NIT outcome measures for MASH should reflect clinical significance

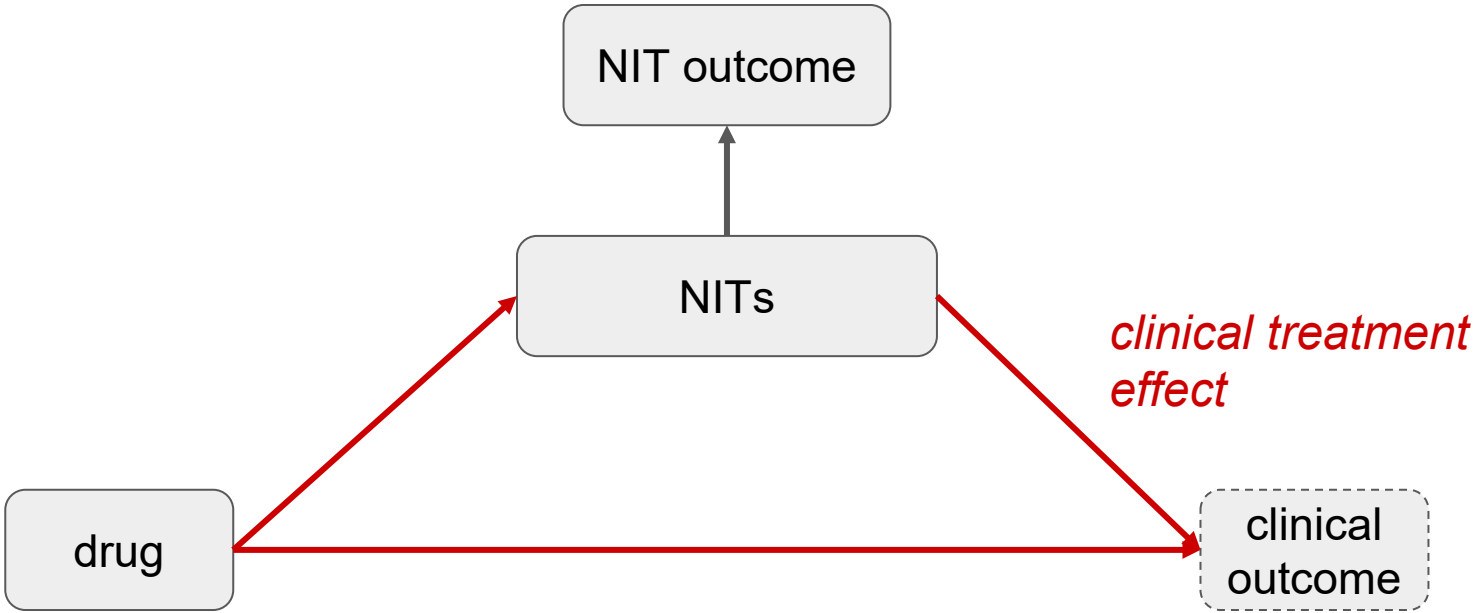
Alejandro Schuler
Asst Prof, UC Berkeley Biostatistics

Brief introduction

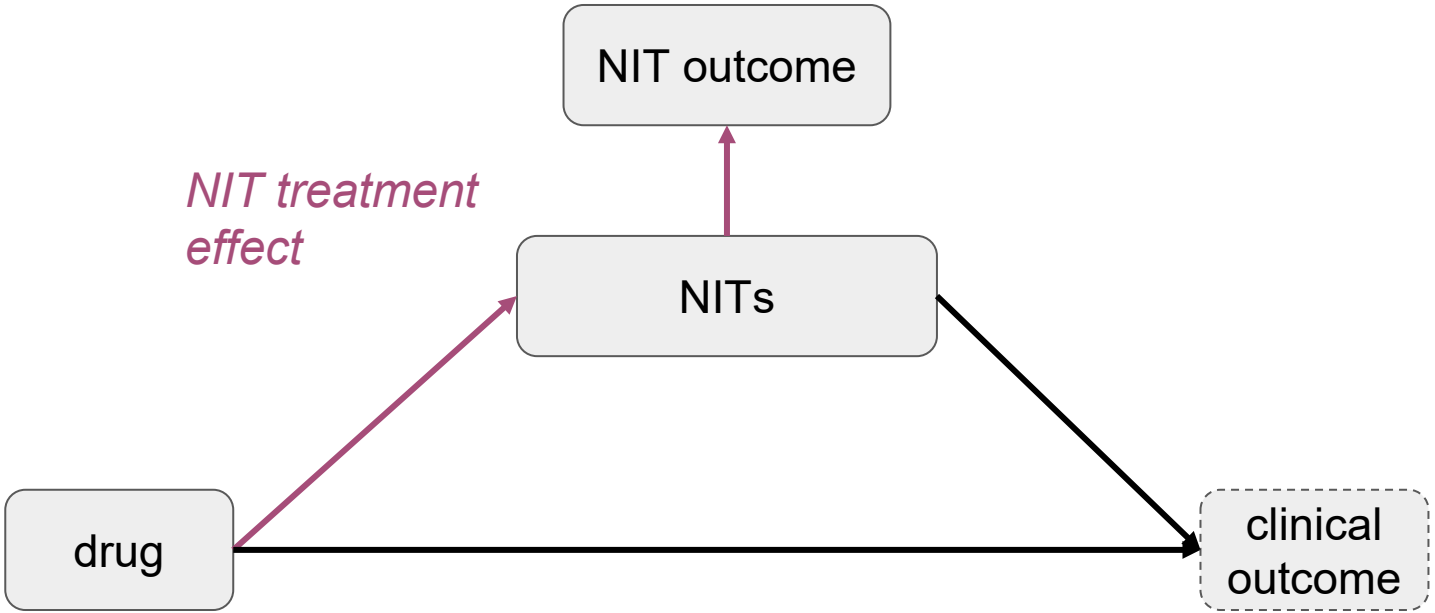
- **Asst Professor, UC Berkeley Biostat 2022+**
  - PhD, Stanford Biomedical Informatics (N. Shah) 2018
  - Data Scientist, Kaiser Division of Research 2018-2020
  - Statistician, Unlearn.AI 2020-2021
  - Postdoc, UC Berkeley CTML (M. van der Laan) 2021-2022

- **Methods:** semiparametric methods for causal inference, heterogeneous treatment effects, gradient boosting, model selection
- **Applications**: clinical trial design, data fusion, predictive models in the clinic, observational studies from EHR data

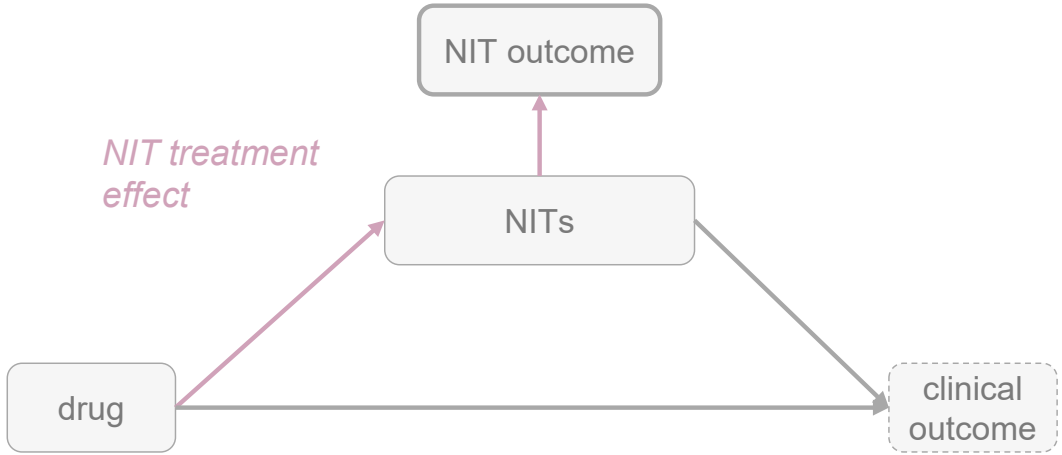
Setting


NIT outcome
NITs
clinical treatment effect
drug
clinical outcome

Setting


NIT outcome
NIT treatment effect
NITs
drug
clinical outcome

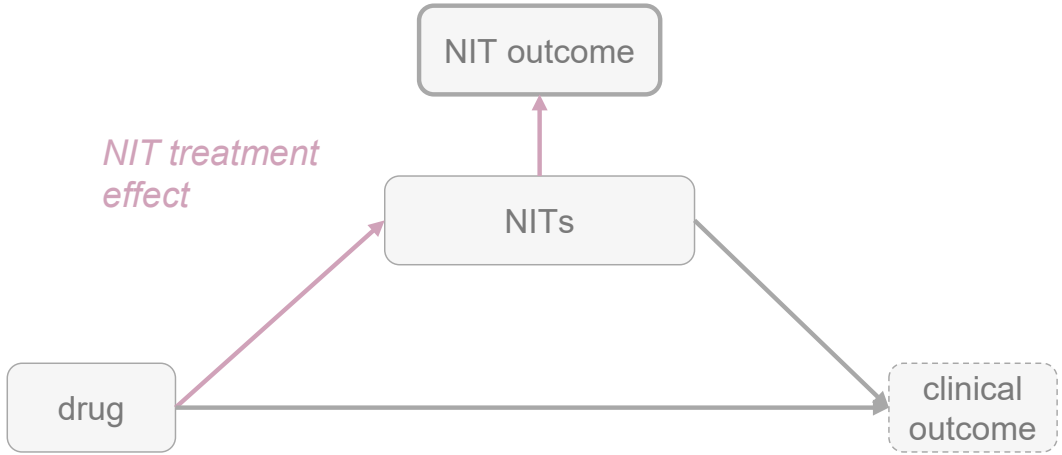
How do we **construct the NIT outcome** so that the trial NIT treatment effect is similar to the population clinical effect?



This is a **well-studied problem**, addressed with modern causal hygiene

- **VanderWeele**, Tyler J. "Surrogate measures and consistent surrogates." Biometrics 69.3 (2013): 561-565.
- Gilbert, Peter B., and Michael G. **Hudgens**. "Evaluating candidate principal surrogate endpoints." Biometrics 64.4 (2008): 1146-1154.
- Price, Brenda L., Peter B. Gilbert, and Mark J. **van der Laan**. "Estimation of the optimal surrogate based on a randomized trial." Biometrics 74.4 (2018): 1271-1281.
- **Athey**, Susan, et al. The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely. No. w26463. National Bureau of Economic Research, 2019.

How do we **construct the NIT outcome** so that the trial NIT treatment effect is similar to the population clinical effect?

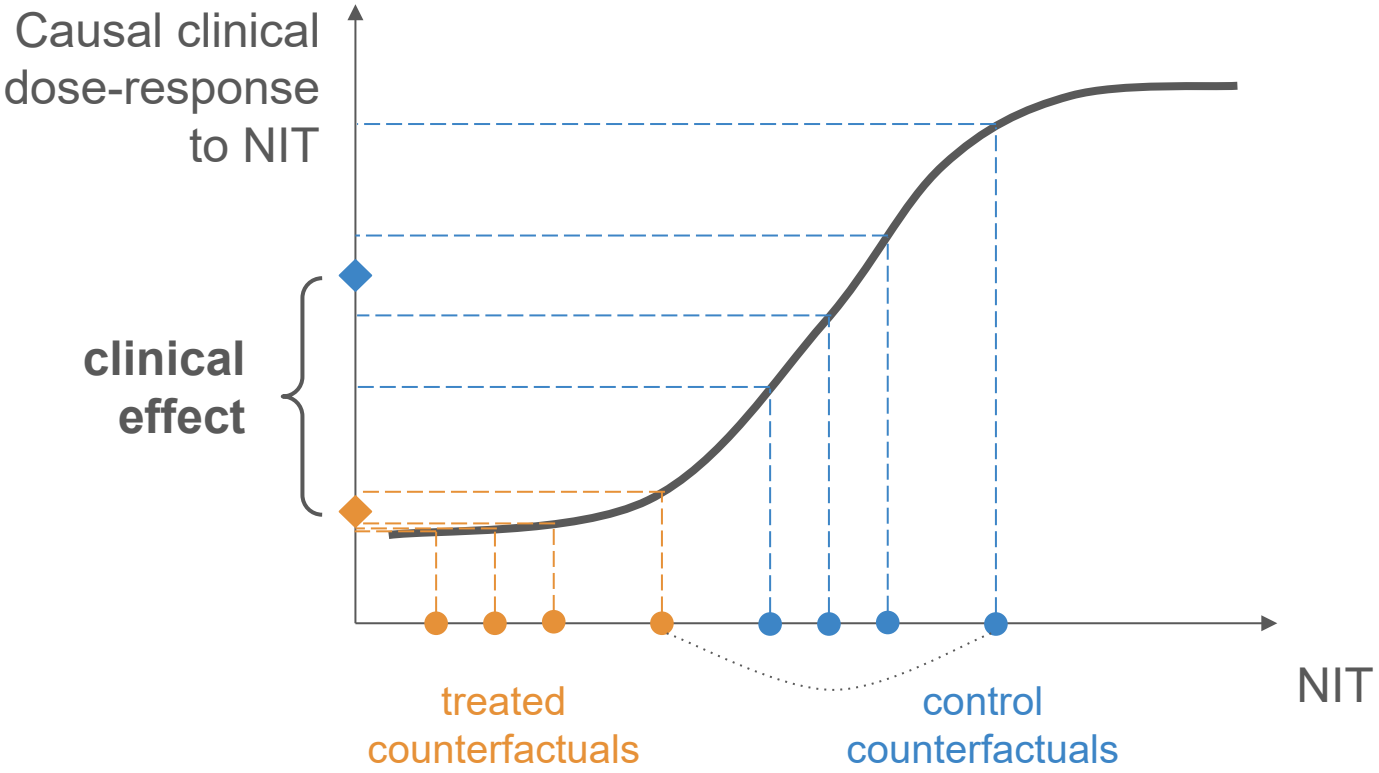


Let's simplify.
if we can't do it when it's easy, we can't do it when it's hard

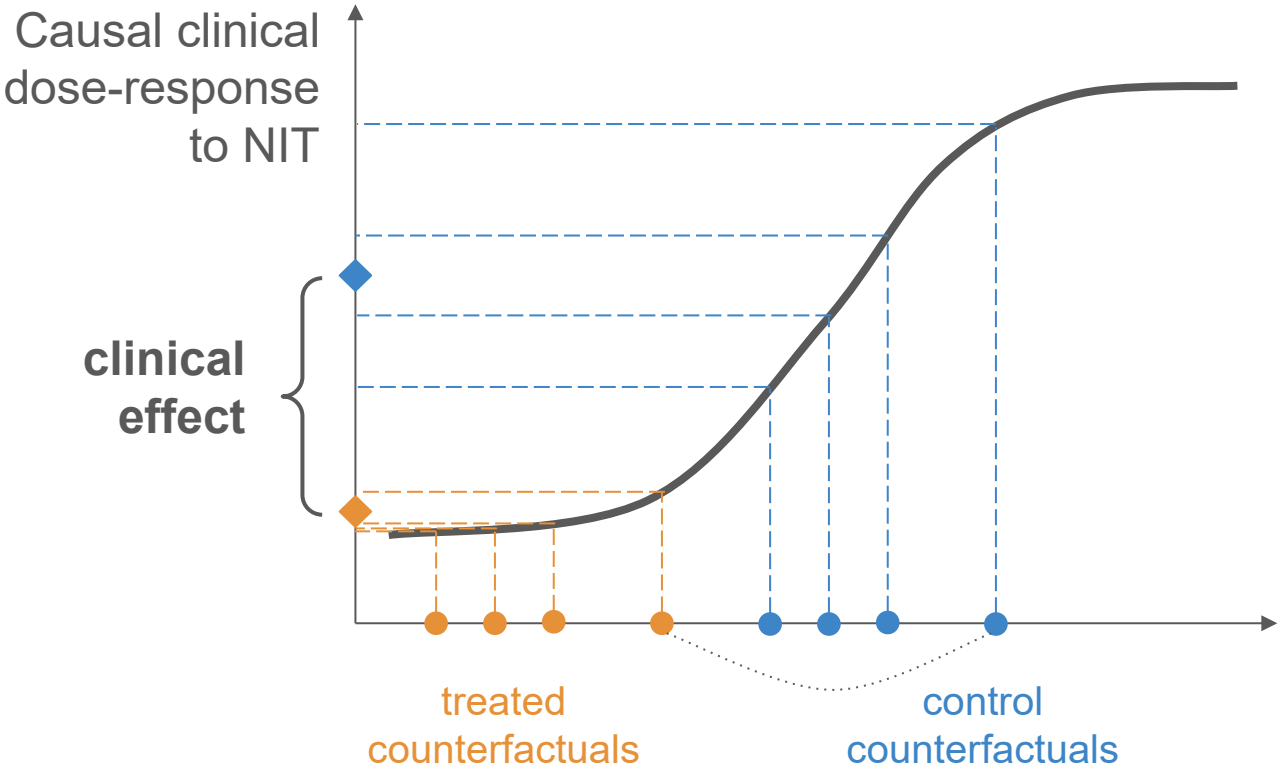
Assume...

1. **Infinite data** (i.e., entire population)
    - if we can't get it right with a huge trial, no point doing a small one either
    - power is not a concern!
2. **No clinical direct effect**
    - all of the clinical effect of the drug is mediated through the NIT
    - again best-case, essentially assumes NIT is a good proxy
3. **Single, continuous NIT**
    - not a real assumption, just makes it easy to draw pictures

... so clinical effect only depends on NIT levels

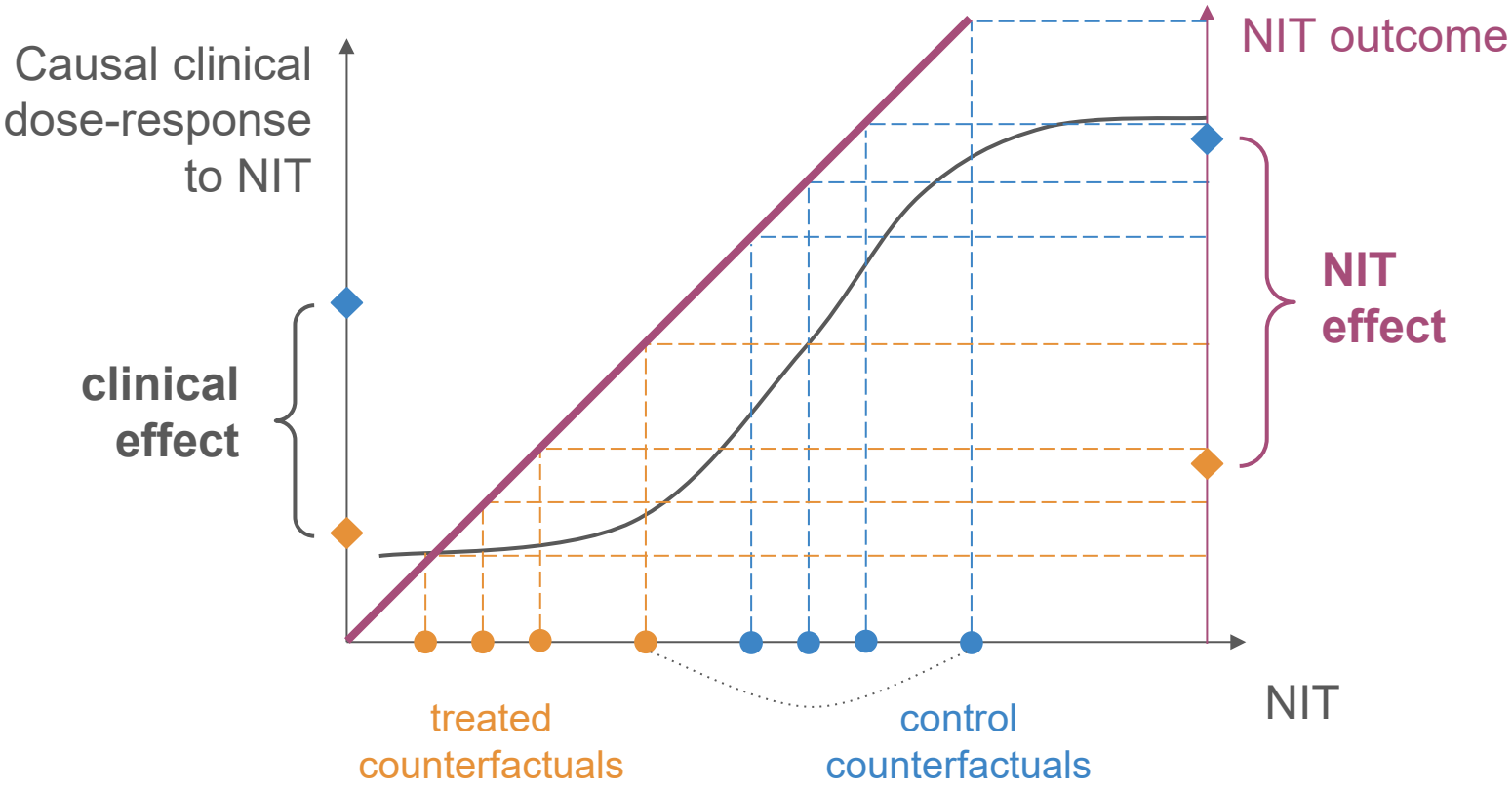


... so clinical effect only depends on NIT levels

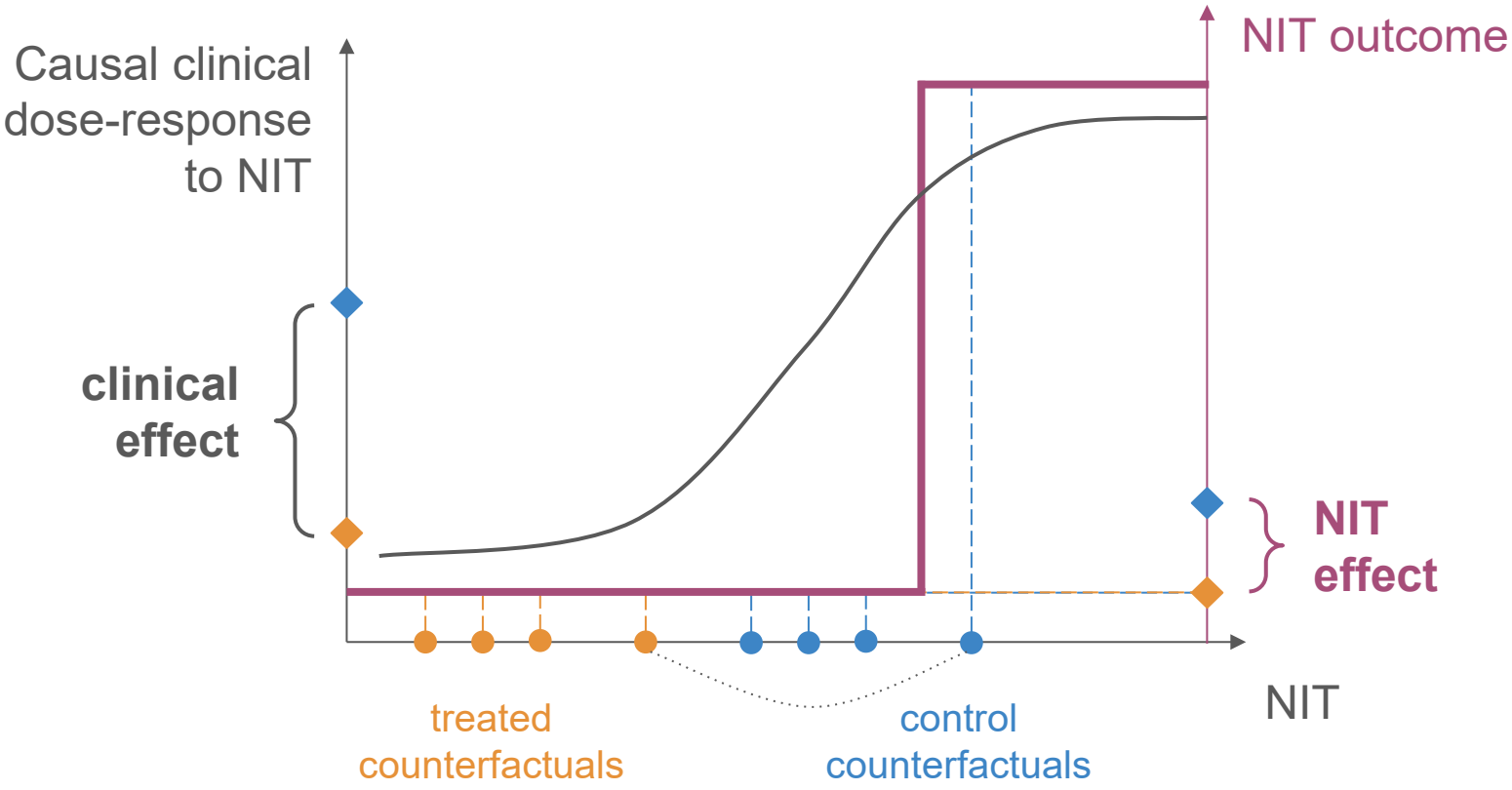


In this scenario, the NIT increases the risk in a particular region - normal variation is fine, and above some point it's all bad

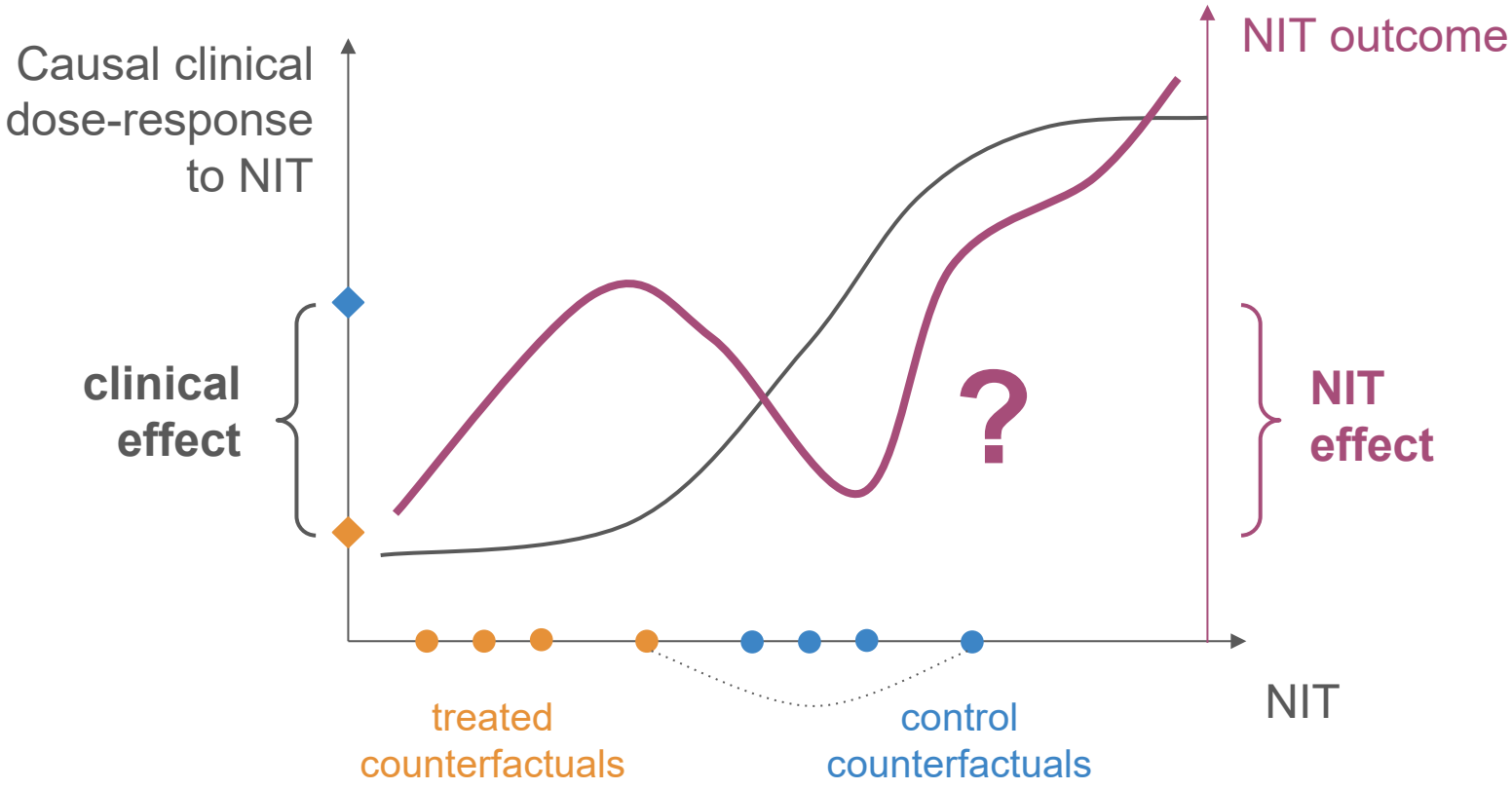
How does the effect compare to the effect on the **continuous** NIT?



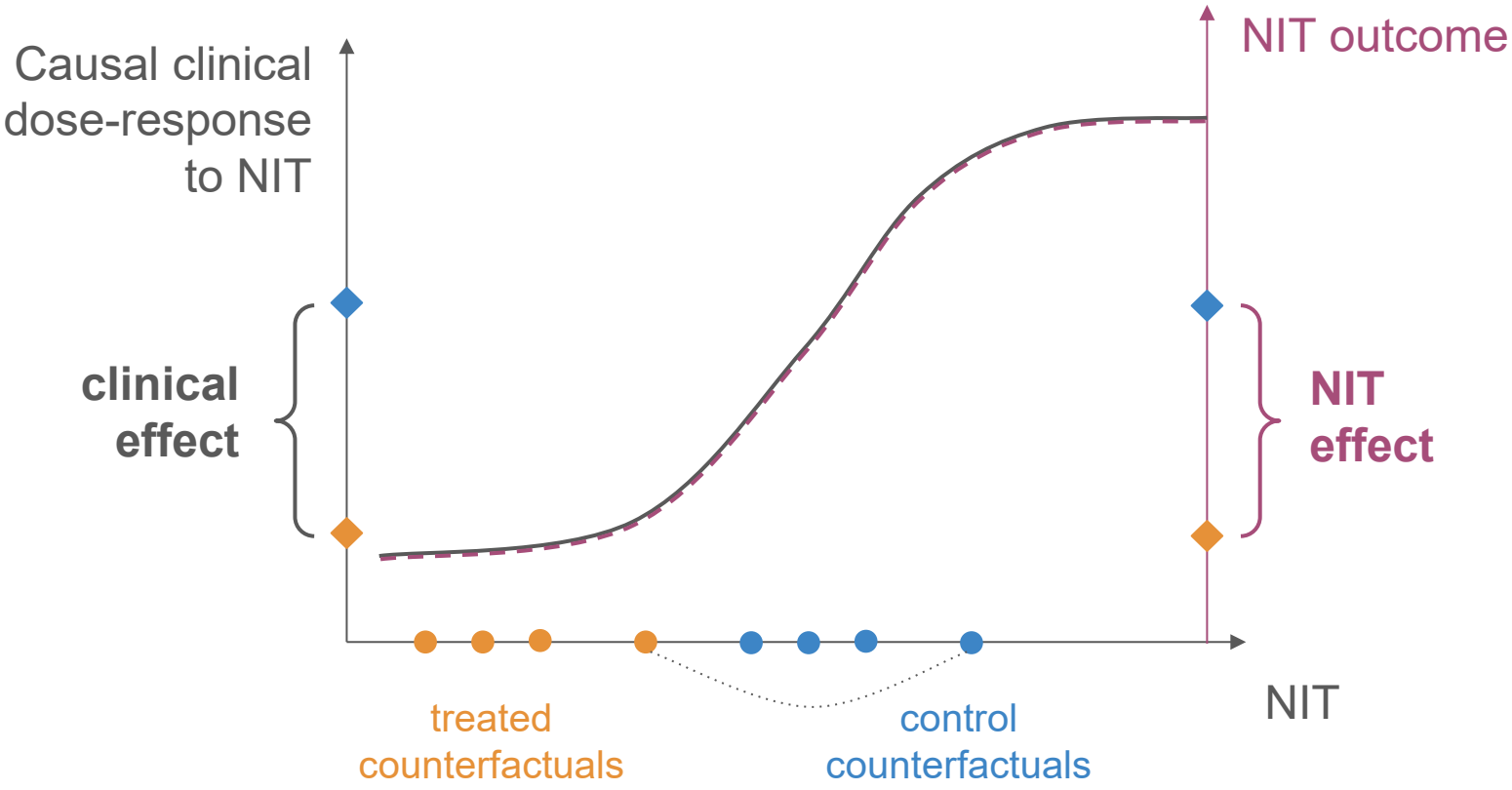
How does the effect compare to the effect on the **binarized** NIT?



What shape do we make the NIT outcome curve so the effects **match**?



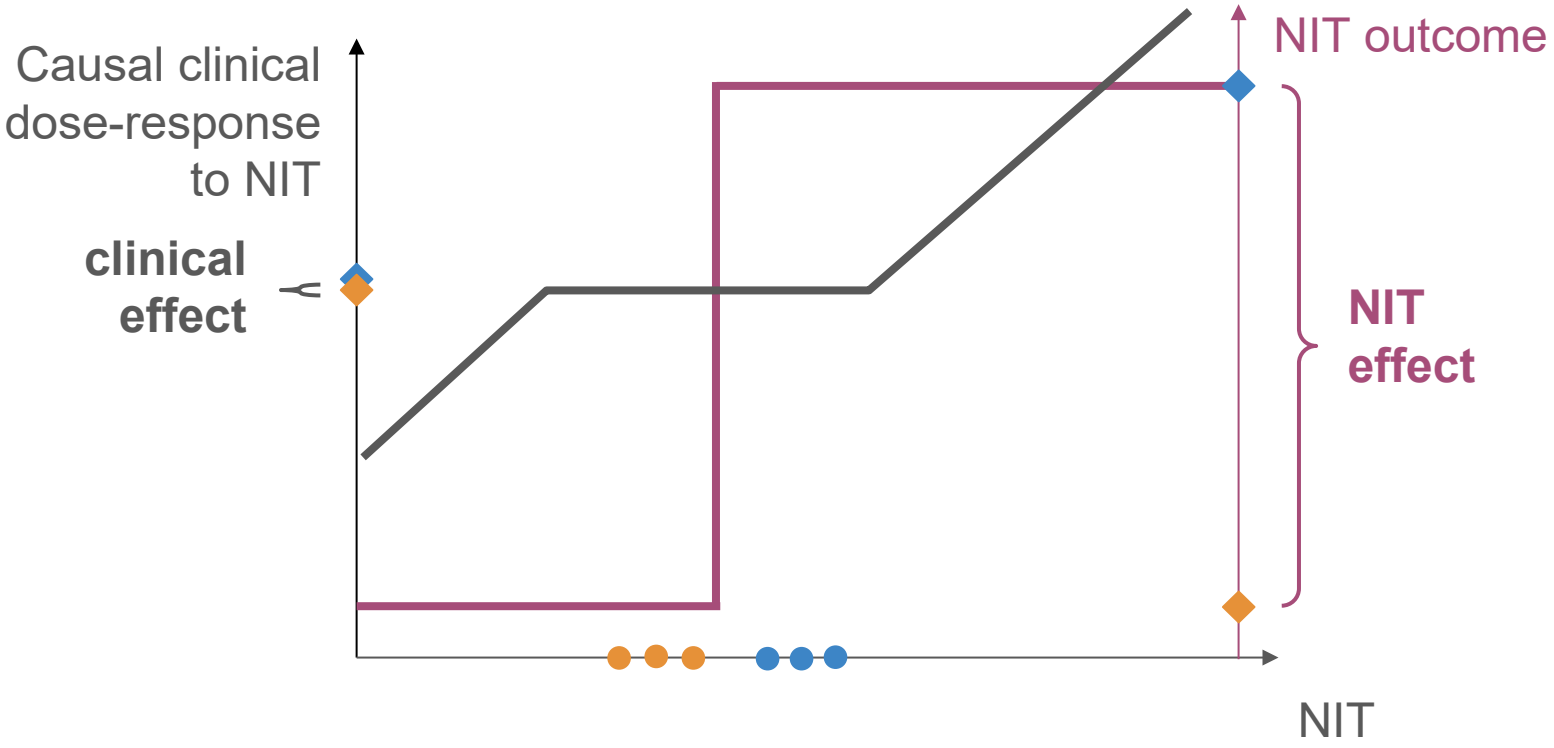
Match the dose-response curve and the effects will match too



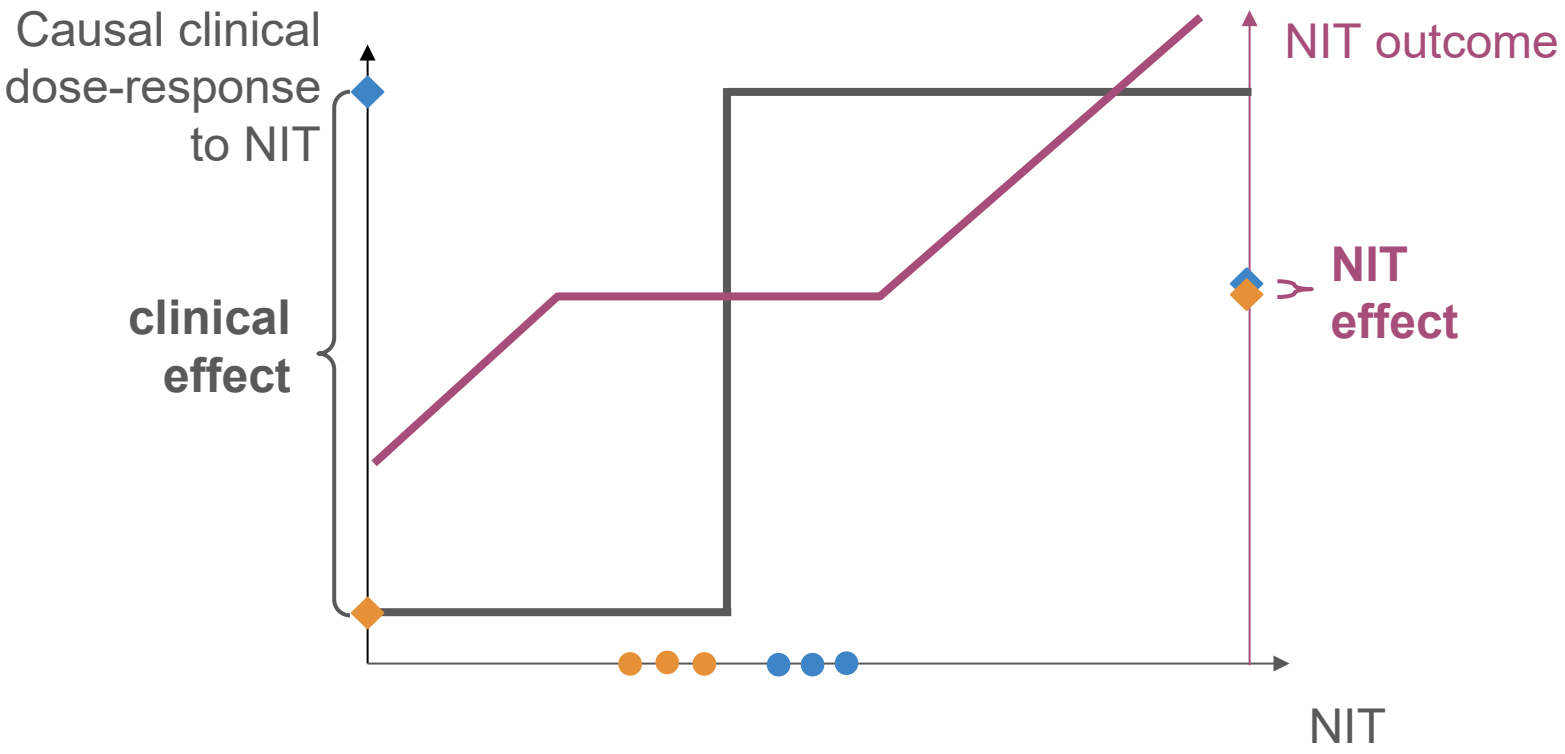
Must define power and error with respect to **clinical effect**

NIT Effect	~~NIT~~ Clinical Effect: present	~~NIT~~ Clinical Effect: not present
detected	True positive (power)	False positive (type I error)
not detected	False negative (failed trial)	True negative

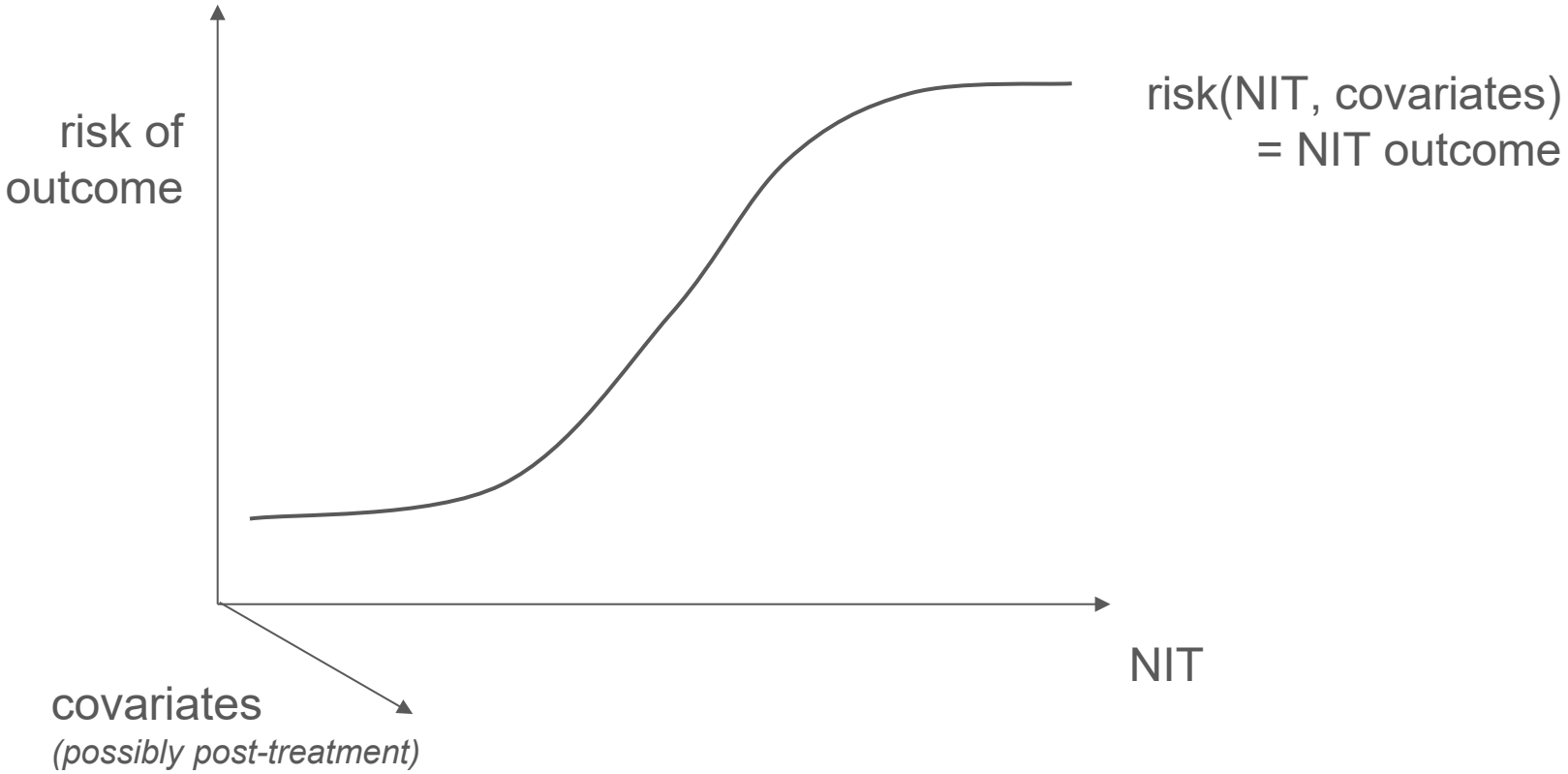
If shapes don't match, **type I error** rate can be 100%



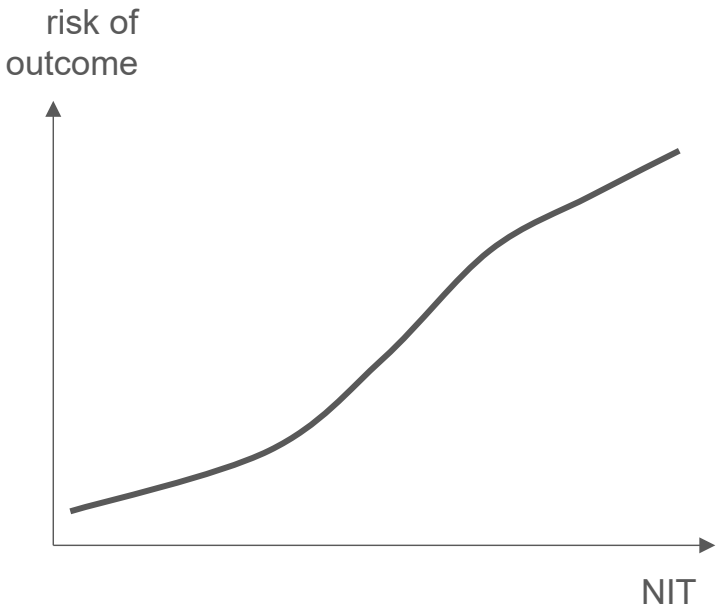
If shapes don't match, **power** can be 0%



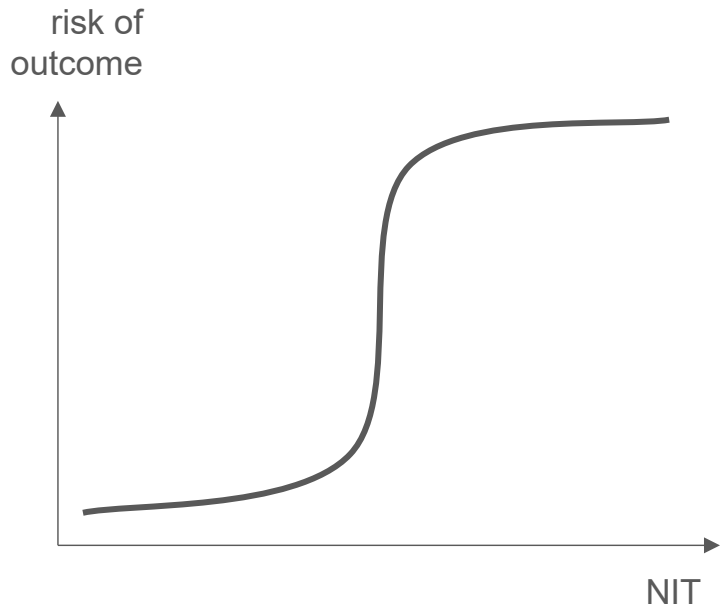
If the NIT-clinical outcome relationship is unconfounded, the dose-response curve is equal to the **observable risk curve** and this is the **ideal** NIT-based outcome



In practice, **theorize** what the shape of this curve roughly looks like and decide based on that



→ Use the continuous NIT as the outcome



→ Binarize NIT at the obvious threshold

Worry a little about assumptions...

- Could there be a **direct clinical effect** in the opposite direction?
  - Think of alternative mediators that could be affected: different biological processes
  - Nothing to do about it except get different NITs

- Is the NIT - clinical outcome relationship **sufficiently unconfounded** to allow for risk to stand in for causal dose response?
  - Again, think of factors (not necessarily affected by treatment) that could create the NIT-outcome association even if NIT does not causally affect the outcome

Takeaways:

- A good NIT outcome depends on how the **NIT relates to the clinical outcome**
  - E.g., binarize if obvious jump in risk, leave continuous if constant rate of increase, clip high/low values if little risk change outside of some region
  - Power for continuous vs. binary outcomes has nothing to do with it
    - no point if you can’t get it right with infinite data!
    - comparing apples and oranges!
- When selecting NITs in the first place, think about **minimizing indirect effects and confounding**