An end-to-end workflow for high-throughput discovery of clinically relevant insights from large biomedical datasets - Additional File 2

Table S1. All communities identified by the Minimal Forest and a short description of their variables.

Community	Q	Desctiption	Q	Description	Q	Description
1	HB	Blood Hemoglobin	MPV	Mean Platelet Volume	Platelet	Blood Platelet Count
	Hct	Hematocrit	Mxd	Blood mixed lymphocyte count	RBC	Blood Red Blood Cell Count
	MCH	Mean Corposcular Hemoglobin	Lymph	Blood Lymphocyte Count	RDW	Red Blood Cell Distribution Width
	MCHC	Mean Corposcular Hemoglobin Concentration	Neut	Blood Neutrophil Count	WBC	Blood White Blood Cell Count
	MCV	Mean Corposcular Volume	PDW	Platelet Distribution Width		
2	180	Freq of brushing teeth	183		186	Dental health
	181	Freq of dental floss	184			
	182	do you have full dentures	185			
3	BMI	Body-Mass Index	6	Education	241	Time spent on chat apps
	Fat	Body Fat	11		253	Smoker?

	HipCircum	Hip Circumference	225	Number of glasses of water per day	WaistCircum	Waist Circumference
	MetaP	Resting Metabolic Rate	230	which supplements in the past month?	Weight	
	NeckCircum	Neck Circumference	233	Use of module spoon to add salt		
	2	Gender	240	Time spent on internet		
4	34	Notice mouth is dry	42	Feel don't have sth to be waiting for	51	Notice unusual heart activity without physical activity
	35	Think you can't feel good experience	43	Feel distressed	52	Fear with no reason
	36	Have trouble breathing	45	Feeling heart dead and heart broken	53	Feel life is meaningless
	37	Difficult to volunteer for work	47	Fear at any moment	98	Asthma
	39	Feel tremor in body	48	Not able to show enthusiasm	101	FH of asthma
	41	Worried about panic or stupidity	49	Not feeling valuable		
5	33	Difficulty in staying calm	40	Feel use of a lot of mental energy	46	Can't keep patience
	38	Wild reactions to situations	44	Calmness is difficult	50	Very irritable and sensitive

6	Chol	Blood cholesterol	HDL	Blood high-density lipoproteins	UAC	Blood uric acid concentration
	Crea	Blood creatinine	LDL	Blood low-density lipoproteins		
	FatPerc	Body fat percentage	TG	Blood triglycerides		
7	20	Sleep duration throughout a day	26	Have nightmares	196	Fastening Seatbelt
	21	Time to falling asleep	32	Number of close friends	212	Method of cooking fish
	22	Consumption of painkillers and tranquilizers	54	Self-assessment of health	213	Candy consumption
	23	Freq of not falling asleep in 30 minutes	58	Self-assessment of energy levels	217	Type of cooking
	24	Wake up at night to go to toilet	107	Depression	221	
	25	Wake up at night or early morning	110	Family history of depression	222	Fruit consumption
	229	Type of dairy product	232	Put salt on the table		
8	12	Freq of vigorous Physical activity	29	Time spent with family and friends	224	Removal of chicken skin before cooking
	14	Freq of moderate PA	30	Time spoke to family and friends on phone	226	Number of glasses of milk per week
	16	Freq of low PA	207	Freq of eating breakfast	227	Number of cups of yogurt per week

	17	Duration of low PA sessions	223	Removal of fat from meat before cooking	228	Number of cheese meals per week
9	27	How many hours of TV	203	Hamburgers, hot dogs	211	Freq of beans
	28	How many hours of movies and satellite	204	French fries, samosa, falafel	216	Dental health
	197	What do you drink with food	205	Olivier salad	218	Freq of muffins and cakes
	198	Freq of carbonated drinks	206	Canned food	219	Freq of snacks
	199	Freq of sweetened drinks	208	Freq of eggs	220	Freq of vegetables
	201	Main reason for buying fast food	209	Freq of meat		
	202	Amount spent on fast food	210	Freq of chicken		
10	55	Limitation in physical activity due to physical problem	78	Nocturnal snoring problem	174	
	56	Limitation in work due to physical problem	166		177	
	57	Amount of body pain in last month	170	PMH past month neck pain	178	
	59	Restriction of social relationships by mental and	171	PMH past month back pain	179	

	Code	Description	Code	Description	Code	Description
		physical problems				
	60	Amount of hurt by mental problems	172	PMH past month knee pain		
	61	Restriction of daily activities by physical and mental problems	173	PMH of gout		
11	FSG	Fasting Serum Glucose	70	FH of heart disease	125	FH of arthritis
	Height	Height	79	FH of sudden death	158	Bladder surgery
	MusclePerc	Muscle Percentage	80	Diabetes	214	Freq of middle-sized sugar cubes
	Pulse	Pulse rate	88	FH of diabetes	215	Freq of fried food in week
	1	Age	89	Blood cholesterol increased	234	Employment
	3	Duration of current residence	92	FH of increased blood cholesterol	237	Type of medical insurance
	4	Native of Yazd	93	Blood coagulation problems	238	Hours sitting at work
	7	Number of household members	97	FH of blood coagulation problems	239	Hours sitting at work or public or private transport
	8	Marital status	115	Osteoporosis	242	Referral to traditional practitioner
	18	Time of going to bed	120	FH of osteoporosis		

	19	Time of waking up	121	Arthritis			
12	BloodPreD	Diastolic blood pressure	10		160	Kidney stone surgery
	BloodPreS	Systolic blood pressure	62	Heart disease?	168	Open heart surgery
	9		71	HTN?	126	Skin cancer
13	102	Thyroid problems	137	FH of lung cancer	145	FH of stomach cancer
	106	FH of thyroid problems	141	FH of colon cancer		
14	75	History of mental illness?	114	FH of Alzheimer's	149	
	111	Alzheimer's disease?	142	Stomach cancer		

Table S2. Predictors for osteoporosis identified by the EN algorithm

Predictor	Description	The observed correlation	Previous mentions in medical literature
HB	(Blood Hemoglobin)	Negative	
RBC	(Blood Red Blood Cell Count)	Negative	
WaistCircum	Waist Circumference	Positive	
MusclePerc	(Muscle Percentage)	Negative	
FatPerc	(Fat Percentage)	Positive	
BloodPreS	Systolic Blood Pressure	Positive	
MetaP	Resting Metabolic Rate	Negative	
Height	Height	Negative	

Table S3. Predictors for resting metabolic rate as identified by the Elastic Net algorithm.

Predictor	Description	The observed pattern	Previous mentions in medical literature
MusclePerc	Muscle Percentage	Positive	
RBC	Blood Red Blood Cell Count	Positive	
BMI	Body Mass Index	Positive	
HB	Blood Hemoglobin	Positive	
UAC	Blood Uric Acid Concentration	Positive	
Fat	Body Fat	Positive	
NeckCircum	Neck Circumference	Positive	
Height	Nocturnal snoring problems - Pain in knee joint	Positive	
Weight	Frequency of sleeping pills - Candy consumption	Positive	