Regulatory & access considerations in mining big data

to advance the development of cancer immune modulator therapeutics

Ahmad A. Tarhini, MD, PhD
Senior Member, Cutaneous Oncology & Immunology
H. Lee Moffitt Cancer Center & Research Institute
Professor, Oncologic Sciences
University of South Florida Morsani College of Medicine
Chair, ORIEN Scientific Committee & ImmunoOncology Research Subcommittee

DISCLOSURES

- Research support to institution: Bristol Myers Squib, Genentech-Roche, Regeneron, Sanofi-Genzyme, Nektar, Clinigen, Merck, Acrotech, Pfizer, Checkmate, OncoSec, Scholar Rock, Agenus.
- Consultant role/Advisory board: Bristol Myers Squibb, Merck, Eisai, Instil Bio, Clinigin, Regeneron, Sanofi-Genzyme, Novartis, Partner Therapeutics, Genentech/Roche, BioNTech, AtraZeneca, ConcertAl, InstilBio.

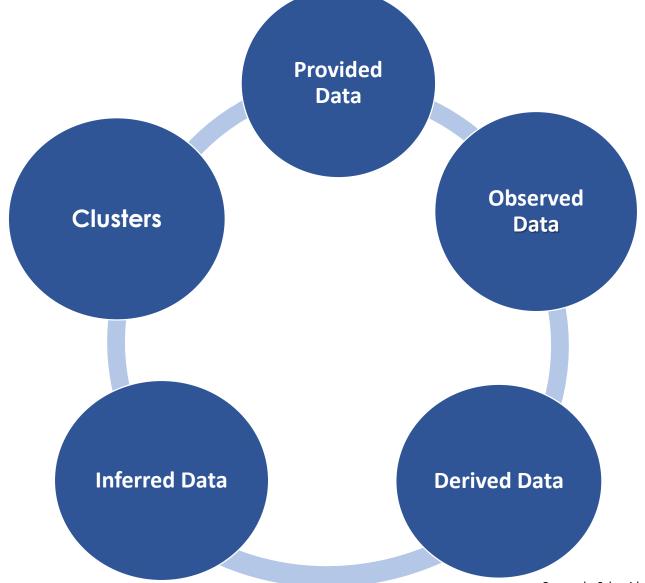
Big Data science increasingly impacts drug development & personalized medicine

Genomics & other biological high throughput data (transcriptomics, epigenomics, proteomics, metabolomics) are expanding at an accelerated pace and are being increasingly integrated with other sources of data

Primarily through new diagnostic and informatics, appropries that on the molecular basis of disease biology and host immunology

- Bio-imaging
- Health claims data
- Data from wearable sensors & mobile health applications
- Other ... (Primary & secondary health data sources)

The never-ending lifecycles of health data



Sources of 'omics' data ever expanding ...

- The Cancer Genome Atlas Program (TCGA)
- Gene Expression Omnibus (GEO)
- Dependency Map (DepMap) Consortium
- The database of Genotypes and Phenotypes (dbGaP)
- The Cancer Imaging Archive (TCIA)
- Clinical Proteomic Tumor Analysis Consortium (CPTAC)
- Sequence Read Archive (SRA)
- Single Cell Expression Atlas
- Other sources at NCBI, EBI, others ...

Health Insurance Portability and Accountability Act (HIPAA)

- Requires patient authorization to disclose potentially identifying data
- Or a decision an institutional reviews erandivars albeard to personal tealth requiremention
- General rule: Personal health information cannot be disclosed without the patient's authorization
- Three primary measures:
 - 1. Informed consent
 - 2. Study review and approval by **IRBs** and Research Ethics Committees (**RECs**)
 - 3. Data use limited to **deidentified** information
- Intent: Participants are fully aware of the research, and given a clear and accurate account of the potential risks and expected benefits

General Data Protection Regulation (GDPR)

- GDPR (May 2018) applies to all personally identifiable information (PII) related to people in U.K. or E.U. but covers organizations even if they operate outside of those jurisdictions
- Health Data covered under 'Special Categories' of Art. 9 of GDPR
- Organizations must appoint a Data Protection Officer (DPO) to oversee the implementation of GDPR & HIPAA

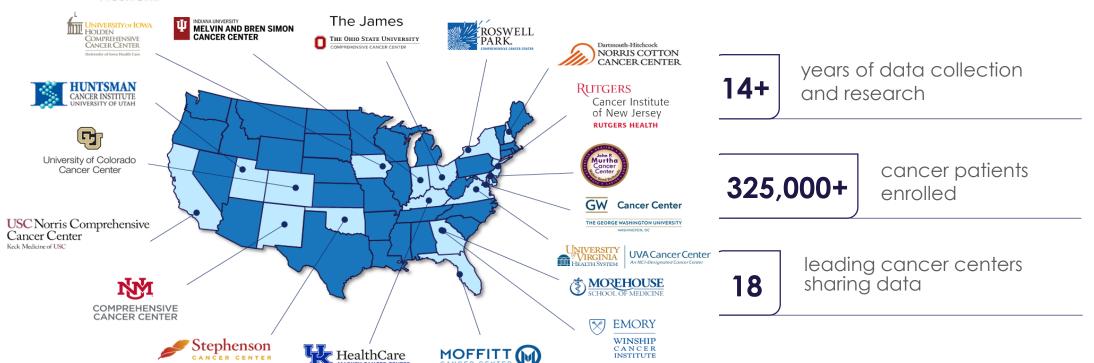
Key Elements in Maximizing Privacy in Big Data Research

- Ensuring HIPAA and GDRP organizational compliance
- Thoughtful governance of Big Data Sharing
- Improve EHR systems
- Patients given the information and opportunity to ask questions needed to give meaningful informed consent to future uses of their data
- Particular attention to racial and ethnic minorities
- Engagement of patients/patient representatives
- Whether to further regulate patient data sharing, who should regulate it, and how it should be regulated ...



78-LINIVERSITY of OKLAHOMA

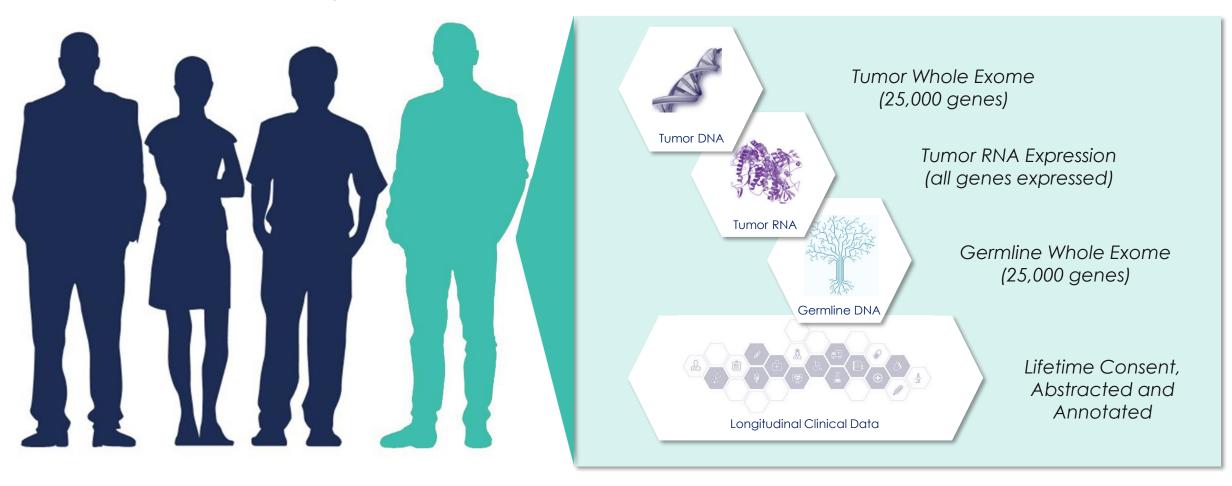
An alliance of 18 U.S. cancer centers, which have implemented a common longitudinal protocol (Total Cancer Care; NCT03977402)



Mission: Accelerating cancer discovery and delivering hope through collaborative learning and partnerships

ORIEN AVATAR

Oncology database with **whole exome**, **whole transcriptome and germline sequencing** data coupled with **longitudinal clinical profiles** from **consented patients**





Commit to using a common protocol to ensure data quality and promote partnerships in research

Total Cancer Care (TCC) Protocol

- IRB Review and Approval: Institutional IRB of record
- **Consent:** Patients individually consented at their respective institutions
- One central honest broker: M2GEN where data is de-identified
- Governing Body: Representatives from the 18-member cancer centers, Patient Advisory Council & M2GEN



Patient Demographics Across ORIEN

Collabo

African American

Morehouse, EMO, CINJ, UVA & GW

o us d prc

Rural & Frontier HCI, SIM & OSU

resea

Appalachian MAR & OSU

Inclusivan

Participation by Concer Centers, Community roviders

nmunit

derse

Adolescents, Young Adults &

Pediatrics

RPC, UCC,

MAR & MCC

American Indian

UNM, SCC, HCI

nts in tatio bns

Active Military & Veterans

MURTHA

Asian and Pacific Island

USC

Hispanic

MCC & USC

Notes:

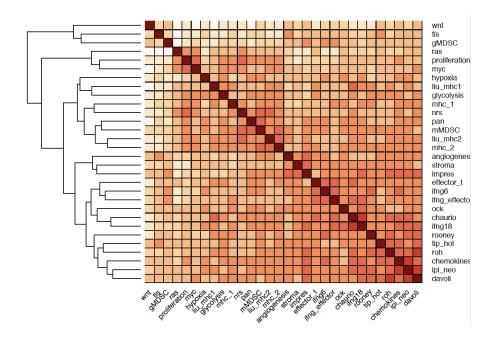
- NCI defined race and ethnicity criteria
- NCI / NIH demographic taxonomy





Predictors of immunotherapeutic benefits in patients with melanoma

- Pearson's correlation of 28 predictive gene signatures in melanoma based on AUROC
- Heatmap illustrates 2 main clusters:
 "immune active" (right/lower) &
 "immune suppressive" (left/upper)



Top 7 GE signatures & correlation with survival (>24 vs. <24 months) in cutaneous melanoma (n=123 pts)

Gene	Description	Reference	AUROC	U-test
Signature				(P-value)
IFNg-6	6-gene IFNg associated	Ayers 2017	0.670	0.0013
IFNg-18	18-gene IFNg	Ayers 2017	0.661	0.0024
Effector T cell	Effector T cell gene set	Bolen 2011	0.676	0.0009
IFNg/Effector T cell	IFNg & effector T cell gene set	Fehrenbacher 2016	0.679	0.0007
Chemokine	Chemokine gene set	Coppola 2011 Mule 2012	0.644	0.0067
MHC II	13 MHC-II genes	Liu 2021	0.629	0.0152
Cytotoxic/Hel per T cell & B cell interplay	T & B cell interplay	Tarhini 2017	0.642	0.0075

- Signatures related to IFNg, effector T cells, chemokines, MHC II & Tertiary Lymphoid Structures were significantly predictive of ICI benefits in advanced melanoma
- Ongoing analyses are investigating these signatures in other malignancies and integrating the GE data with data related to TMB, somatic mutations and germline genetic variations



ORIEN Guiding Principles

Collaboration

Commit to using a common protocol to ensure data quality and promote partnerships in research

Inclusiveness

Participation by Cancer Centers, Community Providers and Patients in their respective communities with representation of diverse and underserved populations

Accessibility of Data

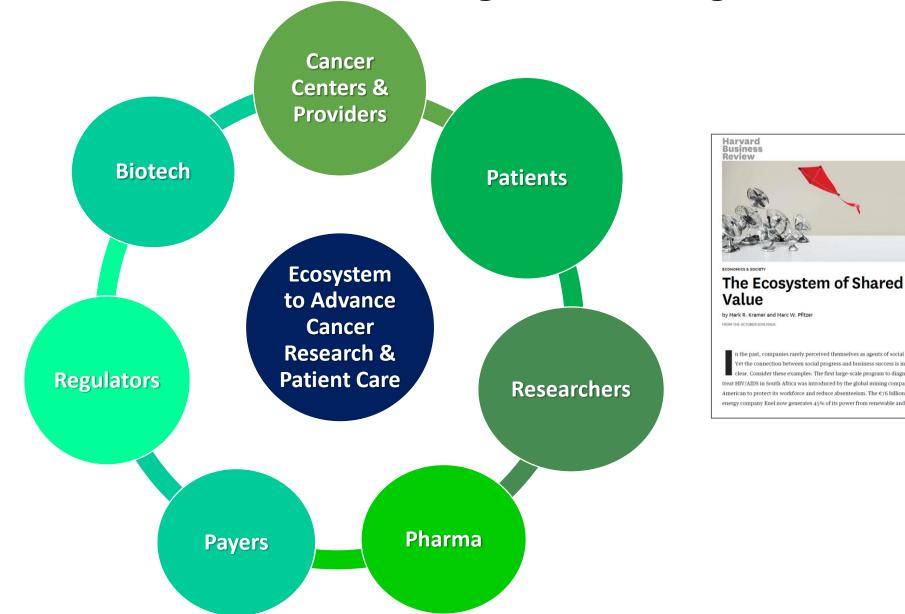
Access to aggregated, harmonized research data; centrally processed and maintained by M2GEN

Partnerships

Establish trusted business relationships within and across sectors to achieve long-term program sustainability

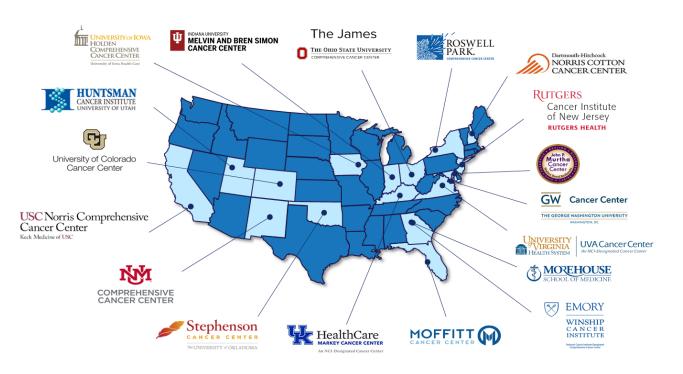


An Ecosystem of Shared Values With the Primary Goal of Advancing Cancer Research while Serving & Protecting the Patient



Acknowledgements

ORIEN & M2GEN









Member Name	City	State
Rutgers Cancer Insitute of New Jersey	New Brunswick	NJ
Emory Winship Cancer Institute	Atlanta	GA
George Washington Cancer Center	Washington	DC
Holden Comprehensive Cancer Center	Iowa City	IA
Huntsman Cancer Institute	Salt Lake City	UT
UK Markey Cancer Center	Lexington	KY
H. Lee Moffitt Cancer Center & Research Institute	Tampa	FL
Morehouse School of Medicine	Atlanta	GA
Murtha Cancer Center	Bethesda	MD
Norris Cotton Cancer Center	Lebanon	NH
The James Cancer Hospital	Columbus	ОН
Roswell Park Comprehensive Cancer Center	Buffalo	NY
Stephenson Cancer Center	Oklahoma City	OK
IU Health Simon Cancer Center	Indianapolis	IN
UCHealth Cancer Care-Anschutz Medical Campus-University of Colorado Cancer Center	Aurora	СО
UNM Comprehensive Cancer Center	Albuquerque	NM
USC Norris Comprehensive Cancer Center	Los Angeles	CA
UVA Emily Couric Clinical Cancer Center	Charlottesville	VA

Thank you!