

- Sammy Yuan (GSK)
- Kara Barwell (GSK)
- Andrew Rainboldt (Intelligencia AI)
- Gerry Liaropoulos (Intelligencia AI)
- Elena Hilario Martinez (Intelligencia AI)
- Gary Summers (GSK)

Moneyball in Drug Development

Learning about clinical trials from curated, publicly available clinical trial data

GSK

Intelligencia AI

Who They Are

- Founded in 2017 in New York
- 100+ people (>90% advanced degrees)
- Patented technology on Probability of Success
- Collaborations with leading pharmaceutical companies

What They Do

Intelligencia AI

 \mathbf{O}

- Expertly-curated & granular data on clinical trials
- Proprietary and harmonized ontologies
- Al-driven PTRS & Explainability
- Modular and granular benchmarks & baselining
- Insights for Competitive Intelligence & Scenario Analysis

🔞 intelligencia.ai	Te Select Drug	Indica Liqui	ition id cancer, Solid c	Modality ancer Select N	lodality	Target		Mol	4 ⊙ lect MaA	Phase Phase 2, Phase	1	Primary Spo	nsor ry Spornor	p
12 Portfolio Optimizer	Search table	9, 8,1	195 orgoing proj	rans 🛈								Ö Custi	enize display	≜ Export as
Oxenview Deep Dive	Drug 💭	Indication 0	Line of Treatment	Population Characteristics	Adjuvant / Neoatjuvant Therapy	Triab	Latent Phase	Primary Sporsor	Modality	Mechanism of Action	Phase Transition () Probability	: PTRS ©	Historical Approval Q Rate	Deep Dive
Clinical Benchmarks Biology	Lisocabtagene Microleccal	Lymphoma, b-cell,	Group 01:	Group 01: Previously	Non adjuvant	NCT04245839	Phase 2	Celgere	Autologous CAR T	B-lymphocyte antigen	87%	- 76%	8N	2
Trial Dusign	Tislelizunab	Solid cancer	Une II, Une		Non adjuvant.	NCT03736889	Phase 2	BeiGere	Monocional	Programmed cell death	85%	73%	7%	2
O BREAKING PROBATE	Crizotinio	Non-Small cell lung	11, 114 114	MET Alterations	Non adjavant	NCT00585195	Phase 1	Pfpr	Other small	Hepatocyte prowth	275	- 73%	45	ε.
	Association	Leukemia, myeloid.	Line I		Non adaptant	NCT02544438.	Phase 2	BioSight	Therapeutic	DNA Disrupting Agent.	85%	- 73%	75	
	Talouetamab	Autore Multiple myeloma	Line III+	Previously Treated	Non adjuvant	NCT03435846	Phase 2	Jamien	Especific T-cell	T-cell surface	855	- 725	14%	2
	Brentukimab vediotin	Classic hodgkin	linell	Previously Treated	Non adjuvant	NCT02572167	Phase 1/2	Seagen	Antibody-drug	Tumor necrosis factor	85%	725	15%	2
	Nivolumab	lymphona http://www.	Line II, Line	With: Unspecified Previously Treated	Nerveland	NCT05396885,	There is a		conjugate/Immun.	Tumor necrosis factor		- 716	100	2
	CART-BENCHA	Leukomia, mveloid.	II, Line II+	With: Unspecified Previously Treated	Non acquirant	NCT04155749 NCT03863100.	Phase 2	Ascentage	Other small	Receptor superfamily	0376	- 125	140	2
	Unverembating	accelerated phase	II, Une II+ Group 01:	With: Unspecified	Non adjovant	IAI78120695 NCT03468751	Phase 2	Pharma	Monoclonal	protein kinase FLT3 Programmed cell death	84%	- 10	1/5	2
	Serplulinab	Solid cancer	Line II, Line		Non adjuvant	NCT03941574	Phase 2	Henlius Biotech	antibody Binascife Turol	protein 1 Antagonist	84%	- 72%	78	
	Blinatumomab	kynphoblastic leukemia	II, Line II+	Providence & Providence &	Non adjuvant.	NCT04521231	Phase 1/2	Ampin March Davis 7	engager antibody	CD19 Binding Agent	85%	- 71%	20%	12
	Pembrolizumab	Bladder cancer		With: BCG vaccine,	Non adjuvant	NCT02625961	Phase 2	Dohme	antibody	protein 1 Antagonist	83%	- 70%	12%	12
			_					,						
		_						,						
								,						
								,						

Ļ

Clinical Trial Data in These Studies

Expertly curated from ClinicalTrial.gov, announcements, conferences & publications

DRUGS

- Mechanisms of action
- Targets
- Modalities
- Genes
- Biological pathways
- Protein classes

PROGRAMS

Defined by the following dimensions:

- Primary drug
- Additional drugs
- Therapeutic area
- Lead indication
- Administration mode
- Intervention
- Primary drug dosage
- Indication
- Line of treatment
- Stage of disease
- Patient selection biomarker
- Other patient characteristics: Sex, Age, Smoking status
- Previously treated with ...
- Adjuvant status
- Sponsor

TRIALS & COHORTS

Defined the program dimensions plus:

- Start date
- End date
- Termination date
- Basket
- Umbrella
- Allocation: random or not
- Masking: Open label, double blind...
- Intervention model: Single-Arm, parallel, crossover,...
- Safety result
- Trial size
- Endpoints



Moneyball Questions

- 1. Impact of Trial Design on Phase PoS
 - 1. Patient selection biomarkers
 - 2. Combo v. monotherapy
 - 3. Trial size
 - 4. Single arm vs. comparator
 - 5. Masked v. open label
 - 6. Randomize v. not randomized
 - 7. Dose & exploration
 - 8. Endpoints
- 2. Impact of prior clinical trials on Phase PoS
 - 7. Same drug in different indication.

Intelligencia AI

- 8. Same target (different drug) in same indication.
- 9. Same target, different drug, different indication

- 3. Basket & platform trial (multiple shots on goal) 10.Impact on PoS
 - 11.Impact on false-positive rate
 - 12. Probability of false-negatives
 - 13.Optimal size of basket trial
 - 14.Correlation of the arms (common tumors, combos with common components)
 - 15.Any MAB boost?
 - 16.Exploration v. Exploitation
 - 17.Rate of accumulating information over time

Introducing Causal Inference: The ladder of causation

Ru	ng	Action	Question
3	Counterfactual	Imagining	If I had done X, what would Y be?
2	Intervention	Doing	What will happen to Y if I do X?
1	Association	Observing	How does observing X change my belief in Y?

Influence Diagram



Causal Diagram





Causation v. Correlation: Observational vs. Experimental Data

- 1. Question: Does birthweight impact infant mortality?
- 2. Conditional probabilities are *different situations*:
 - a) p(Motality|Birth Weight = High)
 - b) p(Motality|Birth Weight = Low)

- 3. We want the results of this *experiment*:
 - c) Randomize birth weight
 - d) p(Motality|Birth Weight = High)
 - e) *p*(Motality|Birth Weight = Low)

5K 🔞 Intelligencia AI





6 24 April 2024

- Getting experiments from observational data



Previous Statistic's Advice on Selecting Controls

(Bad) Advice on controls

- 1. Include relevant variables.
- 2. Include independent variables unaffected by treatment.
- 3. If unsure whether to include a variable, omit it.
- 4. Do not include outcome variables.
- 5. Build models with and without the control variables and contrast the findings.

Causal salad: Tossing various "control" variables into analysis (ex: regression), observing changes in estimates, and telling a story about causation.



Do the situations differ? Yes, a lot.

GS

🝥 Intelligencia AI



• The Building Blocks for Causal Models





Close All the Back-Door Paths







Data Can Help Create the Model – But it can't do everything



If 10 variables

- 45 pairs
- Max # of graphs: $3^{45} \approx 3x10^{21}$
- Min # of graphs: $2^{45} \approx 3.5 \times 10^{13}$
- At 1,000,000/second: 1-93 million years



DAG created via a workshop with a few of GSK's oncology & Biostats experts





1. Focused on what is predictive, not causal.

2. Comments during workshop:

- Set arrows with correlation matrix
- Exclude variables for which we have no data
- Make a small model and expand it later if needed



Finally! A good start



To focus on causality, ask these questions

- A. Does X determine or affect the realization of Y?
- B. Is it necessary to know X before determining Y?
- C. Does knowing/determining X limit the available options for Y?

Apply Domain Knowledge

• 4 Pillars Framework: Compound->Target->MOA->Efficacy

- How decisions affect correlations in data



Type of variables

- **A.** Goal variables
- **B.** Environmental variables
- **C.** Decision variables

Must consider all types of variables: goal, environmental, and decision.



Hypotheses About the Use of Patient Selection Biomarkers

- 1. Biomarker use will vary by indication because of pharmacology & economics
- 2. Hard-to-treat cancers will use biomarkers to boost PoS
- 3. Biomarker use will vary by modality, but maybe conditionally independent of success.
- 4. Targeted therapies will use biomarkers more than non-targeted therapies
- 5. Biomarker use will vary by target class
- 6. Frequency of biomarker use will increase from 3L to 2L to 1L.
- 7. Targets with validated targets will use biomarkers more frequently than others.
- 8. Challenging cancers will use more combos
- 9. Use of combos will increase from 3L to 2L to 1L
- 10. Biomarker use will vary inversely to combo therapy
- 11. Combo therapies will use biomarkers more than mono therapies
- 12. As # of indication success increases, so will the percentage of trials with biomarkers.

13. Use of biomarkers will vary inversely with the support of PCE.





Hypothesized Confounders for Biomarker → Trial Result

Pharmacology & Pharmacology Info

1. Target Class

2. Targeted

- 3. Target Validated
- 4. Mono v. Combo

5. PCE

Trial Design Decisions

- 1. Indication
- 2. Line
- 3. Previous phase's design

Other

- 1. Big pharma v. not big pharma
- 2. Does biomarker assay exist

Population Decisions 1. Indication

2. Line

- Causal Discovery: Causal Model Built from Data



Tier	Variables
1	Sponsor, Modality, Target Class, Targeted, Target Validated
2	Indication, Line
3	Other Design Variables
4	Results

Ph2 Causal Models: PC (left) v. FGES (right)







GSK 💿 Intelligencia AI



Variable	Variable	Chi-Square	p-value
Target Validated	Indication	33	6.6E-5
Modality	Indication	34	4.5xE-9
Target Class	Indication	159	9.1E-19
Target Class	Biomarker	19	8.0E-4



Phase 2: Variables for closing back door paths and studying front door components

Variable	Close Back Door Path	Front Door Components
Biomarker	Indication, Modality	None
Modality	Possibly target class	Indication
Indication	Validated Target, Modality	Biomarker, Line
Target Class	None	Modality, Targeted, Indication, Biomarker
Target Validated	None	Targeted, Indication, Open Label
Line	Indication	Mono v. Combo, Open Label
Open Label v. Blinded	Validated Target, Line/MonoCombo	None
Single Arm v. Multi	MonoCombo, Open Label/Validated Target	None
Mono v. Combo	Line	Single Arm
Randomized v. Not random		
Size		
PCE		

GSK 💿 Intelligencia AI











	Trial Result				
	Success Failure				
True	7	2			
False	8	6			

Odds Ratio: 2.6 *p*-value: 40%

Odds Ratio: #DIV/0 *p*-value: 30%

- Hierarchical Bayesian Inference



CONTRACTOR CONTRACTOR CONTRACT

