

Anylam's Use of Citeline Data

Harrison A. Miller

April 17th, 2024

||| About Me - Harrison Miller

- Machine learning scientist in research at Alnylam
 - Build predictive models for early-stage drug development
- M.S. in Applied Mathematics (2021) and B.S. in Chemical Engineering (2020)
- Starting at Harvard Business School in August (Class of 2026)



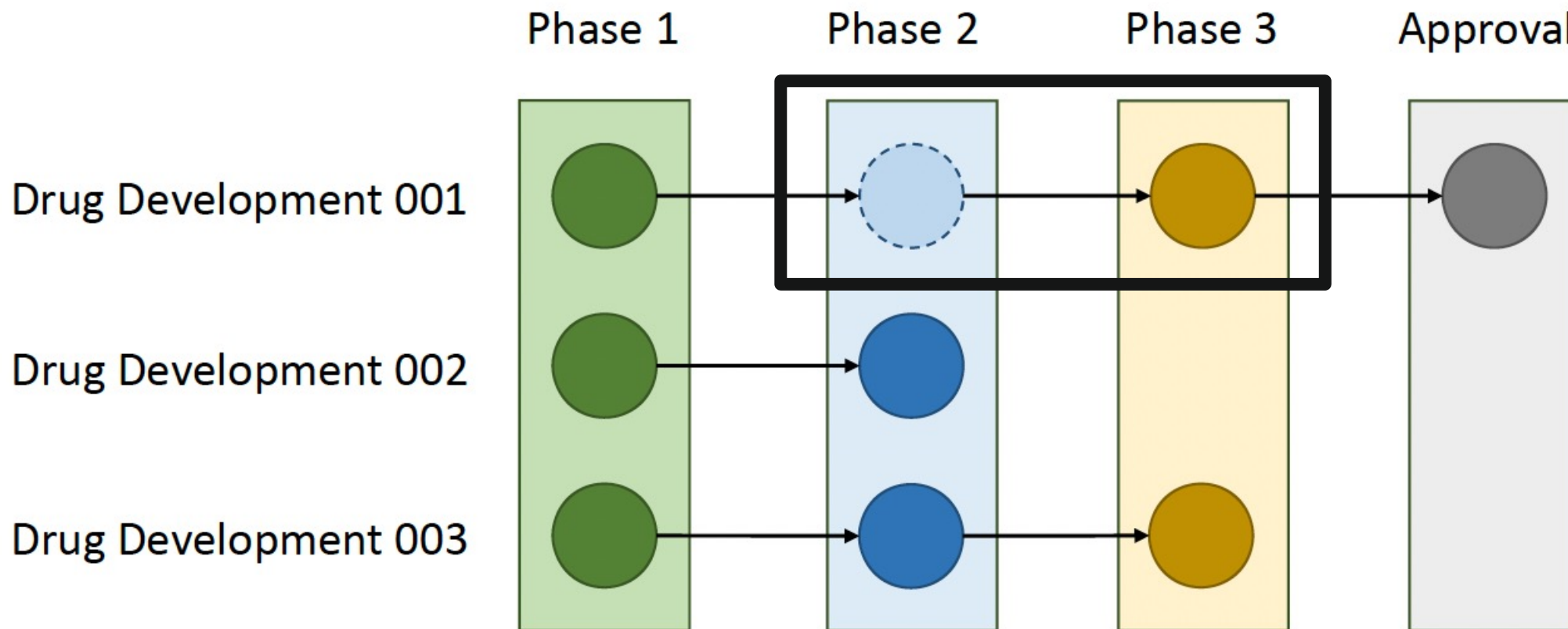
Background on Probability of Success Methodologies

Based on methodologies from *Estimation of clinical trial success rates and related parameters* by Wong et. Al.

In the *Wong* paper, path-by-path and phase-by-phase methodologies are introduced and compared

- Previous literature utilized phase-by-phase due to limitations in other data sets but due to the size of the Citeline data set, *Wong* was able to employ a path-by-path approach that we believe is a better way of viewing success rates of drug programs
- Path-by-path is more robust to missing data/incomplete data sets than phase-by-phase

The Wong Methodology: Comparing Path-By-Path and Phase-By-Phase



Method	POS _{1,2}	POS _{2,3}	POS _{3,APP}	POS _{1,APP}
Path-By-Path	1	2/3	1/2	1/3
Phase-By-Phase	1	1/2	1/2	1/4

Alnylam's Thoughts and Experiences with Citeline Data

- While processing Citeline data, the amount of missing clinical trial data for a given drug/indication was noticeable
 - This was something *Wong* developed an algorithm for to fill out as much of the data set as possible
- We found the most robust information of the dataset was found in the Pharma Projects part of Citeline
- Citeline data does not differentiate well between certain indications

The Progression of PTRS

Ancient History



- Small data sets
- Simple homegrown analytics

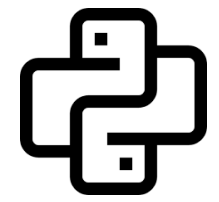
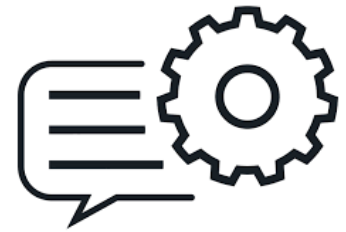
Yesterday's news



CS Tufts Center for the Study of Drug Development
TUFTS UNIVERSITY

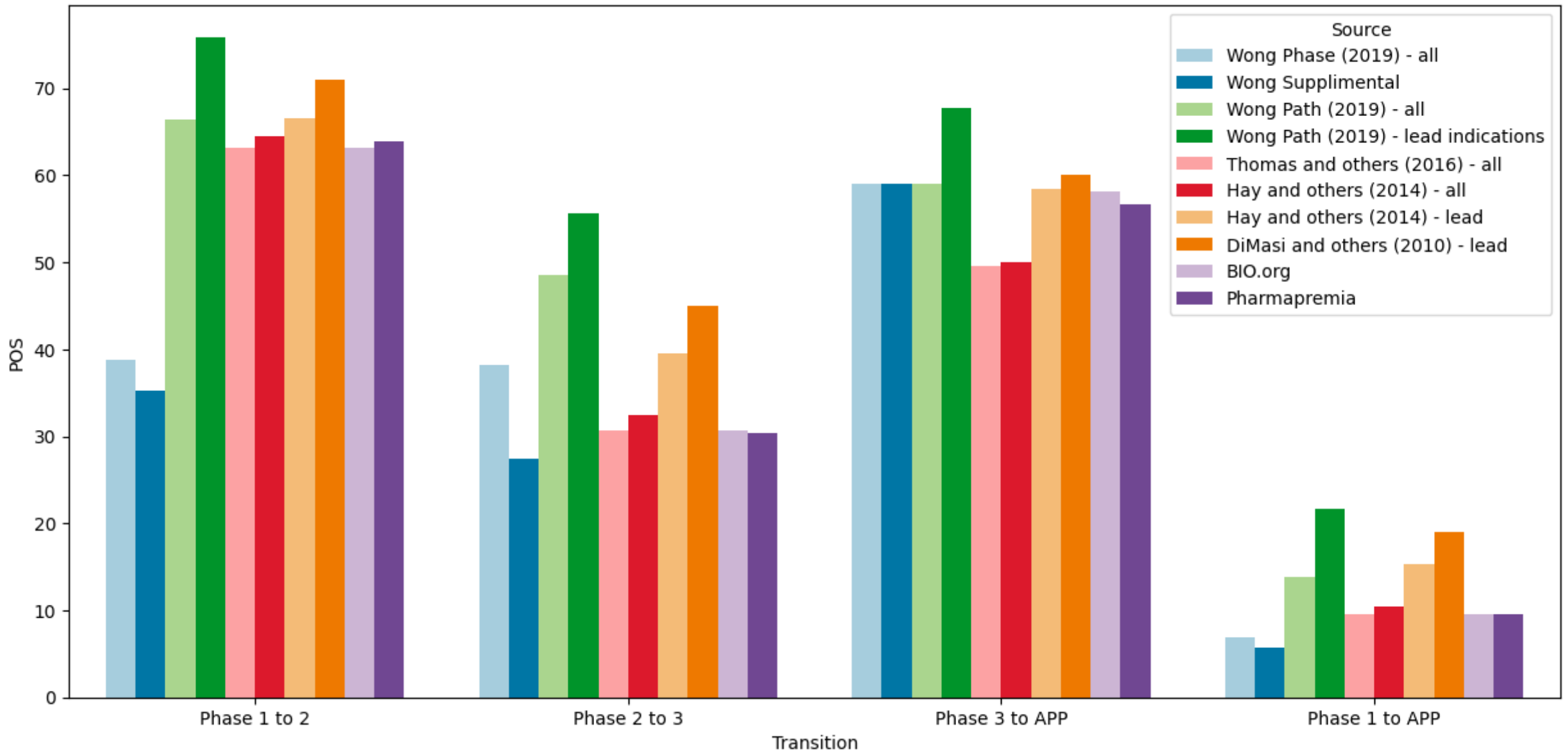
- Larger, curated data sets (Dimasi, KMR, others)
- Proprietary analytics for fee or publication

Today / Tomorrow



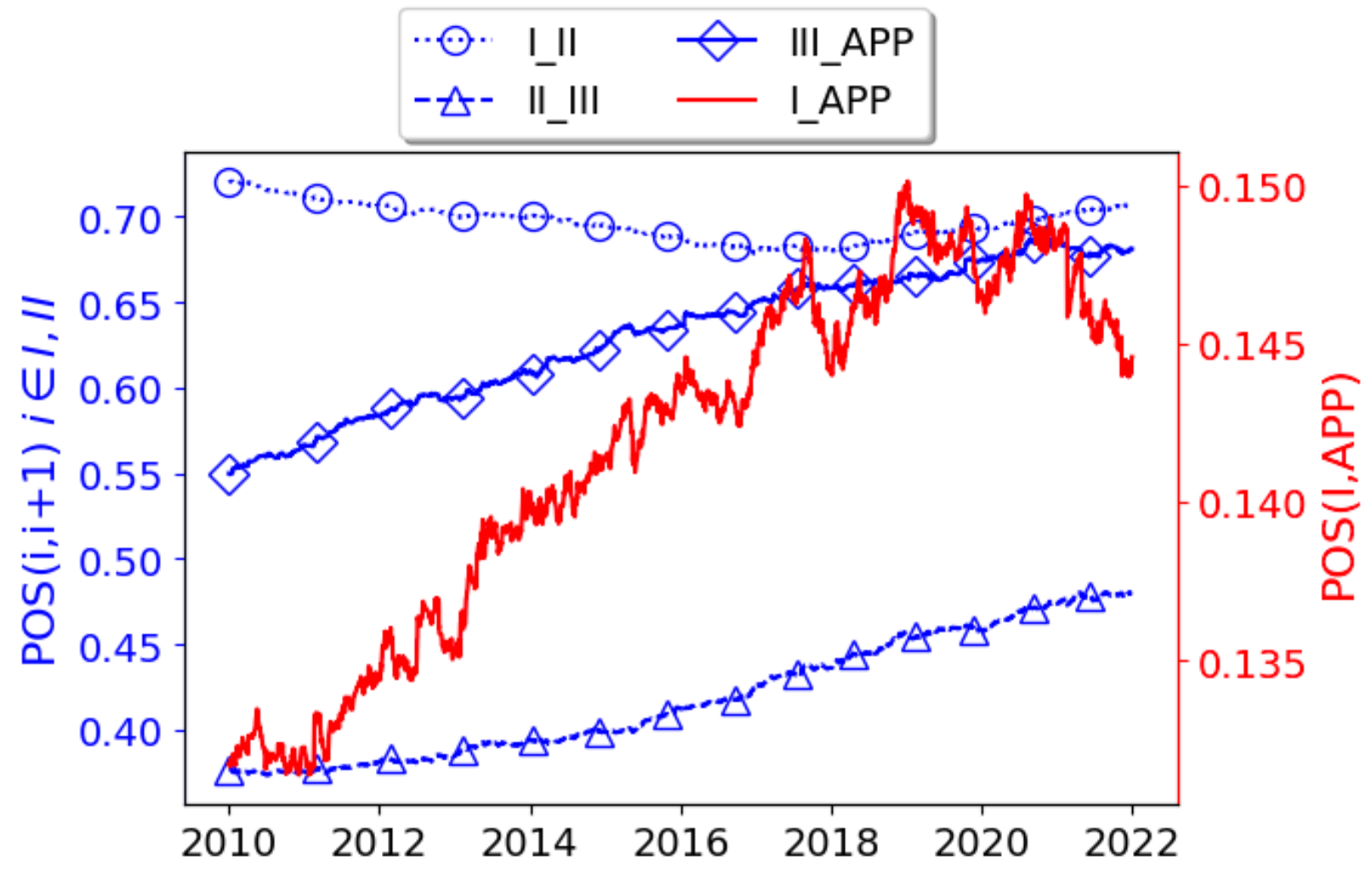
- Large, public cloud-based data
- Server-based analytics that can handle large data sets
- Natural language processing / AI
- Shared code via git hub

In-House Algorithm Aligns with Literature Values



PTRS Over Time

POS going up over time



How Can Clinical Trial Data be Used in Research

- Clinical trial data can be filtered to show drugs that have failed for some reason
 - Can you succeed where someone else failed because you have a safer, more potent technology?
 - Anylam's platform is more specific and potent than many antibodies or small molecules
- By filtering on preclinical data, can you be a fast follower?
 - Companies like EQRx and Checkpoint Therapeutics have used fast following to proliferate while limiting clinical trial costs
 - Being first in class or first in any area is always harder (there are more hurdles to clear)
- By looking at approvals, you can identify drug targets that are not worth going after due to competition issues (how soon will you be competing with generics?)



Other Business Implications

- Benchmarking – comparing our performance versus others
- Identifying outlier performance and understanding drivers
- This can help inform the organization where we need to be more efficient and where we have an edge
 - This can help inform the organization where we need to be more efficient and where we have an edge
 - This can help us show external parties why Alnylam is a company you want to be associated with
- Specific areas of relevance:
 - Success rates for specific drug MoAs, disease areas, modalities, etc.
 - Cycle times (i.e., will programs be more or less expensive?)

