Title

Reproducibility and Variability of Protein Analytes Measured Using a Multiplexed Modified Aptamer Assay.

Journal

The Journal of Applied Laboratory Medicine, 4(1)

ISSN

2576-9456

Authors

Tin, Adrienne
Yu, Bing
Ma, Jianzhong
et al.

Publication Date

2019-07-01

DOI

10.1373/jalm.2018.027086

Peer reviewed

HHS Public Access

Author manuscript

J Appl Lab Med. Author manuscript; available in PMC 2019 October 25.

Published in final edited form as:

J Appl Lab Med. 2019 July ; 4(1): 30–39. doi:10.1373/jalm.2018.027086.

Reproducibility and Variability of Protein Analytes Measured Using a Multiplexed Modified Aptamer Assay

Adrienne Tin[1,2,*,†], **Bing Yu**[3,†], **Jianzhong Ma**[3], **Kunihiro Masushita**[1,2], **Natalie Daya**[1,2], **Ron C. Hoogeveen**[4], **Christie M. Ballantyne**[4], **David Couper**[5], **Casey M. Rebholz**[1,2], **Morgan E. Grams**[1,6], **Alvaro Alonso**[7], **Thomas Mosley**[8], **Gerardo Heiss**[9], **Peter Ganz**[10], **Elizabeth Selvin**[1,2], **Eric Boerwinkle**[3,11], **Josef Coresh**[1,2]

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD;

[2]Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins University, Baltimore MD;

[3]Department of Epidemiology, Human Genetics and Environmental Sciences, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX;

[4]Section of Cardiovascular Research, Department of Medicine, Baylor College of Medicine, Houston TX;

[5]Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC;

[6]Division of Nephrology, Johns Hopkins School of Medicine, Baltimore, MD;

[7]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA;

[8]Department of Neurology, The University of Mississippi Medical Center, Jackson, MS;

[9]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC;

[*]**Address correspondence to this author at:** 615 N Wolfe Street, W6017, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205. Fax 410-955-0863; atin1@jhu.edu.

[†]A. Tin and B. Yu contributed equally to this work.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form.* **Employment or Leadership:** P. Ganz, SomaLogic. **Consultant or Advisory Role:** C.M. Ballantyne, Abbott, Roche; P. Ganz, SomaLogic. **Stock Ownership:** None declared. **Honoraria:** None declared. **Research Funding:** The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HSN268201100010C, HHSN268201100011C, and HHSN268201100012C). Neurocognitive and visit 5–7 data are collected by U01 2U01HL096812, 2U01HL096814, 2U01HL096899, 2U01HL096902, 2U01HL096917 from the NIH (NHLBI, NINDS, NIA and NIDCD). Lab work was partially supported by R01 HL134320 and DK089174. SomaLogic conducted the SOMAscan assay for this study at no charge. Johns Hopkins University has signed a collaboration agreement with SomaLogic to conduct SOMAscan of ARIC stored samples at no charge in exchange for the rights to analyze linked ARIC phenotype data. B. Yu, American Heart Association (17SDG33661228); K. Matsushita, NHLBI; C.M. Rebholz, a mentored research scientist development award from the National Institute of Diabetes and Digestive and Kidney Diseases (K01 DK107782); E. Boerwinkle, National Institutes of Health - ARIC Lipid lap contract. **Expert Testimony:** None declared. **Patents:** C.M. Ballantyne, 61721475.

[10]Division of Cardiology, San Francisco General Hospital, University of California, San Francisco, CA;

[11]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX.

Abstract

Background: There is growing interest in the use of multiplexed aptamer-based assays for large-scale proteomic studies. However, the analytic, short- and long-term variation of the measured proteins is largely uncharacterized.

Methods: We quantified 4001 plasma protein analytes from 42 participants in the Atherosclerosis Risk in Communities (ARIC) Study in split samples and at multiple visits using a multiplexed modified aptamer assay. We calculated the CV, Spearman correlation, and intraclass correlation (ICC) between split samples and evaluated the short-term (4–9 weeks) and long-term (approximately 20 years) variability using paired *t*-tests with log-transformed protein concentrations and Bonferroni-corrected significance thresholds. We performed principal component (PC) analysis of protein analyte concentrations and evaluated their associations with age, sex, race, and estimated glomerular filtration rate (eGFR).

Results: The mean baseline age was 57 years at the first visit, 43% of participants were male and 57% were white. Among 3693 protein analytes that passed quality control, half (n = 1846) had CVs < 5.0%, Spearman correlations > 0.89, and ICCs > 0.96 among the split samples. Over the short term, only 1 analyte had a statistically significant difference between the 2 time points, whereas, over approximately 20 years, 866 analytes (23.4%) had statistically significant differences ($P < 1.4 \times 10^{-5}$, 681 increased, 185 decreased). PC1 had high correlations with age (−0.73) and eGFR (0.60). PC2 had moderate correlation with male sex (0.18) and white race (0.31).

Conclusions: Multiplexed modified aptamer technology can assay thousands of proteins with excellent precision. Our results support the potential for large-scale studies of the plasma proteome over the lifespan.

Circulating concentrations of specific proteins are routinely measured in clinical research and for medical diagnosis and prognosis. High-throughput identification and quantification of proteins with multiplexed technology provide opportunities for mining the human proteome to improve our understanding of disease processes, link genetic variants to blood protein regulatory networks, and potentially discover novel biomarkers for early detection and disease diagnosis. Historically, immunoaffinity-based methods that target a single protein are used most commonly for clinical assays (1). Cross-reactivity of affinity reagents limits the number of proteins that can be evaluated in a multiplexed immunoaffinity-based assay to the low hundreds (2). Single-stranded DNA aptamer-based assays have been developed as an alternative to immunoaffinity-based methods (3) and have been increasingly applied in clinical and epidemiological studies (4–8). For example, a panel of 9 proteins identified from thousands has been shown to predict cardiovascular disease better than traditional risk factors (4).

The human proteome is dynamic and sensitive to external stimuli. There may be considerable short-term (i.e., weeks) and long-term (i.e., years) variation in addition to sample handling and laboratory (methodologic) variability. Understanding the laboratory reproducibility of aptamer-based assays, as well as short-term and long-term biological variation of proteins, is critical for designing longitudinal studies of the proteome. Previous studies have reported overall intra- and interassay median CV of 4% to 8% for multiplexed aptamer-based assays (4, 6, 9). Concentrations of specific proteins may vary by age, race, and sex and be influenced by the function of the kidney, which filters, reabsorbs, and catabolizes many low molecular weight proteins (10). Few studies have investigated short- and long-term variation of a large number of plasma proteins along with laboratory reproducibility. We conducted a study to assess the laboratory reproducibility, short-term, and long-term biological variation in >3500 plasma protein analytes measured by a multiplexed modified aptamer assay. The plasma samples were collected at multiple time points from 42 participants selected from the Atherosclerosis Risk in Communities (ARIC) Study using stratified sampling based on race, sex, and kidney function.

METHODS

Study population

The ARIC study is a prospective cohort study of 15792 adults (45–64 years old at visit 1 in 1987–1989) recruited from 4 US communities: Forsyth County, North Carolina; Jackson, Mississippi; suburban Minneapolis, Minnesota; Washington County, Maryland (11). For the purpose of this study, we randomly selected 42 participants from those with data at visits 2 and 5, and the repeated visit 5 from 7 strata (6 participants each) defined by race (self-reported white or black), sex (male or female), and estimated glomerular filtration rate (eGFR) at visit 5 (30–59 mL/min/1.73m^2 or 60–100 mL/min/1.73m^2; see Table 1 in the Data Supplement that accompanies the online version of this article at http://www.jalm.org/content/vol4/issue1). The sample size of 42 was determined on the basis of having 80% power to detect a minimum Spearman correlation of 0.43 at an α concentration of 0.05 for a 2-sided test. eGFR concentration was calculated from serum creatinine after calibration across visits using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation (12, 13). Protein analyte levels from these 42 participants were measured using previously unthawed EDTA plasma collected at visit 2 (1990–1992), visit 5 (2011–2013), and a repeated visit 5 (4–9 weeks after visit 5). Plasma samples collected during the repeated visit 5 were split as 2 vials for assessing technical reproducibility with their identities masked to personnel processing the modified aptamer assay. The ARIC protocol for blood sample collection and processing was designed to minimize spontaneous biochemical reactions after blood collection and is consistent with recommended practice for epidemiological studies (14, 15). In brief, blood samples were put immediately in an ice water bath after venipuncture. Centrifugation was then performed within 10 min after venipuncture at room temperature (15–25 °C). After centrifugation, the aliquots were stored at −80 °C within 90 min from venipuncture.

Modified aptamer protein assay, quality control, and normalization

Concentrations of protein analytes were quantified using a multiplexed modified DNA-based aptamer technology (SOMAscan assay), which transforms individual protein concentrations into a corresponding modified aptamer (SOMAmer reagent) concentration that can be quantified by standard DNA microarrays in relative fluorescence units (RFUs) (16, 17). The SOMAscan assay included 4001 aptamers that mapped to 3581 unique proteins or protein complexes in the Universal Protein Resource (UniProt) databases. Our unit of analysis included the measures of protein analyte from each aptamer quantified in RFUs.

Protein analyte measurements underwent the regular SOMAscan data standardization and normalization process (18). Briefly, hybridization control normalization was first applied to each sample based on a set of hybridization control sequences to correct for systematic biases during hybridization. Second, median signal normalization was applied to measures within a plate to remove sample or assay biases that may be because of pipetting variation, variation in reagent concentrations, assay timing, and other sources of systematic variability within a single plate run. Finally, each plate contained calibrator samples for each SOMAmer reagent, which was used to correct for plate-to-plate variation based on established global reference standards. Protein analytes with calibration factor greater or less than the median calibration factor (0.4) were excluded from all analyses.

Traditional protein assays

Measurements of 9 proteins assessed in the multiplex aptamer-based assay were previously quantified using clinical assays in ARIC (9 were available at visit 2 and 8 at visit 5): albumin (bromcresol green colorimetric assay), alanine aminotransferase (ALT; a kinetic reaction assay), β_2-microglobulin (B2M; an immunoturbidimetric assay), high-sensitivity C-reactive protein (CRP; an immunoturbidimetric assay), cystatin C (an immunoturbidimetric assay), N-terminal pro-B-type natriuretic peptide (NT-proBNP; a sandwich immunoassay), parathyroid hormone (PTH; a sandwich immunoassay), high-sensitivity cardiac troponin T (hs-cTnT; an immunoassay), and thyroid-stimulating hormone (TSH; a sandwich immunoassay). All clinical assays were conducted using Roche Diagnostics instruments, except for CRP at visit 5, which used a Beckman Coulter instrument. In contrast to the multiplexed assays, all of the clinical assays used serum, except for hs-cTnT at visit 5, which was measured in plasma. All clinical assays had CVs < 10%. Details are reported in Table 2 in the online Data Supplement.

Analysis of reproducibility.—To verify the split sample identity, we performed hierarchical clustering of the protein analyte measures using Ward's distance and compared the Spearman correlation coefficients between the protein analyte measures from the split samples vs those from 2 randomly selected samples.

To quantify reproducibility of the protein measures, we computed the CV, Spearman correlation, and intraclass correlation (ICC) of the measures of protein analytes between the split samples. While CV and ICC quantify the magnitude of the variation between the split samples, Spearman correlation compares the ranking of the measures between the split samples. CV was defined as $\exp(\sqrt{mean(variance(log\ (RFU)))}) - 1$ (19). Compared with the

common definition of CV (average of the SD divided by the mean of blind duplicate samples), this alternative definition avoids underestimation of the CV when the SD of the blind duplicate samples is dependent on the mean. We also evaluated CV stratified by eGFR at 60 mL/min/m^2. The 95% CI for the Spearman correlation was estimated using the Bonett and Wright method (20). ICC was estimated using 1-way ANOVA (21).

To evaluate reproducibility across platforms, we computed the Spearman correlation coefficients for 9 proteins that were measured by both clinical assays and the modified aptamer assay at visit 2 and 8 proteins at visit 5. For the 8 proteins with available measures from clinical assays at both visits 2 and 5, we computed change from visit 2 to visit 5 using values from the clinical assays and the modified aptamer assay, then computed Spearman correlation coefficient between the changes obtained from the clinical assays and the modified aptamer assay.

Analysis of short- and long-term variability

To reduce skewness, log base 2 transformation was applied for the analysis of variability. Statistical significance was determined using the Bonferroni method. To evaluate short-term variability, we compared the measures of the protein analytes at visit 5 with those at the repeated visit 5 (4–9 weeks after visit 5). Given that 2 measures were available for each protein analyte at the repeated visit 5 from the reproducibility study, we randomly selected 1 of the 2 measures to be included in the dataset for short-term variability analysis. To evaluate long-term variability, we used the measures of the protein analytes at visits 2 and 5 (approximately 20 years apart). To contrast assay reproducibility in split samples with biological short- and long-term variability, we evaluated whether the paired measures between split samples, in the short term and long term, differed significantly using paired *t*-tests. To investigate whether heterogeneity in protein analyte concentrations might have changed over the long term, we tested for the difference in variance between visits 2 and 5 for the paired protein concentrations (22).

To explore the extent to which long-term variation may be related to age, sex, race, and kidney function, we performed principal component analysis (PCA) using the measures at visits 2 and 5 and computed the Spearman correlation coefficients of the first 10 principal components (PCs) with these 4 demographic and kidney function variables. All analyses were conducted using R version 3.4.1.

RESULTS

Participant characteristics and summary of protein analytes

Among the participants, 43% were male and 57% were white. The mean age at visit 2 (1990–1992) was 57 years, and the mean age at visit 5 (2011–2013) was 76 years (Table 1). Of the 4001 protein analytes, 92% (3693) passed the calibration filter and were included in our analyses. These 3693 protein analytes mapped to 3323 proteins or protein complexes in UniProt, of which 90.3% (3001) were tagged by 1 modified aptamer. Three were human virus proteins.

Split sample reproducibility of protein analytes from aptamer-based assay

We limited the analysis to 40 of the 42 split sample pairs (2 pairs were excluded because of suspected mislabeling as evidenced by mispairing in a nearest neighbor dendrogram; see Fig. 1 in the online Data Supplement). Among the 40 split sample pairs, the median (25th and 75th percentile) of the reproducibility metrics was as follows: CV, 5.0 (4.1, 6.9); ICC, 0.96 (0.92, 0.98); and Spearman correlation coefficient, 0.89 (0.81, 0.94; Fig. 1). Overall, the 3 metrics yielded consistent results (see Fig. 2 in the online Data Supplement). For example, of the 3287 protein analytes with CVs < 10%, 96.7% had high ICC (>0.80) and 81% had high Spearman correlation coefficients (>0.80). Table 3 in the online Data Supplement presents these reproducibility metrics (CV, ICC, and Spearman correlation coefficients) of all 1830 protein analytes with CV < 5%. Among the split samples from participants with eGFR below and ≥60 mL/min/m^2, the CVs were similar—median CV (25th and 75th percentile), eGFR < 60 mL/min/m^2, 5.0 (4.1, 6.9), and eGFR ≥ 60 mL/min/m^2, 4.9 (4.0, 6.4).

Compared with clinical assays, 6 of the 9 proteins with available clinical measures had high correlations with measures from the modified aptamer assay at visit 2: CRP, TSH, cystatin C, B2M, PTH, and NT pro-BNP (Spearman correlations >0.8; Table 2). Five of these 6 proteins also had high correlations (Spearman correlations > 0.8) at visit 5, and clinical measures for the sixth (PTH) was not available at visit 5. The long-term changes of these 5 proteins measured from clinical assays were also highly correlated with the change in these 5 proteins as measured by the modified aptamer assay (Spearman correlations > 0.75).

Difference between split samples and short- and long-term variabilities

None of the protein analytes had statistically significant differences in concentrations between the split samples. Over the short term (4–9 weeks), only 1 protein analyte had a statistically significant difference in concentrations ($P < 1.35 \times 10^{-5} = 0.05/3693$; see Table 4A in the online Data Supplement). Over the long term (approximately 20 years), 866 protein analytes (23.4%) had statistically significant differences in concentrations: 681 (18.4%) increased and 185 (5.0%) decreased (see Table 4B in the online Data Supplement). The median fold change (25th and 75th percentiles) was 0.52 (0.33, 0.82) among those with a significant increase in concentrations, and −0.37 (−0.60, −0.25) among those with a significant decrease in concentrations. The protein analytes with statistically significant increases included the following low-molecular-weight kidney function biomarkers: B2M (fold change, 0.56, $P = 4.6 \times 10^{-17}$) and cystatin C (fold change, 0.38, $P = 1.1 \times 10^{-13}$). Fig. 2 presents the gaussian kernel smoothed distributions of the differences between the split samples and the fold changes in the short and long terms. Fig. 3 in the online Data Supplement presents the scatter plots of measures of the top 10 protein analytes with the most statistically significant long-term changes. Regarding the difference in variance over the long term, 284 protein analytes (7.7%) had a significant change: 163 (4.4%) increased and 121 (3.3%) decreased (see Table 4C in the online Data Supplement). The top PCs generated from the measures of protein analytes from visits 2 and 5 had moderate-to-high correlations with key demo-graphic and clinical variables. PC1 had high correlations with both age (−0.73) and eGFR (0.60; see Fig. 4 and Table 5A in the online Data Supplement). PC2 had moderate correlations with male sex (0.18) and white race (0.31). The 50 protein

analytes with the highest absolute loadings in PC1 and PC2 are reported in Table 5, B and C in the online Data Supplement.

DISCUSSION

In this study, we evaluated the technical reproducibility and the short- and long-term biological variability of 3693 protein analytes measured by a commercial multiplexed modified aptamer assay (SOMAscan). Measures of the majority of the protein analytes were reproducible, demonstrated by split sample CVs of <10%. Compared to measures of 9 proteins quantified using clinical assays, 6 proteins were highly correlated across platforms, including long-term changes for 5 proteins: CRP, TSH, cystatin C, B2M, and NT pro-BNP. Over the short term, only 1 protein analyte had a statistically significant change, although, over the long term, almost a quarter of the protein analytes (23.4%) had a statistically significant change. The long-term variation of the protein analytes as captured by PCs was correlated with key demographics and clinical variables. These results suggest stability in the protein analytes over the short term and the ability to quantify longer term changes in studying the proteome as measured by this multiplexed modified aptamer technology.

The high reproducibility of this multiplexed modified aptamer assay in our study is consistent with the results from previous studies (4, 6, 9, 23, 24), which reported median CVs of 4% to 8%. Few prior studies have evaluated measures of the protein analytes over time and across platforms (6, 25–27). We compared short-term (4–9 weeks) and long-term (approximately 20 years) changes in the same individuals across platforms. Our comparison included important biomarkers of chronic diseases and hormones: CRP, an acute-phase reactant; NT-proBNP, a marker of heart failure; cystatin C and B2M, markers of kidney function; and TSH and PTH, 2 key hormones involved in multiple metabolic processes (28–30). Three proteins, albumin, ALT, and hs-cTnT, had only moderate correlations between the clinical assays and the modified aptamer assay, suggesting that, for some proteins, the established and the modified aptamer methods may differ meaningfully with implications for estimating disease associations. Over the long term, substantially more protein analytes increased (18%) rather than decreased (5%). The increase in protein concentrations over the long term may be partly attributed to reduced kidney function in older age because the kidney plays an important role in the removal of low-molecular-weight proteins (31, 32). In aggregate, the first PC generated by protein measures over the long term was also negatively correlated with eGFR, a measure of kidney function.

Our study included community-dwelling adults with a range of demographic and clinical characteristics. We evaluated reproducibility using multiple metrics, including a comparison of long-term changes between 2 assay platforms. Nonetheless, some limitations of the study warrant mention. The number of proteins compared with clinical assays was limited. We sought to quantify the overall pattern of reproducibility and variability instead of making inference on any specific protein analyte. Hence, we designed a study of moderate sample size. The split-sample assays were conducted in 1 day on 2 plates and therefore we only quantified intraday, rather than interday, laboratory reproducibility or drift. Our experimental design could not exclude the possibility of protein degradation during long-term storage. The overall low proportion of protein analytes having significant change in variance between

visits 2 and 5 (7.7%) did not support widespread protein degradation. Studies of stability of individual proteins among the thousands quantified after long-term storage will be useful. Further, the modified aptamer technology provides relative quantification instead of absolute quantification, and the specificity of each modified aptamer may need to be validated in future studies.

In conclusion, in this study of 3693 protein analytes at 3 time points over 20 years, we found a commercial modified aptamer-based technology to be highly reproducible. These results support the use of this technology in clinical and epidemiologic studies of the proteome for identifying disease determinants and estimating disease risk.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

The authors thank the staff and participants of the ARIC study for their important contributions.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, and final approval of manuscript.

Nonstandard abbreviations:

ARIC	Atherosclerosis Risk in Communities
eGFR	estimated glomerular filtration rate
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
RFU	relative fluorescence unit
ALT	alanine aminotransferase
B2M	β_2-microglobulin
CRP	C-reactive protein
NT-proBNP	N-terminal pro-B-type natriuretic peptide
PTH	parathyroid hormone
hs-cTnT	high-sensitivity cardiac troponin T
TSH	thyroid-stimulating hormone
ICC	intraclass correlation
PCA	principal component analysis
PC	principal component

REFERENCES

1. Ellington AA, Kullo IJ, Bailey KR, Klee GG. Antibody-based protein multiplex platforms: technical and operational challenges. Clin Chem 2010;56:186–93. [PubMed: 19959625]
2. Smith JG, Gerszten RE. Emerging affinity-based proteomic technologies for large-scale plasma profiling in cardiovascular disease. Circulation 2017;135:1651–64. [PubMed: 28438806]
3. Taussig MJ, Schmidt R, Cook EA, Stoevesandt O. Development of proteome-wide binding reagents for research and diagnostics. Proteomics Clin Appl 2013;7: 756–66. [PubMed: 24178846]
4. Ganz P, Heidecker B, Hveem K, Jonasson C, Kato S, Segal MR, et al. Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. JAMA 2016;315:2532–41. [PubMed: 27327800]
5. Heier CR, Fiorillo AA, Chaisson E, Gordish-Dressman H, Hathout Y, Damsker JM, et al. Identification of pathway-specific serum biomarkers of response to glucocorticoid and infliximab treatment in children with inflammatory bowel disease. Clin Transl Gastroenterol 2016;7:e192. [PubMed: 27628422]
6. Ngo D, Sinha S, Shen D, Kuhn EW, Keyes MJ, Shi X, et al. Aptamer-based proteomic profiling reveals novel candidate biomarkers and pathways in cardiovascular disease. Circulation 2016;134:270–85. [PubMed: 27444932]
7. Sattlecker M, Khondoker M, Proitsi P, Williams S, Soininen H, Kloszewska I, et al. Longitudinal protein changes in blood plasma associated with the rate of cognitive decline in Alzheimer's disease. J Alzheimers Dis 2016;49:1105–14. [PubMed: 26599049]
8. Benson MD, Yang Q, Ngo D, Zhu Y, Shen D, Farrell LA, et al. Genetic architecture of the cardiovascular risk proteome. Circulation 2018;137:1158–72. [PubMed: 29258991]
9. Candia J, Cheung F, Kotliarov Y, Fantoni G, Sellers B, Griesman T, et al. Assessment of variability in the somascan assay. Sci Rep 2017;7:14248. [PubMed: 29079756]
10. Ferguson MA, Waikar SS. Established and emerging markers of kidney function. Clin Chem 2012;58:680–9. [PubMed: 22311920]
11. The Atherosclerosis Risk in Communities (ARIC) study: design and objectives. The ARIC investigators. Am J Epidemiol 1989;129:687–702. [PubMed: 2646917]
12. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150: 604–12. [PubMed: 19414839]
13. Parrinello CM, Grams ME, Couper D, Ballantyne CM, Hoogeveen RC, Eckfeldt JH, et al. Recalibration of blood analytes over 25 years in the Atherosclerosis Risk in Communities study: impact of recalibration on chronic kidney disease prevalence and incidence. Clin Chem 2015;61:938–47. [PubMed: 25952043]
14. Tworoger SS, Hankinson SE. Collection, processing, and storage of biological samples in epidemiologic studies: sex hormones, carotenoids, inflammatory markers, and proteomics as examples. Cancer Epidemiol Biomarkers Prev 2006;15:1578–81. [PubMed: 16985015]
15. Rai AJ, Gelfand CA, Haywood BC, Warunek DJ, Yi J, Schuchard MD, et al. Hupo plasma proteome project specimen collection and handling: towards the standardization of parameters for plasma proteome samples. Proteomics 2005;5:3262–77. [PubMed: 16052621]
16. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, et al. Aptamer-based multiplexed proteomic technology for biomarker discovery. PLoS One 2010;5:e15004. [PubMed: 21165148]
17. UNIPROT. Universal protein resource (uniprot) databases. http://www.Uniprot.org/ (Accessed December 2017).
18. SomaLogic. Somascan® data standardization. http://somalogic.com/wp-content/uploads/2017/06/SSM-071-Rev-0-Technical-Note-SOMAscan-Data-Standardization.pdf (Accessed September 2017).
19. Bland JM, Altman DG. Measurement error proportional to the mean. BMJ 1996;313:106. [PubMed: 8688716]
20. Bonett D, Wright T. Sample size requirements for estimating Pearson, Kendall, and Spearman correlations. Psychometrika 2000;65:23–8.

21. Gamer M, Lemon J, Singh P. irr: Various coefficients of interrater reliability and agreement. https://rdrr.io/cran/irr/ (Accessed December 2018).
22. Pitman EJG. A note on normal correlation. Biometrika 1939;31:9–12.
23. Sun B, Maranville J, Peters J, Stacey D, Staley J, Blackshaw J, Stephen B. Consequences of natural perturbations in the human plasma proteome. Nature 2018;558:73–79. [PubMed: 29875488]
24. Williams SA, Murthy AC, DeLisle RK, Hyde C, Malarstig A, Ostroff R, et al. Improving assessment of drug safety through proteomics: early detection and mechanistic characterization of the unforeseen harmful effects of torcetrapib. Circulation 2018;137:999–1010. [PubMed: 28974520]
25. De Groote MA, Higgins M, Hraha T, Wall K, Wilson ML, Sterling DG, et al. Highly multiplexed proteomic analysis of quantiferon supernatants to identify biomarkers of latent tuberculosis infection. J Clin Microbiol 2017;55:391–402. [PubMed: 27852671]
26. Christiansson L, Mustjoki S, Simonsson B, Olsson-Strömberg U, Loskog ASI, Mangsbo SM. The use of multiplex platforms for absolute and relative protein quantification of clinical material. EuPA Open Proteom 2014;3:37–47.
27. Lim SY, Lee JH, Welsh SJ, Ahn SB, Breen E, Khan A, et al. Evaluation of two high-throughput proteomic technologies for plasma biomarker discovery in immunotherapy-treated melanoma patients. Biomark Res 2017;5:32. [PubMed: 29152306]
28. Grewal J, McKelvie RS, Persson H, Tait P, Carlsson J, Swedberg K, et al. Usefulness of N-terminal pro-brain natriuretic peptide and brain natriuretic peptide to predict cardiovascular outcomes in patients with heart failure and preserved left ventricular ejection fraction. Am J Cardiol 2008;102:733–7. [PubMed: 18773998]
29. Pepys MB. C-reactive protein fifty years on. Lancet 1981; 1:653–7. [PubMed: 6110874]
30. Astor BC, Shafi T, Hoogeveen RC, Ballantyne C, Coresh J. Increasing levels of novel kidney function markers predict outcomes in the general population: the Atherosclerosis Risk in Communities (ARIC) study. Philadelphia: American Society of Nephrology; 2011.
31. Maack T, Johnson V, Kau ST, Figueiredo J, Sigulem D. Renal filtration, transport, and metabolism of low-molecular-weight proteins: a review. Kidney Int 1979;16: 251–70. [PubMed: 393891]
32. Carone FA, Peterson DR, Oparil S, Pullman TN. Renal tubular transport and catabolism of proteins and peptides. Kidney Int 1979;16:271–8. [PubMed: 529676]

IMPACT STATEMENT

Proteomic assay has increasingly been used in epidemiological studies. We quantified plasma proteins from 42 individuals at multiple time points using a modified aptamer assay and assessed reproducibility, short- and long-term variation. Among 3693 protein analytes, half had CV <5.0% between split samples. Over 4–9 weeks, 1 analyte had a statistically significant difference ($P < 1.4 \times 10^{-5}$) vs 23.4% over approximately 20 years. Principal component (PC) 1 had high correlation with age (−0.73) and estimated glomerular filtration rate (0.60). Multiplexed modified aptamer technology can assay thousands of proteins with excellent precision, supporting the potential for large-scale longitudinal studies of the plasma proteome.

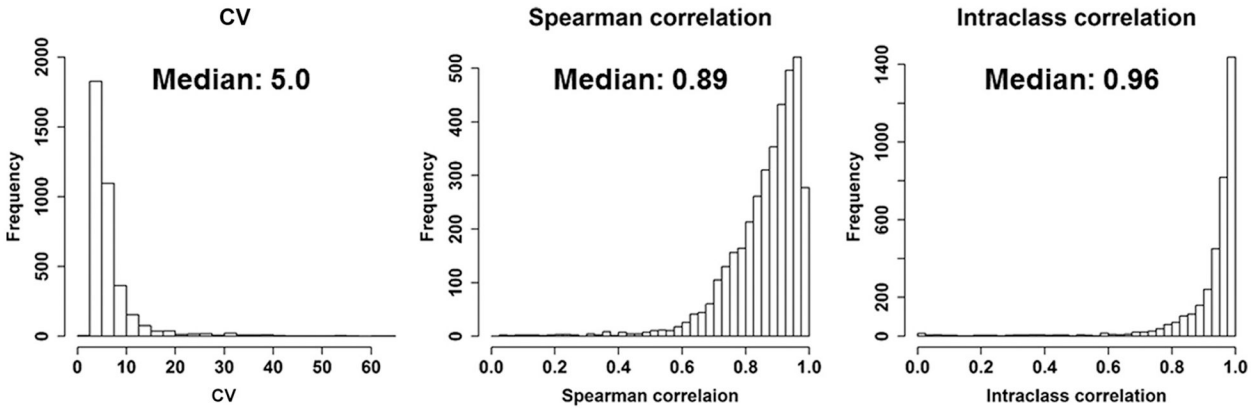

Fig. 1.

Measures of the reproducibility of 3693 protein analytes in 40 split-sample pairs.

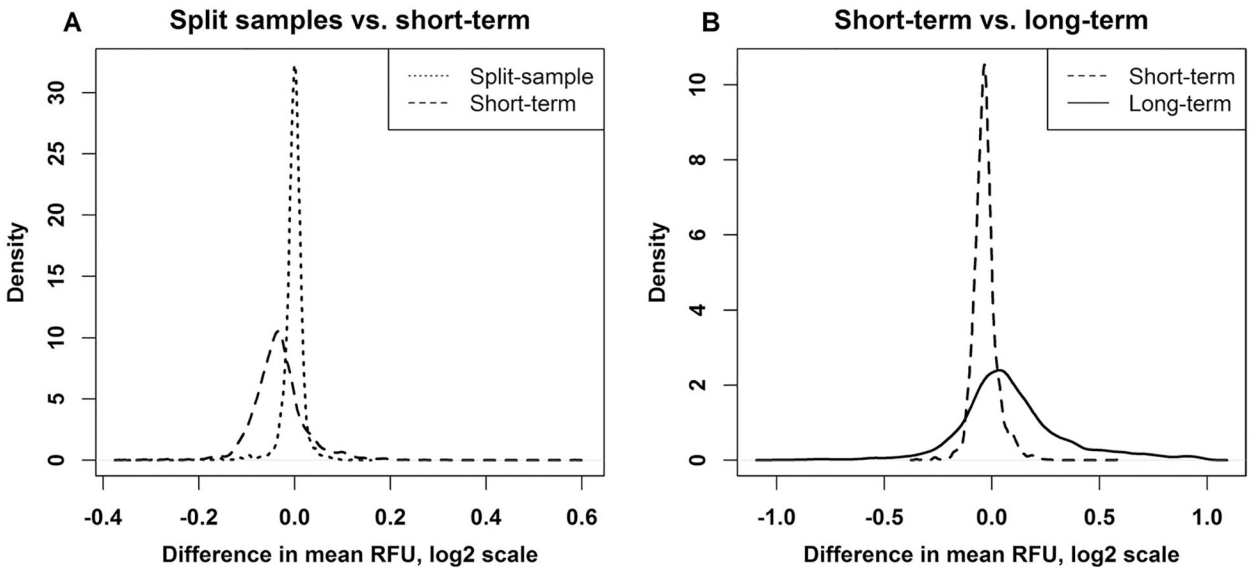

Fig. 2.

Gaussian kernel smoothed distribution of the difference of the protein analytes between time points: split samples vs short term (6–8 weeks apart) in (A) (n = 40) and short term vs long term (~20 years) in (B) (n = 42).

Table 1.

Participant characteristics.[a]

Variable	Visit 2 (1990–1992)	Visit 5 (2011–2013)
Male, n (%)	18(42.9)	
White, n (%)	24(57.1)	
Age, mean (SD)	56.6(4.8)	77.5(5)
BMI,[b] mean (SD)	28.6(5.5)	30.2 (5.2)[c]
Smoking, n (%)		
Current smoker	3(7.1)	1 (2.4)[d]
Former smoker	11 (26.2)	16 (38.1)[d]
Never smoker	28(66.7)	23 (54.8)[d]
SBP, mean (SD)	118.2(17.9)	134.5(21.9)
DBP, mean (SD)	72.2(10.8)	66.2(8.8)
Hb A1c, median (1st, 3rd quartile), %	5.5 (5.2, 5.8)	6.0 (5.6, 6.6)
Total cholesterol, mean (SD)	203.4(26.6)	173.1 (30.3)
Diabetes, n (%)	6(14.3)	18 (42.9)[c]
Hypertension, n (%)	15 (35.7)	29(69.0)
Hypertension medication, n (%)	12(28.6)	34(81.0)
Cholesterol lowering medication, n (%)	4(9.5)	28(66.7)
eGFR, mean (SD), mL/min/1.73m^2	96.4(11.8)	64.8(13.7)

[a] Smoking status was self-reported. HbA1c was measured using NGSP-certified instruments (Tosoh Corporation) standardized to the Diabetes Control and Complications Trial (13). Diabetes was defined as a fasting glucose of ≥126 mg/dL or a nonfasting glucose of ≥200 mg/dL, self-reported of physician diagnosis of diabetes, or taking diabetes medication during the 2 weeks prior to a study visit. Hypertension was defined as the mean of 3 sitting measurements of systolic blood pressure of ≥140 mmHg, of diastolic blood pressure of ≥ 90 mmHg, or self-reported use of hypertension medication during the 2 weeks prior to study visit. Hypertension medication and cholesterol lowering medication were based on the inspection of medication containers during study visit.

[b] BMI, body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; HbA1c, hemoglobin A1c; NGSP, National Glycohemoglobin Standardization Program.

[c] $n = 41$.

[d] $n = 40$.

Table 2.

Spearman correlations of protein measures quantified from modified aptamer assay and other clinical assays.[a]

	Visit 2		Visit 5		Change (log scale) between visits 2 and 5			
Protein	**N**	**Spearman correlation (95% CI)**	**N**	**Spearman correlation (95% CI)**	**N**	**Spearman correlation (95% CI)**	**Aptamer platform CV, %**	**Clinical assay CV, % visit 2, visit 5**
hs-CRP	39	0.96(0.91,0.98)	42	0.97(0.93,0.98)	39	0.90(0.80,0.95)	11.3	4.3,2.2
TSH	39	0.93(0.86,0.97)	42	0.94(0.87,0.97)	39	0.94(0.88,0.97)	4.1	<10,3.6
Cystatin C	40	0.89(0.77,0.95)	42	0.91 (0.81,0.95)	40	0.79(0.61,0.9)	15.2	3.1,3.1
B2M	39	0.87(0.74,0.94)	42	0.92(0.83,0.96)	39	0.76(0.55,0.88)	17	7.3,7.3
PTH	39	0.85(0.71,0.93)	–	–	–	–	10.2	5.1,–
NT-proBNP	35	0.84(0.67,0.92)	40	0.81 (0.64,0.9)	33	0.80(0.59,0.91)	3.9	5.4,7.0
Albumin	39	0.56(0.27,0.75)	42	0.49 (0.2,0.7)	39	0.28 (−0.05, 0.55)	4.2	4.0,4.0
ALT	39	0.53(0.24,0.73)	42	0.36(0.05,0.6)	39	0.38(0.06,0.63)	7.2	5.6,5.6
hs-cTnT[b]	15	0.52 (−0.02, 0.83)	42	0.46(0.16,0.68)	15	0.14(−0.4, 0.61)	6.3	6.0,6.4

[a] When multiple CVs were reported from ARIC clinical assays (low and high concentrations or intraday and interday runs; see Table 2 in the online Data Supplement), we selected the larger CV for this table.

[b] Limited to values above lower limit of detection (LOD) from the targeted (clinical) hs-cTnT assay (Roche Diagnostics).

J Appl Lab Med. Author manuscript; available in PMC 2019 October 25.