Identifying the "right" subgroup

Noah Simon

April, 2022

The "right" subgroup

I think this discussion is less about statistical methodology...

and more about scientific tradeoffs

In trying to develop methodology here, I have found that...

Once you have clarified your goal...

most methods which target that goal perform fine

The following ideas have evolved from discussion with

No, that's not just a picture of me as a child...

The following ideas have evolved from discussion with

He's the one in the center...

giving everyone else a hard time at seminar!

The following ideas have evolved from discussion with...

Disclaimer

While they have helped shape my thinking...

they are potentially in disagreement with what I say here

(especially Scott...[1])

[1] I'm half hoping he's in the audience to just vehemently disagree!

Identifying the "right" subgroup?

I will stick mostly to oncology...

Not because I know much there...

But I definitely know less about everything else!

Identifying the "right" subgroup?

The practice of medicine has *always* been about

- characterizing dysfunction
- treating based on specific characterizations

Identifying the "right" subgroup?

In the beginning this was based on simple observation alone:

You have a breast lump... So... $\longrightarrow$ Leeches!

Now we have more sophisticated methods:

over-expression of HER2 on breast tumor $\longrightarrow$ Herceptin

---

In oncology, tumors are characterized using site/histology

Identifying the "right" subgroup? (Precision Medicine)

My understanding is:

Medicine attempts to differentiate diseases...

to develop treatments that target specific disease characteristics

---

[Biomolecular] Precision medicine attempts to differentiate diseases using biomolecular profiling

to develop treatments that target specific biomolecular disease characteristics

Identifying the "right" subgroup?

Two common scenarios:

Developing a targeted treatment + diagnostic

Developing a new diagnostic, for an existing, non-targeted treatment

Targeted Treatments

In oncology, currently 163 FDA approved targeted therapies[2]

Various different targets

Many without "companion diagnostics"...

Instead approved in specific cancer types

(histology-based personalization!)

[2] from https://www.cancer.gov/about-cancer/treatment/types/targeted-therapies/targeted-therapies-fact-sheet

Companion Diagnostics

50ish approved companion diagnostics[3]

Essentially based on

- Binarized continuous score (most commonly protein expression measured via IHC)
- Nearly binary count data (eg. copy number variation)
- Binary variable (presence/absence of mutation)

Essentially all of these were based on clear targets of therapy

[3] from `https://www.fda.gov/medical-devices/in-vitro-diagnostics/list-cleared-or-approved-companion-diagnostic-devices-in-vitro-and-imaging-tools`

The "wrong" subgroup

It seems relatively fruitless to...

Characterize the [in]effectiveness of non-targeted treatments

---

Why do I tend to miss free throws?

Because I keep forgetting to wear my lucky shirt...?

Or maybe because I'm generally bad at basketball...

Targeted Therapies

Let's focus on Targeted Therapies...

with targets that are, at least, somewhat understood

For example

- Cetuximab (that targets EGFR receptor)
- Anti-PD1 (/L1) therapies

Targeted Therapies

These examples are somewhat subtle

Cetuximab– quantity of EGFR expression is less important than...

If signaling is done through binding with ligand

RAS mutations!

Anti-PD[L]1 therapy (atezo/pembro)– PD[L]1 expression matters

Unclear where it should be measured

via IHC or rna-seq

Perhaps immunogenicity of tumor is also important (eg. TMB)

Targeted Therapies

The most promising problems are characterized by

- Small number of biologically informed features
- Often a continuous/ordinal measure with unknown cutpoint

Two Questions

Q1) Formally, what do we mean by the "right" subgroup?

Q2) How do we identify it?

What do we mean by the "right" subgroup?

Suppose we have a treatment whose effect we believe is modified by a covariate x

Imagine we have a binary outcome y (1 is good!)

Suppose we consider [personalized] tx effect as

$$\Delta(x) = P(y = 1|x, T = 1) - P(y = 1|x, T = 0)$$

Further suppose tx effect is assumed non-decreasing in x

What do we mean by the "right" subgroup?

1 of 3 choices we often implicitly target

- Largest group with average benefit

$$\{\text{all } x > x_0\} \text{ where...}$$
$$x_0 \text{ is the minimum value w/ } E[\Delta(x)|x > x_0] > 0$$

- All people with "individual" expected benefit

$$\{\text{all } x \text{ w/ } \Delta(x) > 0\}$$

- A subset of people with substantial expected benefit

The right answer is context dependent!

What do we mean by the "right" subgroup?

1 of 3 choices we often implicitly target

- Largest group with average benefit

$$\{\text{all } x > x_0\} \text{ where...}$$
$$x_0 \text{ is the minimum value w/ } E[\Delta(x)|x > x_0] > 0$$

- All people with "individual" expected benefit

$$\{\text{all } x \text{ w/ } \Delta(x) > 0\}$$

- A subset of people with substantial expected benefit

I think we often aim for first or second

I think a good case can be made for the third!

What do we mean by the "right" subgroup?

I would say the "right subgroup" is...

> A subgroup for which we can, with high power, identify a positive average treatment effect with a reasonable sample size

A segue about Clinical Trials

We act like clinical trials can tell us precisely about an average treatment effect were we to employ this treatment in the population at large

I think in that respect they are quite imperfect

We only use them because they are vastly superior to every alternative.

A segue about Clinical Trials

In line with that reasoning...

I am not particularly concerned with identifying precisely who benefits from the new therapy...

I would just like strong evidence that there is a subset of patients who benefit on average

and I would like some idea of who they are!

Adaptive Enrichment Designs

Let's think more about this in the context of Adaptive Enrichment

Adaptive Enrichment Trials modify ongoing trial enrollment to

- Better learn about who benefits from new treatment
- Improve power

Adaptive Enrichment Designs

I have seen two primary frameworks for adaptive enrichment:

I will call the first the multiplicity framework...

In this approach, we

1. Prespecify a bunch of strata
2. Drop poorly performing strata as trial progresses
3. At the end, evaluate discrete hypotheses for each remaining stratum

(usually using some closed testing procedure)

Adaptive Enrichment Designs

I have seen two primary frameworks for adaptive enrichment:

I will call the second the single-test framework...

In this approach, we

1. Adaptively modify enrollment as trial progresses to include only people likely to benefit from treatment

2. At the end, calculate a single weighted sum of t-statistics to identify if there is some subset of people who benefit

Adaptive Enrichment Designs

These test two different null hypotheses:

For each strata $S_1, \ldots, S_K$, the first tests hypotheses:

$$H_k : \text{ average treatment effect for ppl with } x \in S_k \text{ is } 0$$

for $k = 1, \ldots, K$

The second tests a somewhat curious global null:

$$H_0 : \text{For any possible strata, characterized by } x\ldots$$
$$\text{average treatment effect is } 0$$

Given realities of what we can/cannot learn from clinical trials...

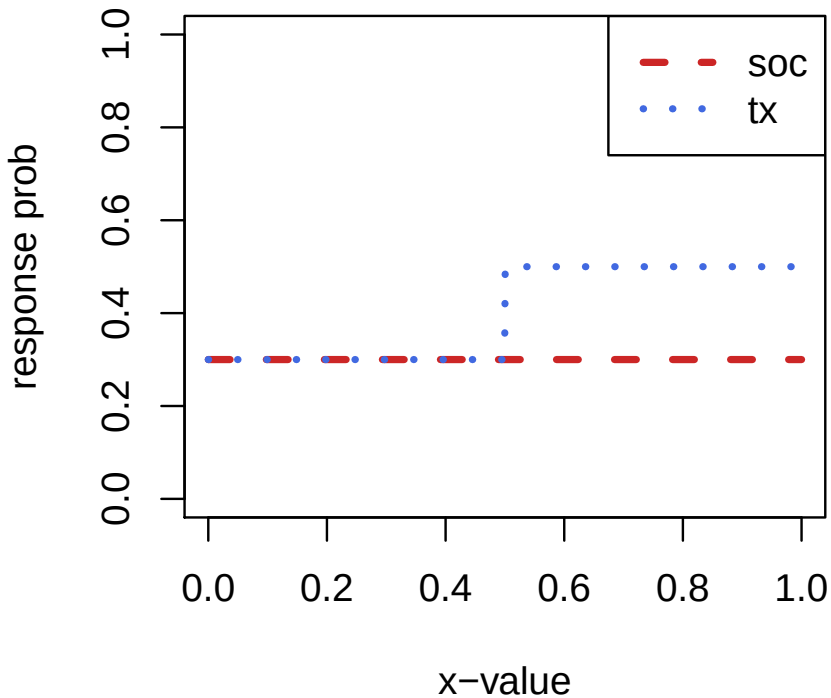
I would argue that the second may be good enough

A Simple Scenario

Suppose we have a randomized treatment T

a single continuous $X \sim U[0, 1]$

A binary outcome w/



A Simple Scenario

I'm going to consider a 2 block design in the "single test" framework ($n = 200$ per block)

We will consider adapting to exclude patients with x-value below a cutpoint after the first block

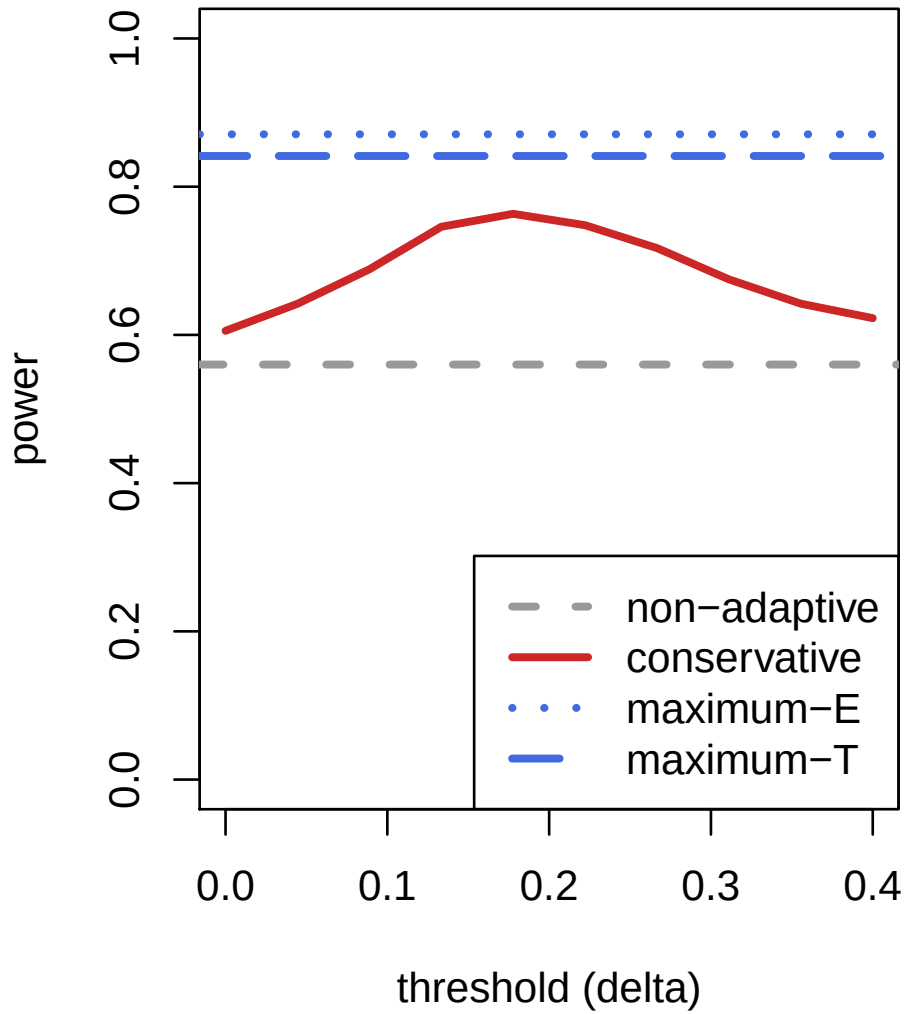
Will consider 9 candidate cutpoints (deciles)

Consider 2 classes of decision rules:

- Choose the lowest cutpoint such that the estimated average treatment effect from stage 1 is at least δ
- Choose the cutpoint with the largest average estimated treatment effect (/z-score)

A Simple Scenario

Aggressive strategies have *much* higher power than conservative



Implications

This might mean that we should more readily restrict enrollment...

Either to run an enrichment design

Or a more aggressive adaptive enrichment design

Choosing the "right" subgroup

Two simple(ish) ways to do this:

- Use heuristics (simulate under various alternatives to understand operating characteristics)
- Use a Bayesian-design; choose a sensible utility; and identify a strategy to optimize it!

Also, thanks to a certain chubby baby

Apparently, we have strong "chubby cheeks" genes in my family

Papers

Simon N. and Simon R. Adaptive Enrichment Designs for Clinical Trials (Bisotatistics 2013).

Simon N. and Simon R.Using Bayesian modeling in frequentist adaptive enrichment designs (Biostatistics 2018).

Simon R. and Simon N. Inference for multimarker adaptive enrichment trials (Statistics in Medicine 2017).