Emulating a target trial to evaluate the clinical utility of a biomarker-based treatment rule

Joint Statistical Meetings

Michael Sachs

2021-08-09

So you want to use machine learning

"I've heard about machine learning and I don't want to miss out, can you help me apply it to my data?"

No! **Start with a hypothesis!**

ML is a suite of tools for prediction, i.e., can I predict future/unobserved values of Y using X?

What is the intended use of the predictions?

- Implement a new policy or screening program on a population level
- **Guide treatments for individual patients**
- Allocate funds/time for further research and development?

A prediction by itself is not clinically useful unless it leads to an action

Rich Simon on black box prognostic models

"There is an **enormous gap between the large literature on prognostic models and the small number of models used in medical practice**. There are several reasons for this discrepancy. **Most prognostic models do not provide actionable information.** That is, they are based on analysis of a heterogeneous set of patients who received a variety of treatments. **Physicians want tools that help them make treatment decisions.** Unless that decision context is clearly and specifically defined at the outset of the study and used to drive the selection of the training set, the resulting model is unlikely to find medical acceptance."

Assume you have a decision rule

1. Doctors are using a risk score and
2. based on it, the definitions of high risk and low risk have been decided

A treatment *decision rule* is that high risk patients receive aggressive combination treatment while low risk patients get a less costly and less toxic monotherapy treatment.

What do we want to estimate?

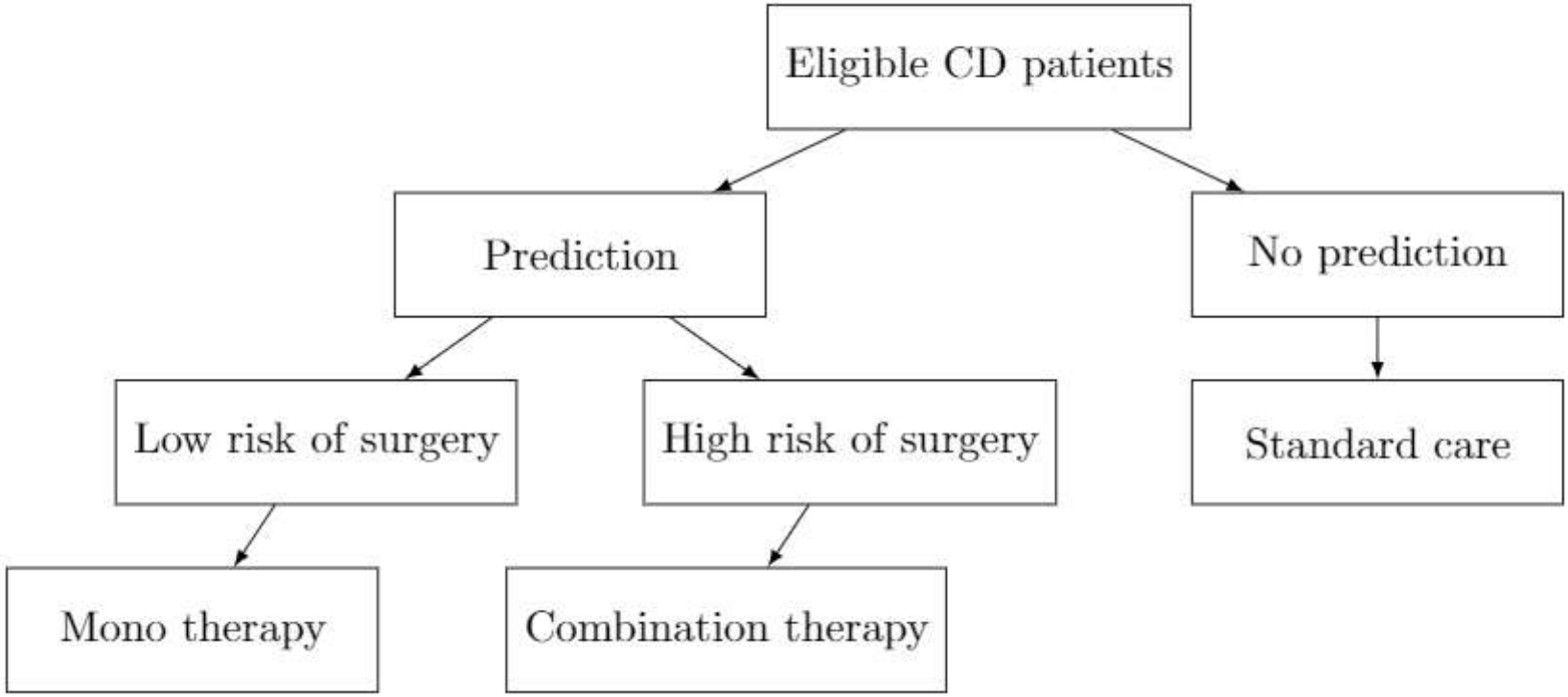
Estimand of interest

$\theta = E\{Y(\text{used decision rule})\} - E\{Y(\text{standard of care})\}$

Where $Y(\cdot)$ denotes the potential outcome under the possibly hypothetical scenario where $\cdot$ was the intervention (decision rule) used in the population.

Standard of care may be physician choice or treat everyone with one of the options A or B

Generate evidence to support use/no use of the model



Where is the money for this going to come from?

1. Very few trials of this nature are being run
2. Most are platform studies, like Probio II
3. What do the rest of us do?

We need preliminary evidence to support such a study.

Using an emulated trial in observational data

Same as Hérnan and Robins (2016)

1. Use the same inclusion/exclusion criteria that a clinical trial might
2. Define a grace period for use of the deterministic decision rule
3. Deal with people that die/have the event prior to the end of the grace period

Different

1. The estimand is $E\{Y(\text{used decision rule})\} - E\{Y(\text{standard of care})\}$
2. since $E\{Y(\text{standard of care})\}$ is literally what we observe, this has no confounding
3. $E\{Y(\text{used decision rule})\}$ can be decomposed to allow for estimation in the observed data

Estimand

$$E\{Y(\text{used decision rule})\} = Pr\{\text{low risk}\}E\{Y(A)|\text{low risk}\} + Pr\{\text{high risk}\}E\{Y(B)|\text{high risk}\}.$$

Under what conditions is this estimable?

How to estimate and do inference?

Assumptions for identification

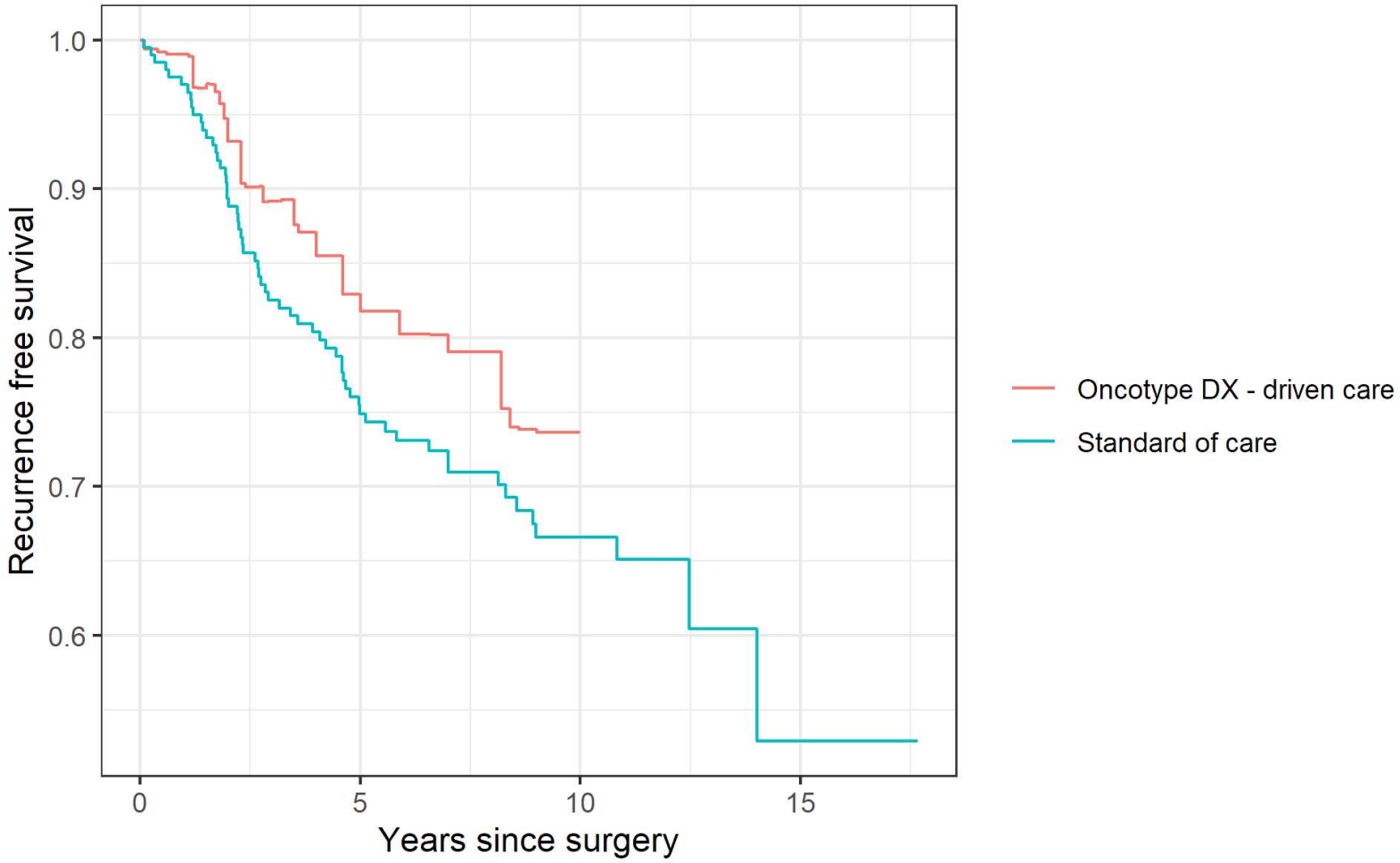
$Pr\{\text{low risk}\}E\{Y(A)|\text{low risk}\} + Pr\{\text{high risk}\}E\{Y(B)|\text{high risk}\}$

Say we observe Y_i the outcome, Z_i the treatment actually received (A or B), H_i an indicator of high risk, and a set of covariates C_i for each subject i.

1. Positivity of treatment assignment: within each subgroup defined by C_i, there is a positive probability of receiving either treatment.
2. No unmeasured confounding, all confounders of $Z_i \to Y_i$ are observed and measured without error.
3. Correct specification of the model $E\{Y_i|Z_i, H_i, C_i\} = g(\beta; Z_i, H_i, C_i)$.

Under those assumptions, each term in the above display can be consistently estimated using g methods

Oncotype DX from observational data



Sensitivity to assumptions

Needless to say, it is implausible that all confounders are observed and measured without error...

In that case, we are working on methods to compute *nonparametric bounds* on the causal effects $E\{Y(A)|\text{low risk}\}$ and $E\{Y(B)|\text{high risk}\}$.

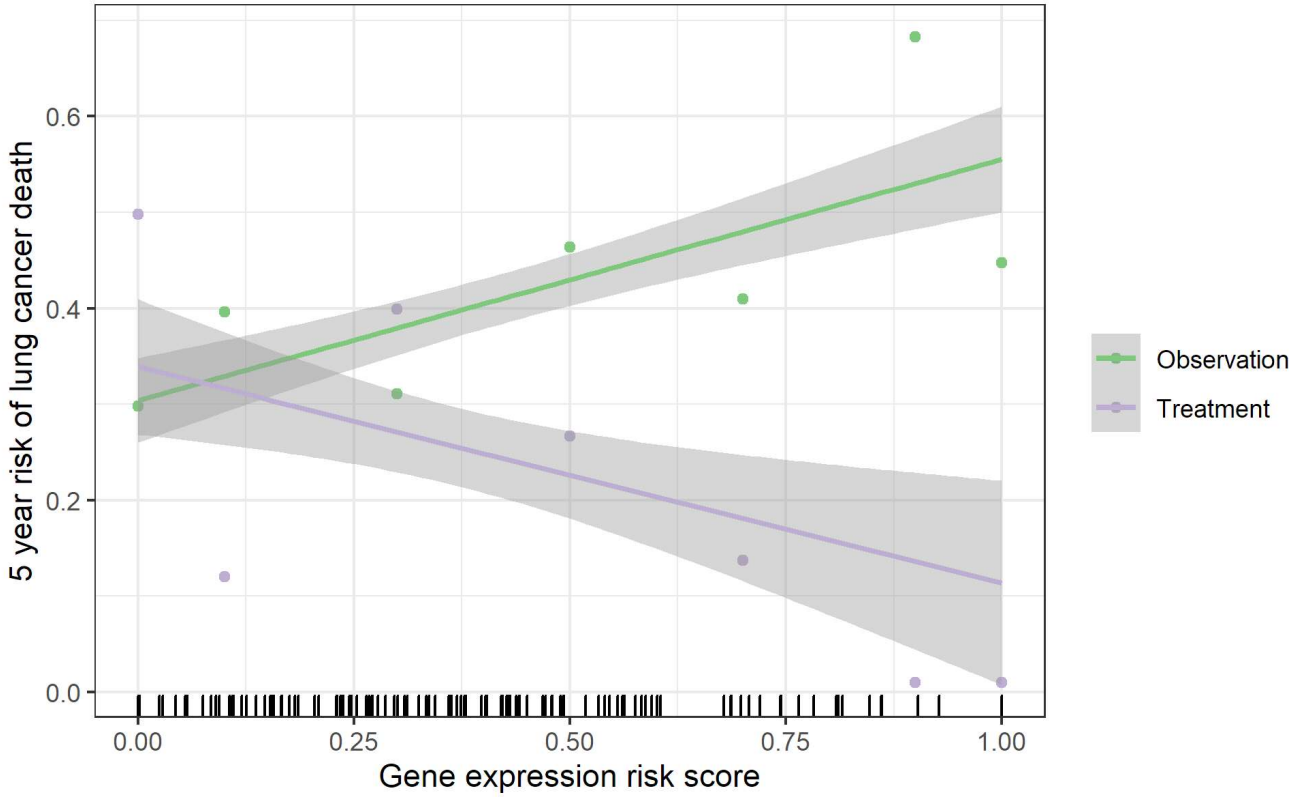
The idea:

- under the assumption of unmeasured confounding,
- the observed distribution and probabilities constrain the possible values of the causal effect
- The constraints can be derived using linear programming and vertex enumeration

See Gabriel et al. (2020) and (2021), both in JASA for some examples

Developing a decision rule

So far we have assumed that the biomarker is binary (high risk or low risk). With a continuous biomarker we can imagine defining a series of high vs low risk groups defined by varying a threshold.



Multiple biomarkers

With multiple biomarkers and high-dimensional data, the optimization is much more difficult.

Gustav Jonzon, Biostatistics PhD student at KI, is working on *reinforcement learning* methods for optimization of treatment decision rules in such settings.

- Develop methods to optimize prediction-based decision rules w.r.t. *individual patient expected return*, that directly target the desired causal outcome
- Given observational data on covariate-history and outcomes, outline assumptions under which the causal effect of the decision rule is estimable
- Under these assumptions and a suitable estimation procedure, design a loss function based on this causal estimator and a choice of method to fit the *action-value function* of the rule

Study designs

The simple trial that randomly assigns individuals to use of the decision rule versus standard of care is the most direct way to estimate θ, but is it the most efficient?

Adam Brand, Biostatistics PhD student at KI is studying these questions and more

- review the literature in order to enumerate the relevant study designs;
- determine the target estimand in each of the designs;
- compare these designs via simulations in various treatment scenarios (e.g., varying standard of care, the nature of the novel treatment and prediction model);
- make recommendations and possibly suggest novel designs.

Conclusion

Some causal thinking can help clarify:

- What we want to estimate/develop to provide actionable information
- Assumptions needed to do that
- Study designs to do that efficiently

Read more about the concept in

Sachs MC, Sjölander A, Gabriel EE. Aim for Clinical Utility, Not Just Predictive Accuracy. *Epidemiology*. 2020;31(3):359-364.

Join us!

The Biostatistics Group at KI is recruiting a tenure-track assistant professor

Take a screenshot (Win+PrtSc on Windows, Cmd+Shift+3 on Mac, PrtSc on Linux)

We are a small but active group, working and living conditions in Sweden are very nice

Get in touch with any questions (michael.sachs@ki.se) or @ki_meb_biostat on Twitter