

National Cancer Policy Forum Workshop: Session 3

Interoperable Electronic Health Records in Cancer Surveillance

Pete Stetson, MD, MA, FAMIA
Chief Health Informatics Officer

Enabling 21st Century Applications for Cancer Surveillance through Enhanced Registries and Beyond
Challenges and Opportunities: Data Collection Methodologies and Technological Advances in Cancer Surveillance
National

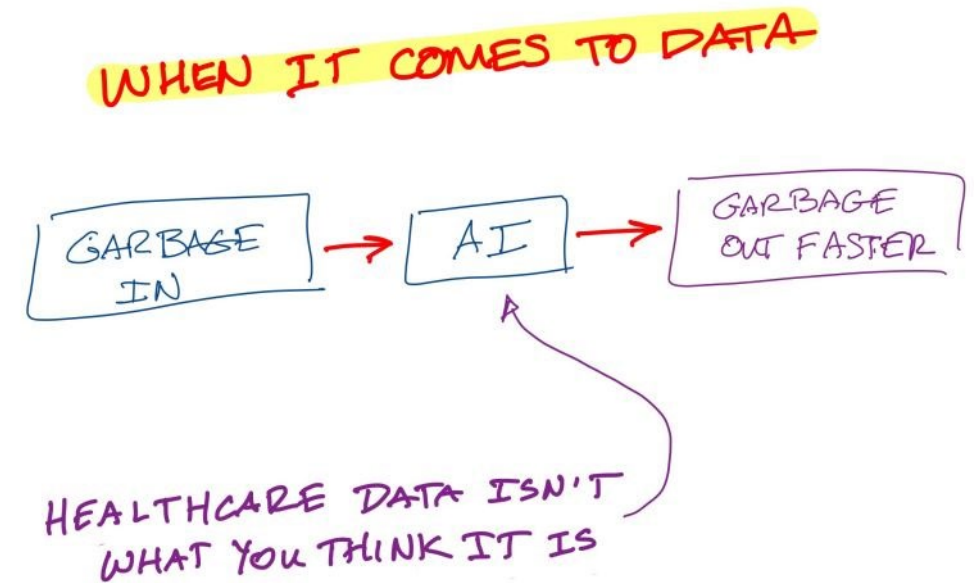


Memorial Sloan Kettering
Cancer Center

Problem:

High-quality, sharable data

- We want to predict cancer onset, tumor evolution, treatment response, toxicity burden so we can improve outcomes and value
- We want to tailor therapies to the whole patient
- Can't do these without **good data – still a constraint on AI**



- Dale Sanders

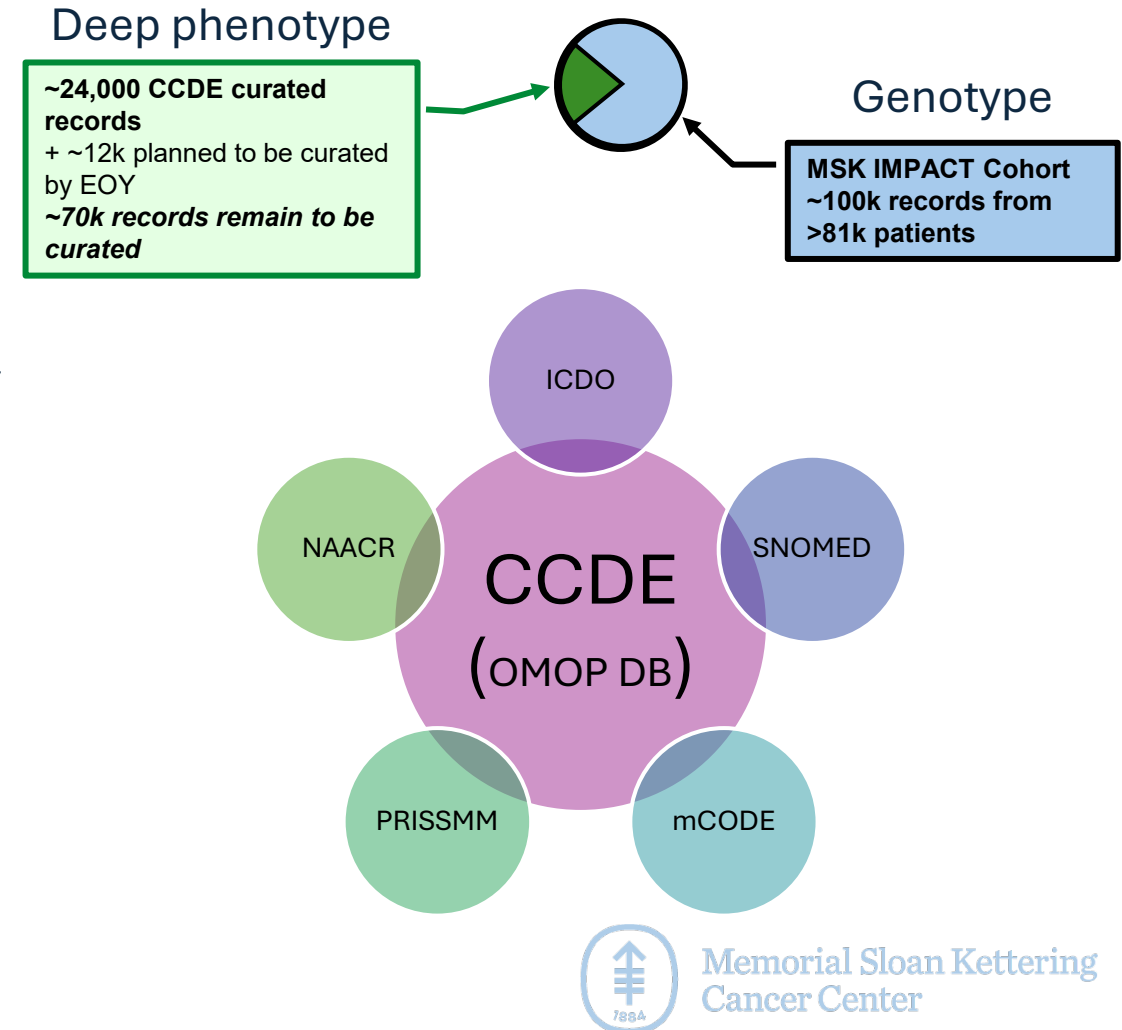


Memorial Sloan Kettering
Cancer Center

Approach:

Tackling real-world problems through quality data

- Deep phenotyping of genotyped cohort
- Standardizing data using Core Clinical Data Elements (CCDE)
 - n=2000 concepts; ~200 typically really hard to get
- Current Use Cases (sample)
 - Curation pipeline for researchers
 - Cancer registry automation pipeline
 - Clinical trial cohort management
 - Predicting tumor evolution
 - Fusion models exploration



Outputs

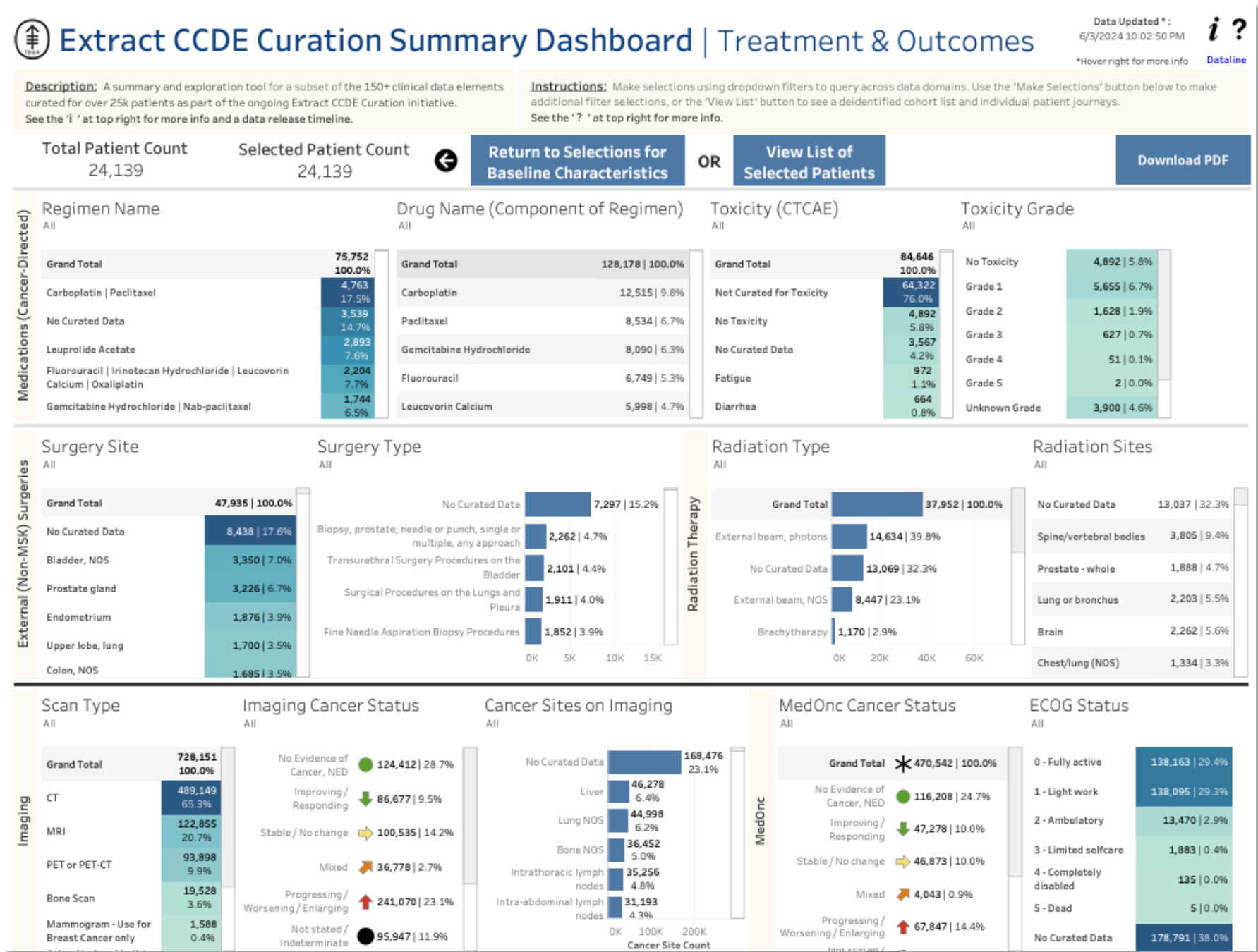
Concept-based phenotypes (and ePROs phenotype-lets)

MSK Extract Pipeline

Effort led by

John Philip, MS

Neil Shah, MBBS

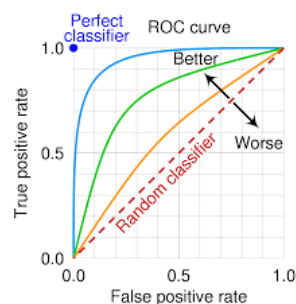


Benefits

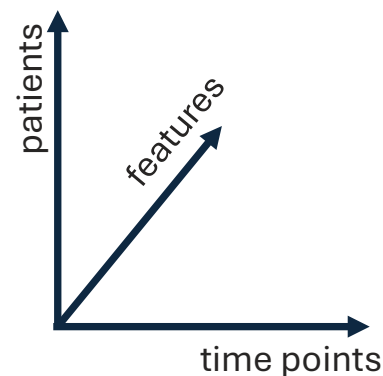
Impact on AI in Oncology; Learning Health Systems



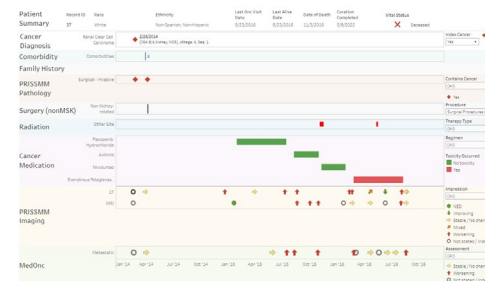
Relieve the
data quality
constraint for
AI



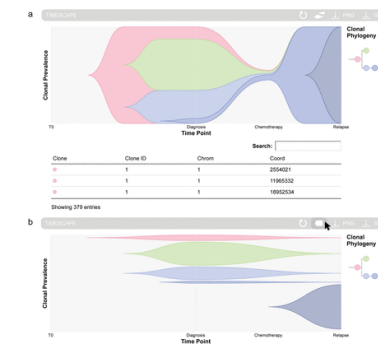
Provide
proving
grounds for
AI models



Data density
further sped
by AI-
enabled
curation



Computable
phenotypes (like
eMerge) and
journeys



Smith MA et al Nature Methods

Combined
Tumor/Clinical
Response
predictions



Memorial Sloan Kettering
Cancer Center

Future State:

Possible roadmap

- **Automated cancer registry** at national scale
- Federated tuning of **AI-enabled curation** (generalized model assurance)
- Industry/Academic partnerships for a **computable phenotype ecosystem**
 - Major EHR and major cloud compute companies enable sharing/publishing of pragmatic computable oncology phenotype (PCOPs) definitions into local EHR instances (Like TurboBooster for research order sets)
 - Your data stays local
 - AI tools know what they're analyzing
 - Obtain semantic interoperability, not just data interoperability
- **Cross-institution computable journeys** (care started at hospital A, continued at hospital B, failed 1st-3rd line Rx now on trial, toxicity at Hospital C, etc)

