Assessing and Communicating Heterogeneity of Treatment Effects (HTE) for Patient Subpopulations: Challenges and Opportunities

Subgroup Identification
The Hardest Problem There Is

Stephen J. Ruberg, PhD
President
Analytix Thinking, LLC
AnalytixThinking@gmail.com

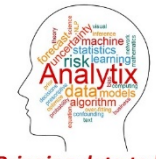
Bringing data to life.

Outline

Clinical Trials

Observational Data

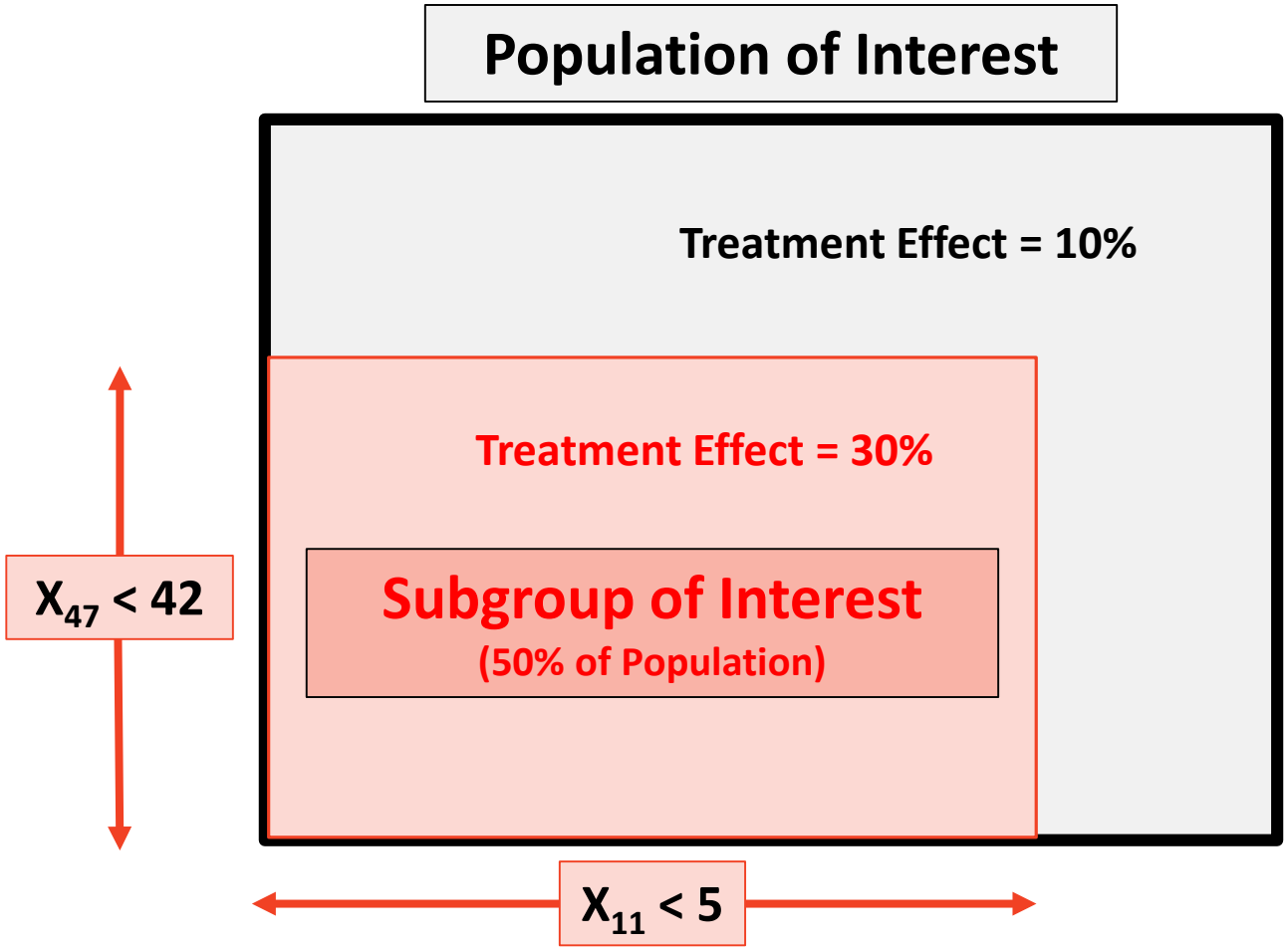
What We Can Do Now



Bringing data to life.

Clinical Trials

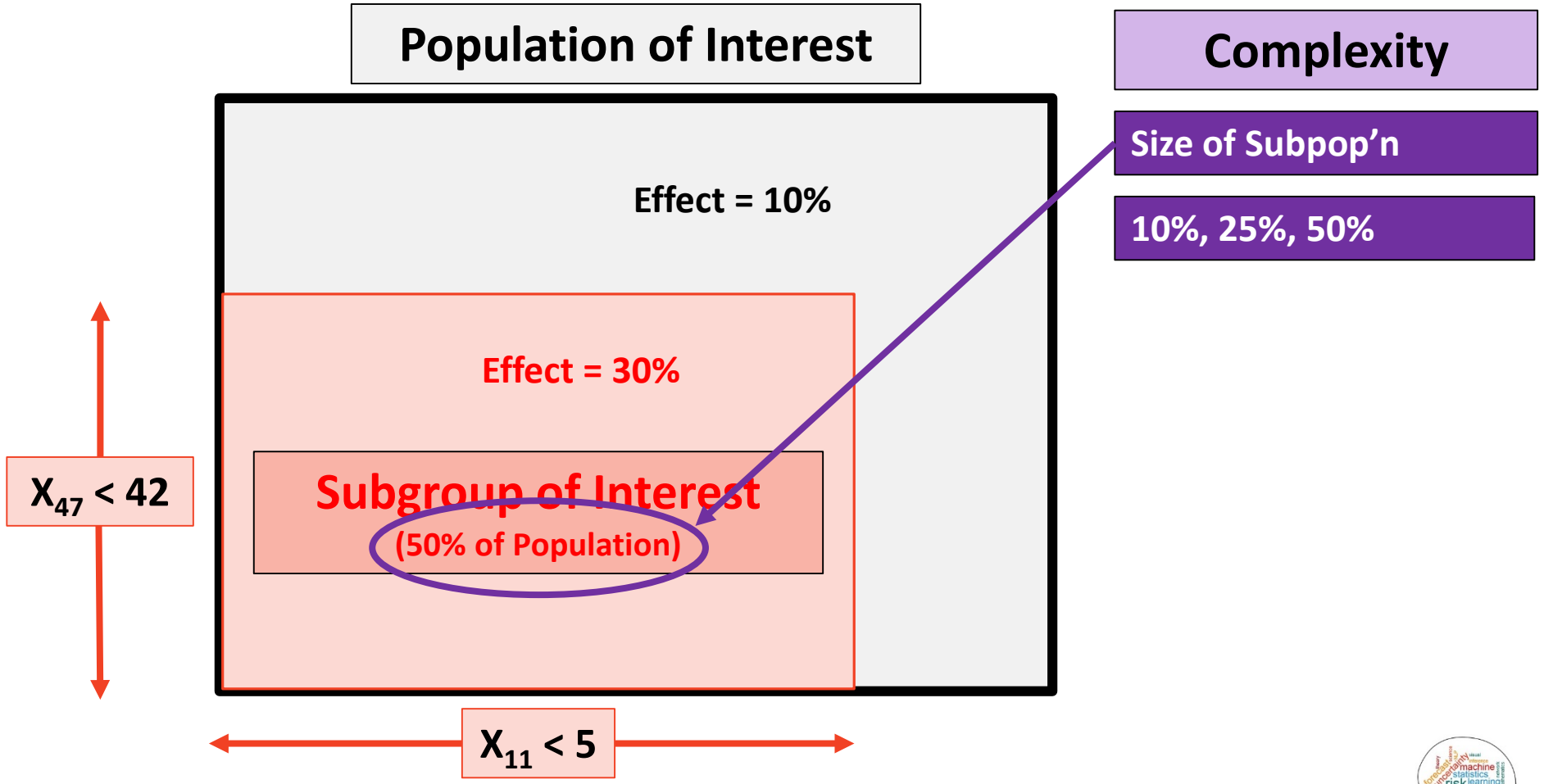
Treatment versus Control



Bringing data to life.

Clinical Trials

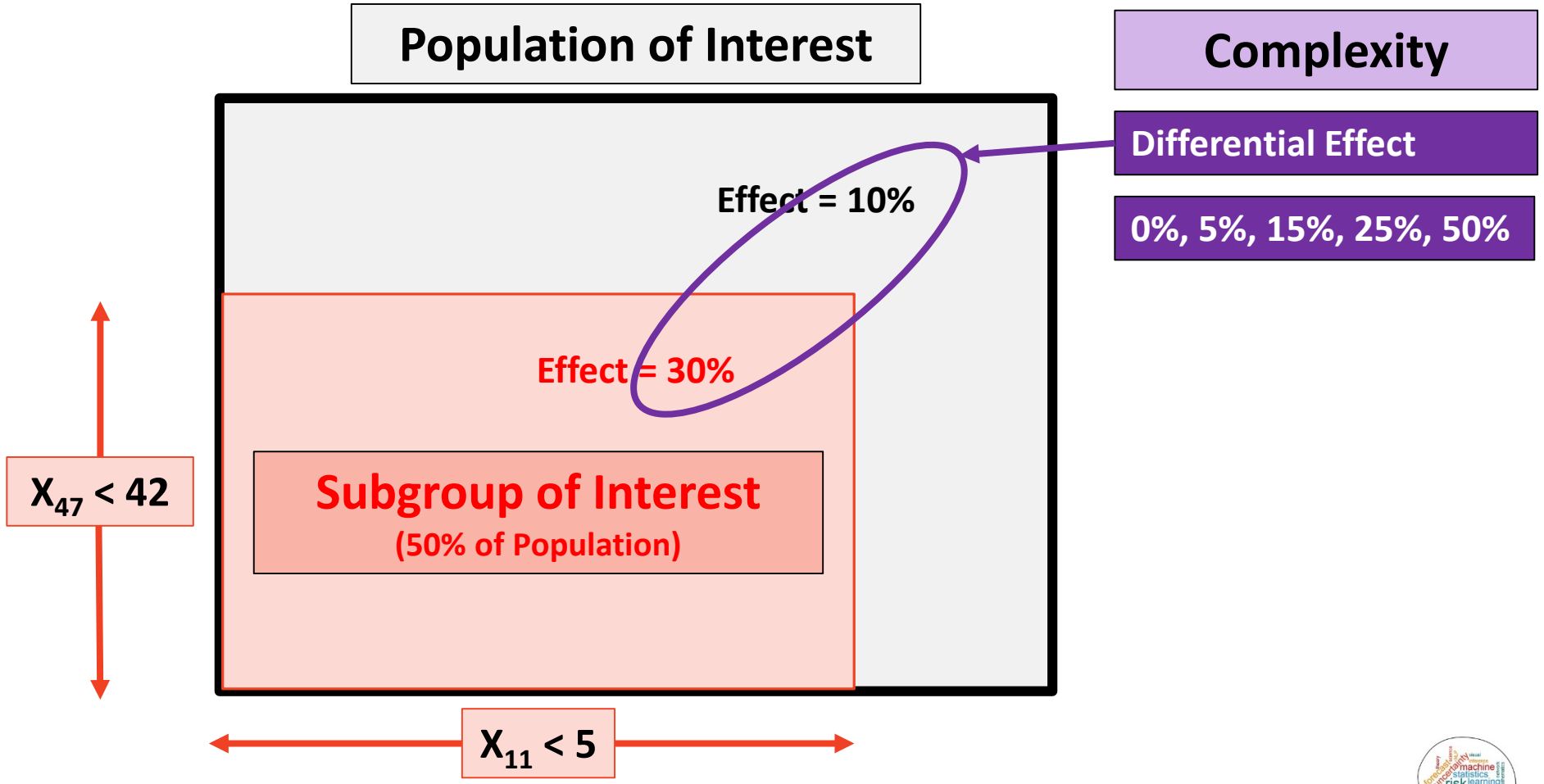
Treatment versus Control



Bringing data to life.

Clinical Trials

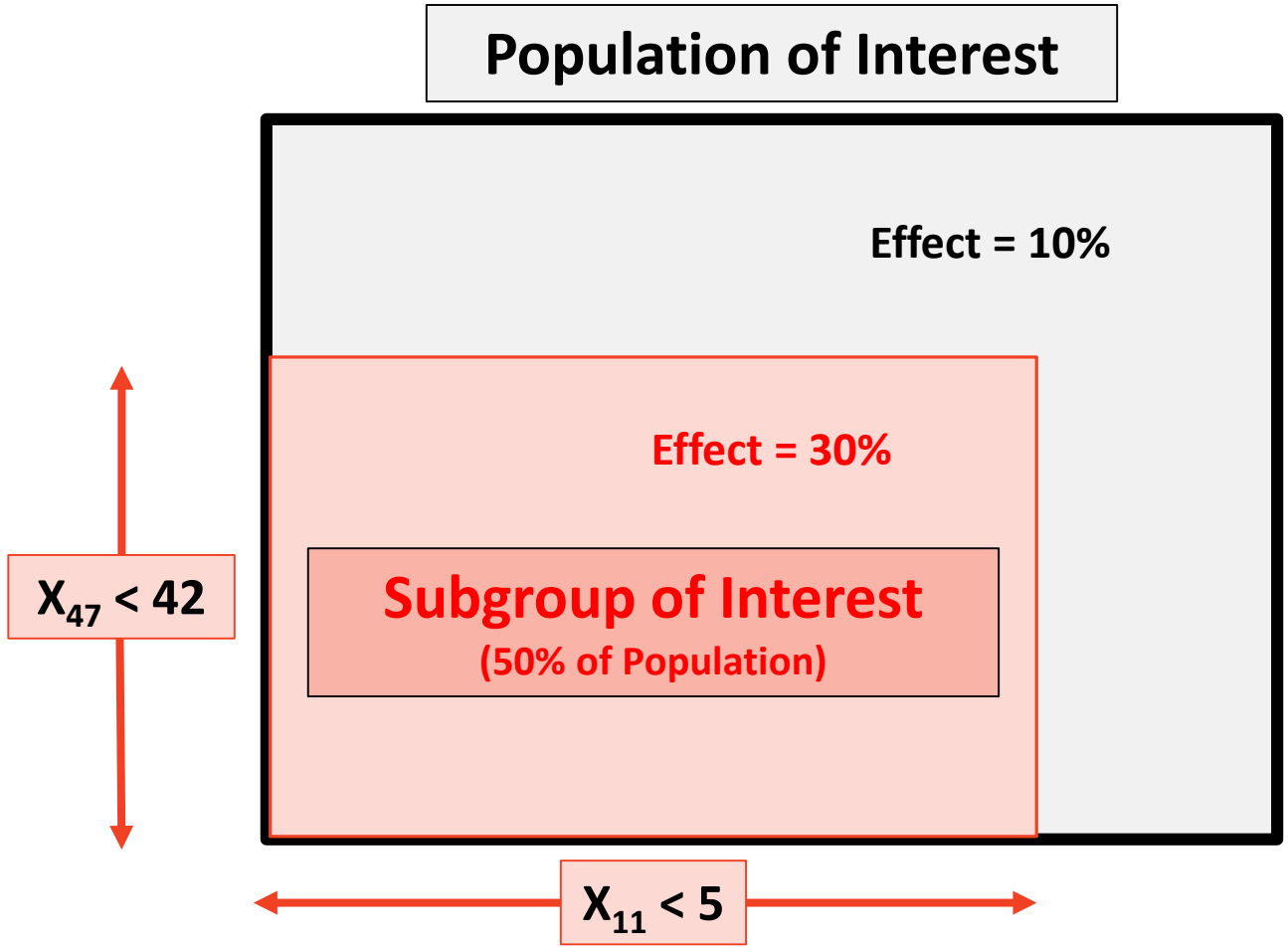
Treatment versus Control



Bringing data to life.

Clinical Trials

Treatment versus Control



Complexity

Number of biomarkers in the dataset

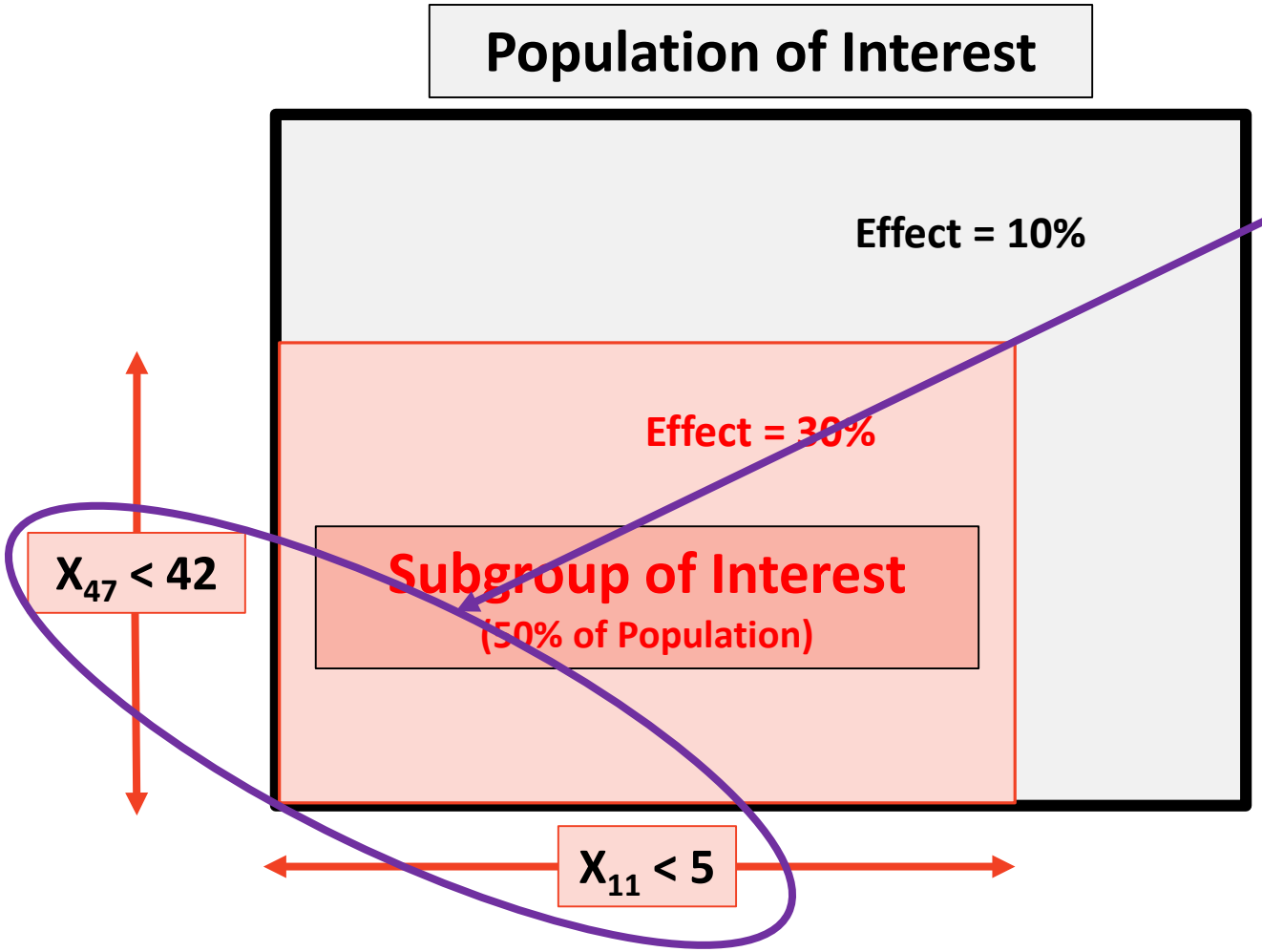
$X_1, X_2, X_3 \ldots X_{110}$

10, 25, 75, 250, 500

Bringing data to life.

Clinical Trials

Treatment versus Control



Complexity

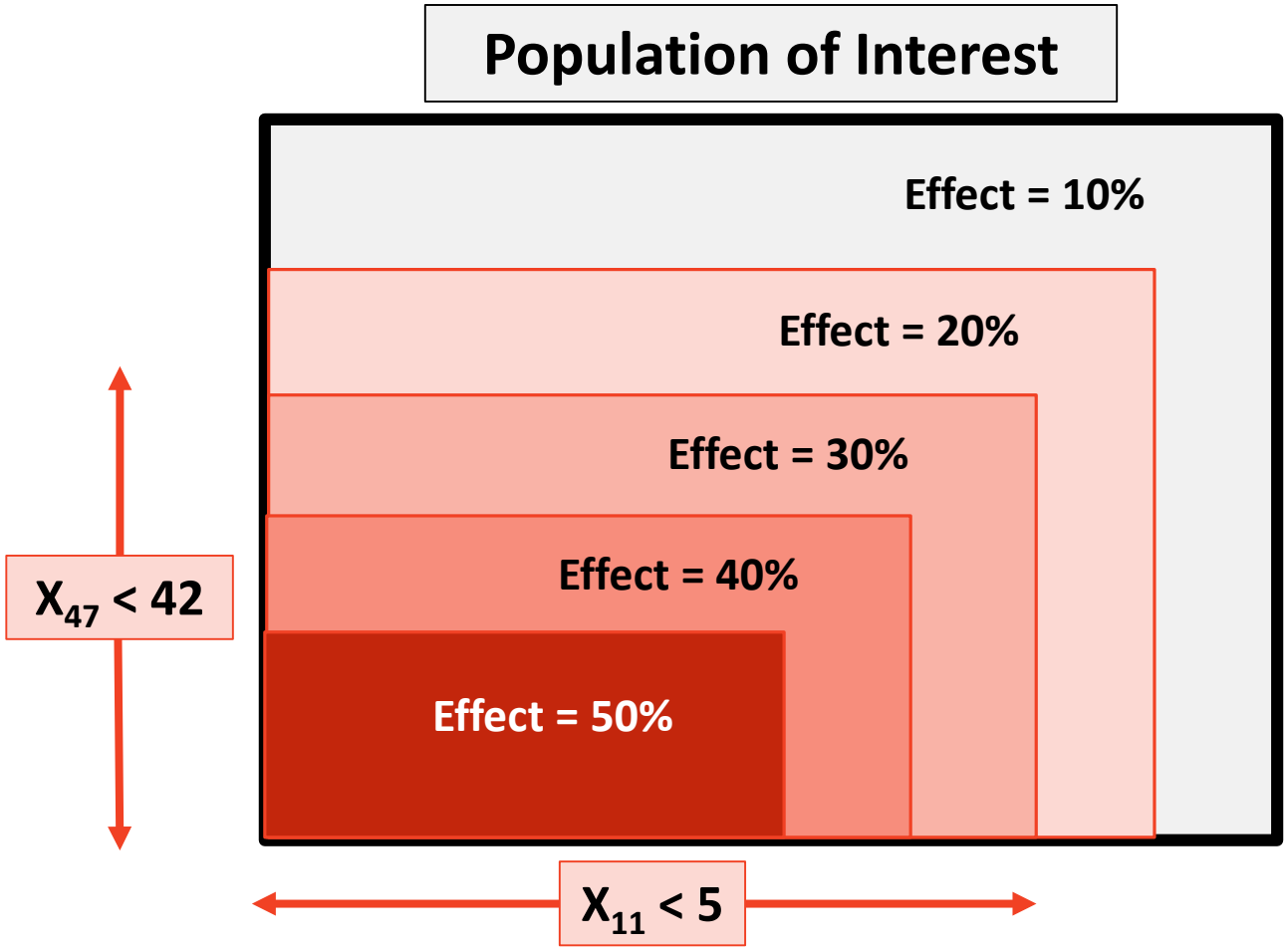
Number of biomarkers defining subpop'n

1, 2

Bringing data to life.

Clinical Trials

Treatment versus Control



Complexity

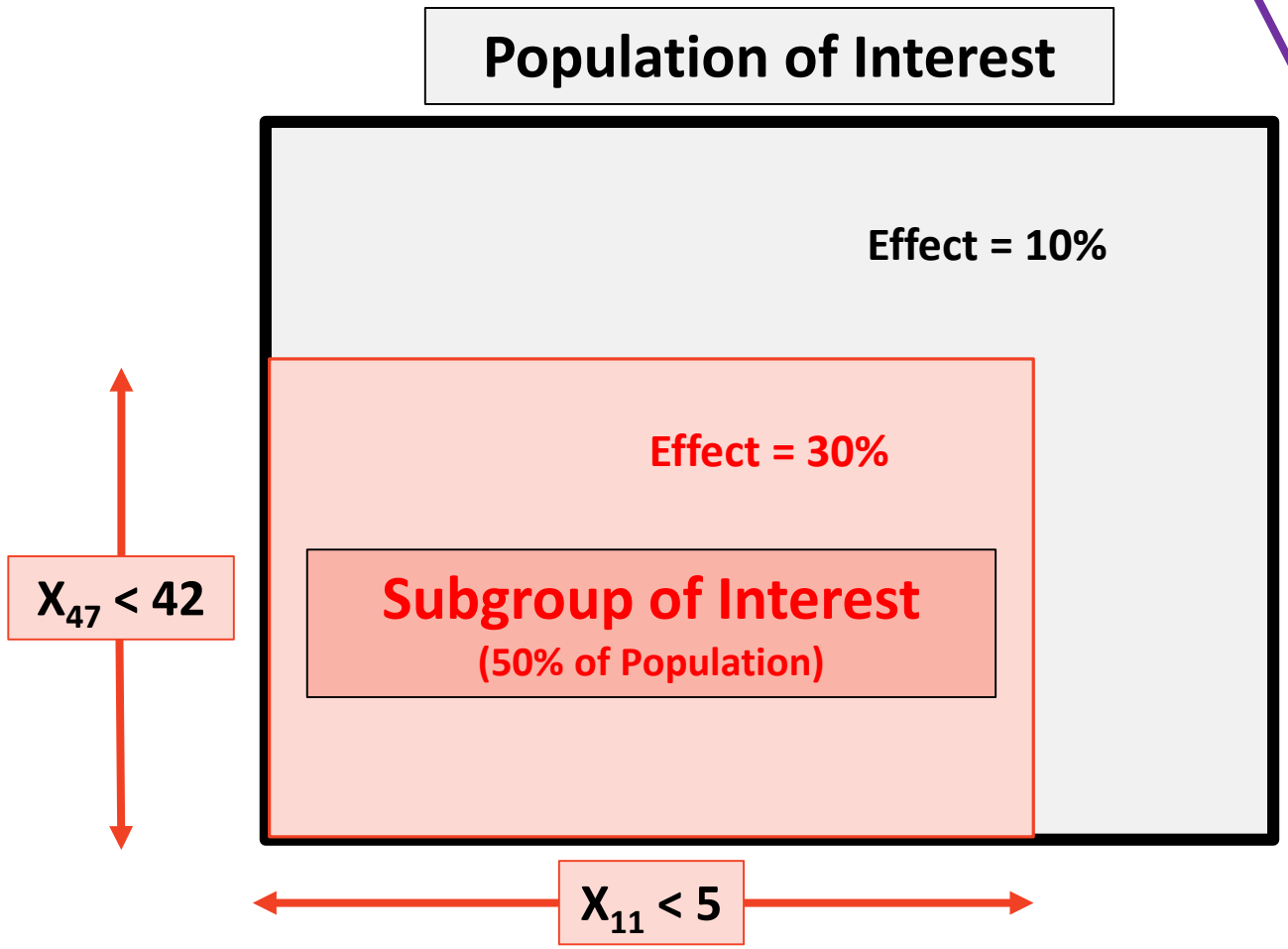
Nature of biomarkers effect

1, 2

Bringing data to life.

Clinical Trials

Treatment versus Control



Complexity

CT Sample Size

100, 250, 500, 1000

Bringing data to life.

Subgroup Identification Challenge

Scenarios (i.e. combinations of possibilities)

3 x 5 x 5 x 2 x 2 x 4 = 1200 !!!

Simulated 1200 datasets with these known parameters.

Posted on ***Innocentive*** and challenged the world …

FIND THE SUBGROUP (i.e. the X's and the cut-offs)

Created a scoring system to rank solutions (0, 100).

Participants could make 1 attempt per day over 3 months.

Subgroup Identification Challenge

Total of 748 entered the competition

- USA 279, India 69, UK 49, Canada 43, Germany 24, Australia 20, Russia 20, Italy 19, Spain 16,
- + 39 other countries (including Seychelles!)

62/120 (52%) did no better than flipping a coin !!!

Only 120 submitted a valid solution (that could be scored)

Bringing data to life.

Subgroup Identification Challenge

Internal benchmark score = 62

- This problem is very hard !!

Only two submissions did marginally better with scores of 64 and 65.

- 118/120 (98%) did worse than the internal benchmark

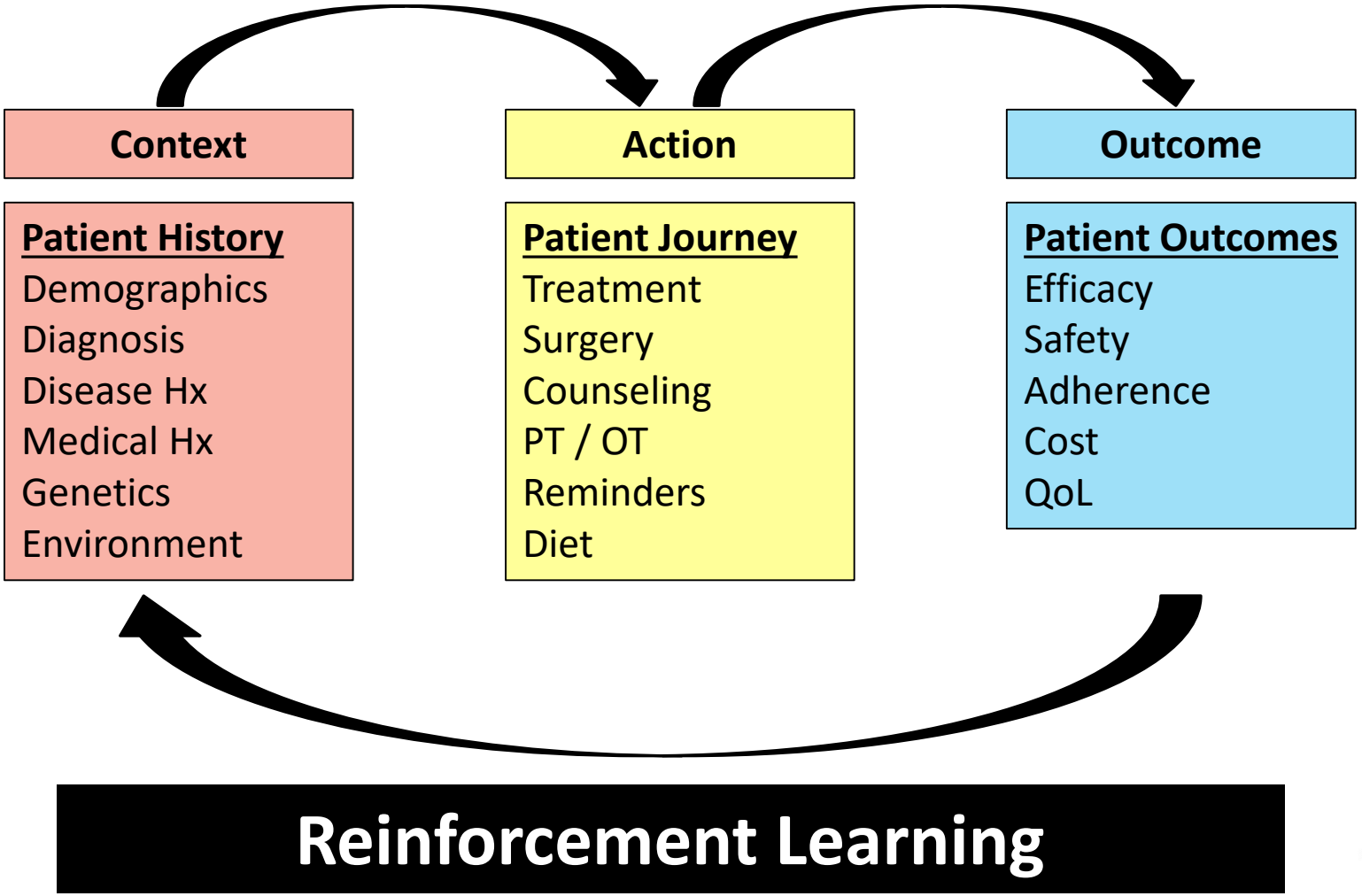
Bringing data to life.

Conclusion

Even with randomized, controlled trials/data, under normal circumstances (i.e. reasonable parameter values) and simple biomarkers relationships to response, **the subgroup is not identified or mis-identified a high percentage of the time.**

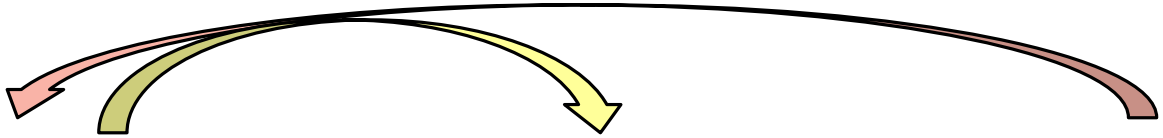
Bringing data to life.

Observational Data



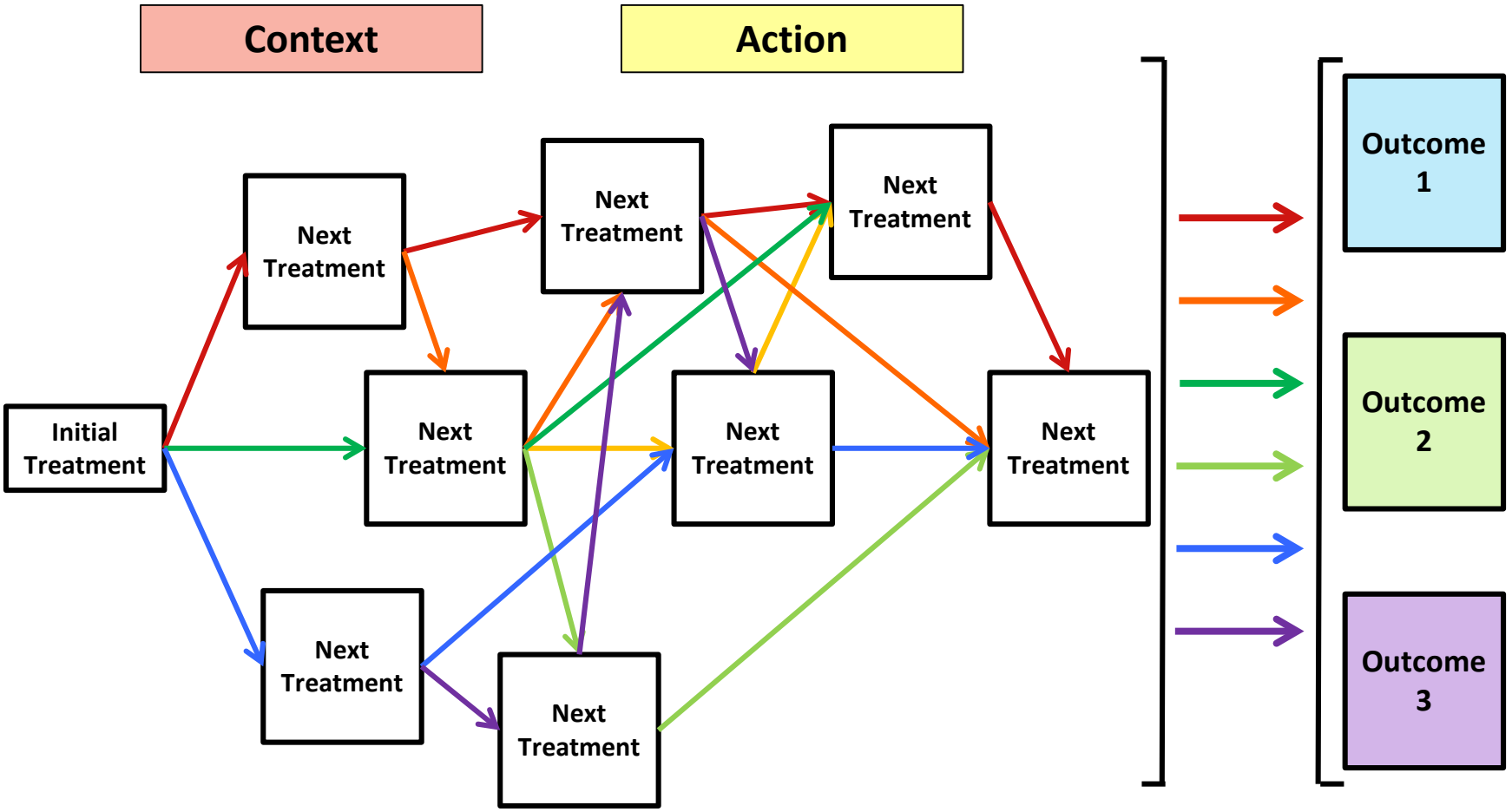
Bringing data to life.

Observational Data



Context	Action	Outcome
Patient History Demographics Diagnosis Disease Hx Medical Hx Genetics Environment	**Patient Journey** Treatment Surgery Counseling PT / OT Reminders Diet	**Patient Outcomes** Good Bad
Dozens to Millions	**Dozens to Thousands**	**Dozens**

Observational Data



Observational Data

Conclusion

- Observational data is messier than controlled clinical trials, although potentially more plentiful.
- Finding the right combination of Context and Action variables (likely to be some complex combination) that predict favorable outcomes is subject to high false positive rates.

Bringing data to life.

Individualized Treatment Regimes

X	$\mathcal{D}(X) \rightarrow A$	A	$\max E^{\mathcal{D}}(Y)$
Context		**Action**	**Outcome**
Patient History Demographics Diagnosis Disease Hx Medical Hx Genetics Environment		**Patient Journey** Treatment Hx Adherence Surgery Counseling PT / OT Reminders Diet	**Patient Outcomes** Efficacy Safety Cost QoL
Dozens to Millions		**Dozens to Thousands**	**Dozens**

Individualized Treatment Regimes

Define,

$$E^{\mathcal{D}}(Y) = \int Y d\mathcal{P}^{\mathcal{D}} = \int Y \frac{d\mathcal{P}^{\mathcal{D}}}{d\mathcal{P}} d\mathcal{P} = E\left[\frac{I\{A = \mathcal{D}(X)\}}{p(A|X)} Y\right],$$

where we use the fact that,

$$\frac{d\mathcal{P}^{\mathcal{D}}}{d\mathcal{P}} = \frac{p(y|x,a) I\{a = \mathcal{D}(x)\} p(x)}{p(y|x,a) p(a|x) p(x)} = \frac{I\{a = \mathcal{D}(x)\}}{p(a|x)}.$$

Our objective is to find $\mathcal{D}(\cdot)$ to maximize the following value function:

Value function

$$\mathcal{D}_o \in \underset{\mathcal{D} \in R}{\operatorname{argmax}} E^{\mathcal{D}}(Y) = E\left[\frac{I\{A = \mathcal{D}(X)\}}{p(A|X)} Y\right], \quad (1)$$

where R is a space of possible treatment recommendations.

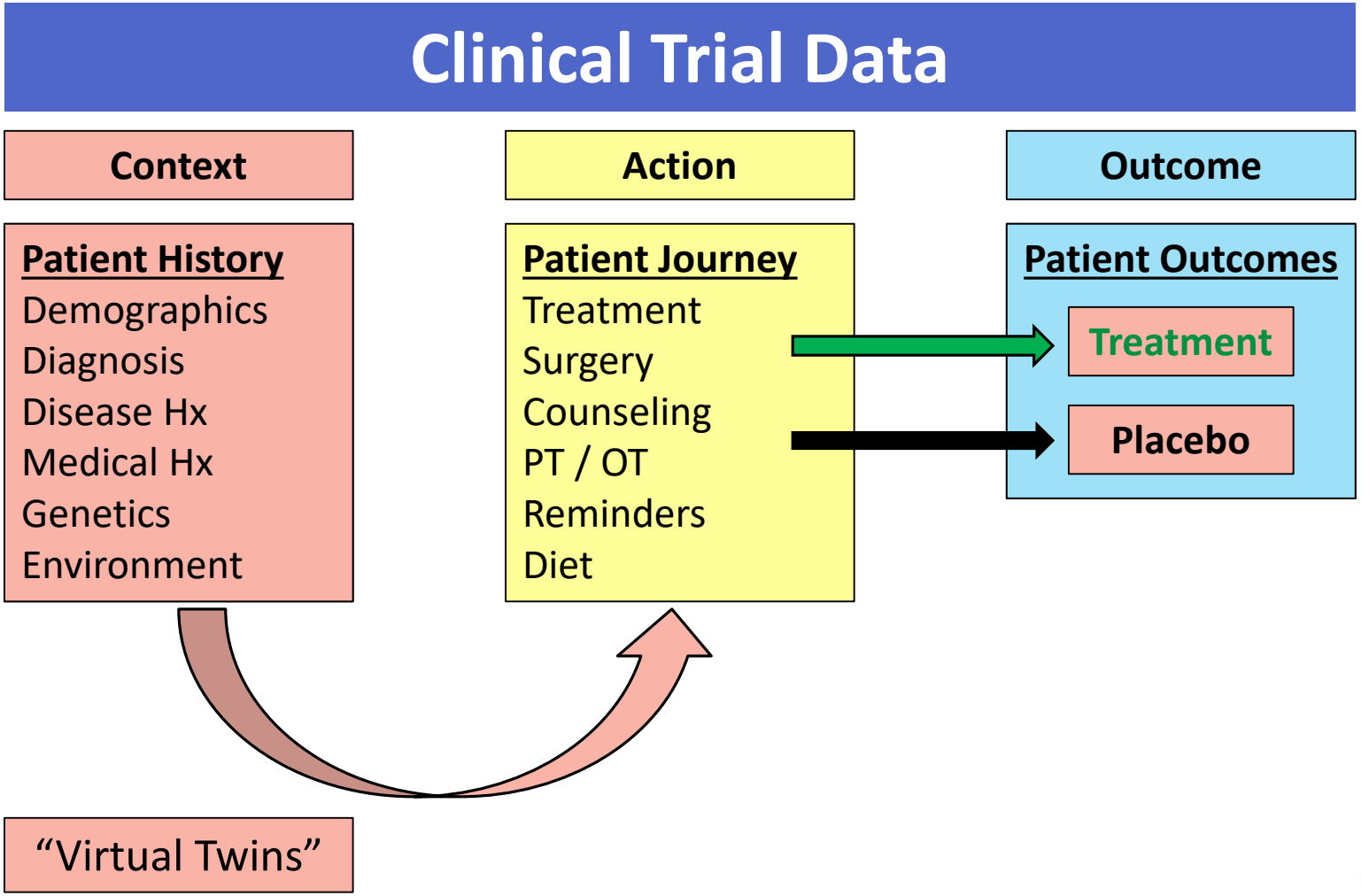
Observational Data

Conclusion

- ITR is an emerging (promising?) methodology.
- It can help find the right $\mathcal{D}(X)$ within the context of available decisions and covariates.
- Use artificial intelligence (i.e. reinforcement learning algorithms) to help solve this problem.

Bringing data to life.

What Can We Do Now?



Bringing data to life.

What Can We Do Now?

Imagine an App

"Here are the 50 patients in the clinical trials database most like you."

Review of summary Context data.

"Here are the treatments and doses that they took in the clinical trials."

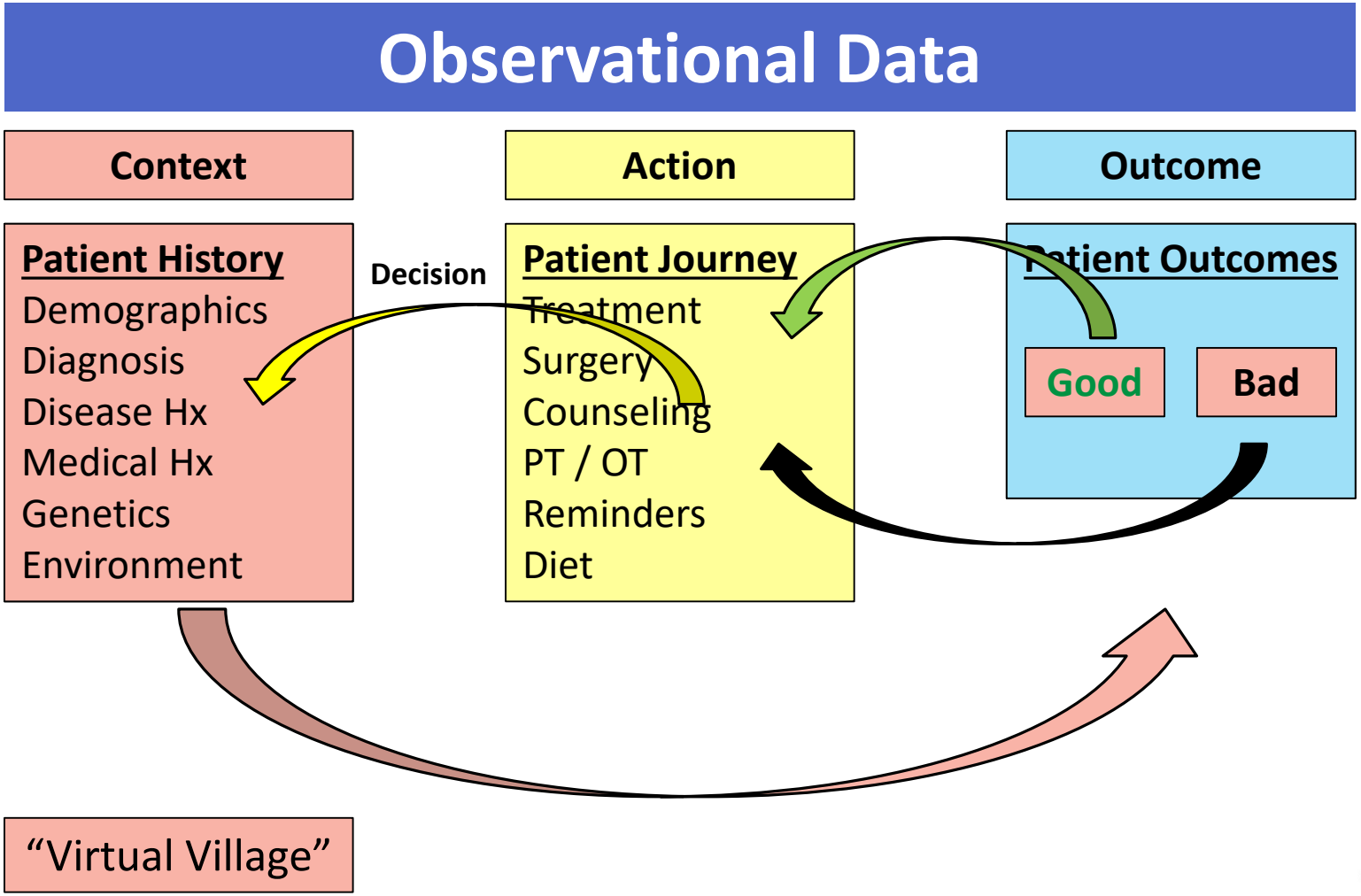
Review of summary Action data.

"Here are the primary efficacy and safety summaries."

Review of summary Outcome data.

Bringing data to life.

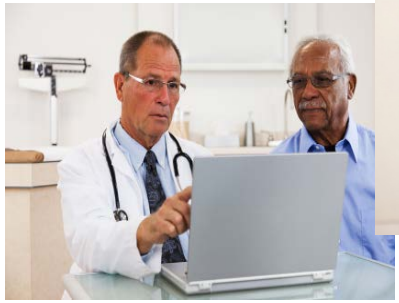
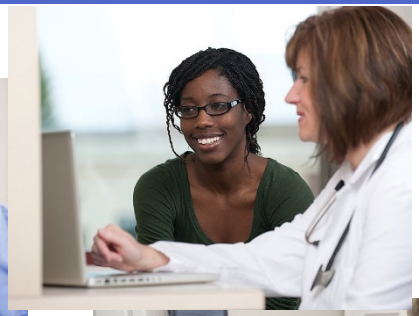
What Can We Do Now?

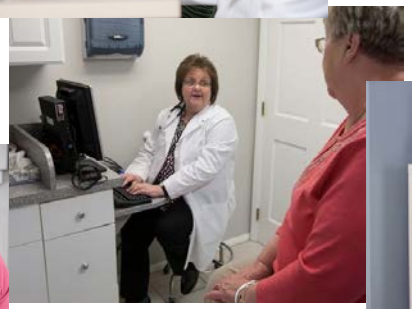


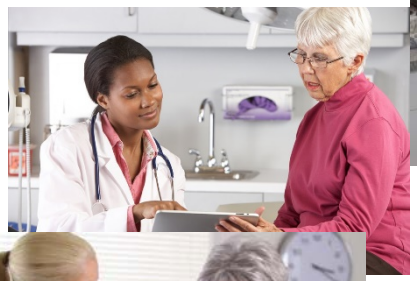
Bringing data to life.

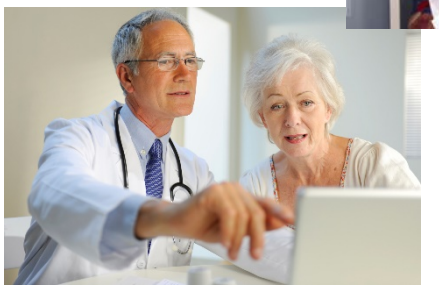
What Can We Do Now?

Imagine an App



Bringing data to life.

What Can We Do Now?

Initially done with CT data from approval.

- Could FDA do this?

Extend to observational data?

- Embed in Electronic Medical Records?
- Create recommender system (e.g. Netflix …)

CREATE PATIENT-CENTERED SUBGROUPS! (?)

Probabilize efficacy and safety outcomes

- Patient use their own utility function

Bringing data to life.

Conclusion

THE problem of this era is …

Finding the few actionable X's (out of many, many X's) that predict favorable outcomes for INDIVIDUAL PATIENTS.

- It is extremely difficult.
- Having the ***right data*** with ***very smart analytics*** is the key.

Bringing data to life.

Conclusion

"My ultimate dream would be to find signatures of high-responding patients through this kind of artificial intelligence, and prospectively identify: Who will be the super-responders? … If we can master that — that would be a great longer-term win."

Vas Narasimhan, MD
CEO, Novartis
Oct 29, 2018

PANEL DISCUSSION SLIDES

Bringing data to life.

Multiplicity and Bayes

Multiplicity

What is the probability that my finding in real or chance (i.e. spurious)?

Bayes

Quantifying the probability of a hypothesis

Bringing data to life.

Multiplicity and Bayes

Solanezumab – EXPEDITION STUDIES

The following is a Steve Ruberg analysis (post hoc).

It explicitly does not use information from Lilly.

It is meant solely to be illustrative and does not represent what analyses were done by Lilly, which were much more sophisticated than presented here, or how decisions were made by Lilly.

Bringing data to life.

Multiplicity and Bayes

Suppose our prior probability that Solanezumab works is 0.30 (i.e. the null hypothesis is false).

There are 5 hypotheses of interest:

- The overall population (primary)
- Mild population
- Moderate population
- *APOE* ε4 carriers
- *APOE* ε4 non-carriers

Multiplicity and Bayes

Split the prior (0.30) across the five hypotheses of interest.

1. Give most of the probability (0.14) to the Overall population.
2. Divvy the remaining probability to the other subpopulations
   a) Equally – i.e. 0.16/4 = 0.04
   b) Weighted – i.e. give larger prior (0.06) to Mild and *APOE* ε4 carriers

Multiplicity and Bayes

Hypothesis	Prior (version A)	Prior (version B)
Overall	0.14	0.14
Mild	0.04	0.06
Moderate	0.04	0.02
APOE ε4 (+)	0.04	0.06
APOE ε4 (-)	0.04	0.02
Total	**0.30**	**0.30**

Bringing data to life.

Multiplicity and Bayes

Hypothesis	Prior (version A)	Prior (version B)
Overall	0.14	0.14
Mild	0.04	0.06
Moderate	0.04	0.02
APOE ε4 (+)	0.04	0.06
APOE ε4 (-)	0.04	0.02
Total	**0.30**	**0.30**

Bringing data to life.

Multiplicity and Bayes

EXPEDITION 1
(Mild and Moderate AD Patients)

ADAS-Cog11
- Overall: p=0.24
- Mild: p=0.008

ADAS-Cog14
- Mild: p=0.003

Is the finding in Mild patients real or spurious?

Doody, Rachelle S. et al. “Phase 3 Trials of Solanezumab for Mild-to-Moderate Alzheimer's Disease.” N Engl J Med 2014; 370:311-321. (with Supplementary Appendix)

Multiplicity and Bayes

EXPEDITION 1
(Mild and Moderate AD Patients)

ADAS-Cog11
- Overall: p=0.24
- Mild: p=0.008

ADAS-Cog14
- Mild: p=0.003

MILD PATIENTS

ADAS-Cog11		
Initial Prior	P-value	Posterior *
0.04	0.008	0.28
0.06	0.008	0.38

*** Using Bayes Factor:** $O_0 \times [-e \times p \times \ln(p)] \leq O_1$

Multiplicity and Bayes

EXPEDITION 1
(Mild and Moderate AD Patients)

ADAS-Cog11
- Overall: p=0.24
- Mild: p=0.008

ADAS-Cog14
- Mild: p=0.003

MILD PATIENTS

ADAS-Cog11					
Initial Prior	P-value	Posterior *	Revised Prior	P-value	Posterior *
0.04	0.008	0.28	0.02	0.008	0.16
0.06	0.008	0.38	0.03	0.008	0.23
ADAS-Cog14					
			0.02	0.003	0.30
			0.03	0.003	0.39

Bringing data to life.