

# Optimization of clinical program

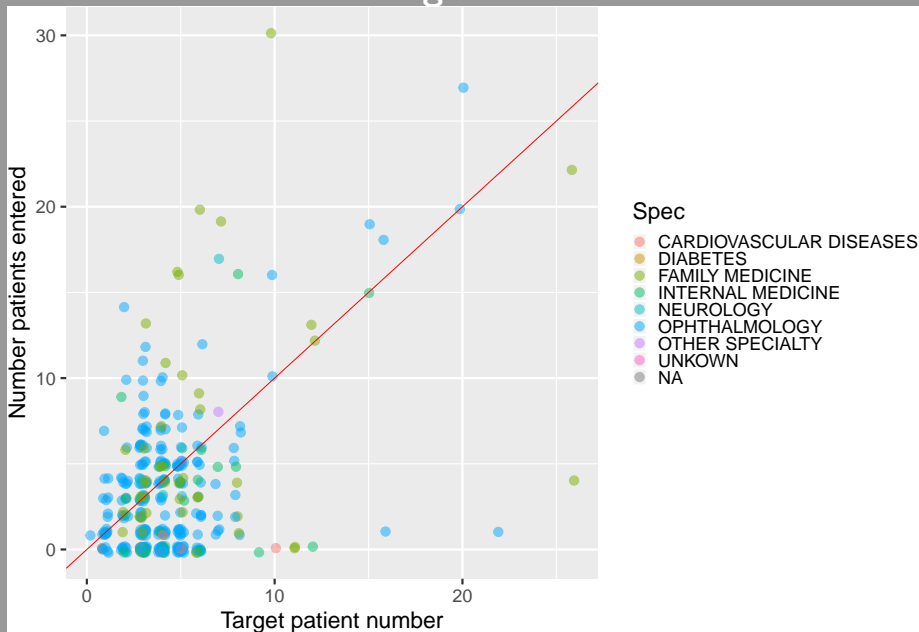
Gilles Guillot

2020-03-05 15:47:26

# Background

- Clinical trials are costly (\$50,000 per patient in oncology) and time consuming
- Often carried out with delays and cost overruns
  - screening failure
  - under-recruitment
  - drop-out
- Clinical trial data of varying quality
  - lack of screening oversight (low screening failure rate)
  - trial patients not representative of target population (high screening failure rate)
  - departure from trial protocol

# Under-recruitment at a glance



# Clinical program optimization problem

Identify physicians and trial sites likely to provide

- high quality trial data
- complying with budget
- on time

*Best predictor of performance in the future is performance in the past.*

# Clinical program optimization problem (cont')

Data analytics objectives:

- Identify *performance indicators* in database
- Extract signal from data
- Identify physicians and trial sites best in terms of those indicators

# Material

## Set A: 139 US physicians

- some replicates in `gpharma_HCP_ID`
- `Patient_Count`
- 11 specialties incl. `OPHTHALMOLOGY`
- admin. / geog. data

# Material (cont')

**Set B:** 365 contributions to studies (PI working in a CENTER)

- Performance indicators:
  - `pat_TARGET_TREATMENT` and `pat_ENTERED_TREATMENT`
  - # of competitor trials by PI (active)
  - # of Supporting Staff
  - PI Tier 1,2,3
    - unclear ordering, not documented
    - not obviously related other variables
  - PI Risk Score H/M/L
    - not documented
    - not obviously related other variables
    - 97% have L
    - can be converted into numerical PI Reliability score 1/2/3
- Admin./geog. data\*
- Uninformative columns:
  - `Trial_STATUS_DESC` all completed\*
  - Performance ratio of `ENTERED/TARGET`
  - PI Availability fully missing

# Method

- Set A
  - Patient\_Count not documented
  - does not seem to relate in a simple way to
    - pat\_TARGET\_TREATMENT or
    - pat\_ENTERED\_TREATMENT in set C
    - even after various summations
- Set C document link between A and B
- Sheet “PI Affiliation”
  - Count of HCP\_Last\_Name not documented
  - does not seem to match physician occurrences in set A or B



# Method (cont')

- Restrict analysis to 274 rows (out of 365) of set B with **ophthalmologists** (requires to cross info from set A and C)
- **Machine Learning / econometrics** approaches based on **time series** model not feasible here
- Purely exploratory/descriptive approach

# Method (cont')

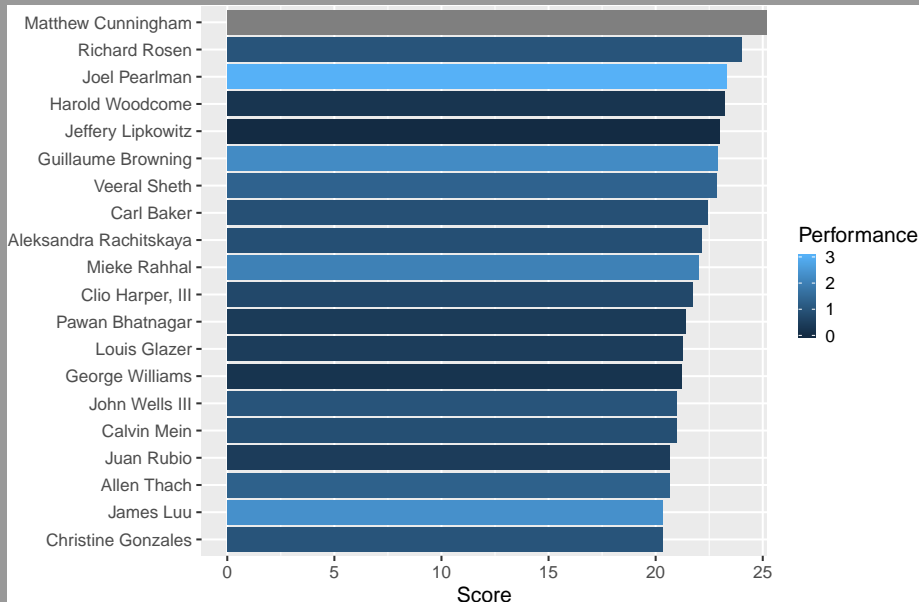
- Consider
  - Performance,
  - # of competitor trials by PI,
  - # of Supporting Staff,
  - PI Tier,
  - PI Reliability Score
- For each physician: compute average of variables above across studies
- Define Score as

Score = mean (Perf, #Trials, #Support staff, Tier, Reliability)

# Results

PI_FORENAME	PI_SURNAME	CENTER_NAME	Score
Matthew	Cunningham	Florida Retina Institute	Inf
Richard	Rosen	NY Eye and Ear Infirmary Mt Sinai	24.00000
Joel	Pearlman	Retinal Consultants Medical Group of Sacramento	23.34375
Harold	Woodcome	Retina Consultants of Providence	23.25000
Jeffery	Lipkowitz	Delaware Valley Retina Associates	23.00000
Guillaume	Browning	Charlotte Eye Ear Nose and Throat Clinic	22.88889
Veeral	Sheth	University Retina and Macula Associ	22.83333
Carl	Baker	The Ophthalmology Group LLC dba Paducah Retinal Center	22.41667
Aleksandra	Rachitskaya	Cleveland Clinic Cole Eye Institute	22.14583
Mieke	Rahhal	Retina- Vitreous Assoc Medical Group	22.00000
Clio	Harper, III	Austin Retina Associates	21.71429
Pawan	Bhatnagar	Retina Consultants	21.40000
Louis	Glazer	Vitreo-Retinal Associates, PC	21.28889
George	Williams	Associated Retinal Consultants, PC	21.22500
John	Wells III	Palmetto Retina Center	21.00000
Calvin	Mein	Retinal Consultants of San Antonio	20.97417
Juan	Rubio	Retina Associates of South Texas PA	20.67500
Allen	Thach	Retina Consultants of Nevada	20.65909
Christine	Gonzales	Retina and Vitreous Center Of S Oregon Pc	20.33333
James	Luu	Retina Consultants of Southern Colorado	20.33333

# Results (cont')



# Data/Method criticism:

- Room for improvement
- Does not lend itself to evaluation/validation
  - no estimate of precision
  - no estimate of uncertainty
- Approach not rooted in a formal **decision theory** framework
  - Score needs to be objectively defined through **utilities**, **penalties** derived from **real costs** and **business objectives**

# Prospect for improvement:

- More exploratory analysis
- Move towards decision theoretic approach
- Move towards use of **time series** data

# Supplementary information

All material temporarily available on github repository

<https://github.com/gilles-guillot/GiantPharma>

- R code
  - data loading and formatting: [load.R](#)
  - data analytics: [analytics.R](#)
- Slides
  - R markdown: [GPharma.Rmd](#)
  - Pdf: [Gpharma.pdf](#)

THANK YOU