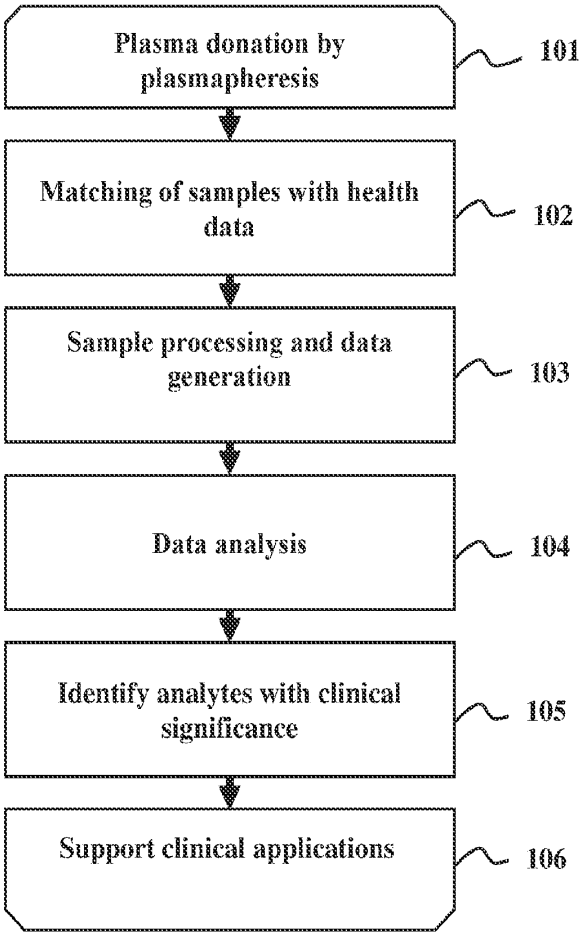

Plasma donation by plasmapheresis
101
Matching of samples with health data
102
Sample processing and data generation
103
Data analysis
104
Identify analytes with clinical significance
105
Support clinical applications
106

FIG. 1

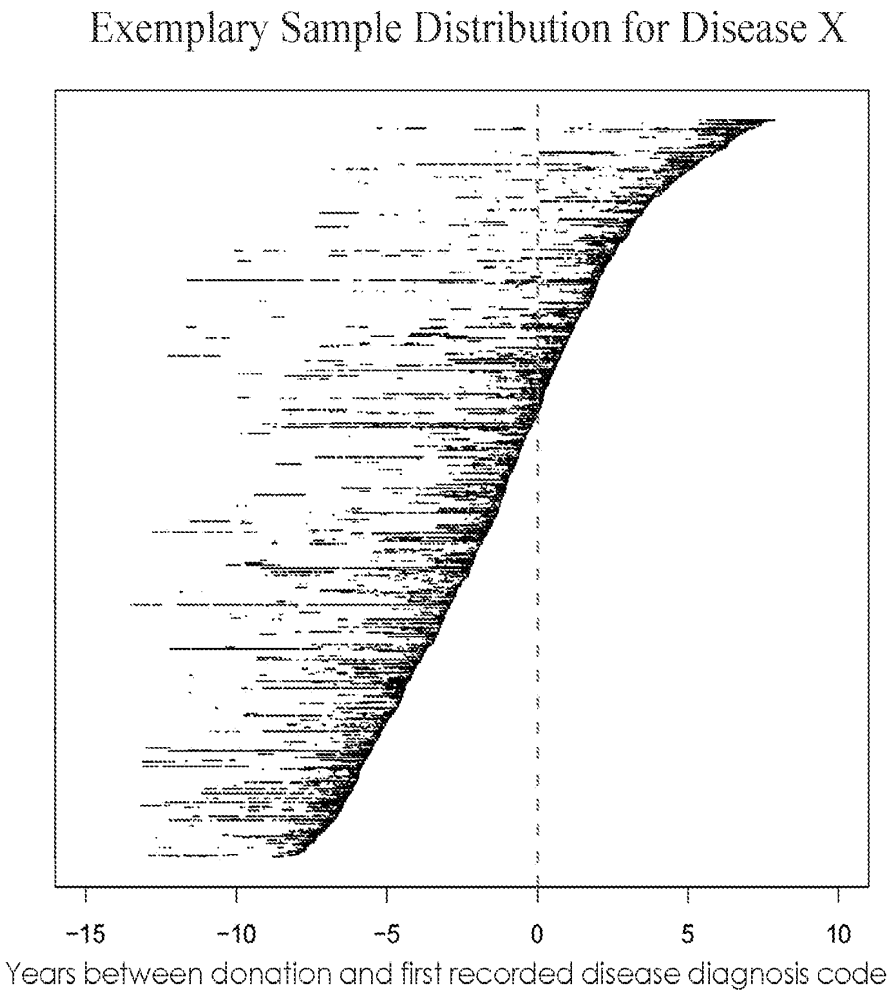

Exemplary Sample Distribution for Disease X
−15
−10
−5
0
5
10
Years between donation and first recorded disease diagnosis code

FIG. 2

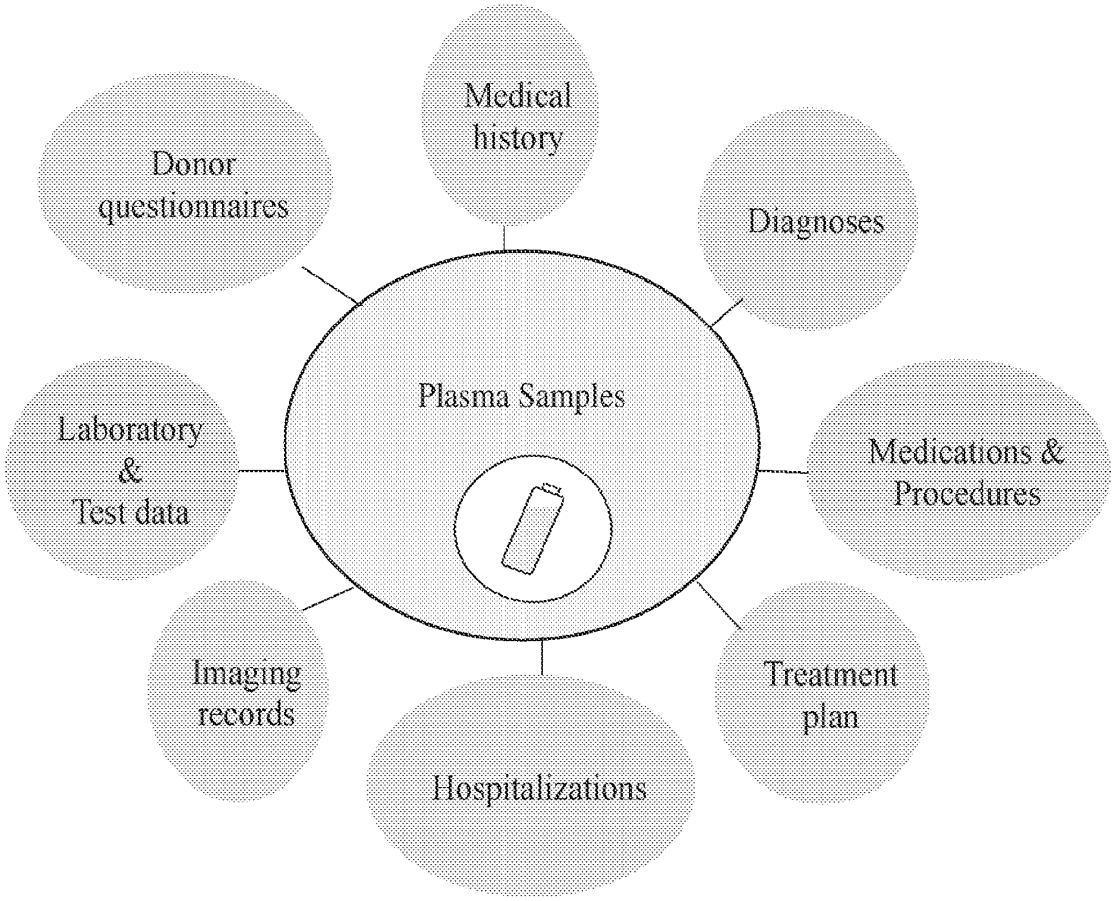

Medical history
Donor questionnaires
Diagnoses
Plasma Samples
Laboratory & Test data
Medications & Procedures
Imaging records
Treatment plan
Hospitalizations

FIG. 3

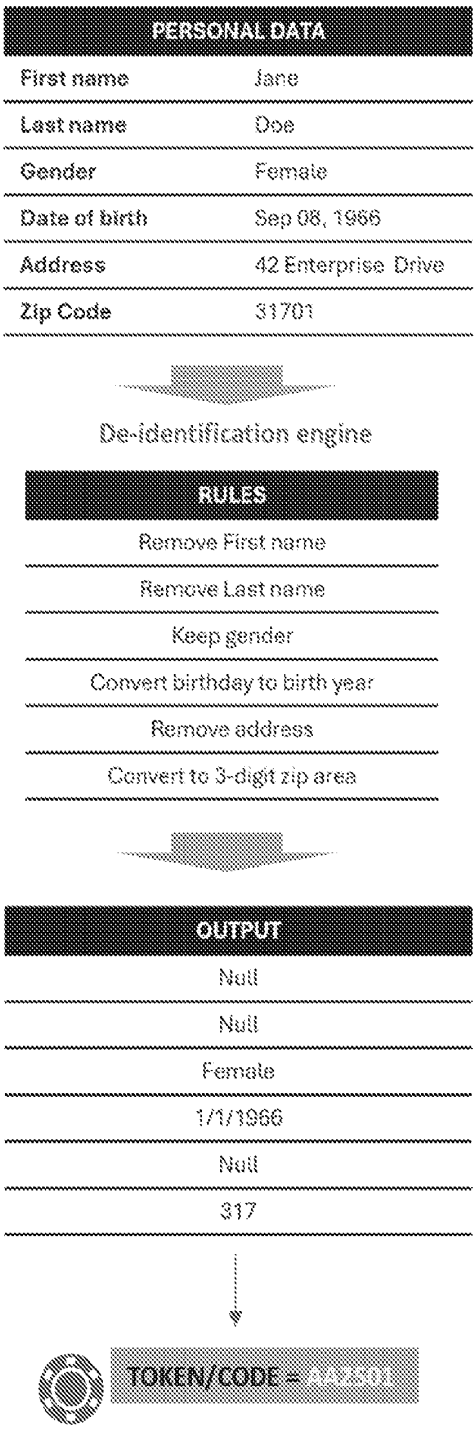

PERSONAL DATA
First name
Jane
Last name
Doe
Gender
Female
Date of birth
Sep 08, 1966
Address
42 Enterprise Drive
Zip Code
31701
De-identification engine
RULES
Remove First name
Remove Last name
Keep gender
Convert birthday to birth year
Remove address
Convert to 3-digit zip area
OUTPUT
Null
Null
Female
1/1/1966
Null
317
TOKEN/CODE =

FIG. 4

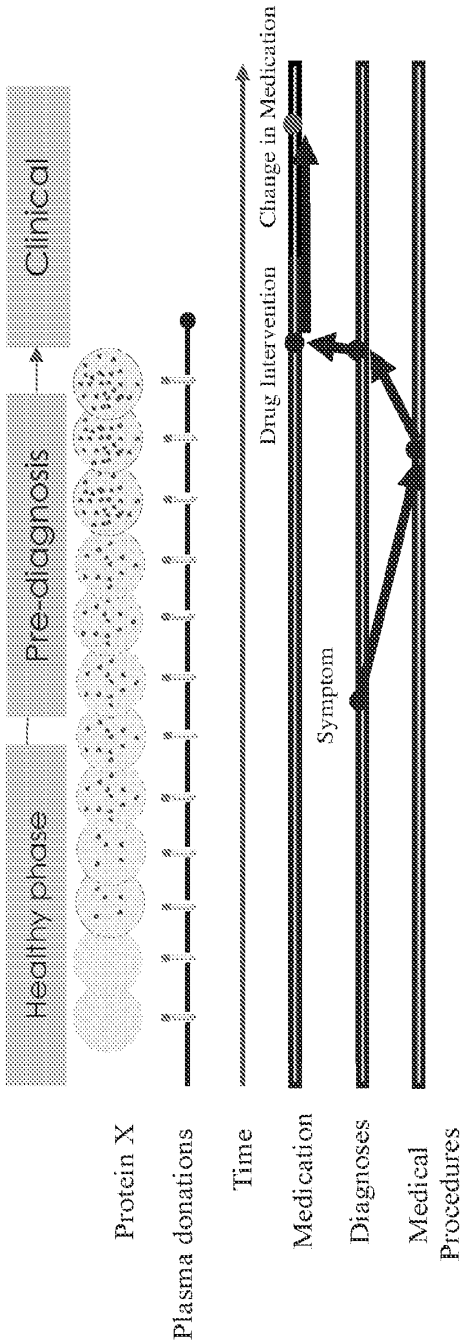

Healthy phase
Pre-diagnosis
Clinical
Protein X
Plasma donations
Time
Medication
Diagnoses
Medical Procedures
Drug Intervention
Change in Medication
Symptom

FIG. 5

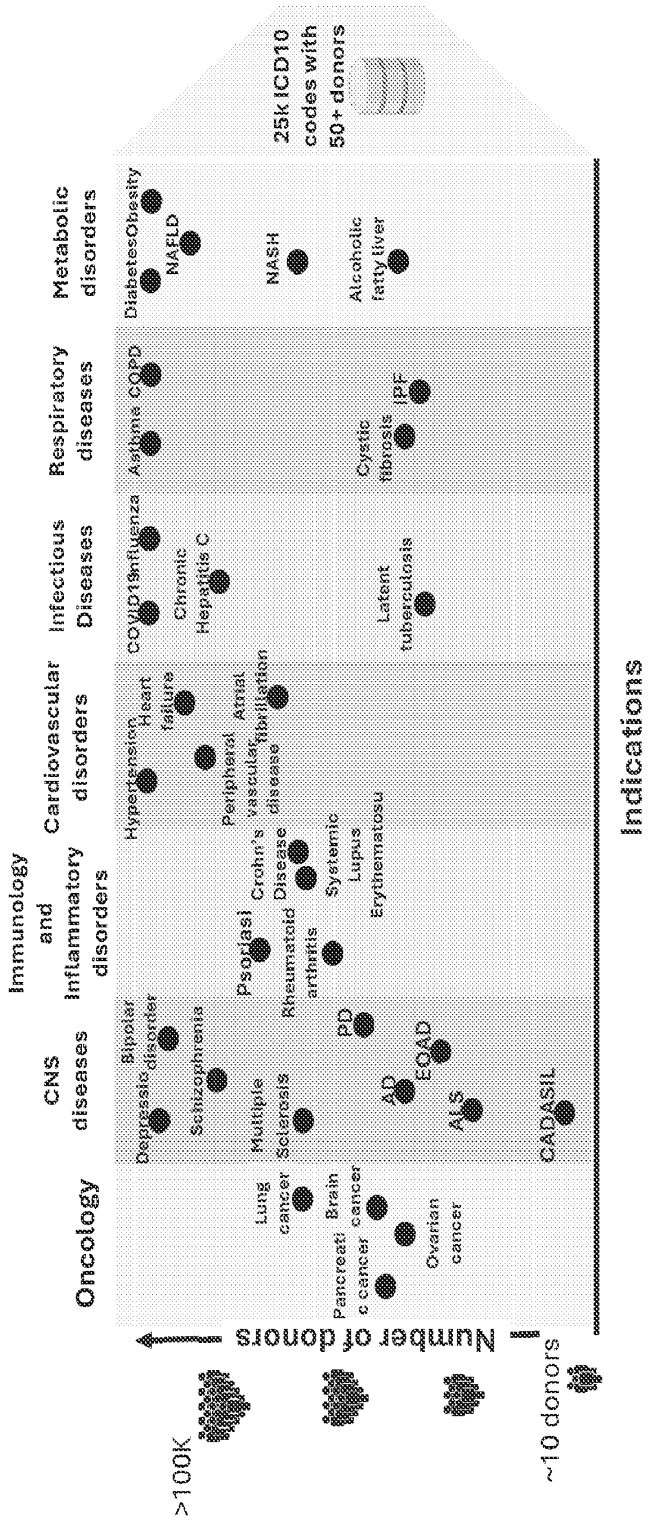

Oncology
CNS diseases
Immunology and Inflammatory disorders
Cardiovascular disorders
Infectious Diseases
Respiratory diseases
Metabolic disorders
Lung cancer
Brain cancer
Pancreatic cancer
Ovarian cancer
Depression
Bipolar disorder
Schizophrenia
Multiple Sclerosis
AD
EOAD
PD
ALS
CADASIL
Psoriasis
Rheumatoid arthritis
Crohn's Disease
Systemic Lupus Erythematosu
Hypertension
Heart failure
Peripheral vascular disease
Atrial fibrillation
COVID19
Influenza
Chronic Hepatitis C
Latent tuberculosis
Asthma
COPD
Cystic fibrosis
IPF
Diabetes
Obesity
NAFLD
NASH
Alcoholic fatty liver
25k ICD10 codes with 50+ donors
Number of donors
>100K
~10 donors
Indications

FIG. 6

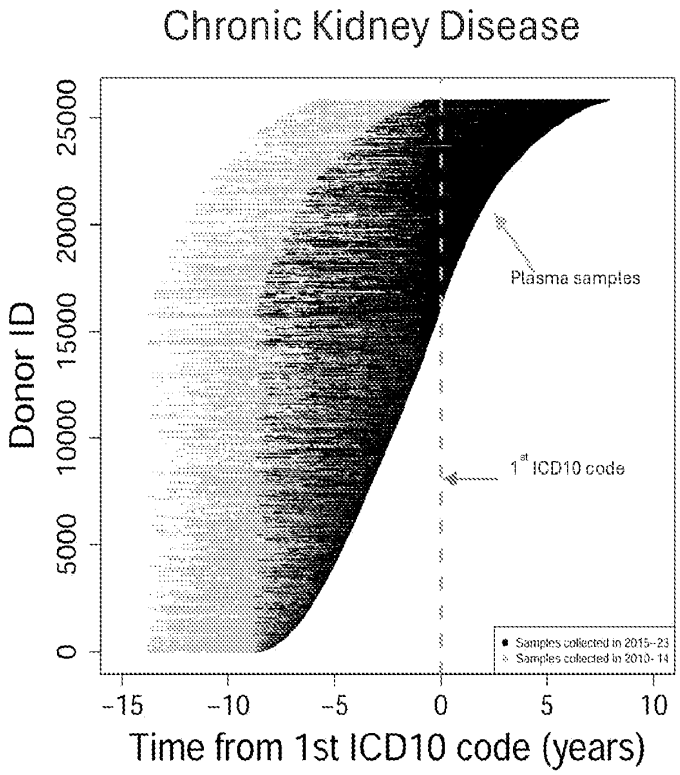

Chronic Kidney Disease
Donor ID
0
5000
10000
15000
20000
25000
Plasma samples
1st ICD10 code
Samples collected in 2015--23
Samples collected in 2010- 14
−15
−10
−5
0
5
10
Time from 1st ICD10 code (years)

FIG. 7

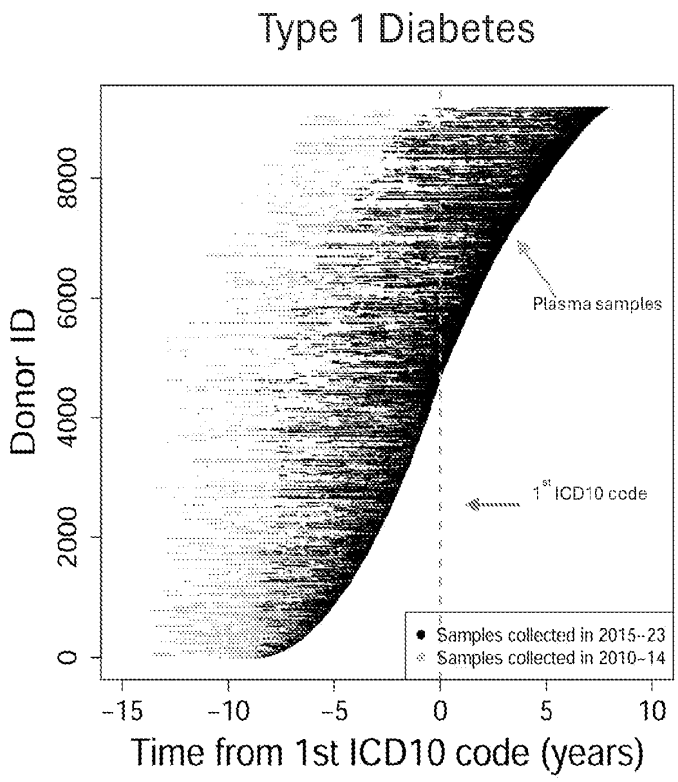

Type 1 Diabetes
Donor ID
0
2000
4000
6000
8000
Plasma samples
1st ICD10 code
Samples collected in 2015--23
Samples collected in 2010--14
−15
−10
−5
0
5
10
Time from 1st ICD10 code (years)

FIG. 8

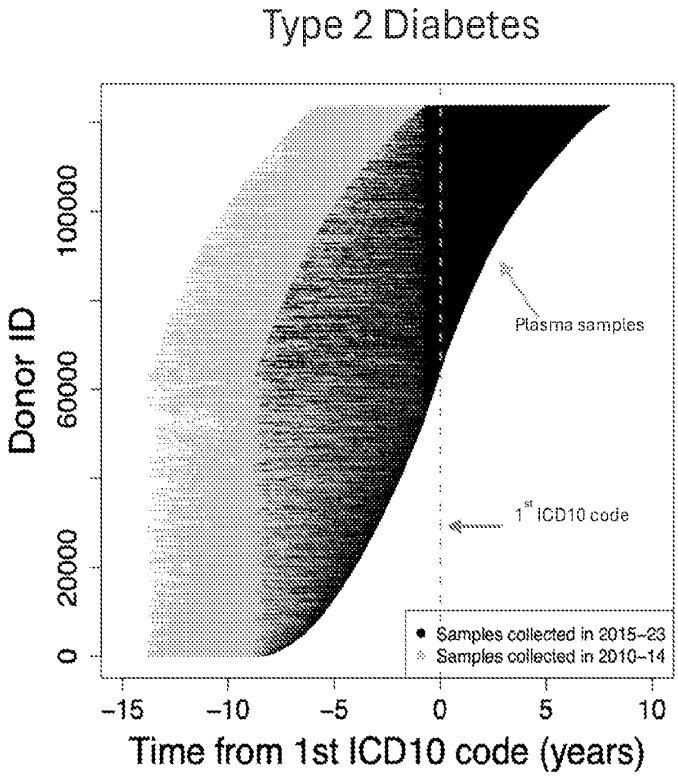

Type 2 Diabetes
Donor ID
0
20000
60000
100000
Plasma samples
1st ICD10 code
Samples collected in 2015–23
Samples collected in 2010–14
−15
−10
−5
0
5
10
Time from 1st ICD10 code (years)

FIG. 9

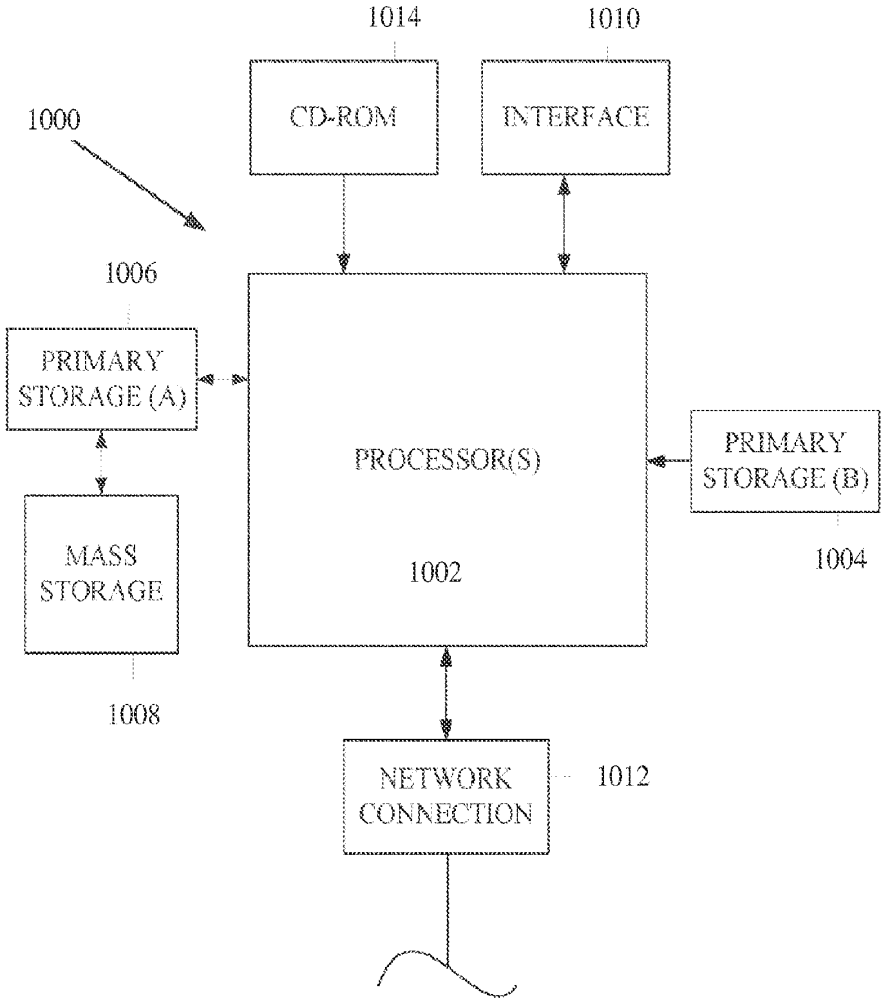

1014
1010
CD-ROM
INTERFACE
1000
1006
PRIMARY STORAGE (A)
PROCESSOR(S)
1002
PRIMARY STORAGE (B)
1004
MASS STORAGE
1008
NETWORK CONNECTION
1012

FIG. 10

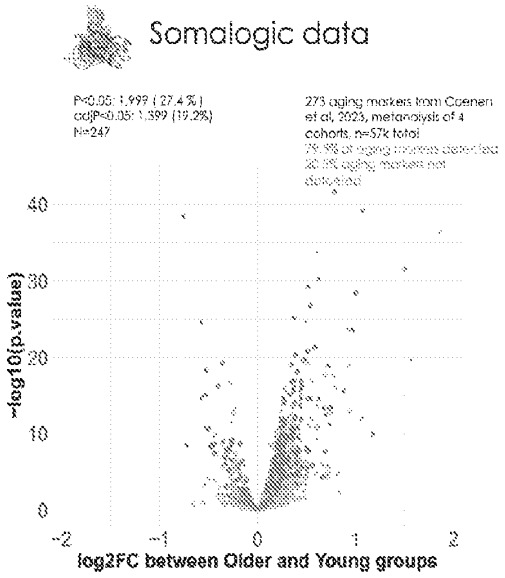


FIG. 11A

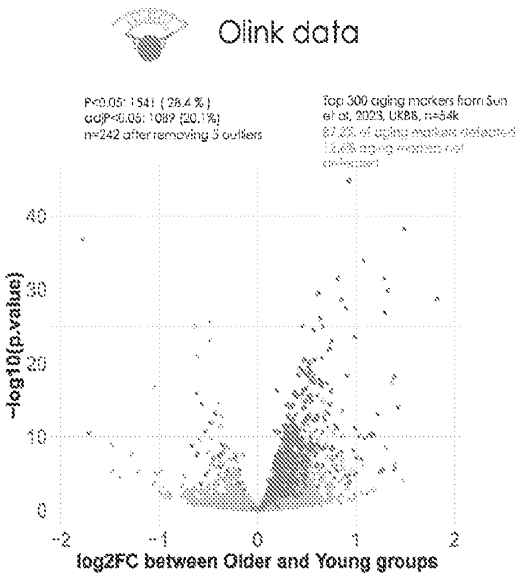


FIG. 11B

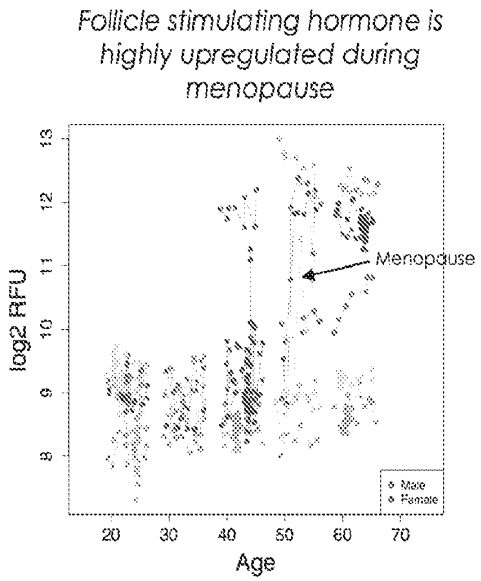


FIG. 12A

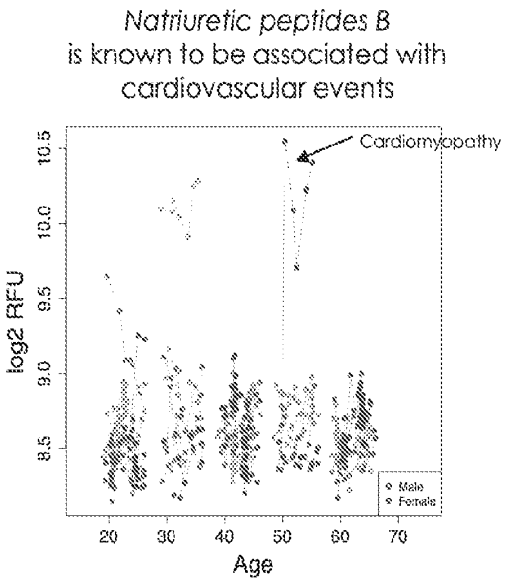


FIG. 12B

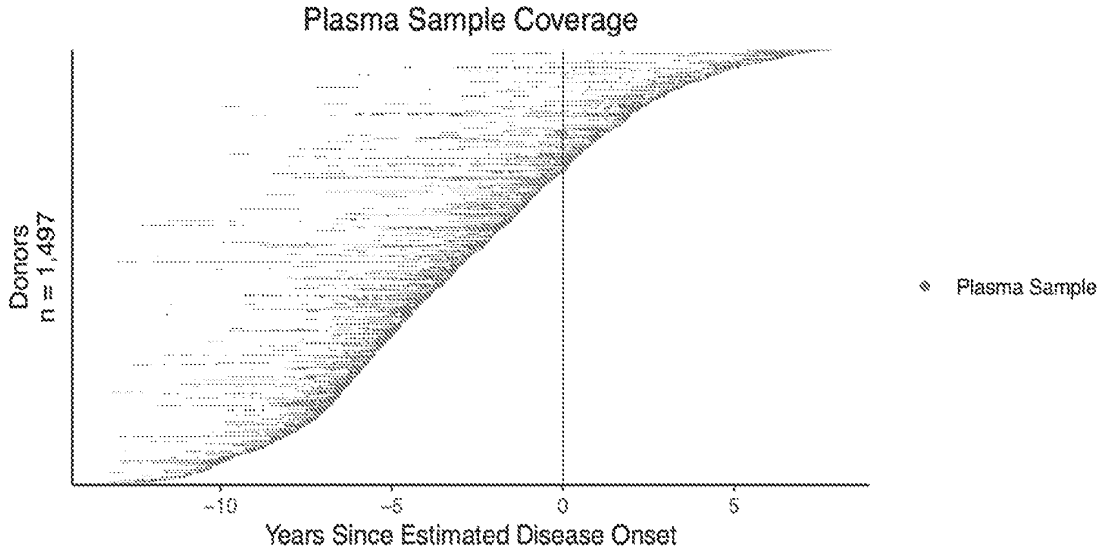

Plasma Sample Coverage
Donors
n = 1,497
−10
−5
0
5
Years Since Estimated Disease Onset
Plasma Sample

FIG. 13

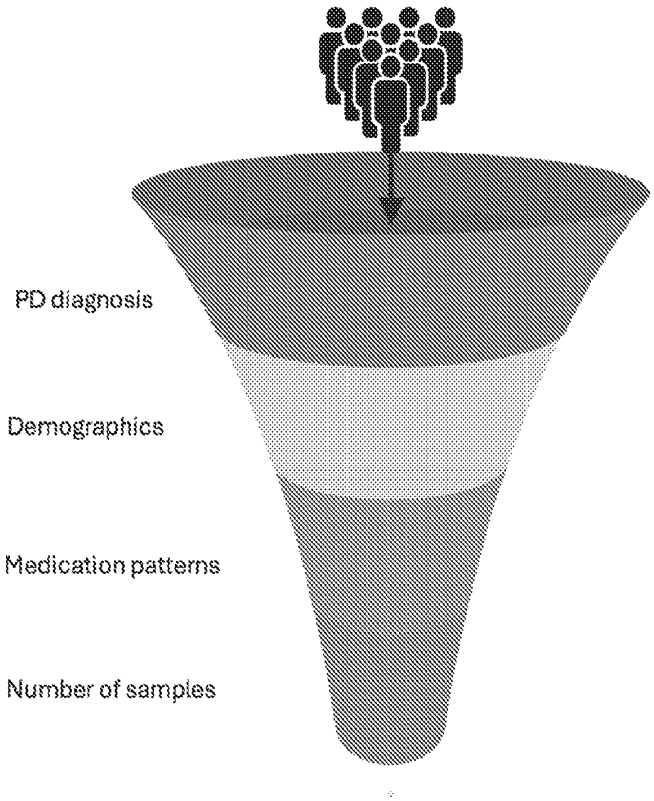

PD diagnosis
Demographics
Medication patterns
Number of samples

FIG. 14

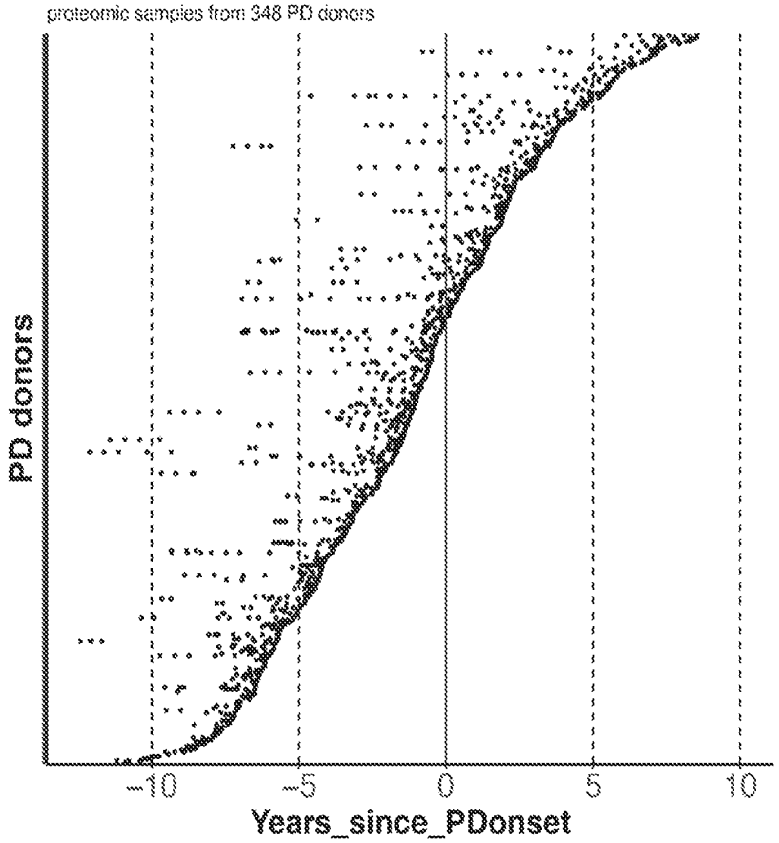

proteomic samples from 348 PD donors
PD donors
−10
−5
0
5
10
Years_since_PDonset

FIG. 15

	PD, N = 1,323	Control, N = 1,286
Years from PD onset (defined based on 1st G20)	-2.0 (4.3,-12.4,8.5) -2.1 (-5.6,1.0)	
Age at donation (years)	55.6 (8.1,24,70) 57.0 (51,61)	55.4 (8.1,22,70) 57.0 (50,61)
Year of sample collection	2016.9 (3.2,2010,2023) 2017 (2015,2019)	2016.9 (3.2,2010,2023) 2017 (2015,2019)
Days since previous sample selected	167.0 (269.0,1.8, 3319.0) 116.7 (36.0,183.0)	176.5 (227.1,1.7, 2214.1) 137.1 (53.8,188.0)
Number of centers per donor	1.1 (0.3,1.0,2.0) 1.0 (1.0,1.0)	1.1 (0.3,1.0,3.0) 1.0 (1.0,1.0)
Storage duration (years)	7.3 (3.2,1.0,14.0) 7.6 (4.9,9.5)	7.3 (3.2,1.0,14.0) 7.6 (4.9,9.4)

Mean (SD,Minimum,Maximum)
Median (25%,75%)

FIG. 16

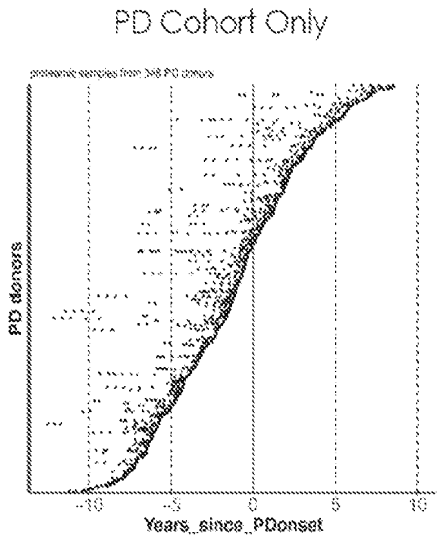


FIG. 17A

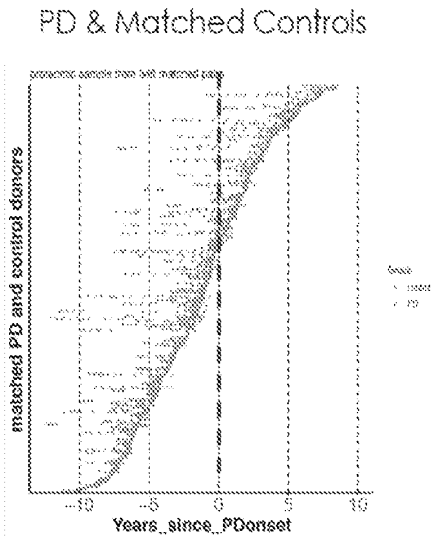


FIG. 17B

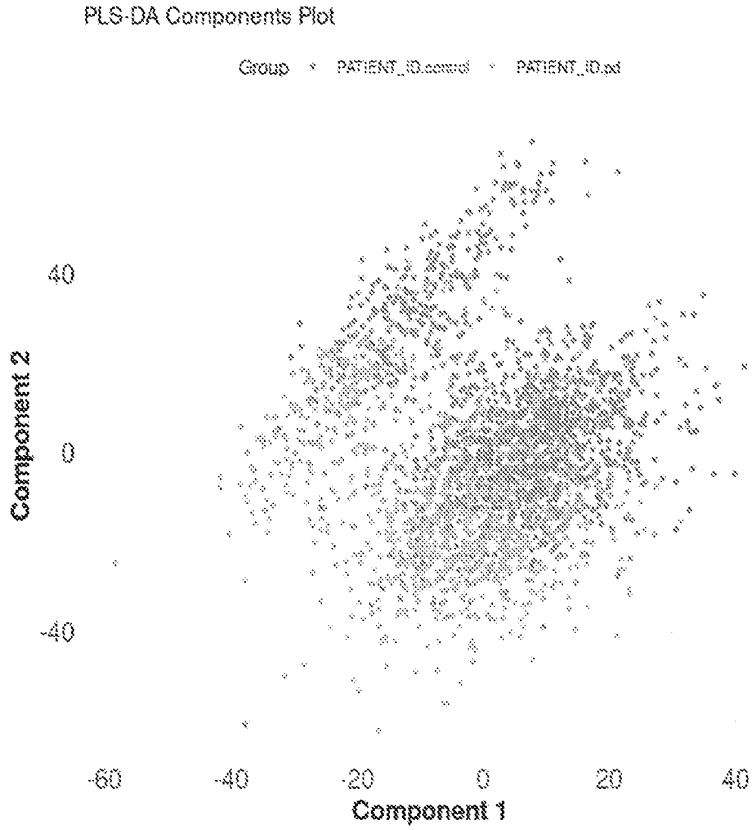

PLS-DA Components Plot
Group
PATIENT_ID.control
PATIENT_ID.pd
Component 2
40
0
-40
-60
-40
-20
0
20
40
Component 1

FIG. 18