy volunteers*.

**Study 2**—Oncologists recruited 70 subjects at the University of Chicago Medical Center for Study 13016, *A pharmacokinetic, pharmacodynamic, and pharmacogenetic study of sorafenib and blood pressure elevation in patients with advanced malignances*. To evaluate candidate plasma/serum biomarkers, samples from the first 15 (9 men, 6 women) of these subjects to have complete 8 week sample collections, ages 34–74, were analyzed. Inclusion

criteria for the initial trial included: life expectancy >12 weeks, age >18 years; Eastern Cooperative Oncology Group Performance Status rating of 0 or 1; acceptable organ and marrow function by prespecified laboratory measures.

**Study 3**—Four healthy adults (2 men, 2 women), ages 24-30, were recruited by the study coordinator of the University of Chicago Study 14911, *Collection of blood from normal healthy volunteers to evaluate methods of biomarker detection and analysis*.

All studies were approved by the University of Chicago Institutional Review Board. Subjects provided written informed consent prior to all study procedures.

### Study Procedures

**Study 1**—Subjects presented to the University of Chicago Clinical Resource Center on 4 occasions over the course of 14–24 days. On 1 of the 4 sample collection dates, subjects had a peripheral intravenous catheter placed and a pre-exercise sample drawn. Subjects then performed a modified Bruce protocol exercise treadmill test for 20 minutes, and samples were drawn at 30, 90, 240, and 480 minutes post exercise. On another occasion, subjects had a peripheral intravenous catheter placed and samples were drawn at 0, 90, 240, and 480 minutes, without any structured exercise intervention. Additional samples were drawn twice daily separated by 8 hours on another 2 dates within the 14-day study period.

**Study 2**—Sorafenib treatment was administered as previously described. Briefly, patients took sorafenib tablets, 400 mg by mouth twice daily. Dose adjustments were made in standard fashion for intolerable toxicities as previously described.<sup>7</sup> Single time point plasma samples were collected from each subject: at a baseline visit within 7 days before study treatment, between treatment Days 6 and 10 when concentrations of drug were first reached, and at Day 35–49 of treatment according to patient/ physician convenience for scheduling other study-related evaluations.

**Study 3**—Healthy volunteers presented for phlebotomy on 3 separate occasions (2 mornings and 1 afternoon session each) over the course of 14 days. Each subject had 4 tubes of blood drawn (2 EDTA-preserved 4 cc and 2 preservative-free 4cc tubes) on each occasion. Immediately after acquisition, 1 EDTA and 1 preservative-free tube were transferred according to typical institutional procedure by pneumatic tube to Lab 1. For processing by Lab 2, the other EDTA tube was incubated on ice for 30 minutes and the preservative-free tube at room temperature for 30 minutes and then both tubes were placed in a refrigerated centrifuge at 1300g for 15 minutes at 4°C. For processing by Lab 1, upon retrieval of blood tubes, tubes were placed in a refrigerated centrifuge at 1300g for 15 minutes at 4°C. Separated plasma and serum were then aspirated and distributed into 200 µL aliquots in screw-top cryovials. Samples were stored in each lab at –80°C until all the serum and plasma samples were collected for each subject's 3 time-points. Then samples from Lab 1 were brought to Lab 2 on dry ice and stored at –80°C until all samples were run simultaneously across the same set of assay plates.

## Sample Processing

All blood samples were drawn from peripheral veins through 21 or larger gauge needles and handled according to the following procedure. Serum was collected into standard preservative-free tubes (“red-top” BD Vacutainer tubes, Franklin Lakes, NJ). Blood was incubated at room temperature for 30 minutes and the serum was separated in a refrigerated centrifuge at 1300g for 15 minutes at 4°C. The separated serum was transferred in small aliquots to labeled polypropylene tubes and stored at –80°C. Plasma samples were drawn in a similar fashion into EDTA-preservative containing tubes (“lavender-top” BD Vacutainer tubes, Franklin Lakes, NJ). The tubes were inverted 3 times and placed on ice. Within 30 minutes of collection, plasma was separated by centrifugation at 1300g for 15 minutes at 4°C. The separated plasma was transferred in aliquots to polypropylene cryovials and stored at –80°C.

## Biomarker Measurements

Plasma and serum samples were thawed on ice and then assayed in triplicate for vascular endothelial growth factor A (VEGFA), angiopoietin-2 (Ang2), and soluble vascular endothelial growth factor receptor-2 (sVEGFR2) by ELISA according to manufacturer's specifications (R&D Systems, Minneapolis, MN). The manufacturer reported intra-assay/inter-assay coefficients of variation (CV) are: VEGFA 7%/9%, sVEGFR2 5%/8%, and Ang2 7%/10% and our analyses were verified to perform within these ranges. The mean lower limit of quantification (LLOQ) for VEGFA is 0.009 ng/mL, for sVEGFR2 is 0.46ng/mL and for Ang2 is 0.82 ng/mL.

Samples were randomly allocated across plates for each study. Control samples consisted of stored serum and plasma from separate aliquots of a single time-point draw from the same individual loaded on separate plates. Each plate contained 2 of these “internal control” samples from 2 individuals used to verify the intra-assay and inter-assay coefficient of variation and to normalize inter-plate variation in measured concentrations.

## Statistical and Data Analysis

Initial sample sizes for each study were determined by convenience and expectation for use of appropriate modeling strategies to estimate the likely magnitude of the relevant effects under study. Sample sizes in each study population were determined by convenience with the expectation that appropriate models for quantitative estimation would be applied to inform decisions on whether or not to proceed with further validation efforts. For Study 1 intraindividual and interindividual variance components analyses were informed by power calculations as described by Bonett where precision for estimates of the confidence interval for the intraclass correlation is determined by the number of subjects and the number of samples collected per subject.<sup>8</sup> For Study 2, we expected to perform repeated measures modeling with the initial 15 subject samples. If marginal effects were observed we would not pursue further study, if marginally significant findings were detected, we would test more subjects and samples. For Study 3 we expected to use small study size analytical strategies.

For study 1, to estimate the magnitude of effect of sex and exercise on biomarkers as well as to determine the between-subject and within-subject variance components we performed analysis of variance (ANOVA). The intraclass correlation coefficient (ICC) was calculated to determine the fraction of total variance attributable to interindividual vs intraindividual

differences ( $ICC = \frac{\sigma_{\alpha}^2}{\sigma_{\alpha}^2 + \sigma_{\epsilon}^2}$ , where  $\sigma_{\alpha}^2$  denotes the interindividual component of variability and  $\sigma_{\epsilon}^2$  the intraindividual component). In study 2, repeated measures linear modeling was used on sVEGFR2 and Ang2 plasma concentrations, a test of the within subject change over time of exposure to sorafenib was performed using a Greenhouse-Geisser adjustment to derive the *P*-value. In study 3, we compared plasma and serum VEGFA and sVEGFR2 values between the 2 laboratories with a Wilcoxon signed-rank test. For plasma VEGFA, any value less than 0.009 ng/mL was assigned a value of 0.009 ng/mL (the LLOQ of the assay).

## Results

### Intra- and Inter-Individual Variance in Ang2 and sVEGFR2

The manufacturer's performance testing for Ang2 and sVEGFR2 had demonstrated detectable concentrations in all subjects with ranges of 1.1–8.9 ng/mL and 6.42–14.50 ng/mL, respectively; but it was unclear whether this range of measurements was due primarily to differences between individuals in these measures (interindividual variance) or to sources of variance (sample processing, exercise, time of day, etc.) that impact the reproducibility of measurements within an individual (intraindividual variance) over time. Although the population range of single time-point assessments was reported in the manufacturer's tests, the differing preservatives and incubation times used in standard serum and plasma sample acquisition could affect some markers' reproducibility more than others. We therefore collected and compared serum and plasma samples concurrently.

In study 1, serial measurements for all subjects were consistently within a range of 1.1–2.8 ng/mL for Ang2 and 6.1–11.2 ng/mL for sVEGFR2 (Figure 1). ANOVA revealed no statistically significant effects of sex or the standardized exercise session. The ICC for Ang2 was: 0.86 for plasma and 0.89 for serum and for sVEGFR2 was: 0.91 for plasma and 0.87 for serum. Thus, for both markers, inter-individual differences in sVEGFR2 and Ang2 were the primary source of variance. The greatest spread between minimum and maximum measurements for any one individual's plasma Ang2 was 0.85 ng/mL, in subjects 1 and 4, and for sVEGFR2 was 1.53 ng/mL, in subject 2.

### Effects of Sorafenib on Ang2 and sVEGFR2 Measurements

Study 1 established a reference range for intraindividual biological reproducibility over a 2-week interval, including standardized cardiovascular stress, for healthy subjects. This provided benchmark reference information for estimating effects of sorafenib on fifteen cancer patients enrolled in Study 2. Now with single time-point measures once before and on 2 occasions after sorafenib treatment, we could estimate the effects of sorafenib exposure on the candidate biomarkers Ang2 and sVEGFR2. A univariate test of the time on sorafenib effect was statistically significant for both markers (Figure 2, Ang2 and sVEGFR2 both with

$P < 0.05$ ). The maximum mean change for Ang2 was 0.5ng/mL which was within the intraindividual variability for healthy subjects of approximately 0.85ng/mL. For sVEGFR2, the maximum mean change was 1.6ng/mL (greater than the measured maximum intraindividual variability for any of the healthy subjects of approximately 1.5ng/mL), suggesting that the change observed in sVEGFR2 is not due to normal physiological fluctuations of the circulating peptide. These findings indicated that sorafenib exposure had a large and consistent enough effect on plasma sVEGFR2 measurements to support further investigation as a candidate pharmacodynamic biomarker for this VEGFR2 inhibiting drug.

### Effects of Sample Processing Site on Biomarker Measurements

To assess the robustness of sVEGFR2 and the previously more widely studied candidate biomarker VEGFA for study in multicenter clinical trials, serum and plasma from 4 healthy adult volunteers were drawn, collected, and sent for processing in 2 different laboratories, designated as Lab 1 and Lab 2, on the University of Chicago campus. Both labs are institutional core facilities well established to support clinical investigations and which followed the same standards of procedure outlined by the study protocol. Plasma concentrations of VEGFA in all 4 volunteers processed by Lab 2 were almost all under the LLOQ of the assay (0.009 ng/mL), while significantly higher, but variable measures were detected in Lab 1. For the simultaneously processed serum samples, the VEGFA measures were higher in both labs (Figure 3). Plasma and serum VEGFA concentrations were significantly different between Lab 1 and Lab 2 ( $P < 0.05$  and  $P < 0.05$ , respectively). In contrast, serum concentrations of sVEGFR2 were not significantly different between the 2 labs ( $P > 0.05$ ). For plasma sVEGFR2 concentrations, a lab-dependent difference met statistical significance criteria ( $P < 0.05$ ) but concentrations for all samples from all subjects measured by both labs fell within the typical intraindividual variability previously observed in study 1 of 1.5 ng/mL. Because of this lab to lab variability in VEGFA measurements in individual subjects, we did not proceed with measurement of frozen samples for VEGFA as originally proposed in the study protocol.

### Discussion

We describe a series of studies of candidate circulating peptide biomarkers that demonstrated biological reproducibility, relative magnitude of drug effect and interlaboratory processing effects on standard immunoassays for Ang2, sVEGFR2, and VEGFA. The explicit purpose of these studies was to perform serial technical evaluations and to estimate whether each marker was sufficiently robust for further use in biomarker development investigations. Measurements in healthy subjects provide an estimation of the largest physiological variability of a biomarker allowing for a reference in patient populations in determining effects of therapeutic agents. Although it is clear, that these studies were too small to establish definitive conclusions about covariates and small effects, they were effective in guiding our subsequent research and project planning. These studies suggest that stepwise evaluation of novel markers with attention to the biological reproducibility and relative magnitude of effect of drug exposure compared to normal measurement imprecision could lead to more efficient pharmacodynamic biomarker development.

Dancey and Wagner<sup>2,9</sup> have detailed the challenges to effective biomarker development in parallel with development of new therapeutics. Now that high throughput and multiplexing technologies are widely available for measuring circulating peptides, a serial, step-wise process of validation and qualification could be a sensible approach to determining which biomarkers are studied further even when a marker is not intended to be developed formally as a companion diagnostic. The appropriate specific steps will vary from one investigational setting to the next. Here, we assessed Ang2 and sVEGFR2 as candidate biomarkers in small numbers of healthy volunteers and then in small numbers of subjects treated with a new drug. With these data we could estimate the magnitude of variance in a population and the effect size of the drug exposure on change in the marker of interest. At the time, there was far more known about the biological significance and function of Ang2 than sVEGFR2. However, because the mean change in sVEGFR2 with sorafenib exposure was greater than any healthy subject's range from minimum to maximum measurement over serial sampling sessions, we were convinced that sVEGFR2 would be a robust, reproducible biomarker and were compelled to study it further. This step-wise process of serial estimation and initial validation requires few subjects, and, this strategy may enable the research team to decide which, among multiple candidate markers, may be most likely to inform further development of a new drug.

This stepwise process also allows the research team to identify modifiable sources of intra-individual variance and to minimize them with the recommended procedures for sample procurement prior to launching large scale protocols. By defining sample handling procedures, optimum time of day at which a sample should be drawn, effects of time at room temperature, etc., the study team can improve the power for the study to detect biomarker-related endpoints. Similarly, if high inter-individual variance is detected, a cross-sectional analysis might be inefficient and a longitudinal within-individual measurement study would be more robust for detecting effects of treatment on that marker.

Our test of laboratory processing effects on VEGFA enabled us to allocate patient samples and funding to alternative endpoints. VEGFA was a popular marker to test even with its recognized pitfalls.<sup>5</sup> Our project verified Jelkmann's concerns with further confirmation by concurrent studies that examined the contribution to variance in VEGFA measurements by lab processing and handling methods.<sup>10,11</sup> Our analysis in study 3 confirmed the robustness of sVEGFR2 and lack thereof for VEGFA and changed our initial plan of studying VEGFA in one of our multicenter trials to studying sVEGFR2.

In conclusion, technical considerations in the acquisition and measurement of circulating peptide biomarkers can significantly affect results and their interpretation. A simple, small scale systematic approach to the initial characterization and validation of candidate biomarkers has the potential to reduce costs of clinical studies and to focus the research team on evaluation of markers with interpretable, reproducible results.

## Acknowledgments

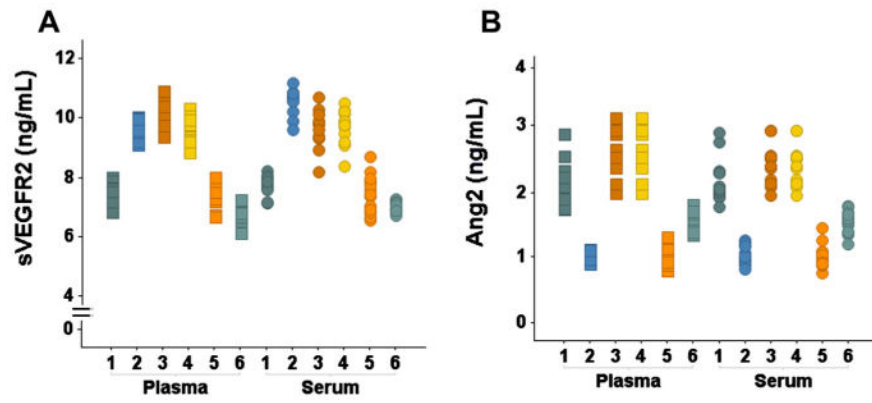
**Funding:** MLM has received research grants from the U.S. National Institutes of Health grants: K23CA124802 NCI Career Development Award, a Preclinical Pilot Translational Science Award from the University of Chicago CTSA UL1RR024999, and a CALGB Foundation Faculty Fellowship, KW and TGK were supported by the

University of Chicago Comprehensive Cancer Center (P30 CA014599), JOG received support from Pritzker School of Medicine Experience in Research program (5R25HL096383). Additional support was provided by research funding from the O'Connor Foundation. MLM has received research support from OncoTherapy Sciences, Inc. and has conducted related unfunded studies in collaboration with GlaxoSmithKline and Abbott Diagnostics, Inc.

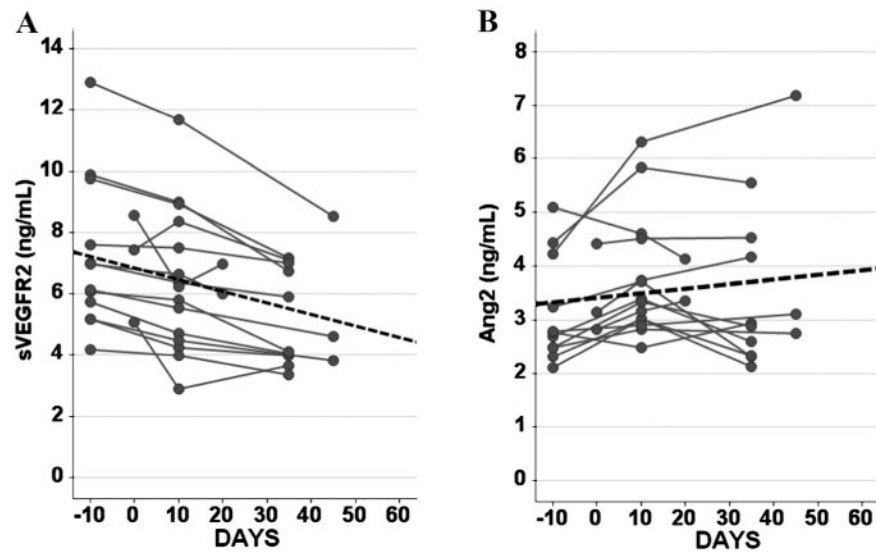
## References

1. Atkinson AJ, Colburn WA, DeGruttola VG, et al. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther.* 2001; 69:89–95. [PubMed: 11240971]
2. Dancey JE, Dobbin KK, Groshen S, et al. Guidelines for the development and incorporation of biomarker studies in early clinical trials of novel agents. *Clin Cancer Res.* 2010; 16:1745–1755. [PubMed: 20215558]
3. Lee JW, Devanarayan V, Barrett YC, et al. Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm Res.* 2006; 23:312–328. [PubMed: 16397743]
4. WestGard, JO. *Basic Qc Practices: Training in Analytical Quality Management for Healthcare Laboratories.* 3rd. Madison, WI: Westgard Quality Corp; 2010.
5. Jelkmann W. Pitfalls in the measurement of circulating vascular endothelial growth factor. *Clin Chem.* 2001; 47:617–623. [PubMed: 11274009]
6. Maitland ML, Kasza KE, Karrison T, et al. Ambulatory monitoring detects sorafenib-induced blood pressure elevations on the first day of treatment. *Clin Cancer Res.* 2009; 15:6250–6257. [PubMed: 19773379]
7. Escudier B, Eisen T, Stadler WM, et al. Sorafenib in advanced clear-cell renal-cell carcinoma. *N Engl J Med.* 2007; 356:125–134. [PubMed: 17215530]
8. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med.* 2002; 21:1331–1335. [PubMed: 12111881]
9. Wagner JA. Strategic approach to fit-for-purpose biomarkers in drug development. *Annu Rev Pharmacol Toxicol.* 2008; 48:631–651. [PubMed: 17937595]
10. Svendsen MN, Brunner N, Christensen IJ, et al. Biological variations in plasma VEGF and VEGFR-1 may compromise their biomarker value in colorectal cancer. *Scand J Clin Lab Invest.* 2010; 70:503–511. [PubMed: 20873967]
11. Hetland ML, Christensen IJ, Lottenburger T, et al. Circulating VEGF as a biological marker in patients with rheumatoid arthritis? Preanalytical and biological variability in healthy persons and in patients. *Dis Markers.* 2008; 24:1–10. [PubMed: 18057530]



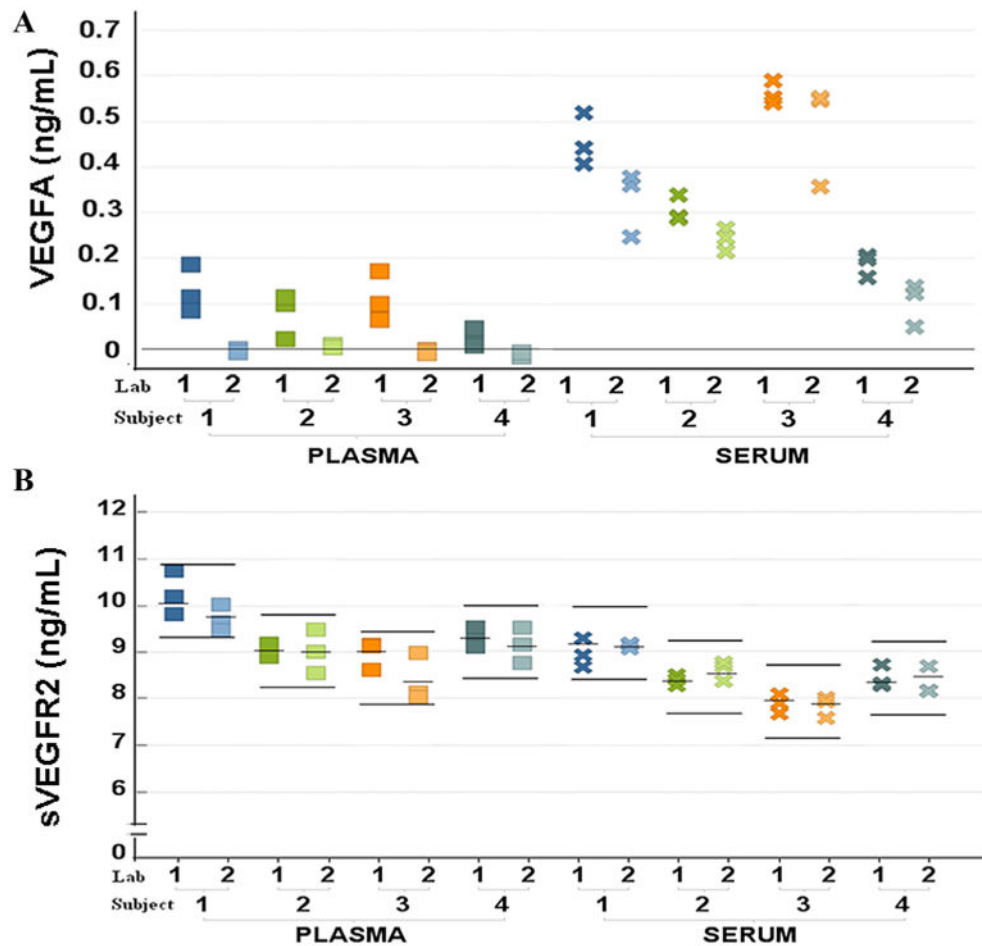


**Figure 1.** The primary source of variance for (A) sVEGFR2 and (B) Ang2 concentrations is interindividual variability. All 13 measurements for each subject are plotted in the same color in both A and B.



**Figure 2.**

Repeated measures linear modeling was used on (A) sVEGFR2 and (B) Ang2 concentrations, a test of the within subject change over time of exposure to sorafenib was performed using a Greenhouse-Geisser adjustment to derive the  $P$ -value ( $P < 0.05$  for both sVEGFR2 and Ang2). Dotted line indicates trend line in both A and B. sVEGFR2 concentrations clearly decline on exposure to sorafenib more than its regular fluctuation in healthy volunteers. Ang2 concentrations increased, but within its regular fluctuation found in healthy volunteers.



**Figure 3.** Comparison of concentrations of VEGFA and sVEGFR2 across processing methods in 2 different labs yields VEGFA as sensitive to handling. All 3 measurements for each subject are plotted in the same color for each subject. The line on (A) VEGFA represents 0.009 ng/mL (the LLOQ of the assay). For (B) sVEGFR2, mean of concentration for each individual is indicated by short bars and the range of 1.5 ng/mL for each individual is indicated by long bars. A Wilcoxon signed-rank test was carried out to compare plasma and serum (A) VEGFA and (B) sVEGFR2 values between the 2 laboratories:  $P = 0.05$  for both plasma and serum VEGFA concentrations; and  $P = 0.05$  and  $P > 0.05$  for plasma and serum sVEGFR2.