It takes a village: An open-science approach to improving quality and efficiency of the real-world evidence generation process

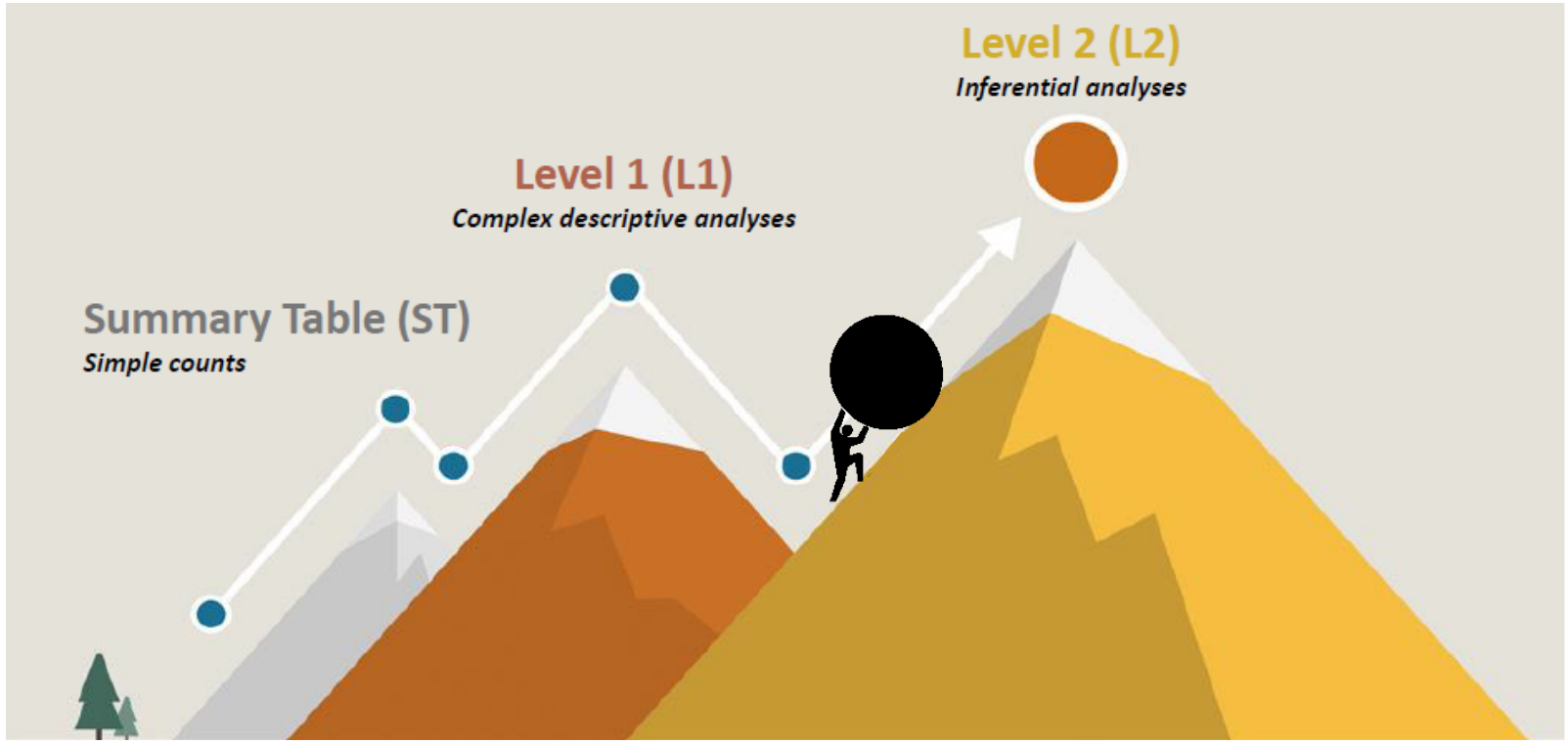
Kristin Feeney Kostka, MPH
Collaborator, OHDSI
Data Science Lead, Deloitte Consulting LLP

Disclaimer

- The opinions expressed here represent my own and not those of my employer.

- Humans are biased. I’m no exception.

Conducting Real World Safety Studies is a Sisyphean Task

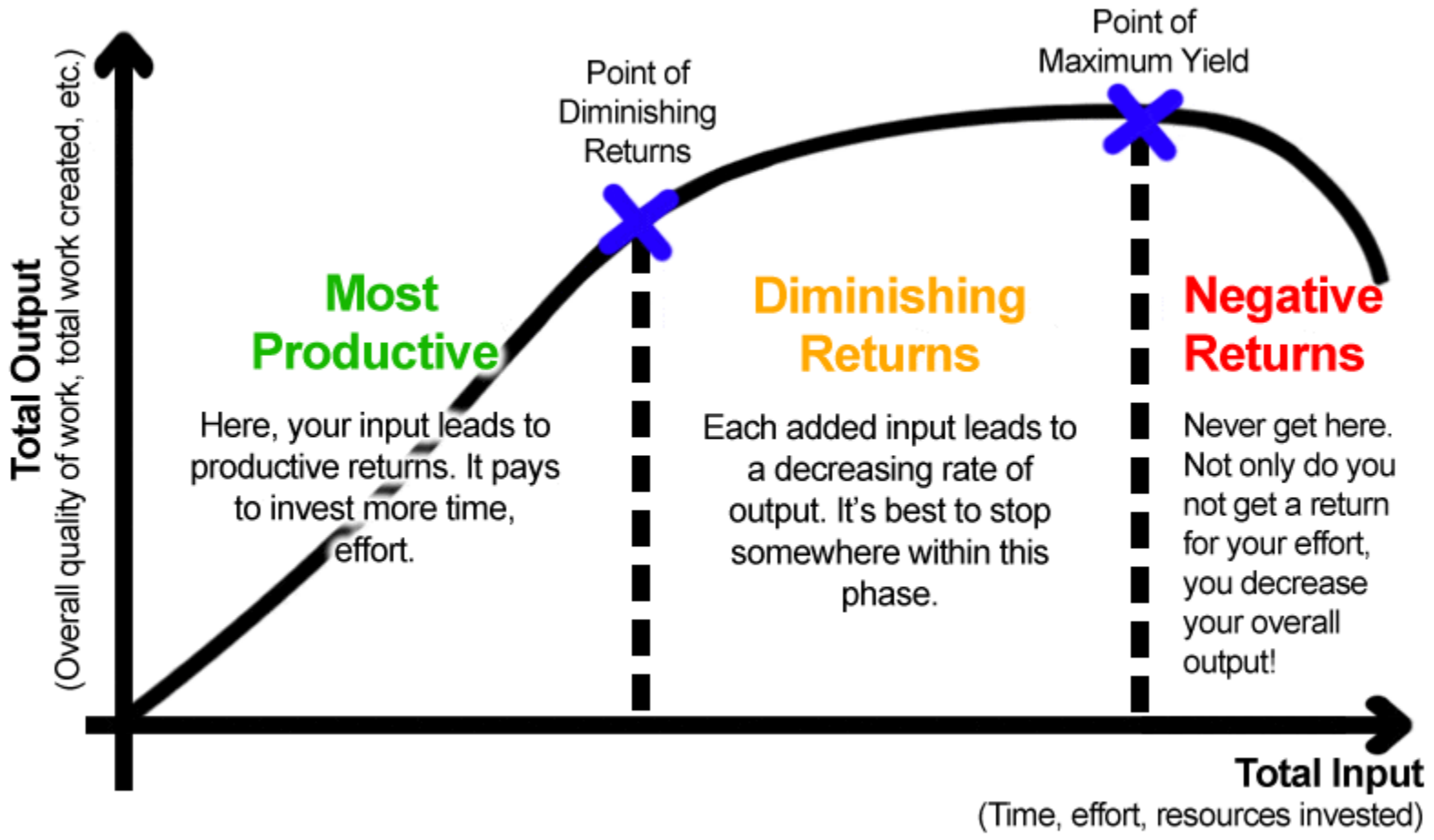


(Modified from Nguyen, 2017)

L1 only: 50 days

L1, L2: 277-741 days

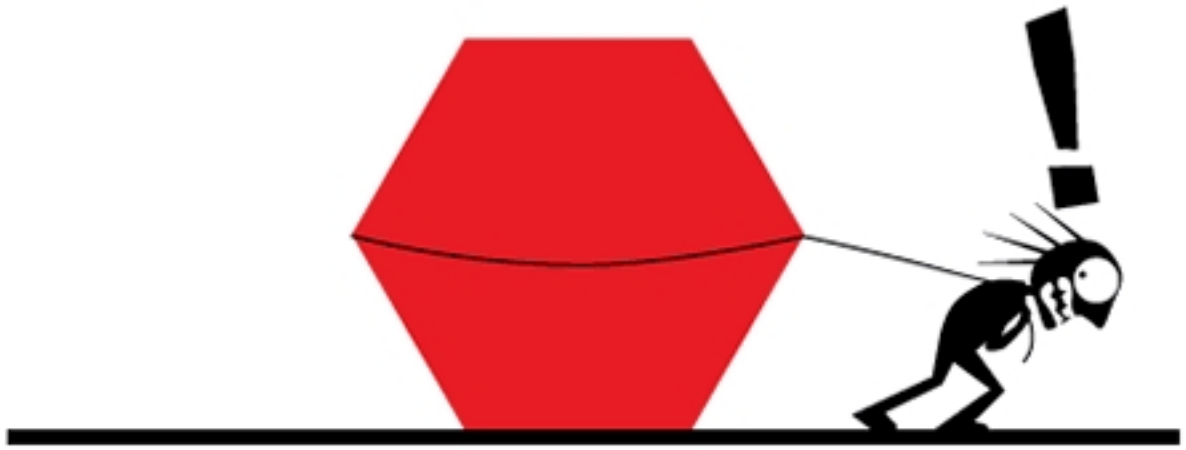
The Current Evidence Generation Paradigm is Flawed



(Source: PersonalExcellence.co)

Borrowing Techniques from Other Industries

THE WATERFALL PROCESS



'This project has got so big, I'm not sure I'll be able to deliver it!'

THE AGILE PROCESS

'It's so much better delivering this project in bite-sized sections'

(Source: Kukhnavets, 2016)

Our Primary Experiment: OHDSI F2F Study-a-thon 2018

Presented on May 18, 2018 OHDSI Community Call

Predicting randomized clinical trial results with real-world evidence: A case study in the comparative safety of tofacitinib, adalimumab and etanercept in patients with rheumatoid arthritis

- Clinical Background, Motivation and my Experience at F2F meeting

Runsheng "Bridget" Wang, MD
Division of Rheumatology, Columbia University Medical Center
DBMI, Columbia University

Study objective

- To compare the safety of tofacitinib with adalimumab and etanercept in patients with rheumatoid arthritis in a retrospective, observational, comparative cohort study.
- We will replicate the design and population inclusion criteria of an ongoing phase 3b/4 randomized clinical trial (NCT02092467), with the aim of predicting the RCT results using real-world evidence from OHDSI.

Common goal: generate reliable evidence on one clinical problem (in less than 48 hours)

Our Secondary Experiment: A Study on Studies

Understanding researcher perceptions on the quality, accuracy and reproducibility of real world evidence generation processes using observational databases

Baseline Perceptions of Observational Database Studies (N=44)

- Average time to complete an observational database study: 257 days (range: 2 days - 1200 days)
- Researchers reported spending an average of 50% of that time on planning (range: 10% - 95%) and the remaining time on execution
- “Somewhat confident” (range: 3-9; median: 7; average: 6.3) about the overall quality of the planning process of their last study
- “Confident” to “very confident” in the expertise of the research teams they conducted the analysis with (range: 3-10; median: 8; average: 7.6)

Baseline Perceptions of Observational Database Studies (N=44)

- “Somewhat confident” in the representativeness of the patient-level data used in their analysis (range: 2-10; median: 6; average: 6.2)
- “Somewhat confident” in the robustness of their analytics methods (range: 2-10; median: 8; average: 7)
- “Somewhat confident” in the reliability of the results generated (range: 1-10; median: 7; average: 6.6)

Post-test Perception of Observational Database Studies (N=33)

- 8 researchers reported this was their first observational database study

- Expectations around moving from study planning → execution:

Expected Time Elapsed	N	%
<1 month	2	6.3%
1-3 months	3	9.4%
4-6 months	8	25%
7-9 months	8	25%
10-12 months	5	15.6%
14-16 months	0	0.0%
16-18 months	3	9.4%
>= 19 months	3	9.4%

75% (n=25) of researchers believed that study planning time should take up at least half of the total time

Post-test Perceptions of Observational Database Studies (N=33)

- 12 researchers attempted to execute the study package on their CDMs

- Of these contributing data sources, researchers felt:
  - "Very confident" (range: 8-10; median: 9; average: 9) in the representativeness of the patient-level data
  - "Confident" to "very confident" (range: 6-10; median: 7.5; average: 7.6) in the robustness of analytical methods
  - "Very confident" (range: 4-10; median: 8.4; average: 8.5) in the overall quality of the study-a-thon planning process and execution
  - "Very confident" (range: 5-10; median: 8; average: 7.9) in the expertise of the study-a-thon community in terms of stakeholder perspective, depth of disciplinary background

Post-test Perceptions of Observational Database Studies (N=33)

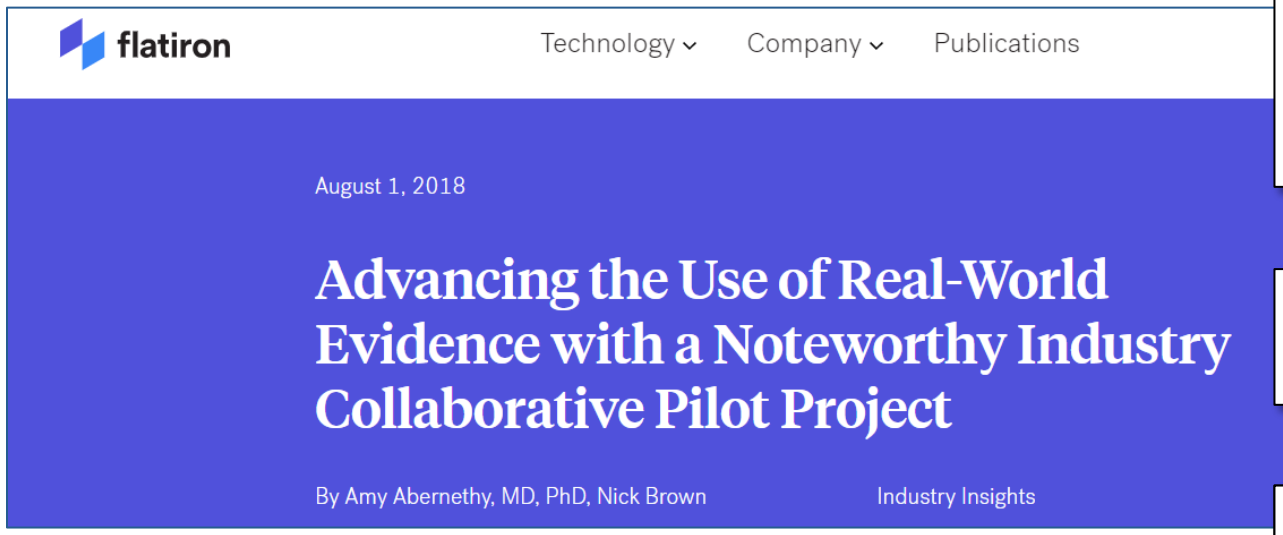
- Would study participants use the study-a-thon process again?
  - "Very likely" (range: 3-5; median: 5; average: 4.5) to want to try to use the approach again for other studies compared to traditional research process

- However, researchers noted that their own organizations may be less likely (range: 2-5; median: 3; average: 3.6) to use this process

Key Takeaways

- To be fair, we didn't actually "finish" the full study
- We still need to do more data characterization of contributing data sources to understand treatment heterogeneity
- You can generate comparable evidence faster, with a larger population, and consuming less resources
- It's also totally possible we might just be drinking the Kool-Aid

But Great Minds Think Alike…



"The ultimate goal of this project was to see how multiple datasets representing different data sources and populations could contribute more complete and reliable understanding to an important current question in cancer care…"

"We chose a common question, and agreed on common analytic plans and common outputs."

"… there was a lot of enthusiasm about the collaboration and how quickly the work itself got done (in just a few months!), which in the cancer research world, as we know, is virtually unheard of. "

Thank You to Our Village!