Public Data

	O1	...	n
D1			
...			
n			

Patient Data

	O1	...	n
P1			
...			
n			

Data Modeling
Clustering
Inference

Hypothesized stratifications

Biologically-relevant stratifications

Clinically-relevant stratifications

serendipidous subspace

Validated stratifications

Biological value

Improved process	Improved outcomes
Hypothesis generation, prioritization	Accelerated discovery of disease genes
	Accelerated discovery of biomarkers, targets

Translational value

Accelerated discovery of biomarkers, targets

Accelerated discovery of drugs

Accelerated discovery of cohorts

Hypothesis generation for epidemiological studies

Clinical value

Improved process	Improved outcomes
Accurate "next" steps (phenotyping, sequencing, labs, treatment)	Lower costs
Faster, more accurate diagnoses	Improved health
Faster, more accurate prognosis	
Accurate prediction of treatment response	