Biomarkers Consortium Project
Use of Targeted Multiplex Proteomic Strategies to Identify Plasma-Based Biomarkers in Alzheimer's Disease

Data Primer

Table of Contents

Additional queries regarding the RBM dataset should be addressed to:

Holly Soares - holly.soares@bms.com

Les Shaw - Les.Shaw@uphs.upenn.edu

Background:

The data described within this document represents the work of the Biomarkers Consortium Project "**Use of Targeted Multiplex Proteomic Strategies to Identify Plasma-Based Biomarkers in Alzheimer's Disease"**. This project was submitted to the Biomarkers Consortium Neuroscience Steering Committee by a subgroup of the Alzheimer's Disease Neuroimaging Initiative (ADNI) Industry Scientific Advisory Board (ISAB) for execution and was managed by a Biomarkers Consortium Project Team that includes members from academia, government and the pharmaceutical industry (for the list of Project Team members, see Appendix I); funding for this project was provided through an overage of funds raised by the Foundation for NIH for the ADNI partnership, as well as Pfizer Inc. This project is intended to be the first part of a multi-phased effort seeking to utilize samples collected by ADNI to qualify multiplex panels in both plasma and cerebrospinal fluid (CSF) to diagnose patients with Alzheimer's Disease (AD) and monitor disease progression.

The study described herein analyzed a subset of plasma samples from the ADNI cohort in a 190 analyte multiplex immunoassay panel. The panel, referred to as the human discovery map, was developed on the Luminex xMAP platform by Rules-Based Medicine (RBM) to contain proteins previously reported in the literature to be altered as a result of cancer, cardiovascular disease, metabolic disorders and inflammation. In addition, RBM partnered with Satoris to include plasma proteins believed to be involved in cell signaling and previously reported to change in patients with Alzheimer's disease (Ray et al., *Nat. Med.*, 2007). The analysis of plasma samples on the human discovery map is available on a commercial fee-for-service basis. The current document describes the technology and experimental design of the plasma multiplex biomarker pilot study.

Description of Multiplex Technology:

The Luminex xMAP technology uses a flow-based laser apparatus to detect fluorescent polystyrene microspheres which are loaded with different ratios of two spectrally distinct fluorochromes (Figure 1A). Using a precise ratio of the fluorochromes, up to 100 different beads can be generated such that each contains a unique color-coded signature. The beads serve as a solid phase matrix that can then be coated with either ligand or capture antibodies (Figure 1B) and then standard sandwich or competitive assay formats applied to detect the analytes. Signal is typically amplified via a reporter streptavidin-phycoerythrin conjugate. The beads are read one at a time as they pass through a flow cell on the Luminex laser instrument using a dual laser system (Figures 1C and D). One laser records the color code for individual beads (e.g. analyte ID) and the other quantitates the reporter signal (e.g. biomarker concentration). In theory, up to 100 different analytes can be measured per well per 250 ul of sample. However, dynamic range, matrix interference and cross-reactivity limit the number of analytes that can be multiplexed in one well. The actual RBM panel consists of several panels with between 3 and 24 multiplexed analytes. The combination of analytes per panel is proprietary to RBM. In addition, the dilution of samples per plate is also proprietary information.

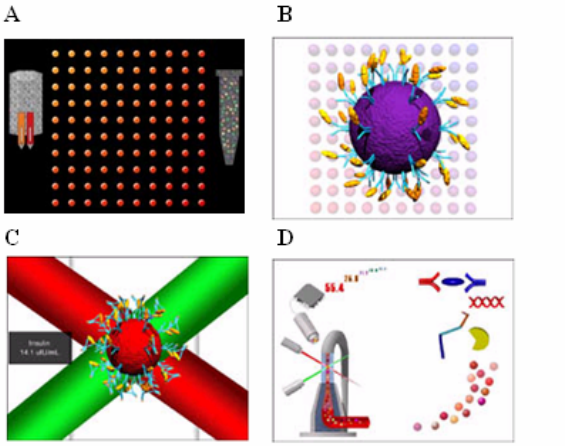


Figure 1: Luminex xMAP platform and basis of technology

RBM has attempted to validate each of the analytes on the 190 analyte panel up to clinical laboratory improvement amendment (CLIA) standards, but the assays themselves are not CLIA approved. Each analyte has an individual standard curve with between 6-8 reference standards. Each plate is run with 3 levels of QCs (low, medium and high) for each analyte. Assays are qualified based on least detectable dose (LDD - see below), precision, cross-reactivity, dilutional linearity, and spike recovery (assessment of accuracy). Cross validation to alternative methods is reported for some assays where feasible. In addition dynamic range of analytes in a healthy population is reported. The assays themselves should be considered exploratory and are not in full compliance with diagnostic standards for assays. For example, reference calibrators are diluted in a buffer and not in matrix (i.e. plasma or serum) and measurement bias is a component of the platform. Linearity of dilution and stability were not evaluated. In addition, the magnitude of batch-batch variation is not defined. RBM uses the following criteria for assay qualification:

Least Detectable Dose

The least detectable dose (LDD) is the concentration of target analyte that produces a signal that can be distinguished from that produced by a blank with 99% confidence. It is determined from the average and standard deviation of the signal for a minimum of 20 replicate determinations of the standard curve blank for each assay. Three standard deviations are added to the average of the signal, and this value is converted to concentration as interpolated from the dose response curve. The LDD is considered the most reliable lowest point for the individual assays.

Lower Assay Limit

The lower assay limit (LAL) is the lowest calibrator for individual assays. The LAL can be below the LDD. On the RBM panels, the LAL are often on the flat part of the calibration curve and are associated with significant error.

Precision
Precision is defined by the agreement between replicate measurements of the same material when measured within Run (intra-assay CV) and over a series of Runs (inter-assay or Total CV). It is determined by measuring 3 levels of controls in duplicate over a minimum of 5 Runs and provides information concerning random error expected in a test result caused by factors that vary under normal laboratory operating conditions such as pipeting, timing, mixing, and temperature.

Cross-reactivity
Cross-reactivity is the ability of an assay to differentiate and quantify the analyte of interest in the presence of other similar analytes in the sample that could have a positive or negative effect on the assay value. It is determined by testing high concentrations of each MAP analyte across all multiplexes. However, true specificity against highly related proteins is not well described in some cases.

Linearity
Linearity is the ability of the assay to obtain test results which are proportional to the concentration of analyte in a sample when serially-diluted to produce values within the dynamic range of the assay. It is determined by testing high positive or spiked samples serially-diluted in standard curve diluent. The average % recovery throughout the dilution series is then calculated as observed vs. expected concentration.

Spike Recovery
Spike recovery is performed as an assessment of accuracy which is often not possible for biological products due to the unavailability of pure "gold" standards. It is used to account for interference caused by compounds introduced from the physical composition of the sample or sample matrix that may affect the accurate measurement of the analyte. It is performed by spiking different amounts of standard spanning the assay range into standard curve diluent (control spike) and known samples. The average % recovery is calculated as the proportion of spiked standard in the sample (observed) to that of the control spike (expected) following analysis.

Correlation
Agreement of RBM multiplexed assay values to other methods is assessed by testing samples in an alternate FDA approved commercial immunoassay system, when available. This comparison of methods is performed to estimate inaccuracy or systematic error. Data from the two methods are graphed in a comparison plot and the correlation coefficient is determined.

Dynamic Range
The dynamic range is defined as the range of standard used to produce the standard curve. It is initially realized during assay development when standards are analyzed in a wide range above and below the expected concentrations using full-log dilutions. The standards are subsequently retested using reduced serial dilutions that target the useful part of the standard curve.

Matrix Interferences

Matrix interference assays are performed to determine whether the presence of interferants commonly found in serum/plasma samples introduce any systematic error in RBM MAPs. Samples spiked with high levels of hemoglobin, bilirubin, or triglyceride, as well as unspiked samples, are tested in each multiplex. % recovery is calculated as observed (spiked sample) result versus expected (unspiked sample) result.

RBM provides reports of analytes with the LDD and range for that particular run. Values that are below LDD are typically reported as LOW. High values may be reported as >top of analyte range number. If there is not sufficient volume, RBM will report as quantity not sufficient (QNS).

Listing of the Multiplex Analytes, LDD and Range:

Each analyte on the panel has a validation report that is available through RBM. Validation reports and dynamic range for serum and plasma in young healthy normal patients are known and can be obtained from Rules-Based Medicine. Table 1 lists the analytes, concentration units, LDD, LAL and range in young healthy volunteer plasma. In addition, Table 1 lists summary statistics from the QCs run during the analysis of the ADNI plasma subset (see below for more detailed discussion).

Table 1: List of Analytes, units, LDD, LAL, Dynamic range in healthy volunteers and summary QC statistics of each analyte. Analytes with one QC (red) with a CV above 25% are highlighted in yellow.

Analyte	Units	Lower Assay Limit	Least Detectable Dose	RBM Low Plasma Range	RBM High Plasma Range	Level 1 qc			Level 2 qc			Level 3 qc		
						mean.1	sd.1	cv.1	mean.2	sd.2	cv.2	mean.3	sd.3	cv.3
Alpha-1-Microglobulin (A1Micro)	ug/ml	0.044	0.014	6.2	16	4.36	0.649	14.9	35.9	6.53	18.2	48.1	8.99	18.7
Alpha-2-Macroglobulin (A2Macro)	mg/mL	0.0044	0.013	0.13	1.0	0.291	0.0212	7.29	0.728	0.0567	7.79	2.79	0.501	17.9
Alpha-1-Antichymotrypsin (AACT)	ug/ml	2.2	0.75	pending	pending	26.3	2.74	10.4	158	11.2	7.04	920	68.5	7.44
Alpha-1-Antitrypsin (AAT)	mg/mL	0.0038	0.0095	1.2	3.1	0.167	0.037	22.2	0.513	0.0522	10.2	1.55	0.141	9.06
Angiotensin-Converting Enzyme (ACE)	ng/ml	0.090	0.20	36	279	2.65	0.196	7.4	13.5	0.99	7.35	108	9.04	8.35
Adrenocorticotropic Hormone (ACTH)	ng/mL	0.035	0.29	<LOW>	2.2	1.31	0.153	11.7	4.06	0.44	10.8	8.13	0.722	8.87
Adiponectin	ug/mL	0.016	0.028	1.6	14	0.133	0.0377	28.3	0.517	0.0583	11.3	1.67	0.154	9.26
Alpha-Fetoprotein (AFP)	ng/mL	0.33	0.34	<LOW>	6.7	1.72	0.353	20.6	26.1	3.81	14.6	163	16.7	10.2
Agouti-Related Protein (AGRP)	pg/mL	22	125	pending	pending	184	37.4	20.3	615	68.4	11.1	1959	193	9.84
Angiopoietin-2 (ANG-2)	ng/mL	0.029	0.32	pending	pending	25.9	1.79	6.91	85.7	4.62	5.4	NA	NA	NA
Angiotensinogen	ng/mL	0.39	<0.5	0.55	670	181	8.33	4.59	429	21.7	5.06	772	41.4	5.36
Apolipoprotein A-I (Apo A-I)	mg/mL	0.00052	0.0055	0.19	0.89	0.0363	0.00526	14.5	0.127	0.014	11	0.436	0.0488	11.2
Apolipoprotein A-II (Apo A-II)	ng/ml	0.76	3.1	pending	pending	35.1	8.66	24.7	119	21.3	17.8	387	60.8	15.7
Apolipoprotein A-IV (Apo A-IV)	ug/ml	0.024	0.19	pending	pending	1.03	0.0933	9.02	3.17	0.191	6.05	8.77	0.508	5.79
Apolipoprotein B (Apo B)	ug/ml	8.2	13	pending	pending	140	13.8	9.87	744	64.2	8.63	2798	260	9.3
Apolipoprotein C-I (Apo C-I)	ng/ml	0.81	1.8	pending	pending	21.9	2.72	12.4	75.2	10.1	13.5	244	23.5	9.63
Apolipoprotein C-III (Apo C-III)	ug/mL	0.34	3.3	28	224	14.5	2.67	18.4	42.9	3.92	9.13	124	12.4	9.97
Apolipoprotein D (Apo D)	ug/ml	1.4	26	pending	pending	49.5	8.6	17.4	125	13.6	10.9	399	27.2	6.8
Apolipoprotein E (Apo E)	ug/ml	0.24	1.3	pending	pending	7.05	0.937	13.3	13.1	1.22	9.36	33.1	2.96	8.96
Apolipoprotein H (Apo H)	ug/mL	3.3	2.6	131	430	51.9	5.88	11.3	82.8	7.48	9.03	177	14.6	8.28
Amphiregulin (AR)	pg/mL	200	168	pending	pending	5267	1348	25.6	37380	8623	23.1	131253	32208	24.5
AXL Receptor Tyrosine Kinase (AXL)	ng/mL	0.061	0.051	pending	pending	1.5	0.107	7.13	4.29	0.318	7.42	11.3	0.841	7.43
Beta-2-Microglobulin (B2M)	ug/mL	0.0048	0.067	1.2	6.2	0.153	0.0154	10.1	0.31	0.0259	8.36	0.735	0.0638	8.68
Brain-Derived Neurotrophic Factor (BDNF)	ng/mL	0.013	0.023	0.32	16	0.859	0.0831	9.68	2.99	0.289	9.65	8.02	1.24	15.4
B Lymphocyte Chemoattractant (BLC)	pg/ml	0.65	55	pending	pending	68.1	6.63	9.74	205	20.2	9.86	524	81.9	15.6
Bone Morphogenetic Protein 6 (BMP-6)	ng/mL	0.034	0.55	pending	pending	1.75	0.25	14.3	5.63	0.631	11.2	15.9	0.99	6.24
Brain Natriuretic Peptide (BNP)	pg/ml	25	102	pending	pending	820	316	38.5	4327	443	10.2	7603	833	11
Betacellulin (BTC)	pg/mL	11	100	pending	pending	593	45.1	7.61	1707	99.2	5.81	5127	335	6.53
Complement C3 (C3)	mg/mL	0.0033	0.0046	0.76	2.1	0.149	0.0181	12.2	0.453	0.043	9.49	1.37	0.0976	7.14
Cancer Antigen 125 (CA-125)	U/mL	3.8	8.7	<LOW>	12	60.5	6.32	10.4	513	42.2	8.24	5213	489	9.38
Cancer Antigen 19-9 (CA-19-9)	U/mL	0.43	1.6	<LOW>	9.2	16.3	3.6	22	106	9.73	9.15	510	43.3	8.49
Calbindin	ng/ml	2.2	2.9	<LOW>	5.0	15	2.27	15.1	54.7	6.76	12.4	213	27.2	12.8
Calcitonin	pg/mL	0.91	12	<LOW>	12	38.9	3.22	8.26	185	13.7	7.37	1001	93.4	9.33
CD 40 antigen (CD40)	ng/mL	0.0095	0.022	0.17	1.5	0.455	0.0259	5.68	1.57	0.118	7.5	6.9	0.51	7.39
CD40 Ligand (CD40-L)	ng/mL	0.0080	0.013	<LOW>	1.1	0.175	0.0239	13.7	1.06	0.16	15.1	3.84	0.565	14.7
CD5 (CD5L)	ng/ml	43	40	pending	pending	301	44.3	14.8	841	78.1	9.28	2153	162	7.54
Carcinoembryonic Antigen (CEA)	ng/mL	0.14	0.19	<LOW>	4.8	3.68	0.328	8.91	20.9	1.58	7.55	128	6.51	5.08
Chromogranin-A (CgA)	ng/mL	4.2	10	pending	pending	60	6.69	11.1	256	22	8.56	755	90.4	12

Table 1 Continued

Analyte	Units	Lower Assay Limit	Least Detectable Dose	RBM Low Plasma Range	RBM High Plasma Range	Level 1 qc			Level 2 qc			Level 3 qc		
						mean.1	sd.1	cv.1	mean.2	sd.2	cv.2	mean.3	sd.3	cv.3
Creatine Kinase-MB (CK-MB)	ng/mL	0.019	0.16	<LOW>	1.1	0.627	0.0793	12.6	3.85	0.577	15	28.9	4.93	17.1
Clusterin (CLU)	ug/ml	0.77	1.3	pending	pending	52.4	3.16	6.03	164	19.6	12	558	50.6	9.08
Ciliary Neurotrophic Factor (CNTF)	pg/mL	7.4	92	pending	pending	96.6	10.6	11	407	27.8	6.83	1369	88.1	6.43
Complement Factor H	ug/ml	8.3	13	pending	pending	413	33.2	8.05	1195	80.7	6.75	3174	210	6.6
Cortisol (Cortisol)	ng/ml	2.7	2.0	48	250	165	26.3	15.9	440	87.2	19.8	1250	263	21
C-peptide	ng/ml	0.021	0.23	pending	pending	1.7	0.0756	4.45	6.92	0.336	4.86	26.1	2.5	9.58
C-Reactive Protein (CRP)	ug/mL	0.010	0.0065	0.18	35	0.234	0.0504	21.5	0.728	0.176	24.2	2.16	0.333	15.4
Connective Tissue Growth Factor (CTGF)	ng/ml	0.36	3.3	2.8	7.5	0.991	0.406	41	5.74	0.652	11.4	19.3	3.06	15.8
Cystatin-C	ng/ml	16	76	608	1170	10.2	0.947	9.31	34.9	2.47	7.1	89.4	8.52	9.53
Epidermal Growth Factor (EGF)	pg/mL	1.7	3.0	<LOW>	505	46	4.9	10.6	246	18.8	7.64	1261	115	9.13
Epidermal Growth Factor Receptor (EGFR)	ng/mL	0.050	0.042	pending	pending	0.674	0.115	17	7.73	0.724	9.36	29.7	3.56	12
Epithelial-Derived Neutrophil-Activating Protein 78 (ENA-78)	ng/mL	0.010	0.033	0.069	5.3	0.688	0.0446	6.48	2.38	0.182	7.65	5.69	0.691	12.1
EN-RAGE	ng/mL	0.35		4.6	592	10.2	1.74	17.1	24.9	2.84	11.4	51.2	4.63	9.05
Eotaxin-1	pg/mL	4.8	19	<LOW>	177	101	15.1	15	595	36.3	6.1	2383	253	10.6
Eotaxin-3	pg/mL	98	70	<LOW>	177	6657	776	11.7	22693	2109	9.29	59533	5694	9.56
Erythropoietin (EPO)	pg/mL	1.7	14	<LOW>	284	97.1	12	12.3	487	28.1	5.77	2516	320	12.7
Epiregulin (EPR)	pg/mL	27	17	pending	pending	209	40.3	19.3	1381	125	9.06	5361	438	8.18
E-Selectin	ng/mL	0.080	0.25	pending	pending	3.87	0.276	7.14	18.8	0.862	4.58	80.5	4.24	5.27
Endothelin-1 (ET-1)	pg/mL	12	14	<LOW>	26	63.8	14.6	22.9	485	31.3	6.45	2799	201	7.17
Fatty Acid-Binding Protein, heart (FABP, heart)	ng/mL	0.17	0.71	<LOW>	10	97.1	10.3	10.6	581	46	7.92	89.3	6.43	7.2
Factor VII	ng/mL	0.44	0.57	106	443	2.45	0.311	12.7	41.7	3.13	7.52	387	29.4	7.6
FASLG Receptor (FAS)	ng/mL	0.40	1.2	pending	pending	49.2	2.78	5.66	161	5.46	3.39	509	22.6	4.44
Fas Ligand (FasL)	pg/mL	1.7	16	pending	pending	54.1	4.61	8.52	174	15.5	8.91	543	36.2	6.67
Fetuin-A	ug/ml	4.7	3.0	pending	pending	254	18.2	7.17	709	58.5	8.25	2339	196	8.36
Fibroblast Growth Factor 4 (FGF-4)	pg/mL	5.8	248	pending	pending	717	76.2	10.6	2541	190	7.47	7619	475	6.23
Fibroblast Growth Factor basic (FGF-basic)	pg/mL	16	249	<LOW>	2000	203	102	50.2	2357	192	8.15	36707	2536	6.91
Fibrinogen	mg/mL	0.017	0.014	2.2	8.0	0.0956	0.0197	20.6	0.295	0.034	11.5	0.89	0.114	12.8
Ferritin (FRTN)	ng/mL	0.90	1.5	5.0	552	28.5	4.27	15	78.6	6.77	8.61	217	23.9	11
Follicle-Stimulating Hormone (FSH)	mIU/mL	0.13	2.7	2.0	98	2.03	0.237	11.7	9.91	0.866	8.74	69.5	5.14	7.39
Granulocyte Colony-Stimulating Factor (G-CSF)	pg/mL	2.5	3.4	<LOW>	37	48.3	7.27	15.1	285	34.2	12	1259	134	10.6
Growth Hormone (GH)	ng/mL	0.020	0.092	<LOW>	4.4	0.414	0.0392	9.48	2.32	0.178	7.68	16.6	0.91	5.48
Glucagon-like Peptide 1, total (GLP-1 total)	pg/ml	0.53	7.5	<LOW>	812	18.7	3.63	19.4	125	19.3	15.5	904	80.9	8.94
Glucagon	pg/ml	12	200	<LOW>	1477	418	76.9	18.4	2335	219	9.4	7880	2021	25.7
Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF)	pg/mL	5.7	39	<LOW>	152	132	17.7	13.3	329	47	14.3	867	68.1	7.86
Growth-Regulated alpha protein (GRO-alpha)	pg/mL	3.5	103	pending	pending	130	10.4	8.02	433	31.3	7.23	1285	84.9	6.61
Glutathione S-Transferase alpha (GST-alpha)	ng/ml	0.42	3.1	6.7	62	16.6	2.38	14.4	55.3	10.4	18.8	194	43.1	22.2
Haptoglobin	mg/mL	0.0098	0.028	0.047	7.6	0.0867	0.00963	11.1	0.324	0.0325	10	1.05	0.0995	9.46
Heparin-Binding EGF-Like Growth Factor (HB-EGF)	pg/mL	29	13	pending	pending	237	67.1	28.3	6831	978	14.3	26440	2935	11.1
Chemokine CC-4 (HCC-4)	ng/mL	0.0077	0.024	pending	pending	0.203	0.0103	5.1	0.637	0.0337	5.3	2.03	0.105	5.15
Hepatocyte Growth Factor (HGF)	ng/mL	0.061	0.20	pending	pending	6.49	0.697	10.7	22.1	2.59	11.7	91.9	17.3	18.9
Heat Shock Protein 60 (HSP-60)	ng/ml	2.0	18	pending	pending	103	12.2	11.8	299	55.2	18.5	967	147	15.2
T Lymphocyte-Secreted Protein I-309 (I-309)	pg/mL	6.1	95	pending	pending	584	53.9	9.23	1985	128	6.42	6289	325	5.17
Intercellular Adhesion Molecule 1 (ICAM-1)	ng/mL	0.087	2.1	42	213	2.48	0.673	27.1	13.7	1.1	8.01	80.3	11.5	14.3
Interferon gamma (IFN-gamma)	pg/mL	0.88	3.2	<LOW>	9.5	52.5	5.01	9.55	261	23	8.8	1343	76.3	5.68
Immunoglobulin A (IgA)	mg/mL	0.0073	0.043	0.58	5.6	0.336	0.0481	14.3	0.842	0.0684	8.12	2.44	0.261	10.7
Immunoglobulin E (IgE)	ng/mL	1.3	3.9	<LOW>	770	2.65	1.2	45.4	57.3	13.1	22.9	418	41.7	9.98
Insulin-like Growth Factor-Binding Protein 2 (IGFBP-2)	ng/mL	0.27	1.2	pending	pending	6.75	0.692	10.2	19.3	1.22	6.35	55.5	2.92	5.27
Insulin-like Growth Factor I (IGF-I)	ng/mL	0.86	6.0	<LOW>	177	50.1	6.63	13.2	143	12.8	8.99	749	54.8	7.32
Immunoglobulin M (IGM)	mg/mL	0.014	0.080	0.24	3.3	0.211	0.063	29.8	0.707	0.104	14.7	2.11	0.253	12
Interleukin-1 alpha (IL-1 alpha)	ng/mL	0.00098	0.0071	<LOW>	0.35	0.0151	0.00136	8.96	0.0478	0.00355	7.43	0.144	0.0091	6.32
Interleukin-1 beta (IL-1 beta)	pg/mL	0.35	1.1	<LOW>	8.7	5.13	0.828	16.2	28.9	3.06	10.6	142	11.2	7.85
Interleukin-10 (IL-10)	pg/mL	2.9	6.3	1.8	38	185	13.7	7.4	632	40.1	6.35	2677	224	8.38
Interleukin-11 (IL-11)	pg/mL	17	56	pending	pending	831	56.2	6.77	2224	213	9.58	6048	348	5.76
Interleukin-12 Subunit p40 (IL-12p40)	ng/mL	0.096	0.23	<LOW>	2.7	5.35	0.504	9.42	19.4	1.45	7.5	77.5	5.53	7.13
Interleukin-12 Subunit p70 (IL-12p70)	pg/mL	21	26	<LOW>	165	1528	119	7.78	5530	362	6.54	20580	1230	5.98
Interleukin-13 (IL-13)	pg/mL	1.4	16	<LOW>	133	38.5	4.7	12.2	85.9	9.23	10.7	236	14.5	6.12
Interleukin-15 (IL-15)	ng/mL	0.025	0.58	<LOW>	4.6	0.506	0.189	37.3	1.71	0.237	13.9	6.21	0.599	9.66
Interleukin-16 (IL-16)	pg/mL	10	16	232	3380	84.6	9.42	11.1	330	18.6	5.64	1217	73.9	6.07
Interleukin-18 (IL-18)	pg/mL	7.0	25	72	1020	356	23.4	6.57	1314	147	11.2	6163	635	10.3
Interleukin-1 receptor antagonist (IL-1ra)	pg/mL	10	44	17	622	238	42.5	17.8	1373	192	14	6619	492	7.44
Interleukin-2 (IL-2)	pg/mL	7.7	19	<LOW>	61	156	15.1	9.71	549	50.6	9.22	2021	172	8.51
Interleukin-25 (IL-25)	pg/mL	1.9	13	<LOW>	124	664	58	8.72	2588	219	8.46	7207	539	7.48
Interleukin-3 (IL-3)	ng/mL	0.0022	0.049	<LOW>	1.2	0.0297	0.00622	20.9	0.0934	0.0153	16.3	0.276	0.0241	8.75
Interleukin-4 (IL-4)	pg/mL	7.6	34	<LOW>	103	647	66.8	10.3	2234	181	8.1	6834	510	7.46
Interleukin-5 (IL-5)	pg/mL	0.83	4.0	<LOW>	62	5.53	1.48	26.7	15.6	1.88	12.1	28.6	5.8	20.3
Interleukin-6 (IL-6)	pg/mL	0.95	1.8	<LOW>	25	34.8	4.55	13.1	184	18	9.79	1019	84.5	8.29
Interleukin-6 receptor (IL-6r)	ng/mL	0.013	0.030	pending	pending	1.22	0.0561	4.6	5.99	0.343	5.73	25.1	0.99	3.94
Interleukin-7 (IL-7)	pg/mL	3.6	59	3.7	125	628	52.6	8.38	2087	109	5.23	6871	509	7.41
Interleukin-8 (IL-8)	pg/mL	2.1	3.8	<LOW>	59	71.5	6.5	9.09	246	20.9	8.51	907	57.3	6.32
Insulin	uIU/mL	0.051	0.66	<LOW>	34	4.74	0.36	7.6	16.6	1.06	6.36	29.8	1.52	5.1
Interferon gamma Induced Protein 10 (IP-10)	pg/ml	9.0	30	pending	pending	444	40.8	9.19	2075	113	5.45	9697	934	9.63
Kidney Injury Molecule-1 (KIM-1)	ng/ml	0.0080	0.017	0.053	0.57	0.113	0.0139	12.3	0.337	0.0324	9.6	1.19	0.119	9.96
Leptin	ng/mL	0.026	0.12	0.41	41	0.8	0.0614	7.68	3.95	0.318	8.05	22.7	1.91	8.45
Luteinizing Hormone (LH)	mIU/mL	0.23	0.80	4.2	151	6.04	0.642	10.6	28.7	2.43	8.47	212	15.1	7.11
Lectin-Like Oxidized LDL Receptor 1 (LOX-1)	ng/mL	0.64	0.82	pending	pending	6.27	0.858	13.7	93.3	12.6	13.5	391	38.8	9.91
Apolipoprotein(a) (Lp(a))	ug/mL	1.3	3.3	10	2888	9.1	1.94	21.3	30.7	4.2	13.7	95.3	10.4	10.9
Lymphotactin	ng/mL	0.029	0.23	<LOW>	0.57	1.87	0.09	4.82	5.79	0.379	6.54	15.7	1.49	9.45

Table 1 Continued

Analyte	Units	Lower Assay Limit	Least Detectable Dose	RBM Low Plasma Range	RBM High Plasma Range	Level 1 qc			Level 2 qc			Level 3 qc		
						mean.1	sd.1	cv.1	mean.2	sd.2	cv.2	mean.3	sd.3	cv.3
Monocyte Chemotactic Protein 1 (MCP-1)	pg/mL	4.8	40	35	401	126	10.8	8.58	489	33.2	6.8	1858	95.1	5.12
Monocyte Chemotactic Protein 2 (MCP-2)	pg/ml	2.1	15	pending	pending	54.8	3.63	6.62	236	10.7	4.52	690	41	5.94
Monocyte Chemotactic Protein 3 (MCP-3)	pg/mL	0.57	17	pending	pending	25.7	2.5	9.73	88.1	5.77	6.54	278	18.6	6.68
Monocyte Chemotactic Protein 4 (MCP-4)	pg/ml	7.7	30	pending	pending	305	34.6	11.4	1517	134	8.84	8102	469	5.78
Macrophage Colony-Stimulating Factor 1 (M-CSF)	ng/mL	0.019	0.10	pending	pending	0.423	0.0498	11.8	2	0.156	7.79	9.67	0.511	5.28
Malondialdehyde-Modified Low-Density Lipoprotein (MDA-LDL)	ng/mL	51	55	pending	pending	334	182	54.6	5014	1390	27.7	10568	2856	27
Macrophage-Derived Chemokine (MDC)	pg/mL	4.0	12	162	774	148	8.04	5.45	528	30.9	5.86	1621	167	10.3
Macrophage Migration Inhibitory Factor (MIF)	ng/mL	0.0040	0.017	pending	pending	0.137	0.0145	10.6	0.431	0.0365	8.48	1.26	0.0986	7.82
Monokine Induced by Gamma Interferon (MIG)	pg/ml	38	100	pending	pending	7513	834	11.1	31953	4435	13.9	161667	13010	8.05
Macrophage Inflammatory Protein-1 alpha (MIP-1 alpha)	pg/mL	7.7	19	<LOW>	89	80.1	16	20	514	33.1	6.45	2803	156	5.58
Macrophage Inflammatory Protein-1 beta (MIP-1 beta)	pg/mL	5.9	21	25	595	261	17.4	6.69	1139	102	8.92	15373	1335	8.69
Macrophage Inflammatory Protein-3 alpha (MIP-3 alpha)	pg/ml	3.7	16	pending	pending	486	38.9	8.01	2161	120	5.55	7231	378	5.22
Matrix Metalloproteinase-1 (MMP-1)	ng/ml	0.11	0.37	pending	pending	3.09	0.309	10	15.2	0.862	5.67	75.9	3.63	4.79
Matrix Metalloproteinase-10 (MMP-10)	ng/ml	0.010	0.010	pending	pending	1.54	0.0828	5.38	7.87	0.446	5.67	38.6	1.99	5.16
Matrix Metalloproteinase-2 (MMP-2)	ng/mL	1.7	15	183	3070	36.1	4.72	13.1	251	25.9	10.3	1506	130	8.66
Matrix Metalloproteinase-3 (MMP-3)	ng/mL	0.016	0.066	<LOW>	1.8	0.27	0.0455	16.9	2.05	0.185	8.99	15	1.13	7.56
Matrix Metalloproteinase-7 (MMP-7)	ng/ml	0.020	0.50	pending	pending	0.212	0.0283	13.4	1.11	0.064	5.75	5.83	0.358	6.14
Matrix Metalloproteinase-9 (MMP-9)	ng/mL	0.89	10	<LOW>	1050	15.3	7.42	48.7	88.1	9.86	11.2	788	52.2	6.62
Matrix Metalloproteinase-9, total (MMP-9, total)	ng/ml	0.72	0.46	pending	pending	2.99	1.1	36.6	23.3	1.79	7.7	119	8	6.7
Myeloid Progenitor Inhibitory Factor 1 (MPIF-1)	ng/mL	0.010	0.39	pending	pending	1.33	0.135	10.1	6.81	0.424	6.22	34.5	3.09	8.95
Myeloperoxidase (MPO)	ng/mL	22	32	<LOW>	1110	809	193	23.8	2285	404	17.7	5643	611	10.8
Myoglobin	ng/mL	0.22	2.0	3.6	37	3.94	0.481	12.2	14.5	1.41	9.68	45.3	3.01	6.65
Neutrophil Gelatinase-Associated Lipocalin (NGAL)	ng/ml	19	400	89	375	17.6	2.47	14	52.3	7.71	14.8	154	10.7	6.98
Nerve Growth Factor beta (NGF-beta)	ng/mL	0.016	0.056	pending	pending	1.09	0.102	9.32	3.55	0.177	4.98	8.8	0.64	7.28
Neuronal Cell Adhesion Molecule (Nr-CAM)	ng/mL	0.073	0.065	pending	pending	4.05	0.233	5.74	11.9	0.704	5.9	21.1	2.26	10.7
Osteopontin	ng/ml	0.17	1.2	4.1	25	399	29.8	7.45	1507	80.8	5.36	5482	374	6.81
Plasminogen Activator Inhibitor 1 (PAI-1)	ng/mL	1.0	1.1	10	87	11.8	1.24	10.4	36.5	3.78	10.3	108	8.64	7.99
Prostatic Acid Phosphatase (PAP)	ng/mL	0.017	0.011	0.058	0.54	0.255	0.029	11.4	1.34	0.14	10.5	6.33	0.735	11.6
Pregnancy-Associated Plasma Protein A (PAPP-A)	mIU/mL	0.0028	0.011	<LOW>	0.48	0.0353	0.0116	32.8	0.283	0.0415	14.7	1.56	0.172	11
Pulmonary and Activation-Regulated Chemokine (PARC)	ng/mL	4.3	0.070	pending	pending	0.659	0.0431	6.54	1.97	0.149	7.53	4.93	0.237	4.82
Platelet-Derived Growth Factor BB (PDGF-BB)	pg/ml	43	139	pending	pending	447	111	24.9	7920	609	7.69	28353	2594	9.15
Placenta Growth Factor (PLGF)	pg/ml	13	9.3	pending	pending	149	19.6	13.2	1836	126	6.86	5773	535	9.27
Pancreatic Polypeptide (PPP)	pg/ml	0.68	5.0	38	1531	11.5	1.15	10	59.6	4.66	7.81	502	40.6	8.09
Prolactin (PRL)	ng/ml	0.023	0.50	0.88	42	0.768	0.0731	9.52	3.95	0.27	6.83	31	1.93	6.22
Progesterone	ng/ml	0.55	0.72	<LOW>	60	30.1	9.16	30.5	112	24.3	21.6	397	65.1	16.4
Proinsulin, Intact	pM	1.00	0.60	pending	pending	20.3	1.29	6.35	75.7	4.11	5.43	253	9.55	3.78
Proinsulin, Total	pM	4.2	2.5	pending	pending	86.4	4.72	5.46	322	19.7	6.12	1075	49.4	4.59
Prostate-Specific Antigen, Free (PSA-f)	ng/mL	0.0070	0.028	<LOW>	1.6	0.205	0.0226	11	0.713	0.0752	10.5	1.91	0.158	8.28
Peptide YY (PYY)	pg/mL	17	<50	<LOW>	396	1488	138	9.27	4623	454	9.82	13680	1215	8.88
Receptor for advanced glycosylation end products (RAGE)	ng/mL	0.084	0.10	pending	pending	0.943	0.0672	7.12	3.54	0.302	8.53	16	1.41	8.84
T-Cell-Specific Protein RANTES (RANTES)	ng/mL	0.019	0.033	2.6	83	0.828	0.127	15.4	2.91	0.32	11	8.98	0.963	10.7
Resistin	ng/ml	0.020	0.030	0.85	10.0	0.0445	0.00554	12.5	0.601	0.057	9.49	4.6	0.417	9.08
S100 calcium-binding protein B (S100-B)	ng/mL	0.070	0.30	pending	pending	5.41	0.668	12.4	19.3	1.58	8.2	102	11.3	11.1
Serum Amyloid P-Component (SAP)	ug/mL	0.074	0.16	15	50	1.26	0.25	19.9	3.83	0.359	9.39	11	1.02	9.28
Stem Cell Factor (SCF)	pg/mL	15	43	<LOW>	281	826	53.8	6.52	2647	235	8.88	7111	648	9.11
Secretin	ng/mL	0.021	<0.015	<LOW>	45	0.2	NA	NA	5.96	0.974	16.3	24.7	2.06	8.35
Serum Glutamic Oxaloacetic Transaminase (SGOT)	ug/mL	0.11	2.5	3.9	28	7.35	1.61	21.9	23.9	2.76	11.5	87.3	16.4	18.8
Sex Hormone-Binding Globulin (SHBG)	nmol/L	1.4	2.3	12	106	2.19	0.399	18.2	7.23	1.04	14.4	23.3	1.88	8.05
Superoxide Dismutase 1, Soluble (SOD-1)	ng/mL	0.059	0.66	pending	pending	6.17	0.913	14.8	19.5	1.51	7.71	64.2	23	35.8
Sortilin	ng/mL	0.085	0.18	pending	pending	2.95	0.3	10.1	9.63	0.329	3.41	22.3	2.53	11.3
Thyroxine-Binding Globulin (TBG)	ug/mL	0.053	0.33	40	104	2.73	0.422	15.5	8.32	0.616	7.4	26.9	2.45	9.1
Thymus-Expressed Chemokine (TECK)	ng/mL	1.9	18	pending	pending	365	43.7	12	3708	1092	29.4	NA	NA	NA
Testosterone, Total	ng/ml	0.029	0.17	0.11	6.5	2.05	0.862	42.1	7.74	1.38	17.8	30.2	5.76	19.1
Tissue Factor (TF)	ng/mL	0.52	0.31	<LOW>	2.4	2.52	0.254	10.1	15	1.46	9.76	138	15.9	11.6
Trefoil Factor 3 (TFF3)	ug/ml	0.00010	0.00019	0.040	0.49	0.805	0.117	14.5	5.11	0.589	11.5	75.5	7.61	10.1

Table 1 Continued

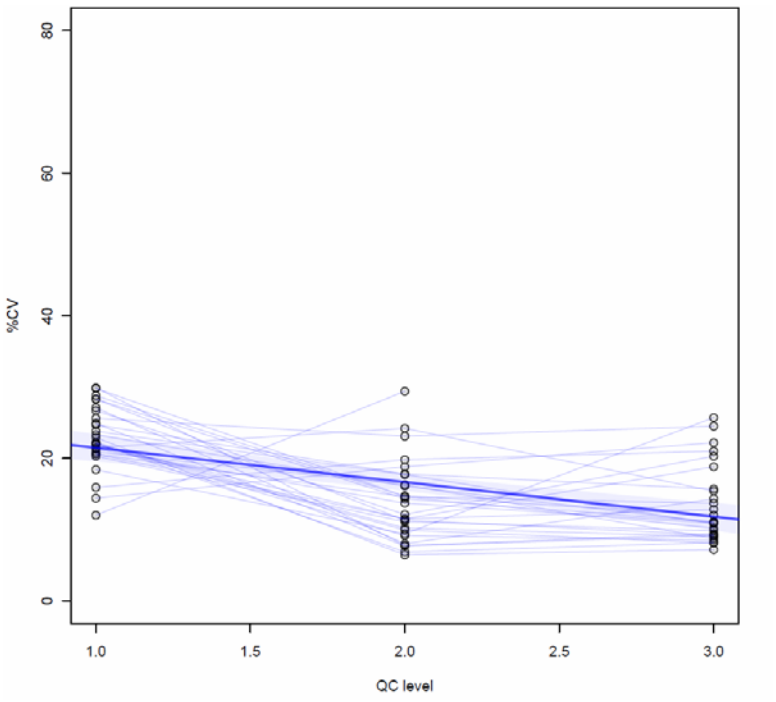
Analyte	Units	Lower Assay Limit	Least Detectable Dose	RBM Low Plasma Range	RBM High Plasma Range	Level 1 qc			Level 2 qc			Level 3 qc		
						mean.1	sd.1	cv.1	mean.2	sd.2	cv.2	mean.3	sd.3	cv.3
Transforming Growth Factor alpha (TGF-alpha)	pg/mL	12	6.7	pending	pending	118	26.3	22.2	2211	152	6.85	9117	745	8.18
Transforming Growth Factor beta-3 (TGF-beta-3)	pg/mL	9.1	20	pending	pending	370	24.1	6.53	1274	120	9.42	4025	379	9.41
Tamm-Horsfall Urinary Glycoprotein (THP)	ug/ml	0.00015	0.00019	0.0084	0.052	4.51	3.49	77.3	47.3	7.13	15.1	153	34	22.2
Thrombospondin-1	ng/mL	57	5.7	pending	pending	1017	154	15.2	3465	663	19.1	10341	1489	14.4
Tissue Inhibitor of Metalloproteinases 1 (TIMP-1)	ng/mL	0.28	4.1	59	192	9.96	1.04	10.4	21.7	1.71	7.87	52.7	3.22	6.1
Thrombomodulin (TM)	ng/ml	0.12	0.10	pending	pending	0.888	0.266	29.9	9.81	0.761	7.75	51.8	4.89	9.43
Tenascin-C (TN-C)	ng/mL	7.4	14	pending	pending	53.2	8.17	15.4	490	40	8.16	2363	173	7.32
Tumor Necrosis Factor alpha (TNF-alpha)	pg/mL	1.0	7.0	<LOW>	27	6.1	2.3	37.8	37.3	9.62	25.8	241	20.3	8.44
Tumor Necrosis Factor beta (TNF-beta)	pg/mL	1.5	18	<LOW>	120	74.8	11.9	15.9	286	26.4	9.25	1088	93.5	8.59
Tumor Necrosis Factor Receptor-Like 2 (TNFR2)	ng/mL	0.16	0.22	3.1	79	5.39	0.579	10.8	15.5	1.25	8.02	42.3	4.3	10.2
Thrombopoietin	ng/mL	0.11	4.4	<LOW>	6.2	5.25	0.534	10.2	17.6	1.12	6.37	56.8	4.09	7.2
TNF-Related Apoptosis-Inducing Ligand Receptor 3 (TRAIL-R3)	ng/mL	0.51	0.24	pending	pending	9.6	0.411	4.28	30.1	2.03	6.74	85.3	4.46	5.23
Serotransferrin (Transferrin)	mg/dl	9.5	93	pending	pending	671	54.1	8.07	2267	174	7.67	8009	780	9.74
Thyroid-Stimulating Hormone (TSH)	ulU/mL	0.017	0.021	0.18	3.7	0.464	0.049	10.6	2.53	0.209	8.26	15.8	1.26	8.01
Transthyretin (TTR)	mg/dl	0.25	0.49	pending	pending	11	2.57	23.3	41.8	6.73	16.1	150	16.6	11
Vascular Cell Adhesion Molecule-1 (VCAM-1)	ng/mL	1.6	1.5	284	1310	33.4	4.21	12.6	104	10.5	10.1	299	26.8	8.97
Vascular Endothelial Growth Factor (VEGF)	pg/mL	7.8	17	91	1790	437	33.6	7.69	1793	123	6.86	8321	724	8.7
Vitronectin	ug/ml	6.7	4.0	pending	pending	433	95.1	22	745	74	9.93	1523	124	8.11
Vitamin K-Dependent Protein S (VKDPS)	ug/ml	0.12	0.39	pending	pending	4.88	0.367	7.52	14.3	1.29	9.01	37.1	2.76	7.46
von Willebrand Factor (vWF)	ug/mL	0.70	2.8	5.3	74	25.5	7.34	28.8	76.3	11.1	14.6	234	32.1	13.7

Analyte Quality Control (QC) results from the 2010 ADNI Plasma Analysis:

QCs were generated by spiking blank human plasma with extracts of cell cultures expressing the individual analytes. Low, medium and high QCs were run on each plate for almost all the analytes. The total ADNI plasma cohort was run on 15 plates. The QC results for each analyte are included in Table 1. QCs were performed in duplicate, but samples were run in singlicate. As a result the first QC result from each plate was used to derive the summary QC statistics for each analyte. It should be noted that analytes with one or more QC CV values above 25% should be treated with caution and are high-lighted in yellow. Figure 1 highlights (A) the 28 analytes with QC CVs within the 20-30% range and (B) the 14 analytes with QC CVs >30%. In addition, analytes with numerous sample values close to or below LDD should be treated with caution.

A)

Endothelin-1 (ET-1)
Transforming Growth Factor alpha (TGF-alpha)
Cancer Antigen 19-9 (CA-19-9)
Vitronectin
Agouti-Related Protein (AGRP)
Alpha-1-Antitrypsin (AAT)
Platelet-Derived Growth Factor BB (PDGF-BB)
Fibrinogen
Alpha-Fetoprotein (AFP)
Apolipoprotein(a) (Lp(a))
Interleukin-3 (IL-3)
Thrombomodulin (TM)
Adiponectin
Intercellular Adhesion Molecule 1 (ICAM-1)
Transthyretin (TTR)
Serum Glutamic Oxaloacetic Transaminase (SGOT)
Myeloperoxidase (MPO)
Glucagon
Heparin-Binding EGF-Like Growth Factor (HB-EGF)
Glutathione S-Transferase alpha (GST-alpha)
Immunoglobulin M (IGM)
Cortisol (Cortisol)
von Willebrand Factor (vWF)
Apolipoprotein A-II (Apo A-II)
Interleukin-5 (IL-5)
C-Reactive Protein (CRP)
Thymus-Expressed Chemokine (TECK)
Amphiregulin (AR)



B)

Matrix Metalloproteinase-9, total (MMP-9, total)
Superoxide Dismutase 1, Soluble (SOD-1)
Pregnancy-Associated Plasma Protein A (PAPP-A)
Brain Natriuretic Peptide (BNP)
Interleukin-15 (IL-15)
Fibroblast Growth Factor basic (FGF-basic)
Matrix Metalloproteinase-9 (MMP-9)
Connective Tissue Growth Factor (CTGF)
Progesterone
Tumor Necrosis Factor alpha (TNF-alpha)
Immunoglobulin E (IgE)
Testosterone, Total
Malondialdehyde-Modified Low-Density Lipoprotein (MDA-LDL)
Tamm-Horsfall Urinary Glycoprotein (THP)

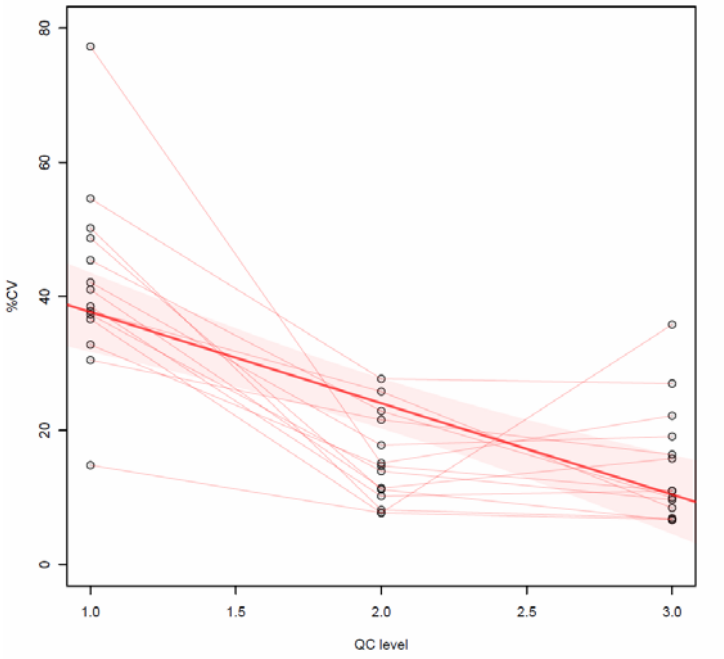


Figure 1: Summary of QC data for (A) the 28 analytes with QCs within the 20-30%CV range and (B) the 14 analytes with QCs >30%CV range.

Methodology:

One thousand sixty five 500ul EDTA plasma samples from the ADNI cohort were selected from a larger set of samples and shipped to RBM for analysis. Plasma samples were obtained in the morning following an overnight fast. For the majority of samples, the time from collection to freezing was within 120 minutes. Processing and aliquoting were per ADNI lab SOPs. The data reported on the ADNI website lists data for 1062 of the 1065 samples. One patient was excluded from the final analysis due to a screen failure. In addition, one patient had a redundant sample submitted for analysis. Only one set of data from that data point was reported. These 3 samples were excluded from further analysis and from the ADNI website. The final samples reported represented baseline and one year collections from 396 MCI, 112 AD and 58 control subjects. Samples were selected based upon availability of additional biomarker endpoints. For example, samples from AD subjects with associated CSF Abeta42 and/or Pittsburgh Compound B (PIB) one year data were included in the AD subset. Samples from normal subjects selected for inclusion were from subjects who had baseline Abeta 42 levels above the median baseline Abeta42 level for the normal cohort. Not all patients selected had a plasma sample available at the one year follow-up. Table 2 represents the demographics of the population selected. The selection process introduced a bias in ApoE4 prevalence in the control population, presumably due to association between ApoE4 levels and CSF Abeta42 levels (i.e. patients with low CSF Abeta42 levels were more likely to have one or more ApoE4 allele).

Table 2: Demographics of the plasma multiplex biomarker cohort

	Control	MCI	AD
N baseline (12 m)	58 (54)	396 (345)	112 (97)
Age	75.3 (62-90)	74.9 (55-90)	75.4 (55-89)
Gender M/F (baseline)	30/28	256/140	65/47
ApoE4% (baseline)	9%	53%	68%
MMSE (range)	28.9 (25-30)	27.0 (23-30)	23.6 (20-27)

It should be noted that age was calculated based upon date of birth and upon date of sample draw from baseline visit. Samples were randomized for processing at RBM and RBM was blind to the clinical information. A Statistical Analysis Plan (see Appendix II) was prepared prior to analysis.

It is recognized that the ADNI cohort is not optimized to identify novel diagnostic markers of disease for AD. As a result, a pilot study using plasma samples from the University of Pennsylvania cohort was completed to test the feasibility of the technology and to generate hypotheses that could be tested in the ADNI cohort. The preliminary findings are presented as hypothesis in the statistical analysis plan (Appendix II).

RBM made a comment on RBM sample numbers 191 (RID 750), 348 (RID 1182), 413 (RID 194), 435 (RID 221), 845 (RID 604) and 1041 (RID 1057). Analyte [Fibrinogen]: Result suggests sample type is serum. Caution is advised when interpreting result as analyte concentrations may vary between serum and plasma. An outlier rule was developed for the dataset and is described in the SAP.

What is posted on the ADNI Website and cautionary notes to data analysis:

There are two datasets posted on the ADNI website relating to the plasma multiplex pilot from the Biomarkers Consortium Project. The first dataset coded ADNI Plasma Multiplex Raw Data includes the original raw data from the run to be intended as reference. The second dataset title ADNI Plasma QC Multiplex data is the cleaned, quality controlled data according to methodology described in the statistical analysis plan. The third file is a detailed listing of what variables were changed in the QC process. It is recommended that raw data not be used to derive summary statistics as most of the analytes are not normally distributed and there are some analytes with quite a few LOW reported. Summary statistics should not be run on data that is not normally distributed. It is recommended that analytes with numerous LOW values listed or with majority of values listed between the LDD/LAL be treated with caution as deriving reliable results may be challenging. Consultancy with a trained statistician is highly recommended prior to reporting results based upon multiple comparisons and upon repeated measures (i.e. baseline and 12 months).

Table 3: Column header definitions in the ADNI Plasma Multiplex Imputation log

Variable Name	Description and Coding
ID	record ID
RID	ADNI subject ID
sampleID	Date of QC
Sample_Received_date	Date sample received at UPenn
Visit_Code	Visit Designator (bl = baseline; m12 = 1 year)
analyte	Name of Analyte with Units
LDD	Least Detectable Dose (see primer for details)
avalue	Recorded Value
analval	Numeric Value after possible imputation (see primer for details)
belowLDD	Is analval < LDD? Note: this flag pertains to both recorded value and imputed value (0=no ; 1=yes)
readLOW	Is recorded value <LOW> or numeric? (0=numeric; 1=<LOW> - see primer for details)
outler	Is recorded value an outlier? (0=no; 1=yes) - outliers imputed to 5SD from mean

Table 4. Column header definitions in the ADNI Plasma QC Multiplex data

Variable Name	Description
ID	record ID
RID	ADNI subject ID
sampleID	Sample ID from UPenn
Sample_Recieved_date	Date sample received at UPenn
Visit_Code	Visit Designator (bl = baseline; m12 = 1 year)

References:

Ray S, Britschgi M, Herbert, C et al., (2007) Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. Nat Med 13(11):1359-62

APPENDIX I

Biomarkers Consortium Project Team Members:

Holly Soares, BMS – Chair
William Potter – Co-Chair
Sophie Allauzen – Novartis
Neil Buckholtz – NIH/NIA
Patricia Cole – ImagePace
Bob Dean - Lilly
Ashok Dongre – BMS
Kevin Duffin – Lilly
Mats Ferm – AstraZeneca
Ron Hendrickson – Merck
Fred Immerman – Pfizer Statistics
Walter Koroshetz – NINDS/NIH
Max Kuhn – Pfizer Statistics
David Lee – FNIH
Brad Navia - Eisai
George Nomikos - Takeda
Eve Pickering – Pfizer Statistics
Mary Savage - Merck
Jeff Seeburger – Merck
Les Shaw – UPenn
David Shera – Merck Statistics
Eric Siemers – Lilly
Adam Simon – Independent Contributor
Judy Siuciak – FNIH
Dan Spellman - Merck
Frank Swenson - Pfizer
Madhav Thambisetty – NIH/NIA Statistics
John Trojanowski – UPenn
Marc Walton – FDA
Hong Wan – Pfizer
Li-San Wang – UPenn Statistics
Jeff Waring - Abbott
Sharon Xie – UPenn Statistics
Panos Zagouras – Pfizer Systems Bio

APPENDIX II
Biomarkers Consortium Project
Use of Targeted Multiplex Proteomic Strategies to Identify Plasma-Based Biomarkers in Alzheimer's Disease

Statistical Analysis Plan

Biomarkers Consortium Project
Use of Targeted Multiplex Proteomic Strategies to Identify Plasma-Based Biomarkers in Alzheimer's Disease

Statistical Analysis Plan

1 Introduction

The Analysis Plan described within this document represents the work of the Biomarkers Consortium Project "**Use of Targeted Multiplex Proteomic Strategies to Identify Plasma-Based Biomarkers in Alzheimer's Disease"**. This project was submitted to the Biomarkers Consortium Neuroscience Steering Committee by a subgroup of the Alzheimer's Disease Neuroimaging Initiative (ADNI) Industry Scientific Advisory Board (ISAB) for execution and was managed by a Biomarkers Consortium Project Team that includes members from academia, government and the pharmaceutical industry; funding for this project was provided through an overage of funds raised by the Foundation for NIH for the ADNI partnership, as well as Pfizer Inc. This project is intended to be the first part of a multi-phased effort seeking to utilize samples collected by ADNI to qualify multiplex panels in both plasma and cerebrospinal fluid (CSF) to diagnose patients with Alzheimer's Disease (AD) and monitor disease progression.

Biomarker tools for early diagnosis and disease progression in Alzheimer's disease (AD) remain key issues in AD drug development. Identification and validation of cost-effective, non-invasive methods to identify early AD and to monitor treatment effects in mild-moderate AD patients could revolutionize current clinical trial practice. Treatment prior to the onset of dementia may also ensure intervention occurs before irreversible neuropathology.

The aim of the project is to determine the ability of a multiplex plasma based immunoassay panel to discriminate among disease states and to monitor disease progression over a one year period. The multiplex panel is based upon luminex immunoassay technology and a 151 analyte panel has been developed by Rules Based Medicine (RBM) to measure a range of inflammatory, metabolic, lipid and other disease relevant endpoints. Prior studies using an older version of the RBM panel (an 89 analyte version) suggested some analytes on the panel differed between AD and controls. The panel has been expanded to include analytes from a recent article describing plasma based biomarkers of AD.

The analyses described in this statistical analysis plan should be regarded as exploratory and meant for hypothesis and model generation, rather than for hypothesis confirmation and model validation. Findings will need to be confirmed and expanded upon in subsequent studies using other, independent data sets.

2 Study Design and Objectives

2.1 Study Design

Samples from baseline and one year ADNI plasma sample set will also be assessed (N=229 Controls, 192 AD, 398 for amnestic mild cognitive impairment (MCI)). Of the 398 MCI subjects, 153 subjects have progressed to dementia as of March 2010. This statistical analysis plan addresses the analysis of data from a subset of these samples.

Previously, a small pilot study was run using University of Pennsylvania banked plasma samples from AD (N~98) and control (N~72) subjects. Data from the pilot study have already been analyzed **(Hu et al., in preparation)**. Assessing the utility of the expanded RBM panel will incorporate some of the findings from the pilot study.

2.2 Study Objectives

- To determine whether baseline levels or changes from baseline levels for individual analytes are associated with patient demographics (age, gender) or disease status.
- To determine whether baseline levels for a combination of analytes derived from either a biological pre-selection based method and/or from a statistically based/machine learning language approach will provide a panel with distinctly different profiles for the ADNI normal controls (NC), MCI or AD.
- To determine whether baseline levels for a combination of analytes derived from either a biological pre-selection based method and/or from a statistically based/machine learning approach will provide a panel that discriminates pre-demented subjects who will progress to dementia in one year and/or two years.
- To determine whether change from baseline levels for a combination of analytes (derived as above) predict cognitive decline in AD or correlate with disease progression.

3 Univariate Analysis

Univariate analyses will be performed first. The results of the univariate analyses may be used to inform and select analytes to be used in the pathway analyses and multivariate predictive model-building. Results from the univariate and multivariate sets of analyses will be compared for overlap and a final panel selected based on optimal overlap.

3.1 Classification Endpoints

Clinical diagnosis at time of enrollment/collection will be used to classify AD, MCI and control groups. Clinical diagnosis of amnestic MCI followed by diagnosis of AD will be used to classify pre-demented progressors.

3.2 Data Quality Control (QC)

Up to 190 analytes may be measured in the plasma updated RBM panel. Plasma data will be analyzed separately and compared for each analyte dependent upon sample availability. The data will be prepared for all analysis as follows:

- Review of the quality control samples data for each run to determine the variability characteristics of the spiked plasma (or serum) QC samples. Characteristics examined for the LOW, MEDIUM and HIGH QC samples for each biomarker will include mean, standard deviation (SD) and the percent coefficient of variation (%CV) for each analyte to determine not only the variability at each concentration but whether or not there is a major change in variability across the concentration range for each analyte.
- Analytes with more than 10% missing data will not be analyzed further. Missing data are generally indicated by "QNS" (quantity not sufficient for analysis) by RBM.
- Analytes with more than 10% recorded as "LOW" will not be included in the multivariate analysis. These analytes will be assessed to compare the proportion of measurable samples in each disease status category. If proportions differ substantially among disease status categories for some analytes, alternative approaches may be explored for incorporating such analytes in the multivariate analyses described below.
- For expression values preceded by a "<" or ">" modifier, the numeric portion of the value will be used for all subsequent data preparation and analyses.
- For each analyte, the distribution of measured values within each diagnostic group will be examined. If the distributions are not normal, the team will seek appropriate transformations (e.g., Box-Cox transformations **(Box and Cox, 1964)** so the transformed markers approximate normality. All subsequent data preparation and analyses will then be conducted on the transformed values.
- Analytes with less than 10% missing/"LOW" values will have the non-numeric values imputed as follows:
    - Values recorded as "LOW" will be imputed to LLD/2
    - Missing values will be imputed to be the mean of the non-missing values for that analyte.
    - Samples with imputed values for more than 25% of the analytes will be excluded from the analysis
- Multidimensional scaling and/or Mahalanobis distances will be used to detect sample outliers and misclassified subjects.
- For univariate analysis, outliers that are more than 5 STD from mean will be assigned the value of the nearest non-outlier point. For outliers apparent in

multivariate reviews, outliers will be imputed using a nearest neighbor or other appropriate algorithm.

The imputation and outlier definition strategy defined above is only one of many possible strategies that could be used. If resources permit a limited number of alternative strategies may be used to assess the robustness of the analytical conclusions obtained using the strategy defined above.

As part of data QC, patient, visit, and sample identifiers will be checked for uniqueness and logical consistency. Graphical displays will be used to check for systematic patterns related to batch, run date, sample quality measures, and QC sample characteristics.

Cleaning, outlier detection, and distribution displays of all samples will be performed prior to merging phenotype data with the biomarker data. Misclassification assessment will be performed prior to statistical analysis.

3.3 General approach

Analysis of variance (ANOVA) and analysis of covariance (ANOCOVA) models will be used to compare mean analyte levels among groups of interest. These ANOVA/ANOCOVA models will include the diagnosis/disease status group and other covariates including age, gender and apoE4 genotype/status, as well as possible interactions among these factors. The interactive effect between group and other covariates will be tested. Depending on the outcome of these tests, the differences between groups will be tested either by the main effect of diagnosis or the effect of diagnosis at a fixed level of other covariates (i.e., apoE4 status) or through the adjusted least square means.

A major analytic concern in these tests is the control of overall type I error rate due to the relatively large number of CSF and plasma proteins tested in this aim. The team will address this concern using false discovery rate (FDR) methodology.

3.4 Hypotheses to Be Tested

The following univariate hypotheses will be addressed for each analyte:

HO1i: Analyte i is not associated with age [age treated as a continuous variable]
HO2i: Analyte i is not associated with gender
HO3i: Analyte i is not associated with ApoE status

HO4i: Analyte i is not associated with disease status or change in disease status (adjusted for age, gender, and/or ApoE status as necessary)

An initial set of analyses will look at whether the mean baseline level of each individual marker differs among disease groups (normal, MCI, AD) via an ANOVA or ANCOVA and t-test analysis. "Disease status" will be based on the clinical calls recorded in the ADNI database. Additional analyses may be conducted using disease status defined using one or more alternative definitions based on cognitive and/or functional tests.

False discovery (FDR) corrections will be applied to p-values and will be reported along with raw p-values. When adjusting for and evaluating the impact of multiple tests, a distinction will be made between:

- a set of proteins defined a priori as being of particular biological interest based on review of pathway annotations (see Table 1)
- the remaining assayed proteins.

FDR corrections based on the Benjamini-Hochberg method will be calculated separately for the two sets.

A second set of analyses will be performed using data only from MCI subjects. ANOVA/ANCOVAs similar to the above will be run to assess whether mean baseline levels of the analytes differ among MCI non-converters and converters.

A third set of analyses will be run to determine whether change from baseline analyte levels at one year are associated with change in disease status.

A fourth set of analyses will be run to determine whether any of the analytes correlate with significant changes in Clinical Dementia Rating Scale-Sum of the Boxes (CDR-SB) or Auditory-Verbal Learning Test (AVLT).

A fifth set of analyses will determine whether levels of any of the analytes are associated with low CSF abeta/high tau, high amyloid brain burden and significant brain atrophy.

Analyses to examine relationships between analyte levels and use of acetyl cholinesterase inhibitors or other medications by subjects may also be performed.

4 Pathway Analysis of Biomarkers

Although statistical machine learning-based approaches can generate a short list of discriminatory proteins, such analyses reveal little about biological relevance. In addition to machine learning approaches, the current proposal will use a systems biology approach to better understand pathway relationships between identified proteins. The Project Team will use pathway mining tools, such as those offered by Ingenuity and Pathway studio, to find the functional connections between the markers from plasma samples. This will provide direct evidence to support key hypotheses. To further increase the biological

relevance of the protein markers in the predictive models, biomarkers will be selected based on their presence in distinct biological pathways.

In addition empirical characterizations of marker data such as pair-wise correlations or higher-order relations (e.g. principal components analysis (PCA)) will be used. This analysis will derive an initial short list that will then be analyzed using multivariate and machine learning language approaches.

5 Multiple Marker Models

Multivariate statistical methods and multiple machine learning approaches will be used to identify the optimal combination of a group of proteins to predict disease status. The problem of classification and prediction has received a great deal of attention in mining "-omics" data. In the case of this project, the task will be to classify and predict the diagnostic category of a sample on the basis of protein quantitative profiles. The main type of statistical problem is the identification of "marker" genes that characterize the difference between diagnosis groups (e.g. AD, MCI) – the so called "variable/feature selection" problem. One challenge is to find the optimal combination of uncorrelated proteins. This factor not only is very important to improve prediction accuracy but also contributes to the merits of a good classifier: the simplicity and insight gained into the predictive structure of the data.

In all multivariate model building, feature selection will be done using data only from the training set. Feature selection based on a completely independent data set is not feasible for this project due to sample size and the fact that this is the first study to use this version of the RBM panel.

Multiple marker analysis will be used to build relationships to the disease groups. The candidate models include: logistic regression, linear discriminant analysis, nearest shrunken centroid, random forests, support vector machines and partial least squares. The technique of **Xiong et al. (2004)** may be applied to search for the linear combination of informative proteins that optimally discriminates between the diagnostic groups. Models generated by the various methods will be compared and the "best" model will be chosen based on model fit, robustness, and parsimony considerations.

Models will be fit with two sets of covariates, 1) assay results only and 2) assays results plus additional patient information including gender, age, and ApoE4 allele status Other biomarkers such as amyloid PIB load, hippocampal atrophy, baseline mini-mental state examination (MMSE), and/or baseline Alzheimer's Disease Assessment Scale-Cognitive Subscale 11 (ADAS-cog 11) may also be used. For a specific model, differences in performance between models fit using the two classes of predictors variables should be characterized to understand the predictive ability of the assays beyond that of routine clinical information on the patients. If possible, formal inference should be made regarding the statistical significance of including the assay variables above and beyond that of the clinical data. Analysis will focus on the following:

- good characterizations of error rates; poor fitting models should not be interpreted.
- any feature selection routines should be extensively cross-validated **(see Ambroise and McLachlan, 2002)**
- measures of marker importance should be biased towards those that use uncertainty (e.g. logistic regression slope tests) as opposed to those that do not (e.g. random forest variable importance, etc).

The multivariate results will be compared to the single marker analysis and (especially) biological relevance.

In addition to the modeling efforts outlined above and described in detail in sections 5.1 and 5.2, an additional set of analyses will be based upon confirming a predictive model developed from the University of Pennsylvania pilot dataset. In brief, a predictive model using pilot data for the 24 analytes shown in Table 1 was identified as providing good discrimination between AD vs Control.

Table 1. Analytes in Predictive Model from UPenn Pilot Study

Analytes	Analyte	Analyte
Alpha-1 Microglobulin ↑	FAS ↑	Resistin ↑
Angiopoietin 2 ↑	HGF ↑	Stem Cell Factor ↑
Apolipoprotein E ↓	IGF-BP2 ↑	Tenascin C ↑
Beta2-Microglobulin ↑	IL-10 ↑	Thrombomodulin ↑
B Lymophocyte Chemoattractant ↑	NT-proBNP ↑	TIMP-1 ↑
Cortisol ↑	Osteopontin ↑	VCAM-1 ↑
E-Selectin ↓	Pancreatic Polypeptide ↑	VEGF ↑
FABP ↑	PAPP-A ↑	Von Willebrand Fact ↑

The model based on pilot data for these 24 analytes will be tested using the ADNI dataset for its predictive abilities to discriminate AD from Control, MCI from Control and MCI rapid progressors to dementia (within 2 years) vs slow or stable MCI The algorithm for the pilot study prediction model will be provided by the model developers.

5.1 Analyte Filtering

Several approaches to filtering and feature selection may be examined. Results of the univariate analyses described above may be used to define a starting set of markers for the analysis. Results of the pathway analysis may also be used to define a starting set. In addition, pre-filtering of markers in an unsupervised fashion prior to building models based on empirical measures may also be applied.

5.2 Model Building Approach

For each type of model, predictive model building will be based on an iterative resampling approach.

For each of the K resampling iterations, the steps will include:

- Splitting the data into training and test sets
- Applying an unsupervised filter on the predictors based on data in the training set only.
- Building and tuning the predictive model on the current training set
- Predicting the current test set
- Calculating and saving the performance (classification accuracy, Kappa)
- End resampling iteration
- Assess performance of the model over the K sets of performance metrics

In the above algorithm, the resampling schemes can include cross-validation, the bootstrap and repeated training/test set splits (see Appendix). Methods for unsupervised feature selection can include filters on variance of individual predictors, high pair-wise predictor correlations, etc.

In addition to the iterative resampling approach above, an additional set of predictive models may be built based on the training-test 60-40 split data sets and 10-fold cross-validation set defined by the ADNI Biostatics Core. This set of models will permit easier comparisons with other modeling efforts performed using the ADNI Biostat core-defined approach.

6 Power Calculations

The sample size for this project and resulting analyses are based upon and limited by the availability of samples from both the pilot and ADNI samples. Additional post-hoc analysis will be completed based upon variability characteristics of the current study to understand power requirements for subsequent analysis of future datasets, in discussion with the Project Team.

7 References

Ambroise C. and McLachlan G.J. (2002), 2002, Selection bias in gene extraction on the basis of microarray gene-expression data. Proc. Natl. Acad. Sci. U.S.A., 99, 6562-6. Epub 2002 Apr 30.

Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. J. Royal Stat. Soc. B26, 211-246.

Hu, W.T., Chen-Plotkin, A., Shaw, L.M., Arnold, S., Grossman, M., Clark, C.M., Leight, S., McCluskey, L., Ellman, L., Karlawish, J., Hurtig, H.I., Siderowf, A., Lee, V.M-Y., Trojanowski, J.Q., Swensen, F., Pickering, E., Kuhn, M., Soares, H.D. A study to assess the diagnostic accuracy of a plasma-based multiplexed immunoassay panel to identify Alzheimer's disease. In preparation.

Xiong, C., McKeel, D.W., Jr., Miller, J.P. and Morris, J.C. (2004). Combining correlated diagnostic tests: application to neuropathologic diagnosis of Alzheimer's disease. Med. Decis. Making 24, 659-69.

8 Appendix I

In order to make sure that there is balance among the pre-specified cross-validation groups, we present a table of the 10 partitions vs. the baseline-year1 status of ADNI subjects. No info from month 6 or other visits were used.

	AD-AD	AD-MCI	AD-NA	MCI-AD	MCI-MCI	MCI-NA	MCI-NL	NL-MCI	NL-NA	NL-NL
a	16	0	3	7	29	5	0	0	0	21
b	16	1	2	10	25	3	1	0	1	23
c	16	0	1	6	32	2	1	0	1	21
d	17	0	3	6	29	4	1	0	1	21
e	15	0	5	5	28	8	0	1	2	21
f	15	0	2	4	29	3	3	2	1	20
g	17	0	5	6	30	2	0	0	2	21
h	18	0	2	6	27	4	1	0	3	19
i	16	1	3	6	31	3	0	0	2	23
j	15	0	4	8	27	3	2	0	3	20

The following table restricts to the primary groups of interest.

	AD-AD	MCI-AD	MCI-MCI	NL-NL
a	16	7	29	21
b	16	10	25	23
c	16	6	32	21
d	17	6	29	21
e	15	5	28	21
f	15	4	29	20
g	17	6	30	21
h	18	6	27	19
i	16	6	31	23
j	15	8	27	20