Published in final edited form as:
Clin Chem Lab Med. 2011 May ; 49(5): 817–824. doi:10.1515/CCLM.2011.122.

Performance Evaluation of a Multiplex Assay for Future Use in Biomarker Discovery Efforts to Predict Body Composition

Jennifer W. Bea[1,*], **Nicole C. Wright**[2], **Patricia Thompson**[1,2], **Chengcheng Hu**[2], **Stefano Guerra**[3,4], and **Zhao Chen**[2]

[1]Arizona Cancer Center, University of Arizona, Tucson, AZ

[2]Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ

[3]Arizona Respiratory Center, University of Arizona, Tucson, AZ

[4]CREAL Centre, IMIM-Hospital del Mar, CIBERESP, Barcelona, Spain

Abstract

Background—Interest in biomarker patterns and disease has led to the development of immunoassays that evaluate multiple analytes in parallel with little sample. However, there are no current standards for multiplex configuration, validation, and quality, thus, validation by platform, population, and question of interest is recommended. We sought to determine the best blood fraction for multiplex evaluation of circulating biomarkers in postmenopausal women and to explore body composition phenotype discrimination by biomarkers.

Methods—Archived serum and plasma samples from a sample of healthy postmenopausal women with the highest (n=9) and lowest (n=11) percent lean mass, by dual-energy X-ray absorptiometry, were used to measure 90 analytes by bead-based, suspension multiplex assays. Replicates of serum and plasma were analyzed in a random selection of 4 of these individuals.

Results—Ninety-percent of the analytes were detectable for ≥50% of samples; when limited to these well detected analytes, mean replicate correlations for serum and plasma were 0.87and 0.85 respectively. Serum had lower error rates discriminating phenotypes; 7 serum versus 2 plasma analytes discriminated extreme body phenotypes.

Conclusion—Serum and plasma performed similarly for the majority of the analytes. Serum demonstrated a slight advantage in predicting extreme body composition phenotypes in postmenopausal women using parallel evaluation of analytes.

Keywords

biomarkers; lean mass; multi-analyte; multiplex; phenotype discrimination

Corresponding Author: Jennifer W. Bea, Arizona Cancer Center, University of Arizona, 1515 N. Campbell Ave., Tucson, AZ 85724, 520-626-0912 phone, 520-626-5348 FAX, jbea@azcc.arizona.edu.

Conflicts of Interest

The authors have no commercial, proprietary, or financial interest in the products or instruments described in this article.

Statement of Ethics

We certify that all applicable institutional and governmental regulations concerning the ethical use of human volunteers were followed during this research. The University of Arizona Institutional Review Board approved the study protocol and all participants included in the analysis provided written informed consent prior to their participation.

Introduction

The desire to measure blood-based biomarkers as components or indicators of chronic disease states (i.e., cytokines, growth factors), has resulted in a number of industry-developed multiplex platforms.(1–3) The availability of these prefabricated multiplex technologies to measure circulating analytes in parallel, on limited sample, offer a unique opportunity to derive complex information in association studies and possibly clinically useful patterns. However, the parallel approach and use of these multiplex platforms pose specific challenges for epidemiologists conducting large-scale discovery based biomarker studies. These include consideration of the sensitivity and specificity of the analyte measures, reproducibility of the measurement, performance of the type of biologic (i.e., serum or plasma), as well as consideration of study design approaches that promote reduction of complex panels to a limited set of high interest candidates for validation in large sample sets.(3)

Here we describe a biomarker validation study to evaluate the performance of matched serum and plasma samples in a complex panel of analytes. We hypothesized that serum and plasma would perform equally well on the platform to evaluate the 90 analytes. We also conducted exploratory analyses to assess the performance of serum and plasma analytes included in the panel for discriminating between individuals with high and low skeletal muscle mass phenotypes.

Materials and methods

Study design

Archived serum and plasma samples from the Women's Breast and Bone Density (WBBD) study were used for this validation study.(4) The WBBD study was conducted at the University of Arizona between 2001 and 2004, and enrolled 238 pre (41–50 years) and postmenopausal (56–70 years) White and Hispanic women from the greater Tucson community. Participants were generally healthy, and did not have a history of chronic diseases or other health conditions related to bone health including autoimmune disorders, diabetes, cancer, or thyroid disease.

Age adjusted percent lean mass was generated from full body dual energy X-ray absorptiometry (DXA) scans (Hologic QDR 4500w v8.26, Hologic Inc, Waltham, MA, USA, www.hologic.com) taken during the parent study (WBBD). DXA-derived skeletal muscle mass has been validated against lean body mass and skeletal muscle mass measured on MRI. (5) Postmenopausal women in the highest (n=9) and lowest (n=11) tails of percent lean mass were selected for biomarker analysis. Both serum and plasma samples were run for each of the women selected. In addition, replicate serum and plasma samples were analyzed in a random subset of the same women (n=4). Institutional review board approval and participant written informed consent were obtained prior to study initiation. All blood samples used in the validation study described here were identified as either serum or plasma (sample type), but were otherwise blinded for analyte analysis.

Serum and plasma collection

Following a 12-hour fast, blood by venipuncture was drawn by a trained phlebotomist. Serum separator tubes and EDTA treated tubes were used to collect serum and plasma respectively. Samples were allowed to coagulate for 30 minutes and were then centrifuged for 10 minutes (1,300 × g) at 4°C. Serum and plasma fractions were transferred to cryovials for storage in 0.5 ml aliquots and stored at −70° C until shipped on dry-ice for cytokine measurements detailed below; all samples remained frozen and arrived at the laboratory in good condition.

Cytokine measurements

Ninety analytes (Supplemental Data Table 1) were measured by bead based, suspension multiplex assays (Human Multi-Analyte Profile (MAP) version 1.6, Rules Based Medicine (RBM), Austin, TX, USA, www.rulesbasedmedicine.com) using serum and plasma samples. The multiplex assays were developed using the principles of immunoassay,(6) including: selection of respective antigens and antibodies per analyte, optimization of sensitivity, specificity, and dynamic range for each analyte as a single test, and then the incorporation of single assays into a multiplex. Assay development also included validation studies for each multiplexed analyte including determination of least detectable dose (LDD), lower assay limit (LAL), dynamic range, imprecision, spiked recovery, cross-reactivity, and matrix interferences. LDD reflects the mean background (blank readings) value for 20 determinations plus three standard deviations. LAL reflects each assay's sensitivity based on the lowest concentration calibrator used for quantitation; it is often referred to as the assay limit of detection. Values above the LDD were designed to be highly precise and to have coefficients of variation (CV) <20%. Values between the LDD and LAL are considered less precise than values above the LDD, but are reported. Values below the LAL indicate that a sample cannot be measured on the standard curve and are reported as “low”. To acquire values for each analyte evaluated in our study, parallel cytometric quantitation on 8 multiplexes, with calibrators and controls on each plate, was run from 100uL sample using microtiter techniques and Luminex technology. Values were interpreted and summarized by RBM proprietary data analysis software. For each analyte and sample type, RBM provided the LDD, LAL, dynamic range, and sample value, if detectable.

Statistical analysis

Descriptive statistics by percent lean tail were performed using Wilcoxon-rank sum test for continuous variables and chi-squared test for categorical variables. Analyte descriptives including, the percent below LAL (% undetectable) and the percent between the LAL and LDD were calculated. Analytes with greater than 50% undetectable rates were excluded in correlative analyses. Spearman correlation coefficient was used to assess replicate correlations and to calculate the correlation between serum and plasma for each analyte.

Recursive partitioning (7) and random forest (8) models were used to examine the power of serum and plasma markers to discriminate subjects with high percent lean mass from those with low percent lean mass. A total of 1,000 bootstrap datasets were generated from the serum and plasma marker data separately, and both recursive partitioning and random forest models were fit to each of them separately. For the recursive partitioning method, five-fold cross-validation was used to calculate the discriminating error for each bootstrap dataset, while the error of the random forest method is calculated using out-of-bag (OOB) samples. For each method, the 1,000 discriminating errors calculated from the bootstrap datasets can be considered as a sample drawn from the distribution of error in identifying high vs. low lean mass individuals. These errors are then summarized by descriptive statistics and exploratory graphical tools, and are also compared using Wilcoxon rank sum test. To estimate the strength of discriminating power for each individual serum and plasma marker, a random forest with 10,000 trees was built on the original marker dataset and the importance measure for each marker was estimated along with its standard error.

Values for any marker in the high lean mass group were compared to those in the low lean mass group also by the Wilcoxon rank sum test. The Benjamin-Hochberg (9) procedure was used to control the false discovery rate (FDR) at the 0.1 level. For recursive partitioning and random forest models, undetectable values for any marker were imputed as half of the lowest observed level for the same marker.

Results

Study population

The low and high percent lean groups were similar with respect to demographic comparisons. The mean age of the study population was not statistically different (58.9 vs. 59.3 years respectively). Both groups had similar years of estrogen exposure, and years of hormone therapy was not statistically different between the groups. Smoking history was also equally distributed between the two groups, with approximately 35% of both groups reporting never smoking, 45% former smoking, and 20% current smoking. Medical conditions such as, high cholesterol, hypertension, osteoarthritis, and osteoporosis were also not statistically different between the two groups (Table 1).

As expected, there were significant differences between groups for body composition measures due to selection criteria. The low percent lean tail had significantly higher weight (97.3 vs 56.7 kg), body mass index (BMI) (36.8 vs. 22.5 kg/m^2), waist to hip ratio (0.84 vs. 0.76), total fat mass (45.6 vs. 15.4 kg), total percent fat (48.5 vs. 27.7%), appendicular fat mass (22.1 vs. 7.68), and marginally higher total lean mass and appendicular lean mass. Both total body ($p<0.001$) and appendicular ($p=0.053$) lean to fat ratio were lower in the low percent lean tail compared to the high percent lean tail. Complete demographic information and body composition information by percent lean mass is summarized in Table 1.

Detection and replicate performance of analytes

Analyte performance was assessed using detection rates and the serum and plasma replicate correlation. Overall analyte detection was high in both fractions, 87% of serum and 92% of plasma analytes had less than a 50% undetectable rate. Four analytes did not perform well in both serum and plasma (i.e. >50% undetectable) including: calcitonin, GM-CSF, IL-1α, and lymphotacin. Serum had an additional 8 analytes (endothelin-1, IL-2, IL-4, IL-6, MMP-2, MMP-9, PAPP-A, and PSA free) with greater than 50% undetectable rates, whereas plasma only had an addition 3 analytes (IGF-1, MMP-3, and TNF-β) with greater than 50% undetectable rates. Seven analytes in both serum and plasma had >50% of their values between the LAL and LDD (creatine kinase-MB, fatty acid binding protein, fibrinogen, IL-5, IL-10, IL-12p70, IL-15). One additional serum analyte (IL-12p40) and seven other plasma analytes had >50% of their values between LAL and LDD (IL-2, IL-3, IL-4, IL-7, PAPP-A, tissue factor, thrombopoietin). Analyte performance by sample type can be found in Supplemental Data, Table 1.

Excluding the analytes for which >50% of sample were below the LAL, the mean (95% confidence interval (CI)) spearman correlation for the replicate measures was 0.87 (CI; 0.83, 0.91) in serum and 0.85 (CI; 0.81, 0.89) in plasma. The overall replicate performance was high with 64% of analytes in serum and 63% of analytes in plasma having a correlation 0.9 or higher for replicate samples. Correlation<0.5 between replicates were observed for 6 serum (apolipoprotien CIII, cancer antigen 125, FGF basic, IL-12p70, SGOT, and TNF-β) and 6 plasma (cancer antigen 125, endothelin-1, glutathione s-transferase, IL-12p70, IL-7, and SGOT) analytes. When these analytes were removed, the mean spearman correlation of the replicate samples increased to 0.92 (95% CI; 0.89, 0.94) in serum and 0.89 (95% CI; 0.87, 0.92) in plasma.

Serum and plasma concordance

Excluding those analytes for which 50% or more of subjects were below the LAL, the mean serum to plasma spearman correlation was 0.65 (95% CI; 0.59, 0.71). Fifteen analytes demonstrated serum to plasma correlations ≥0.90 and an additional 14 analytes demonstrated serum to plasma correlations between 0.80 and 0.89; descriptive statistics

(mean, SD, range) for these analytes are presented in Table 2. Twenty-four analytes had a serum to plasma correlation < 0.50; descriptive statistics for these analytes may be found in Table 3. After removing analytes with serum to plasma correlations of <0.50, the mean serum to plasma correlation increased to 0.80 (95% CI; 0.77, 0.84). The individual analyte concordance between sample types is detailed in Supplemental Data, Table 1 for all 90 analytes.

Exploring phenotype discrimination by sample type

Recursive partitioning and random forest models were used to assess the percent lean mass discriminating power in each sample type. The discriminating errors of serum and plasma markers calculated from cross-validation on 1,000 bootstrap datasets in both model types are summarized in Table 4. The mean error of serum markers was 0.119 (SD=0.085), while the mean error of plasma markers was 0.170 (SD=0.10). Wilcoxon rank sum test shows that the model based on serum markers had significantly lower error rate than the model based on plasma markers ($p<0.0001$). The median and maximum, as well as the 1st and 3rd quartiles of the serum marker error are all lower than the corresponding parameters of the plasma errors.

For random forest models, the mean error of serum markers was 0.081 (SD=0.05), while the mean error of plasma markers was 0.097 (SD=0.06). Note that the errors of random forest models were generally lower than errors of the recursive partitioning models. As above, the Wilcoxon rank sum test showed that the model based serum markers had significantly lower error rate than that based on the plasma markers ($p<0.0001$). The 3rd quartile and the maximum of the serum marker error are all lower than the corresponding parameters of the plasma errors, while the minimum, 1st quartile and median are the same.

The Wilcoxon rank sum test was used to compare biomarker values between the high lean mass group and the low lean mass group. The Benjamin-Hochberg (9) procedure was used to control the FDR at the 0.10 level among all serum markers and the same procedure was applied also to the plasma markers. Seven serum markers had significantly or marginally significantly different levels between the two groups of subjects (adjusted p-values are given in parentheses). Those in the lower percent lean mass group demonstrated higher leptin (0.018), insulin (0.049), complement 3 (0.051), and serum amyloid P (0.074) levels, and lower adiponectin (0.074), growth hormone (0.049), and eotaxin (0.096) levels than those in the upper percent lean mass group. Two plasma derived markers followed a similar pattern; leptin (0.048) and insulin (0.094) were higher in the lower percent lean mass group.

Discussion

Mulitplexed cytokine technologies are increasingly available as interest shifts towards the evaluation of patterns of multiple biomarkers associated with physiologic processes and disease states. Multiplex platforms offer unprecedented efficiency for evaluating several analytes simultaneously and the parallel screening of multiple biomarkers requires relatively little sample, reagent, and time. (1–3, 10, 11) Luminex based platforms have been found to have generally high reliability and reproducibility (10), but cytokine values may vary greatly depending on analyte, anticoagulant, and inter-individual differences (12). In addition, sample type may play a role in cytokine variability (12). This emerging field is in need of standards to guide assay configuration, analytical validation, and quality control before applying these methods to large longitudinal cohorts, clinical trials, diagnostics, or prognostics.(2, 3) As such, our goal was to evaluate a Luminex based multiplexed assay platform for several outcomes in order to judge its suitability for evaluating multiple biomarkers in a much larger clinical trial of a similar population of women and to optimize analyte detection and quantification in the population. Therefore, performance differences

between sample types of serum and EDTA treated plasma on a complex panel of analytes were evaluated. We also sought to gain insight on the performance characteristics of the specific analytes on the multiplex to discriminate between known lean mass phenotypes in women. Analytes that were best able to distinguish between extreme lean mass phenotypes may also be used to narrow down the analyte panel to more specific candidates for our planned evaluation of biomarker associations with lean mass in the setting of a large clinical trial.

We were able to minimize several sources of potential variability in analyte quantification by standardized sample collection, anticoagulant, storage time, freeze/thaw cycles, and temperature in this study.(13, 14) However, serum and plasma are not the same. While the cellular fraction of the whole blood has been eliminated from both, serum protein levels may be lower than plasma protein levels by 0.24–0.29 gm/dL.(14) Although albumin may be higher in serum,(14, 15) others have noted that the absence of fibrinogen and other coagulant associated high-molecular weight proteins may create a technical advantage over plasma in quantifying less abundant proteins. Particular analytes may also be more stable, and thus produce more reliable measures, in serum versus plasma depending on the analyte. However, immediate separation of both serum and plasma from cellular elements, as in this protocol, enhances analyte stability in both. (14, 16, 17)

Based on the results from this validation study, we rejected the null hypothesis that serum and plasma perform equally well for the biomarkers of interest on this platform. Although analyte detection was high for both sample types and replicates in both serum and plasma were in general agreement, serum had slightly better replicate correlations and better reproducibility overall. In addition, analyte values derived from serum were better able to discriminate between the high and low skeletal muscle mass groups in our exploratory analyses, as evidenced by a lower error rate in identifying high versus low percent skeletal muscle mass in recursive partitioning and random forest models compared to plasma. There were also a higher number of serum derived markers (n=7) that distinguished between the high and low percent skeletal muscle mass groups compared to plasma (n=2) at a FDR<0.1 level. Therefore, based on the comparison of serum to EDTA treated plasma, serum would be the recommended blood fraction from healthy postmenopausal women to evaluate skeletal muscle mass associations with biomarkers on the RBM Human MAP version 1.6.

Some studies suggest that various sample preparations introduce assay interference dependent on the analyte of interest. (18, 19) For example, EDTA is a known chelator of divalent cations and may interfere with assays dependent on calcium or other divalent cations, while heparin may dilute samples during collection, which must be taken into account. (20) Citrate and heparin treated plasma samples were not available for this validation study to more globally compare serum to common plasma preparations (EDTA, citrate, and heparin treated plasma) by analyte. These samples are also not available for the large clinical trial to which the validation study results will be applied. Collecting biospecimens using multiple sample preparations is cost prohibitive and overly burdensome to participants in large clinical trials and observational studies where biomarker studies are often secondary endpoints. Additionally, multiple sample preparations were not applied in many large archived studies that continue to be utilized to investigate new research questions. This serves as a specific limitation, but also speaks to the relevance of this validation study. Since the optimal preparation method may vary by analyte, sample preparation should be carefully considered during assay planning in order to target optimization of the greatest number of analytes of interest when evaluating biomarker patterns without compromising primary study objectives.

Secondarily, we explored the ability of biomarkers to discriminate between the low and high percent lean mass phenotypes selected for the validation study. Only a small number of analytes on the panel were able to discriminate between high and low skeletal muscle mass phenotypes and were predominated by leptin, which is most likely segregating on adiposity in our extreme sampling that not only enriched on lean mass but also on BMI. Thus, factors previously associated with adiposity, including leptin (21), insulin (22), and some inflammatory factors, including serum amyloid A (SAA) (23, 24), complement 3 (25–28), and eotaxin (29, 30), were expectedly predominant in this pilot.

The main purpose of this validation study was to determine the best blood fraction for use in a larger epidemiologic study for which only serum or EDTA plasma are available to evaluate multiple candidate biomarkers related to skeletal muscle mass and potentially to reduce the number of candidate markers to be evaluated in the planned study. Results presented herein suggest that serum has the better reliability and reproducibility on this platform, in addition to the greatest number of discriminating biomarkers for body composition phenotype. Thus, serum will be used as our sample type for a larger planned association study of anabolic/catabolic factors in blood and the skeletal muscle mass phenotype in postmenopausal women. Biomarkers for which replicate correlations and detection were low in serum will be considered for elimination on the final biomarker panel planned for the large clinical trial.

Strengths and limitations

A strength of the study is the generalizability of our evaluation procedures to other studies of circulating biomarkers regarding best sample (serum or plasma), in spite of limited anticoagulant testing. Our study also reflects the normal aging population, which we expect to be studied intensively in the coming years with the aging of “baby boomers”. Additionally, the use of bootstrapping statistical techniques illustrates the utility of this approach for robust statistical interpretation of the small sample set while protecting biospecimens against waste in early planning of biomarker studies. Most current comparative papers evaluating multiplex assay techniques suggest early validation studies to assess dynamic range, linearity, CV, and percentage recovery by each analyte on a particular platform prior to large scale implementation for diagnostic or prognostic studies (1, 2). Our strategy here imparts significant cost savings for larger studies planning to evaluate biomarker associations by early culling of poorly measured biomarkers to achieve a smaller multiplex panel of analytes. Additionally, for those biomarkers for which the detection rate or performance characteristics are low, but for which compelling evidence suggests a role in a specific disease state (in our case sarcopenia or low skeletal muscle mass), such early biomarker analyses will direct early decision making regarding the need to use an alternative platform or acceptability of dichotomous (present/absent) analysis in place of continuous variables that may potentially limit interpretation. However, we must emphasize that the small sample size and multiple analyses performed on the sample set is also a limitation of this study. The discriminatory ability of cytokines for the prediction of body composition phenotype should be considered very preliminary and should be used for study design purposes only. The biomarkers identified as potential body composition discriminators may not be used for diagnostic purposes or identifying high risk groups at this time. Our planned application of serum for multiplex biomarker analyses to a large cohort is needed to replicate the discriminatory findings herein and to provide important information regarding biomarker predictors of body composition across the continuum of phenotypes.

Conclusion

We found that 60% of the analytes in both serum and plasma were largely equivalent, with a correlation ≥80%; however, our results slightly favor the use of serum for the parallel

evaluation of candidate biomarker predictors of lean mass phenotypes in postmenopausal women.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the participants and staff of the Women's Breast and Bone Density Study, Julie Buckmeier, and Leslie Arendell. Support for The WBBD study was provided by the Susan G. Komen Foundation, biomarker assessment by the National Institute on Aging (R01AG027373-01A2), and Dr. Bea by the National Cancer Institute (R25T, CA-78447). Publication of these results will not benefit the supporters.

List of abbreviations, in order cited

WBBD	Women's Breast and Bone Density Study
DXA	dual energy X-ray absorptiometry
EDTA	
MAP	multi-analyte panel
LDD	least detectable dose
LAL	lower assay limit
CV	coefficient of variation
FDR	false discovery rate
CI	confidence interval
EGF	Epidermal Growth Factor
ENA	Erythrocytic Nuclear Abnormalities
RAGE	Receptor for Advanced Glycation End products
EN-RAGE	Extracellular Newly identified RAGE-binding protein
FGF basic	Fibroblast Growth Factor
G-CSF	Granulocyte Colony-Stimulating Factor
GM-CSF	Granulocyte-Macrophage Colony Stimulating Factor
ICAM-1	Intercellular Adhesion Molecule
IFN	Interferon
Ig	Immunoglobulin
IGF	Insulin-like Growth Factor
IL	Interleukin
MCP	Monocyte Chemotactic Protein
MDC	Macrophage-Derived Chemokine
MIP	Macrophage Inflammatory Protein
MMP	Matrix Metalloproteinase
PAI	Plasminogen Activator Inhibitor

PAPP	Pregnancy Associated Plasma Protein
RANTES	Regulated upon Activation Normal T-cell Expressed, and presumably Secreted
SGOT	Serum Glutamic Oxaloacetic Transaminase
SHBG	Sex Hormone Binding Globulin
TIMP	Tissue Inhibitor of Metalloproteinases
TNF	Tumor Necrosis Factor
TNF R	TNF-Receptor
VCAM	Vascular Cell Adhesion Molecule
VEGF	Vascular Endothelial Growth Factor

References

1. Fu Q, Zhu J, Van Eyk JE. Comparison of multiplex immunoassay platforms. Clin Chem. 2010; 56:314–318. [PubMed: 20022982]
2. Ellington AA, Kullo IJ, Bailey KR, Klee GG. Antibody-based protein multiplex platforms: technical and operational challenges. Clin Chem. 2010; 56:186–193. [PubMed: 19959625]
3. Djoba Siawaya JF, Roberts T, Babb C, Black G, Golakai HJ, Stanley K, et al. An evaluation of commercial fluorescent bead-based luminex cytokine assays. PLoS One. 2008; 3:e2535. [PubMed: 18596971]
4. Thomson CA, Arendell LA, Bruhn RL, Maskarinec G, Lopez AM, Wright NC, et al. Pilot study of dietary influences on mammographic density in pre- and postmenopausal Hispanic and non-Hispanic white women. Menopause. 2007; 14:243–250. [PubMed: 17091096]
5. Chen Z, Wang Z, Lohman T, Heymsfield SB, Outwater E, Nicholas JS, et al. Dual-energy X-ray absorptiometry is a valid tool for assessing skeletal muscle mass in older women. J Nutr. 2007; 137:2775–2780. [PubMed: 18029498]
6. Vignali DA. Multiplexed particle-based flow cytometric assays. J Immunol Methods. 2000; 243:243–255. [PubMed: 10986418]
7. Breiman, L.; Friedman, J.; Olshen, R.; Stone, C. Classification and Regression Trees. 1st ed.. Wadsworth, New York: Chapman and Hall/CRC; 1984. 368 p.
8. Breiman L. Random forests. Machine Learning. 2001; 45:5–32.
9. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. 1995; 57:289–300.
10. Richens JL, Urbanowicz RA, Metcalf R, Corne J, O'Shea P, Fairclough L. Quantitative Validation and Comparison of Multiplex Cytokine Kits. J Biomol Screen. 2010 Feb 22. [Epub ahead of print].
11. Young SH, Antonini JM, Roberts JR, Erdely AD, Zeidler-Erdely PC. Performance evaluation of cytometric bead assays for the measurement of lung cytokines in two rodent models. J Immunol Methods. 2008; 331:59–68. [PubMed: 18089291]
12. Wong HL, Pfeiffer RM, Fears TR, Vermeulen R, Ji S, Rabkin CS. Reproducibility and correlations of multiplex cytokine levels in asymptomatic persons. Cancer Epidemiol Biomarkers Prev. 2008; 17:3450–3456. [PubMed: 19064561]
13. Hosnijeh FS, Krop EJ, Portengen L, Rabkin CS, Linseisen J, Vineis P, Vermeulen R. Stability and reproducibility of simultaneously detected plasma and serum cytokine levels in asymptomatic subjects. Biomarkers. 15:140–148. [PubMed: 19848603]
14. Lundblad RL. Considerations for the Use of Blood Plasma and Serum for Proteomic Analysis. The Internet Journal of Genomics and Proteomics. 2004; Vol. 1
15. Miles RR, Roberts RF, Putnam AR, Roberts WL. Comparison of serum and heparinized plasma samples for measurement of chemistry analytes. Clin Chem. 2004; 50:1704–1706. [PubMed: 15331516]

16. Boyanton BL Jr, Blick KE. Stability studies of twenty-four analytes in human plasma and serum. Clin Chem. 2002; 48:2242–2247. [PubMed: 12446483]

17. O'Keane MP, Cunningham SK. Evaluation of three different specimen types (serum, plasma lithium heparin and serum gel separator) for analysis of certain analytes: clinical significance of differences in results and efficiency in use. Clin Chem Lab Med. 2006; 44:662–668. [PubMed: 16681442]

18. Evans MJ, Livesey JH, Ellis MJ, Yandle TG. Effect of anticoagulants and storage temperatures on stability of plasma and serum hormones. Clinical Biochemistry. 2001; 34:107–112. [PubMed: 11311219]

19. Haab BB, Geierstanger BH, Michailidis G, Vitzthum F, Forrester S, Okon R, et al. Immunoassay and antibody microarray analysis of the HUPO Plasma Proteome Project reference specimens: systematic variation between sample types and calibration of mass spectrometry data. Proteomics. 2005; 5:3278–3291. [PubMed: 16038022]

20. Apweiler R, Aslanidis C, Deufel T, Gerstner A, Hansen J, Hochstrasser D, et al. Approaching clinical proteomics: current state and future fields of application in fluid proteomics. Clin Chem Lab Med. 2009; 47:724–744. [PubMed: 19527139]

21. Eriksson J, Valle T, Lindstrom J, Haffner S, Louheranta A, Uusitupa M, Tuomilehto J. Leptin concentrations and their relation to body fat distribution and weight loss--a prospective study in individuals with impaired glucose tolerance. DPS-study group. Horm Metab Res. 1999; 31:616–619. [PubMed: 10598830]

22. Kahn BB, Flier JS. Obesity and insulin resistance. J Clin Invest. 2000; 106:473–481. [PubMed: 10953022]

23. Gomez-Ambrosi J, Azcona C, Patino-Garcia A, Fruhbeck G. Serum Amyloid A concentration is increased in obese children and adolescents. J Pediatr. 2008; 153:71–75. [PubMed: 18571540]

24. Salas-Salvado J, Bullo M, Garcia-Lorda P, Figueredo R, Del Castillo D, Bonada A, Balanza R. Subcutaneous adipose tissue cytokine production is not responsible for the restoration of systemic inflammation markers during weight loss. Int J Obes (Lond). 2006; 30:1714–1720. [PubMed: 16652132]

25. Koistinen HA, Vidal H, Karonen SL, Dusserre E, Vallier P, Koivisto VA, Ebeling P. Plasma acylation stimulating protein concentration and subcutaneous adipose tissue C3 mRNA expression in nondiabetic and type 2 diabetic men. Arterioscler Thromb Vasc Biol. 2001; 21:1034–1039. [PubMed: 11397716]

26. Boggs RD, McCumbee WD, Cobbs SL, Todd DG, Kahle EB, Stewart NL, et al. Increased expression of complement component C3 in the plasma of obese Zucker fa and LA/N fa(f) rats compared with their lean counterparts. Obes Res. 1998; 6:361–367. [PubMed: 9738552]

27. Sniderman AD, Cianflone KM, Eckel RH. Levels of acylation stimulating protein in obese women before and after moderate weight loss. Int J Obes. 1991; 15:333–336. [PubMed: 1885258]

28. St-Pierre DH, Cianflone K, Smith J, Coderre L, Karelis AD, Imbeault P, et al. Change in plasma acylation stimulating protein during euglycaemic-hyperinsulinaemic clamp in overweight and obese postmenopausal women: a MONET study. Clin Endocrinol (Oxf). 2009; 70:539–546. [PubMed: 18702682]

29. Huber J, Kiefer FW, Zeyda M, Ludvik B, Silberhumer GR, Prager G, et al. CC chemokine and CC chemokine receptor profiles in visceral and subcutaneous adipose tissue are altered in human obesity. J Clin Endocrinol Metab. 2008; 93:3215–3221. [PubMed: 18492752]

30. Vasudevan AR, Wu H, Xydakis AM, Jones PH, Smith EO, Sweeney JF, et al. Eotaxin and obesity. J Clin Endocrinol Metab. 2006; 91:256–261. [PubMed: 16263831]

Table 1

Demographics of Study Participants by Percent Lean Mass Tail

	Lower		Upper		
	Mean	**SD**	**Mean**	**SD**	**p-value***
Age, years	58.91	2.77	59.33	2.24	0.672
Height, m	1.62	0.06	1.59	0.08	0.342
Weight, kg	97.34	16.35	56.71	5.36	<0.001
BMI, kg/m^2	36.82	4.98	22.51	2.44	<0.001
Years of Estrogen Exposure	33.55	9.8	37.22	10.08	0.237
Years of Hormone Therapy	3.22	3.27	1.4	2.77	0.747
Waist to Hip Ratio, cm*	0.84	0.05	0.76	0.05	0.011
Energy, kcal	2008.66	894.37	2162.57	716.33	0.676
Total Lean Mass, kg	46.18	8.05	37.9	2.57	0.053
Total Fat Mass, kg	45.57	8.33	15.41	3.29	<0.001
Total % Fat	48.45	1.44	27.67	3.76	<0.001
Lean to Fat Ratio	1.02	0.06	2.55	0.52	<0.001
Appendicular Lean Mass, kg	19.04	4.08	15.58	1.72	0.074
Appendicular Fat Mass, kg	22.06	5.26	7.68	1.65	<0.001
Appendicular Lean to Fat Ratio	0.87	0.12	2.1	0.47	0.053
Appendicular Skeletal Muscle Index, kg/m^2	7.18	1.22	6.17	0.61	0.035
	Lower		**Upper**		
	N	**%**	**N**	**%**	**p-value**
Smoking History					1.000
Never	4	36.4	3	33.3	
Former	5	45.5	4	44.4	
Current	2	18.2	2	22.2	
High Cholesterol					0.406
No	5	45.5	6	66.7	
Yes	6	54.6	3	33.3	

	Lower		Upper		
	Mean	**SD**	**Mean**	**SD**	**p-value***
High Blood Pressure					0.218
No	8	72.7	9	100.0	
Yes	3	27.3	0	0.0	
Osteoarthritis					0.094
No	7	63.4	9	100.0	
Yes	4	36.4	0	0.0	
Osteoporosis					1.000
No	10	90.9	9	100.0	
Yes	1	9.1	0	0.0	

* Statistical differences between groups was tested using Wilcoxon-rank sum test for continuous variables and Fisher's exact test for categorical variables.

Table 2

Analytes with serum to plasma correlations ≥ 80%.

	Serum				Plasma				
Analyte name	**Mean**	**SD**	**Range**		**Mean**	**SD**	**Range**		**r**
Serum to Plasma Correlation ≥ 0.90									
Thyroid Stimulating Hormone, uIU/mL	1.76	0.96	0.36	3.90	1.78	1.01	0.35	4.10	0.99
IgE, ng/mL	58.73	111.43	0.98	487.00	72.81	128.53	3.80	605.00	0.97
Lipoprotein (a), ug/mL	108.69	96.74	8.00	400.00	108.07	99.26	5.20	383.00	0.97
Insulin, uIU/mL	2.20	1.30	0.35	5.00	2.45	1.31	0.42	4.90	0.96
Adiponectin, ug/mL	5.66	2.43	1.70	10.40	5.25	2.32	1.50	9.20	0.95
Ferritin, ng/mL	144.52	96.76	22.70	364.00	132.63	90.30	22.50	337.00	0.95
Cancer Antigen 19-9, U/mL	7.42	7.81	0.59	35.30	10.14	9.72	0.42	45.10	0.93
IgM, mg/mL	0.87	0.40	0.08	1.60	0.86	0.39	0.08	1.90	0.93
TNF RII, ng/mL	5.85	4.61	2.90	26.90	5.44	4.38	2.80	25.70	0.92
α-1 Antitrypsin, mg/mL	2.18	0.57	1.10	3.90	2.14	0.56	1.20	3.70	0.91
Growth Hormone, ng/mL	2.27	3.18	0.08	12.90	1.94	2.40	0.08	9.10	0.91
Haptoglobin, mg/mL	1.93	1.13	0.46	4.00	2.00	1.22	0.47	4.40	0.91
IL-18, pg/mL	349.24	304.08	156.00	1740.00	340.60	310.10	137.00	1760.00	0.91
IL-16, pg/mL	432.88	151.16	224.00	756.00	350.88	117.58	193.00	586.00	0.90
MDC, pg/mL [a]	613.76	232.51	376.00	1320.00	495.60	195.67	277.00	1220.00	0.90
Serum to Plasma Correlation between 0.80 and 0.89									
IgA, mg/mL	1.12	0.40	0.54	2.00	1.14	0.41	0.53	2.00	0.89
Myoglobin, ng/mL	12.69	6.19	5.90	32.20	11.94	5.30	6.60	30.00	0.89
Apolipoprotein A1, mg/mL	0.28	0.09	0.13	0.45	0.29	0.10	0.13	0.49	0.88
Carcinoembryonic Antigen , ng/mL	1.27	0.78	0.28	3.20	1.53	0.92	0.44	3.40	0.88
SHBG, nmol/L	49.09	28.22	17.50	128.00	42.65	20.09	18.50	99.30	0.88
Apolipoprotein H, ug/mL [a, d]	230.25	44.60	91.20	298.00	231.89	50.46	70.20	319.00	0.87
Eotaxin, pg/mL [c]	273.35	273.46	94.20	1290.00	137.52	142.25	36.40	645.00	0.86
VCAM-1, ng/mL [c]	685.72	263.17	393.00	1810.00	650.76	218.19	396.00	1570.00	0.86
α-Fetoprotein, ng/mL	2.40	1.52	0.64	5.90	3.13	1.39	1.40	6.50	0.85
Factor VII, ng/mL	676.12	199.15	212.00	1110.00	803.88	223.32	266.00	1190.00	0.85

	Serum				Plasma				
Analyte name	**Mean**	**SD**	**Range**		**Mean**	**SD**	**Range**		**r**
β-2 Microglobulin, ug/mL	2.05	1.12	1.30	6.80	2.02	0.99	1.20	6.00	0.83
Leptin, ng/mL	26.18	26.16	1.30	90.80	30.17	28.31	1.30	102.00	0.83
α-2 Macroglobulin, mg/mL	1.03	0.16	0.75	1.30	1.05	0.15	0.76	1.30	0.82
Fatty Acid Binding Protein, ng/mL *b, c*	1.55	1.48	0.40	4.50	1.84	1.71	0.47	6.50	0.80

Replicate correlations within sample type for each analyte are ≥0.90 unless otherwise noted by superscript a–d.
Replicate correlations within sample type: serum - a) 0.80 – 0.89; b) 0.70 – 0.79; plasma - c) 0.80 – 0.89; d) < 0.60

All analytes were detected in at least 50% of the samples for both serum and plasma.

Table 3

Analytes with serum to plasma correlations < 50%.

	Serum				Plasma				
Analyte name	**Mean**	**SD**	**Range**		**Mean**	**SD**	**Range**		**r**
Stem Cell Factor, pg/mL[f]	374.08	148.74	113.00	738.00	366.05	159.48	86.20	947.00	0.49
IL-12p40, ng/mL[b, d]	0.25	0.18	0.10	0.92	0.30	0.25	0.13	1.10	0.47
Prostatic Acid Phosphatase, ng/mL	0.33	0.10	0.23	0.55	0.18	0.03	0.14	0.29	0.44
FGF basic, pg/mL[c, e]	645.68	2366.22	64.90	12000.00	803.04	2937.76	103.00	14900.00	0.43
PAI-1, ng/mL	212.44	51.78	104.00	305.00	48.07	21.47	10.00	101.00	0.41
TNF-α, pg/mL[a, e]	6.21	3.78	2.20	16.50	6.62	5.62	2.80	30.80	0.39
VEGF, pg/mL	2048.44	5563.40	417.00	28715.00	1745.12	5621.53	432.00	28715.00	0.37
IL-5, pg/mL[a, e]	9.79	16.76	1.80	81.30	9.77	15.62	1.80	74.60	0.36
CD40 Ligand, ng/mL	2.51	1.12	0.70	4.70	0.32	0.39	0.04	2.10	0.35
EGF, pg/mL	324.40	158.30	67.20	651.00	217.41	325.70	57.90	1670.00	0.32
Fibrinogen, mg/mL[f]	0.02	0.01	0.01	0.04	4.82	1.46	2.60	8.50	0.32
EN-RAGE, ng/mL	82.78	96.27	28.70	526.00	7.78	27.56	0.44	140.00	0.29
IL-8, pg/mL	46.70	50.77	13.30	242.00	16.55	12.56	7.40	72.60	0.29
Brain-Derived Neurotrophic Factor, ng/mL	30.00	7.56	14.10	42.50	2.79	2.02	0.80	10.20	0.26
Erythropoietin, pg/mL	257.44	529.98	25.20	1970.00	199.93	184.55	25.20	667.00	0.26
IFN-γ, pg/mLe	2.11	1.80	0.82	8.20	5.64	7.92	1.00	32.80	0.26
IL-1β, pg/mL[c, e]	0.59	0.43	0.29	1.60	0.54	0.41	0.29	2.00	0.26
RANTES, ng/mL	26.41	12.26	7.50	51.20	10.16	10.61	2.20	53.80	0.25
IL-12p70, pg/mL[c, f]	54.48	8.07	42.10	76.50	65.40	16.87	45.90	124.00	0.24
Myeloperoxidase, ng/mL[e]	1713.20	757.39	617.00	3750.00	151.53	105.58	24.70	600.00	0.24
SGOT, ug/mL[c, f]	6.38	1.74	2.40	10.90	8.85	1.91	5.00	13.90	0.22
IL-7, pg/mL [f]	101.54	29.84	45.20	169.00	48.24	17.49	20.00	95.80	0.20
Thrombopoietin, ng/mL[c, e]	4.00	0.68	2.20	5.50	1.82	1.22	0.50	6.40	0.20
Glutathione S-Transferase, ng/mL[a, f]	0.73	0.19	0.54	1.30	0.62	0.09	0.49	0.83	0.16

Replicate correlations within sample type for each analyte are ≥0.90 unless otherwise noted by superscript a–e.
Replicate correlations within sample type: serum - a) 0.80 – 0.89, b) 0.70 – 0.79, c) <0.70; plasma - d) 0.80 – 0.89; e) 0.70 – 0.79, f) <0.70

All analytes were detected in at least 50% of the samples for both serum and plasma.

Table 4

Summary of Discriminating Errors for Serum and Plasma Markers.

	Mean	Standard Deviation	Min	1st Quartile	Median	3rd Quartile	Max	p-value
Recursive Partitioning Models								<0.001
Serum	.119	.085	0	.05	.10	.15	.40	
Plasma	.170	.100	0	.10	.15	.25	.45	
Random Forest Models								<0.001
Serum	.081	.050	0	.05	.10	.10	.25	
Plasma	.097	.060	0	.05	.10	.15	.30	

Significant differences between serum and plasma for both the recursive partitioning and random forest models were tested using the Wilcoxon Rank Sum test.