Paradigm for Modern Drug Development

Girish Aras
Statistical Consultant

The Scope

- The paradigm is about getting an approval for a specific indication for a drug (in the USA)
- Focus is only the late stage of development, namely Phase 3 Confirmatory trials and beyond
- The history of how the current paradigm evolved via Food Drug and Cosmetic Act is a topic for another discussion

The Paradigm

- Randomized Control Trials (RCTs) to demonstrate drug efficacy (and ‘safety’)
- Followed by several observational studies, non-randomized prospective studies to generate hypotheses for other indications, and, for important subgroups (say elderly, children, severe patients etc.)
- New hypotheses are evaluated via new RCTs, and the cycle continues

Modern Drug Development Key concepts

- Causal Inference
  - Counterfactual
- Randomized Clinical Trial (RCT)
  - Randomized double-blind trial as a gold standard to establish drug efficacy and to a lesser extend, safety
- Observational (Real World Evidence) Studies
  - Causation vs. Correlation
  - Confounding
- Role of regulation & drug approval policy

Causal Inference

- Counterfactuals
  - Ideally, one would like to observe both the real and the counterfactual, i.e., what could have happened, but didn't
    - Ideally, I being on the drug for 6 month and a clone of me not on the drug for the same 6 months!
- At the individual level counterfactual is never available, by definition

Causal Inference (2)

- RCT is the next best thing
  - Hopefully, subjects **like** me are on the drug for 6 months, and many other subjects **like** me are **not** on the drug for the same 6 months
  - Then I hope that my effect size is close to the estimated effect size of those subjects who are like me

Estimated effect size is the difference between the two mean or the two proportions.

Gold Standard for Causal Inference

- Randomized Clinical Trials
  - Going forward in time
  - Randomized
  - Confirmatory (A priori Hypothesis is tested)
  - Placebo (or standard of care) Controlled
  - Double-blind
  - Adequate (statistically powered)
  - Pre-specified Protocol, objective, primary and other endpoints, statistical methods

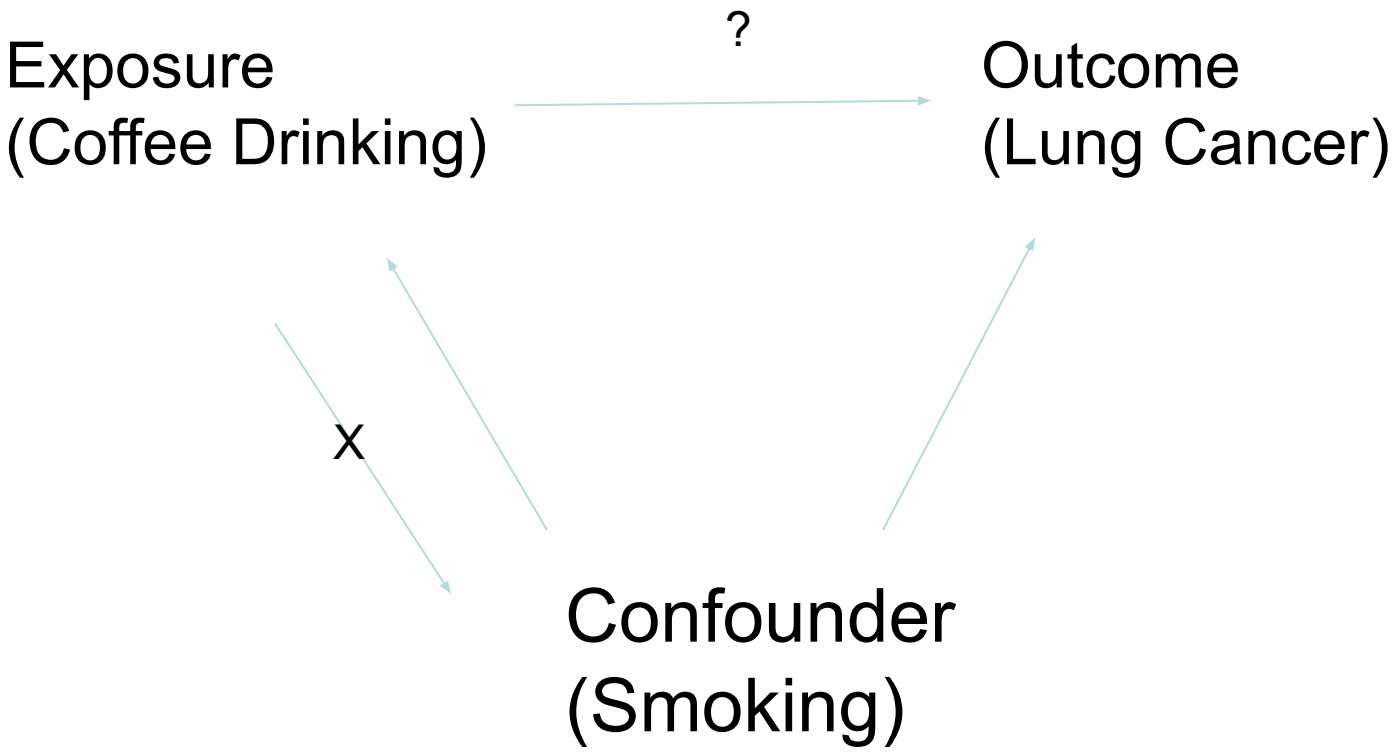
Observational Studies

- Key words associated with Observational Studies
  - Retrospective (Case-Control Studies)
  - Cross sectional (a snapshot in time)
  - Correlation (as opposed to causation)
  - Case Studies (Formal or Anecdotal)
  - **Confounders** (selection bias)
  - Confounding by Indication

Confounding

- Say one is investigating a causal linkage between an exposure (an independent variable, potentially a cause such as treatment) and an outcome (a dependent variable, potentially an effect)
- There could be a third variable, a confounder, lurking in the background which causes both
- In a study that is investigating the causation between coffee drinking and lung cancer, smoking is a confounding variable

Example


?
Exposure
(Coffee Drinking)
Outcome
(Lung Cancer)
X
Confounder
(Smoking)

Confounding Variable (Definition)

- Confounding variable is a
  - causal factor for the outcome as well as the exposure (treatment)
  - And not an effect of the exposure (treatment)
- Confounding variable creates Selection Bias
  - It tends to inflate (or deflate) proportion of its kind in exposure groups, hence also affects the outcome

Confounding (2)

- A confounding variables compete with the exposure in explaining the outcome of a study
- One needs to tease out effect of the exposure, if any, from that of confounder to perform the correct inference on exposure outcome relationship
- The effect of the confounder can go in either direction

A Case-Control Study On Potential Coffee Lung Cancer linkage

Group	Coffee Drinkers	Non-Coffee Drinkers	Total
Lung Cancer Cases	90	70	160
No Lung Cancer Cases	210	630	840
Total	300	700	1,000

- **Lung Cancer Rate (Unadjusted):**
  - Coffee Drinkers: 90/300=30%
  - Non-Coffee Drinkers: 70/700=10%

The risk difference (treatment effect) is 20%, a significantly higher lung cancer rate in coffee-drinkers than non-coffee drinkers.

Biases

- Selection (confounding) Bias
  - Survival Bias (may miss severe cases)
  - **Berkson's Bias**: Hospital-based studies may overrepresent certain groups.
- Recall Bias
  - smoking and lung cancer, cancer patients might overreport past smoking habits compared to healthy controls
- cc

Biases (2)

- **Observer (Interviewer) Bias**
  - If an interviewer expects a certain exposure to be linked to a disease, they may unconsciously ask more leading questions to cases
- **Misclassification Bias**
  - Errors in classifying individuals as exposed or unexposed (or as cases or controls).

Biases (3)

- **Surveillance (Detection) Bias**
  - Women undergoing frequent medical check-ups may be more likely to be diagnosed with breast cancer early, compared to those without regular visits

How to Reduce Bias in Case-Control Studies

- **Careful selection of controls** from the same population as cases
- **Matching** cases and controls on key variables to reduce confounding
- **Blinding interviewers** to case/control status to avoid observer bias

How to Reduce Bias in Case-Control Studies (2)

- **Standardized questionnaires** to minimize recall bias
- **Using multiple sources of exposure data** to validate self-reported information
- In short, Case-Control Studies are a mess!

Stratified data: Among smokers

Group	Coffee Drinkers	Non-Coffee Drinkers	Total
Lung Cancer Cases	80	40	120
No Lung Cancer Cases	120	60	180
Total	200	100	300

- Lung Cancer Rate for Coffee Drinkers = 80/200=40%
- Lung Cancer Rate for Non-Coffee Drinkers = 40/100=40%
- Risk difference = 40%−40%=0%
- No effect in smokers

Stratified data: Among non-smokers

Group	Coffee Drinkers	Non-Coffee Drinkers	Total
Lung Cancer Cases	10	30	40
No Lung Cancer Cases	90	570	660
Total	100	600	700

- Lung Cancer Rate for Coffee Drinkers = 10/100=10%
- Lung Cancer Rate for Non-Coffee Drinkers = 30/600=5%
- Risk Difference = 10%−5%=5%
- Small numerical effect in non-smokers

Adjusted risk difference (Weighted Average)

- Weights based on stratum size
  - For smokers: 300/1000 = 0.3
  - For non-smokers: 700/1000 = 0.7
- Adjusted Risk Difference

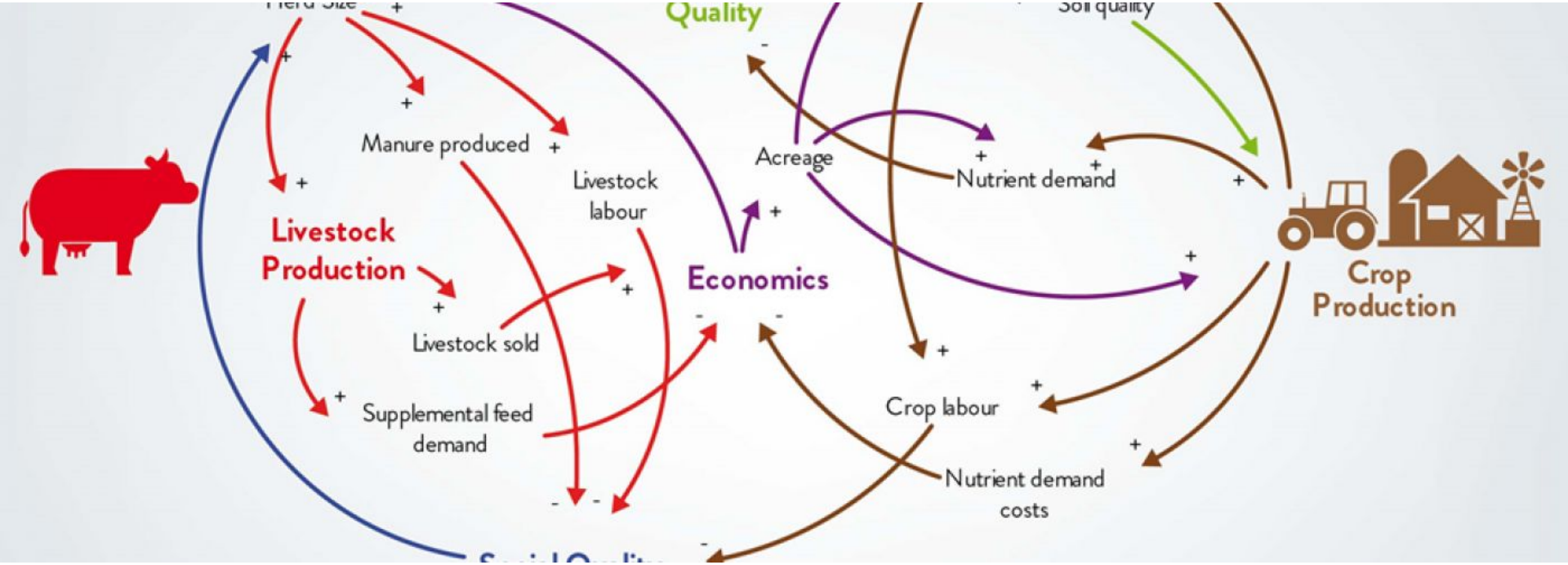
  =(0.3x0)+(0.7x5) = 3.5%
- 20% was the naive estimate whereas 3.5% is the '**causal'** estimate (after adjusting for the confounder)

Observational Studies (2) Covariate adjustment

- Adjust for the known confounders (artifacts of scientific field, and are determined by the experts) via stratification and/or as covariates
- Some known confounders may not be available in the dataset
- There could be **unknown** confounders

How serious is the confounding problem?

- In complex systems such as drug effect on human body, there are large number of variables (potential confounders) that may affect each other, plus the variables of interest, namely exposure and outcome
- It took a long time to establish that bypass surgery is indeed effective (it is not just a plumbing problem!)


Quality
Soil quality
Manure produced
Livestock labour
Acreage
Nutrient demand
Livestock Production
Economics
Crop Production
Livestock sold
Supplemental feed demand
Crop labour
Nutrient demand costs

How serious is the problem? (2)

- Observational data is abundant, cheap (relative to RCT) and available right away
- Observational studies are a valuable component of establishing causal linkage between exposure and outcome
- But relies on a crucial non-verifiable assumption that there are no unknown confounders (in the mathematical framework of Causal Inference it is called ignorability assumption)

Observational Studies (3) Propensity scores

- Stratification and covariate analysis
  - Number of strata increase exponentially
- Propensity Score Estimation:
  - A propensity score is the probability of an individual being assigned to the treatment group based on their observed covariates
  - Typically, logistic regression or machine learning models (e.g., random forests) are used to estimate these probabilities

Randomized Clinical Trials

- RCTs are a definitive, but a limited solution
- Long-term trials are especially expensive and sometimes not feasible in practice
- Long term safety data are mostly not available via RCTs
  - Hence concurrent observational studies are necessary

A note of caution

- Drug development is replete with numerous examples where drug efficacy is propagated based only on observational studies, followed by RCT proving otherwise
  - *Reflection on interdependence among academia, regulatory agencies and pharma industry*, Girish Aras. Journal of Indian Statistical Association, Vol 60 (2), 171-186, December 2022

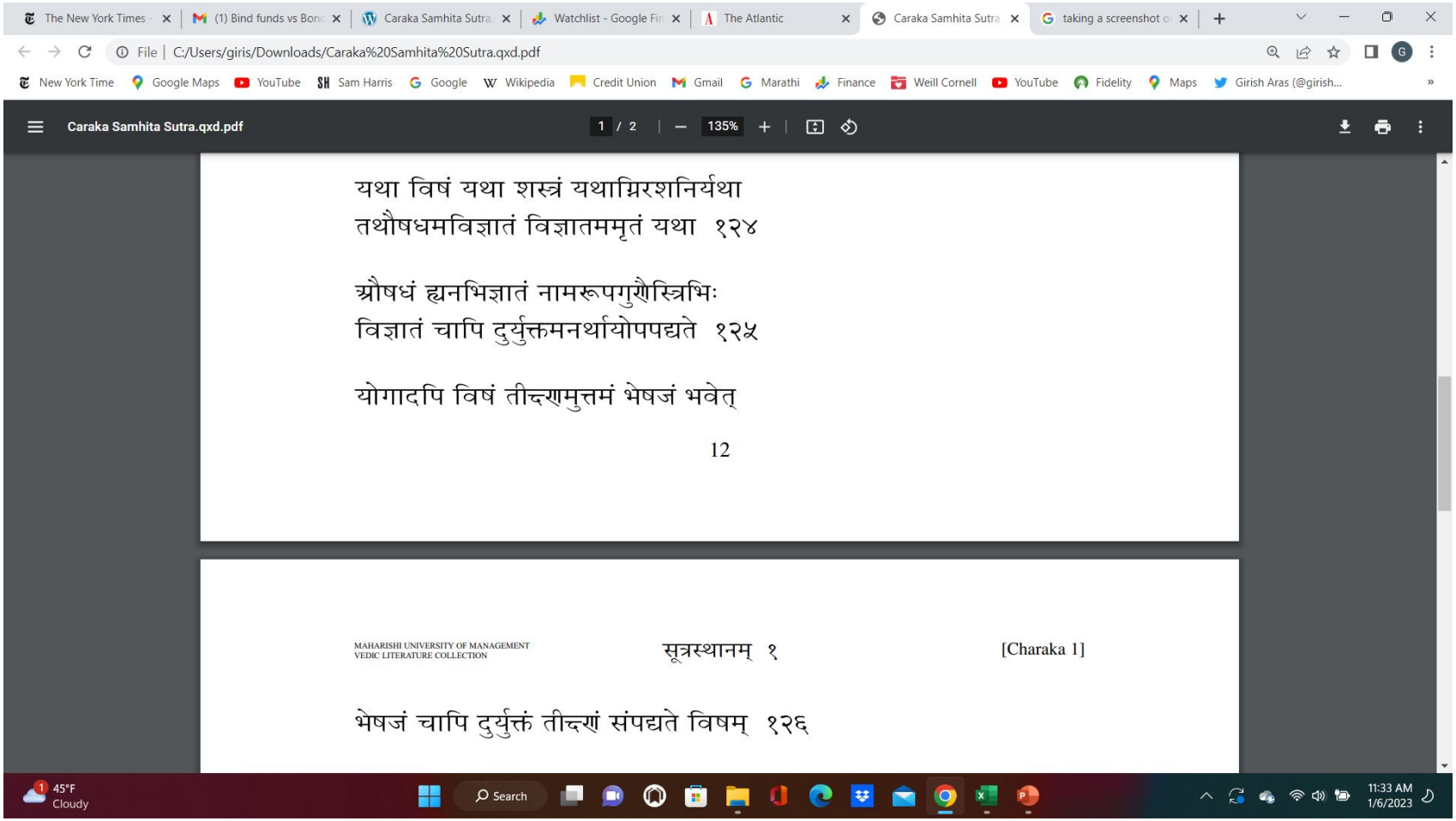
Some notes

- The first randomized clinical trial (RCT) is widely considered to be the study conducted by a statistician Austin Bradford Hill in 1948, investigating the effects of streptomycin on pulmonary tuberculosis
- Observational studies are of course practiced from the dawn of civilization
- Propensity score matching is widely used in fields such as medicine, economics, and social sciences to improve causal inference in non-experimental data
- Nobel price in Economics (2019), Abhijit Banerjee, Esther Duflo, and Michael Kremer for innovative use of RCT-like methods

Pharmacology and Toxicity Charaka Samhita

- A drug, if unknown, is fatal like poison, weapon, fire and thunderbolt while, if known, is vitalizer like nectar. A drug unknown by these– name, form and properties (including actions)- and badly administered even if known are responsible for complications Su1#124-125
- A sharp poison also becomes the best drug by proper administration, but on the other hand even the best drug is reduced to sharp poison, if administered badly Su1#126

The original text in Sanskrit



References to Charakasamhita

- http://www.astrojyoti.com/charkasamkitasanskrit.htm
- http://www.astrojyoti.com/pdfs/DevanagariFiles/charaka_sutra.pdf

References

- *Modern Epidemiology,* Rothman, Greenland, & Lash,
- *Reflection on interdependence among academia, regulatory agencies and pharma industry*, Girish Aras. Journal of Indian Statistical Association, Vol 60 (2), 171-186, December 2022.
- *The International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) Guidelines:*

https://www.ich.org/page/efficacy-guidelines