Clinical Characterization of Trial Participants Using Electronic Health Record Data
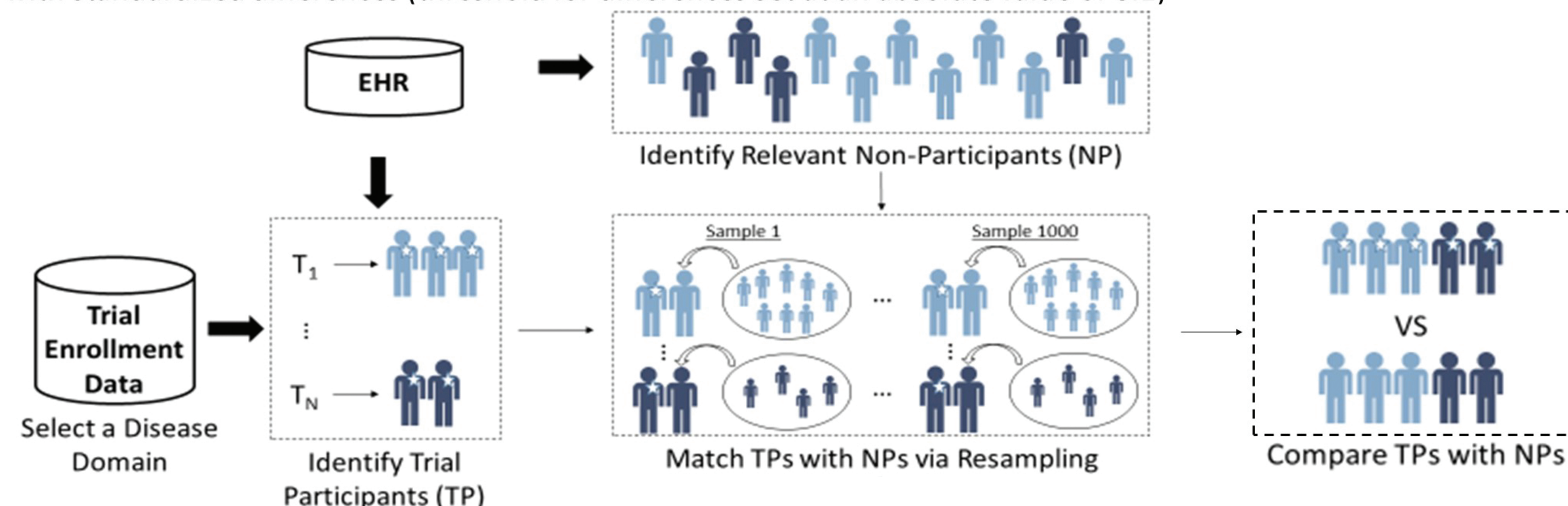
OHDSI · PRESENTER: **James Rogers**

IMPORTANCE

- Clinical trials are integral for medical evidence generation, but generalizability concerns persist
- We aim to provide clinical characterization of trial participants using the electronic health record (EHR) data, and to compare those profiles with those of non-participants

METHODS

- Using 3 disease domains, we created trial participant (TP) cohorts with corresponding matching trial non-participant (NP) cohorts via resampling, then compared underlying clinical characteristics using the FeatureExtraction package with standardized differences (threshold for differences set at an absolute value of 0.1)
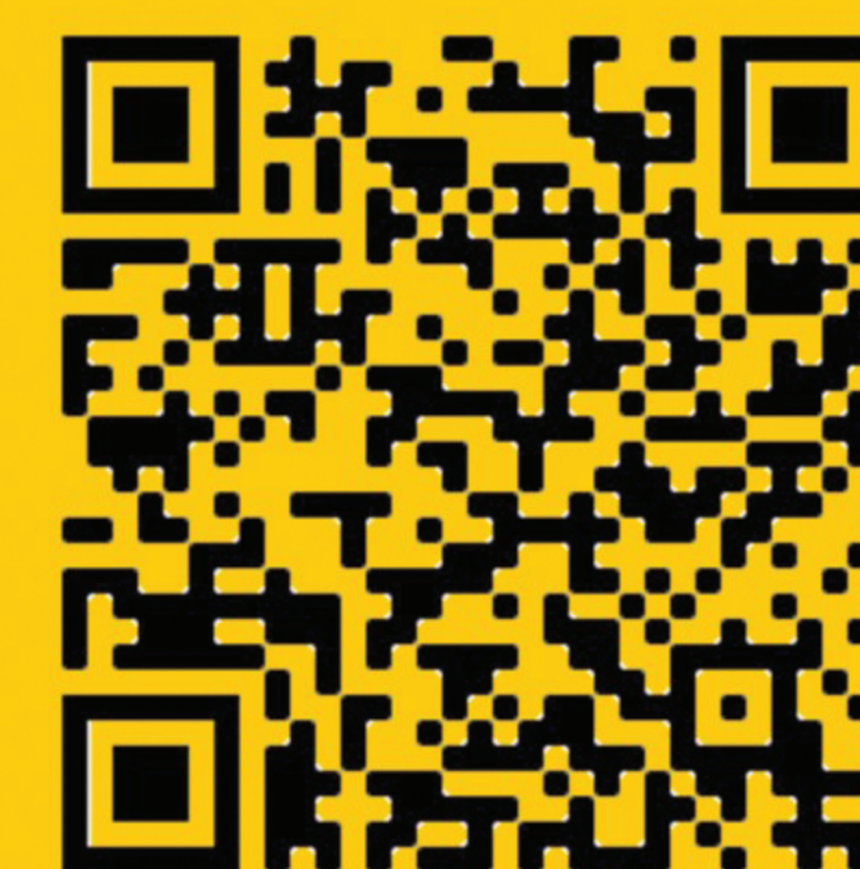

RESULTS

- Below is a summary table of the standardized difference for **underlying conditions** between TPs and NPs in each disease domain, where blue means there were more NPs than TPs while orange means there were more NPs than TPs. The QR code provides more in-depth details and other covariates

	Neoplastic disease	Disorder of digestive system	Inflammatory disorder
Medical history: General			
Acute respiratory disease	-0.106	-0.146	-0.163
Chronic liver disease	-0.087	0.012	0.012
Chronic obstructive lung disease	-0.141	-0.129	-0.191
Depressive disorder	-0.145	-0.186	-0.161
Diabetes mellitus	-0.131	-0.178	-0.252
Gastroesophageal reflux disease	-0.151	-0.215	-0.241
Hyperlipidemia	-0.171	-0.280	-0.150
Hypertensive disorder	-0.227	-0.241	-0.246
Lesion of liver	0.232	0.210	0.001
Obesity	-0.134	-0.208	-0.223
Osteoarthritis	-0.121	-0.122	-0.127
Pneumonia	-0.175	-0.248	-0.150
Renal impairment	-0.221	-0.177	-0.172
Urinary tract infectious disease	-0.178	-0.210	-0.183
Viral hepatitis C	-0.080	0.041	0.043
Visual system disorder	-0.125	-0.103	-0.167
Medical history: Cardiovascular disease			
Atrial fibrillation	-0.162	-0.154	-0.130
Cerebrovascular disease	-0.084	-0.081	-0.115
Coronary arteriosclerosis	-0.192	-0.141	-0.088
Heart disease	-0.223	-0.072	-0.177
Heart failure	-0.189	-0.150	-0.268
Ischemic heart disease	-0.199	-0.252	-0.241
Peripheral vascular disease	-0.113	-0.070	0.119
Venous thrombosis	-0.141	-0.105	-0.080
Medical history: Neoplasms			
Hematologic neoplasm	0.012	-0.057	0.000
Malignant lymphoma	-0.016	-0.144	0.119
Malignant neoplastic disease	0.015	0.028	0.033
Malignant tumor of breast	-0.075	-0.078	0.119
Malignant tumor of lung	0.133	0.077	0.090
Malignant tumor of urinary bladder	-0.015	0.110	0.000
Primary malignant neoplasm of prostate	0.015	0.075	0.090

IMPLICATION

- Our study based on integrated EHR and trial data reaffirms prior generalizability concerns across a diverse set of trials while also providing a more comprehensive approach to evaluating trials

Based on a diverse sample of trials, we found **trials tend to enroll patients that are healthier** than their non-enrolled counterparts

Take a picture to download the full paper

James R Rogers, MS[1]; Cong Liu, PhD[1]; Ying Kuen Cheung, PhD[2]; Chunhua Weng, PhD[1] PhD
[1]Department of Biomedical Informatics, Columbia University, New York, NY, USA; [2]Department of Biostatistics, Columbia University, New York, NY, USA