Interpreting data for acute respiratory infections: What can we infer from viral testing data?

Jonathan Dushoff, McMaster University

Brin MRC: Disease Dynamics and Human Behavior

Goals for post-COVID era

- Better short-term predictions for health-care demand of ARIs
- Better understanding of mortality and morbidity burdens for prioritization
- Readiness to detect and respond to the next new ARI threat

Data streams

- Virological tests
- Serological tests
- Coded physician visits, hospital admissions, deaths

Serological testing data

- What do we think if testing goes up but positivity remains level?
  - * Nothing has probably changed in the population

Virological testing data

- What do we think if testing goes up but positivity remains level?
  - * It depends!
- Maybe tests have become available in a wider geographic area
  - * no evidence for increase in incidence
- Maybe there's a huge demand for tests because of symptoms
  - * if positivity is level, this means incidence has increased

The talk that goes wrong

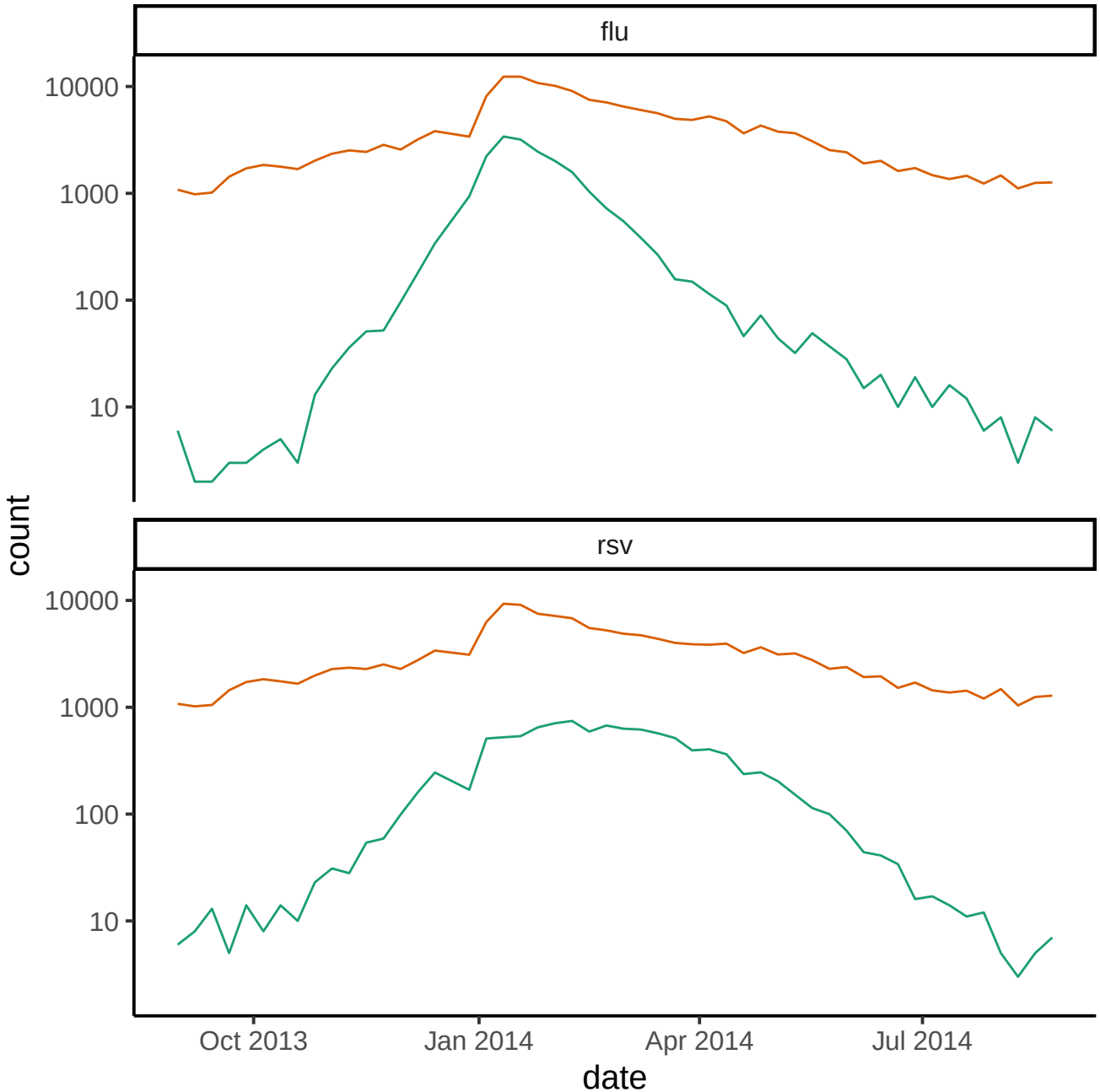
What is the best proxy for incidence?

- Observed cases?
  - Number of positive tests
  - Early in the COVID alpha wave, in some places
- Test positivity
  - Proportion of positive tests
  - Omicron wave
- Some combination

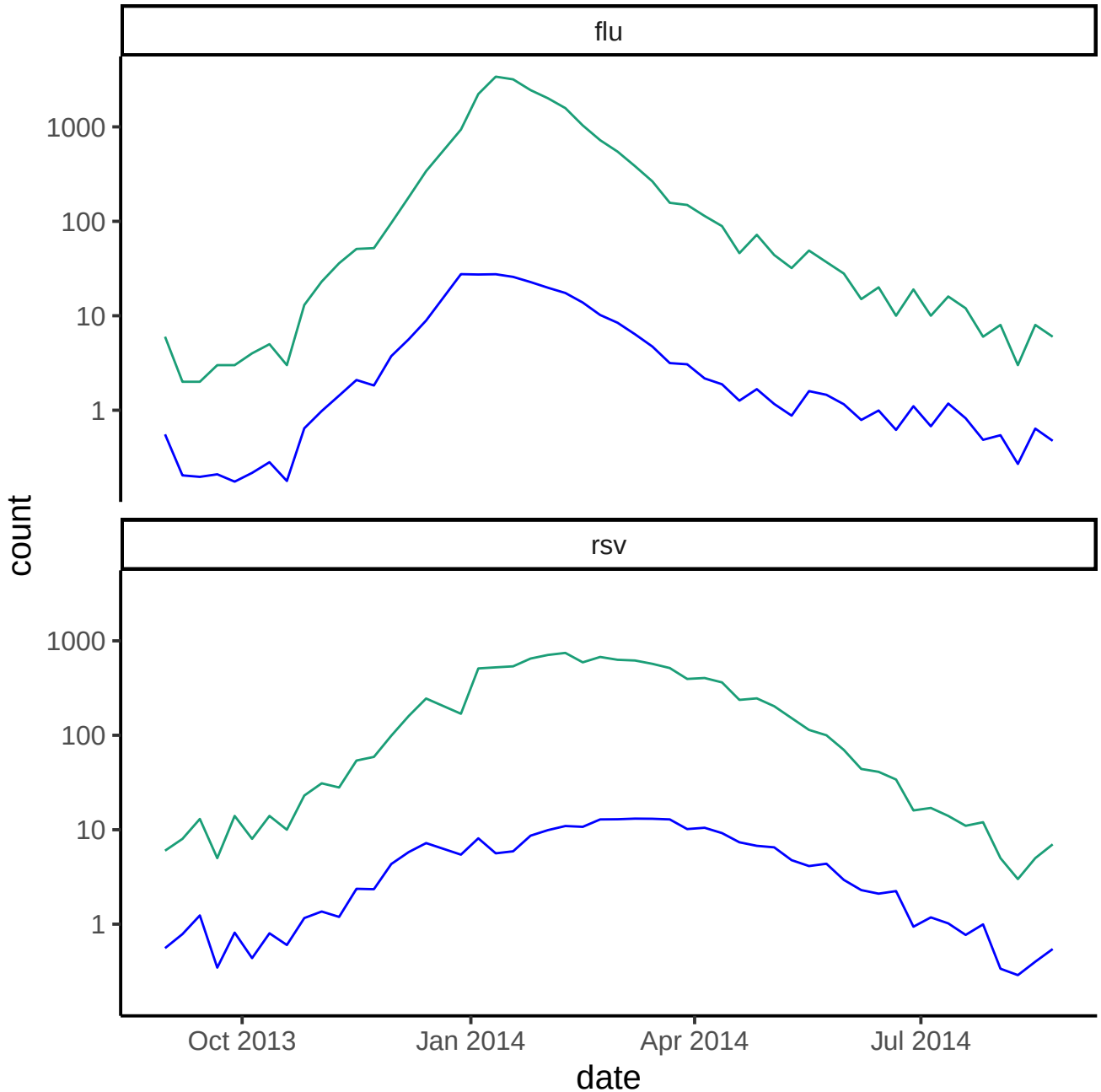
Patterns in data

- https://www.canada.ca/en/public-health/services/surveillance/respiratory-virus-detections-canada.html
- https://github.com/dajmcdon/rvdss-canada

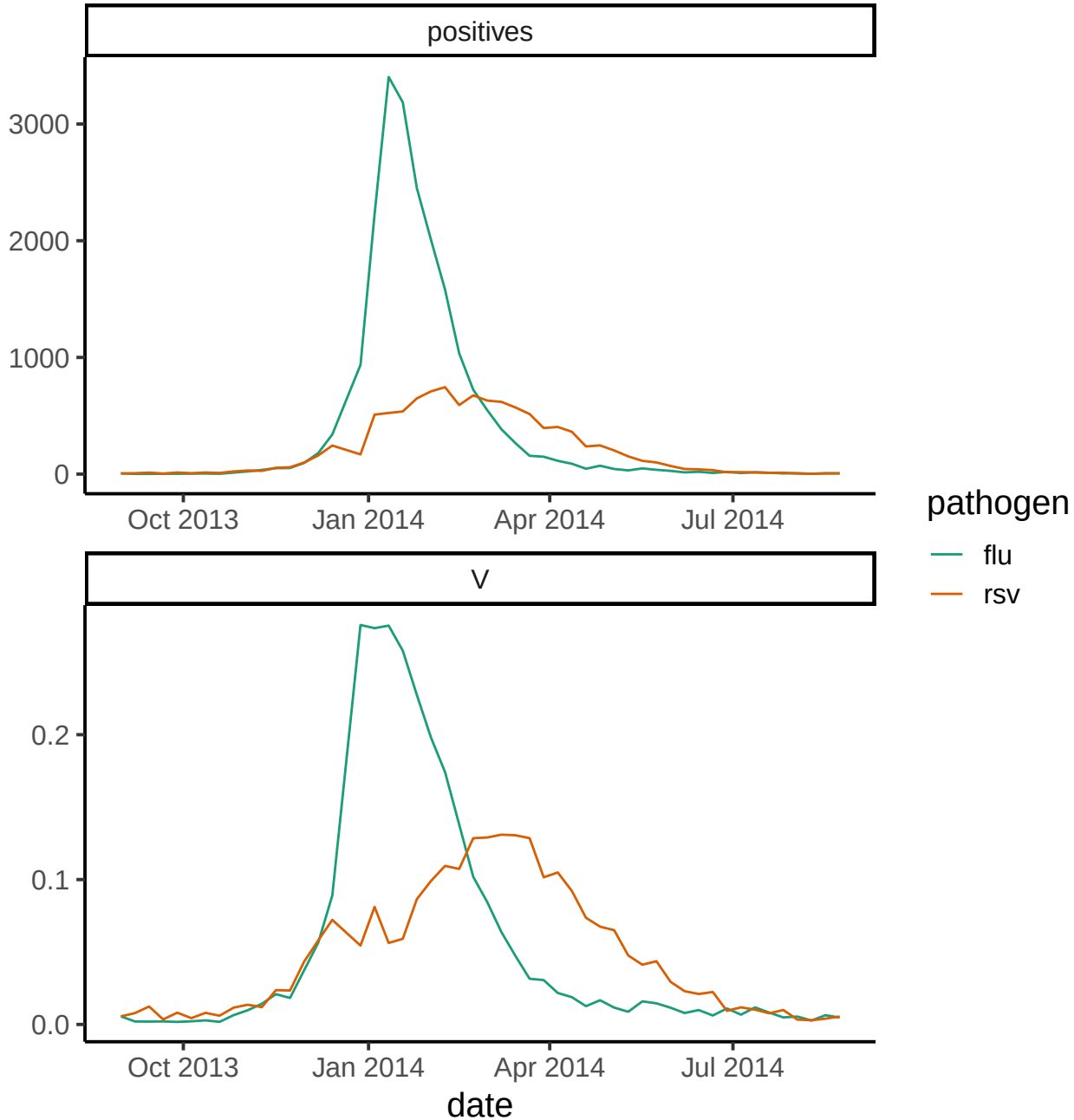
Example: 2014



Example: 2014



Example: 2014



Why did I get a flu test?

- Because I had flu-like symptoms
  - Due to flu or other virus?
- Because I had a close contact diagnosed with flu
  - Virologically or otherwise?
- Because I took a multiplex test!
- Modifiers
  - Is there a flu scare going on?
  - Is it flu season?

Interactions between pathogens

- The flu outbreak increases the number of RSV *tests*
  - Decreases positivity
  - Is it expected to increase the number of positives??
    - * Maybe I get tested because my household has flu, but I come out positive for RSV.
- Flu outbreak may also decrease the actual amount of RSV!
  - Non-specific immunity
  - Staying home

Guidance

Indicator	Description/Rationale	Major limitations
New confirmed cases per 100 000 population per week*	Direct measure of incidence. Reporting delays can be accounted for to improve identification of projected surges (33). Monitoring the percent weekly change in new cases is particularly important to anticipate surges in transmission.	Heavily influenced by surveillance system performance, testing policy and laboratory capacity and reporting policies. At low levels and in small geographical regions, can be sensitive to minor fluctuations in case counts, particularly due to batch reporting. Most countries have now drastically reduced testing and reporting of incident cases, but sentinel surveillance may still provide robust estimates of transmission trends (34). Percent changes may be unstable in situations where there are very few cases.
Test positivity rate per week*	Allows understanding of transmission intensity even in the absence of universal testing/reporting. It may capture a typical case better than syndromic surveillance. Particularly useful for monitoring trends. This indicator can be monitored at sentinel sites or from any facility.	Heavily influenced by testing strategy (i.e., who gets tested) and capacity and changes therein. May be artificially reduced during co-circulation of other pathogens with overlapping symptoms (35)
New COVID-19 hospitalizations per 100 000 population per week*	A predictable (in the absence of shifts in circulating variants) subset of all incident cases requiring hospitalization. Thus, this is an indirect indicator of incidence. Unlikely to be subject to surveillance policy changes/differences.	May be influenced by hospitalization policy, e.g., if even mild cases are hospitalized for isolation purposes. Delayed measure of incidence. May be influenced by changes in severity of variants, even in setting of stable transmission intensity.
New ILI or ARI cases (per 100 000 population or per fixed sentinel site catchment) per week*	May be helpful where COVID-19-specific surveillance is not robust. Allows comparison with historical ILI/ARI baseline data. Ideally a subset or all should be tested for SARS-CoV-2 and other pathogens to understand what is driving the ILI or ARI rates.	Indirect measure of COVID-19 incidence; need to understand relative levels of other respiratory pathogens (e.g., influenza, RSV).
Product of weekly ILI or ARI rates and weekly percentage positivity for SARS-CoV-2*	Yields estimate of actual COVID-19 incidence. May be helpful where COVID-19-specific surveillance is not robust	Indirect measure of COVID-19 incidence. Requires ILI/ARI rates and SARS-CoV-2 positivity to come from same catchment population.

https://www.who.int/publications/i/item/who-2019-ncov-adjusting-ph-measures-2023.1

Incidence

- Incidence is not an end in itself
  - Incidence $\times$ severity to predict burden
  - Incidence $\times$ immunogenicity to predict short-term protection, dynamics
  - Incidence * immunity kernel to predict longer-term protection

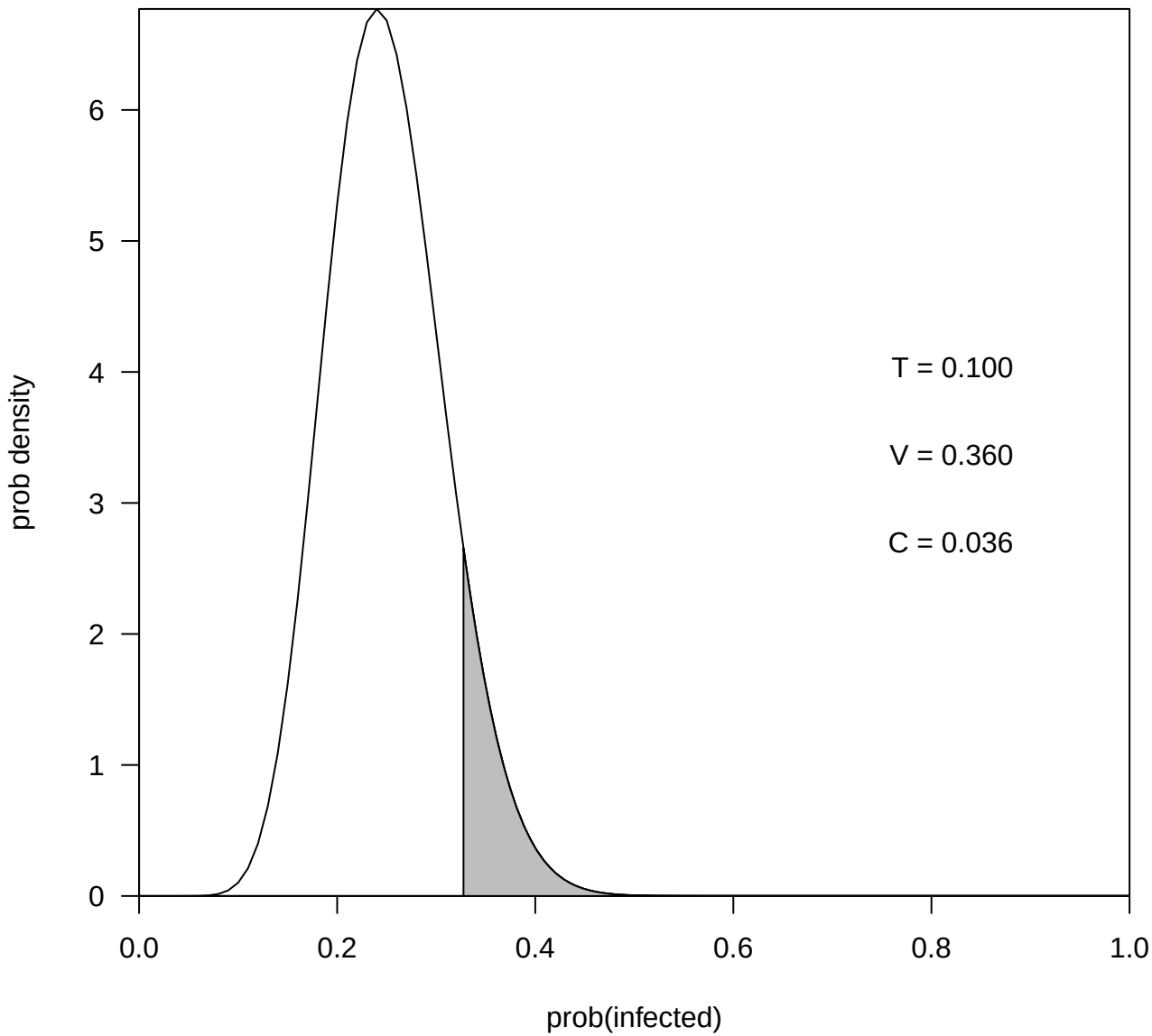
Some modeling approaches

- Direct estimation
  - Infer incidence from positivity and cases each week
- Phenomenological fitting
  - Make use of smooth latent curves through time
- Mechanistic fitting
  - Make use of dynamical models underlying latent variables
  - SIR, information flow, policy changes

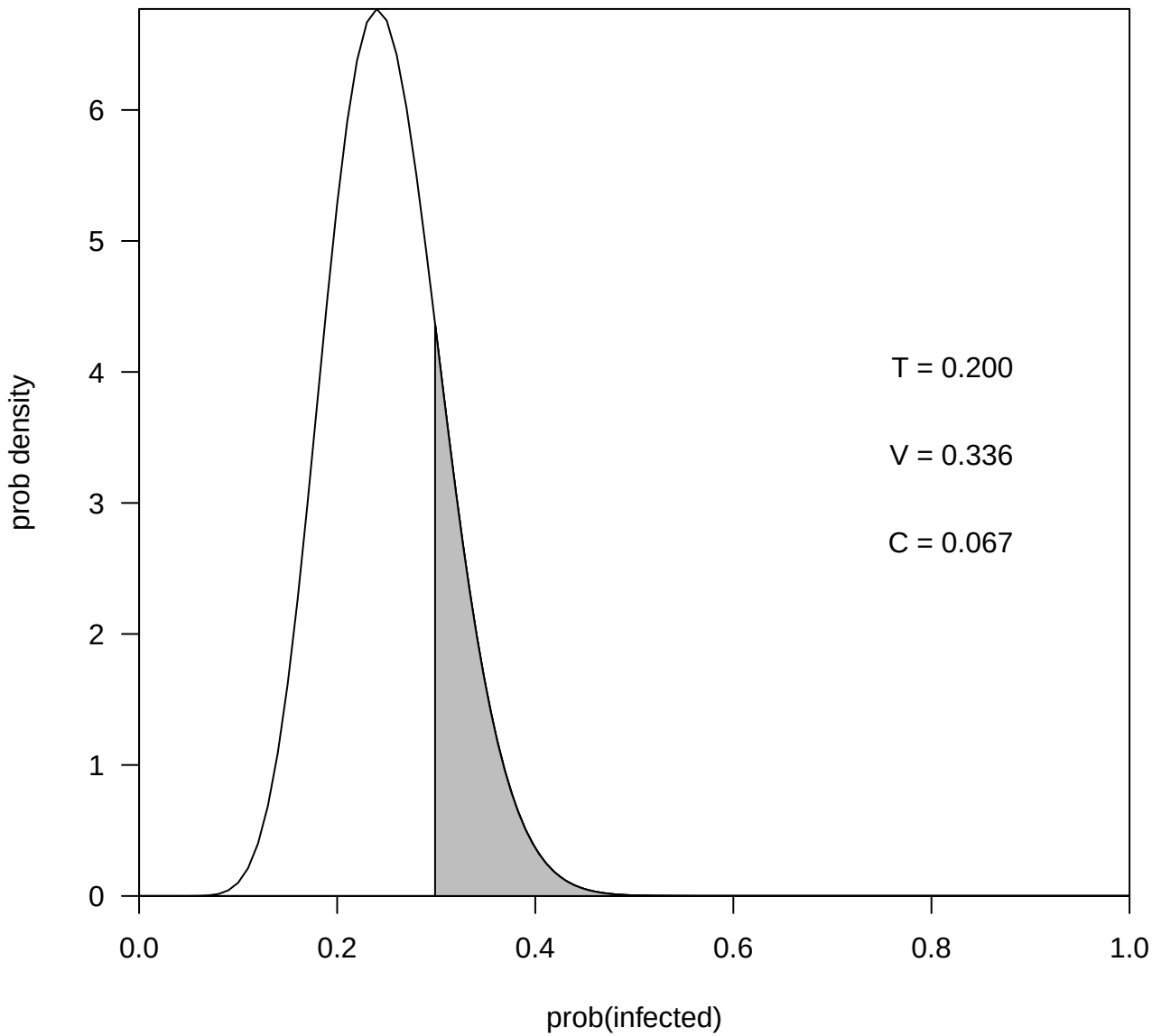
Top-down approach

- Inspired by early COVID; limited tests, active discussion of how to use them
- Imagine risk prioritization; people in each risk class have a certain probability of testing positive
  - The *mean* of this distribution corresponds to prevalence in the population
  - *Variation* corresponds to the information gained by risk prioritization

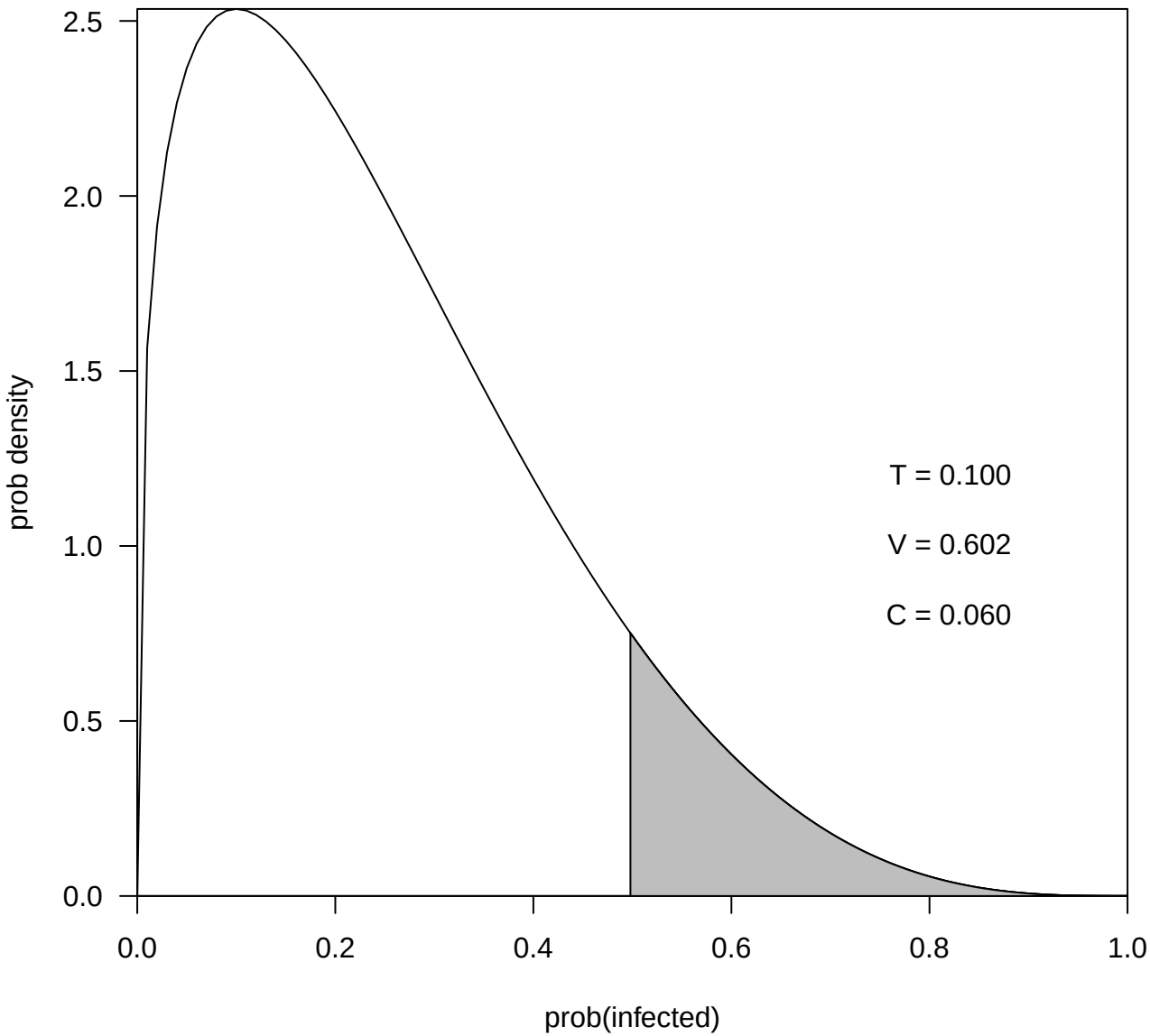
Beta-distributed risk



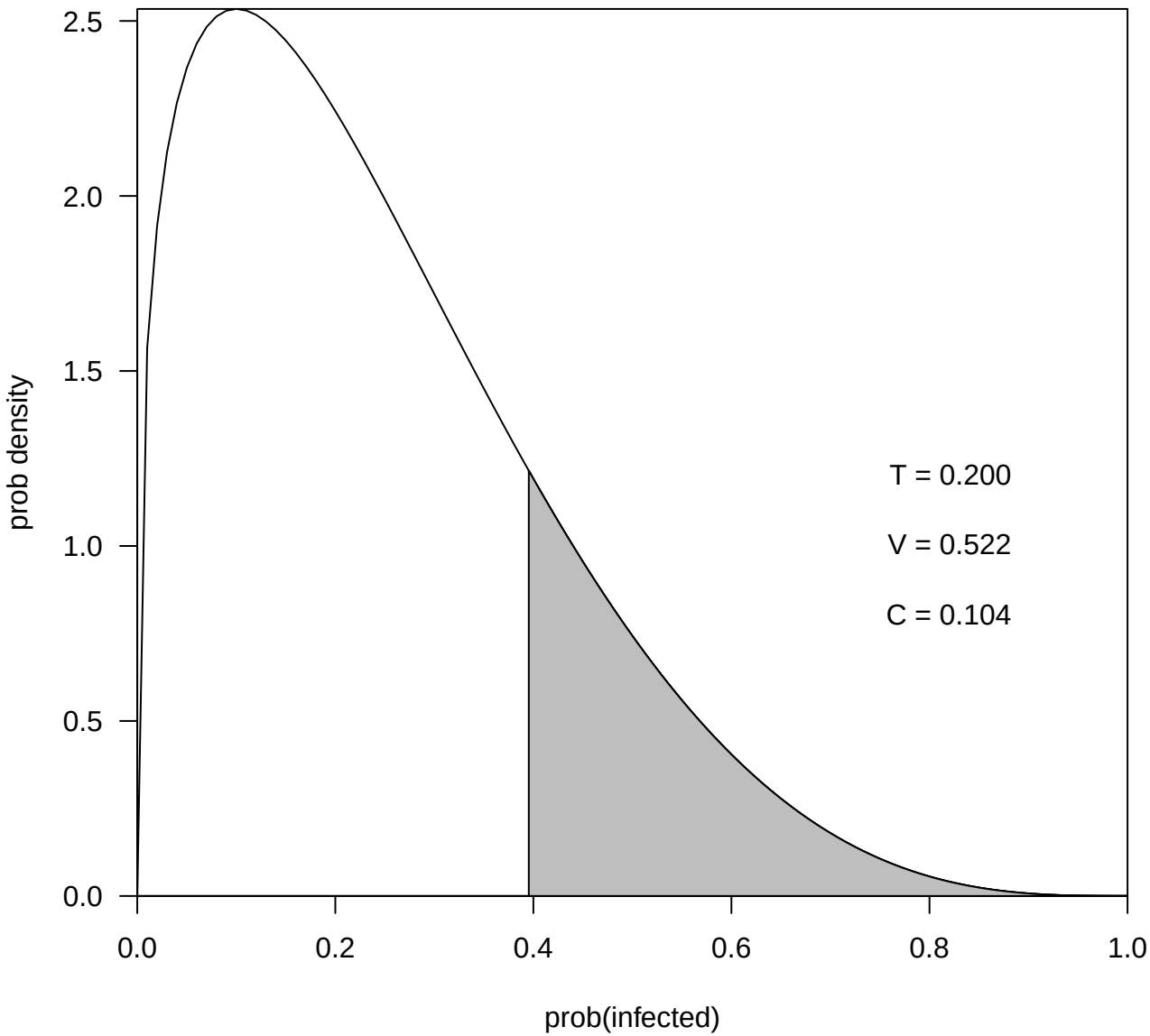
Beta-distributed risk



Beta-distributed risk



Beta-distributed risk



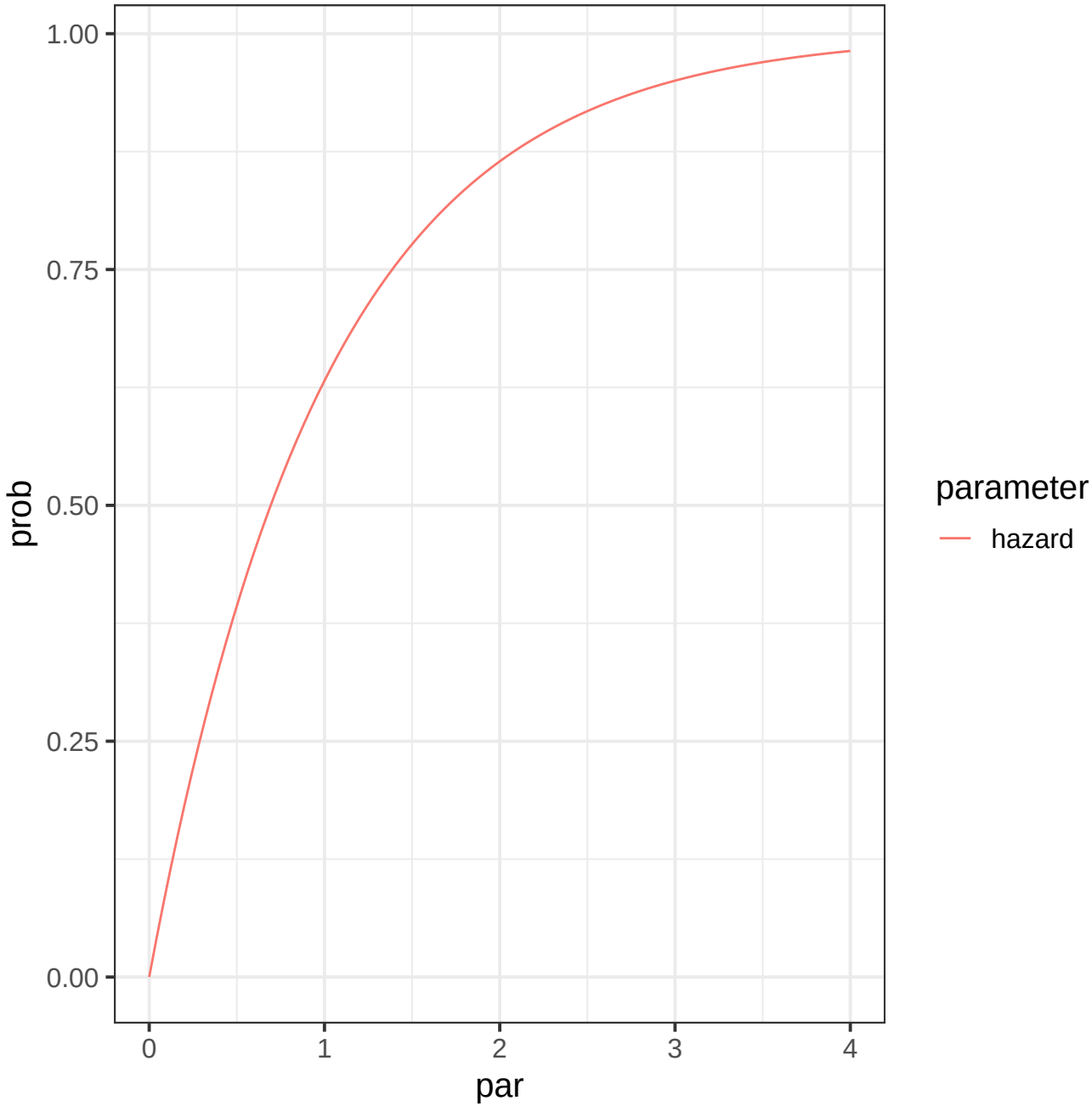
Bottom-up approach

- Model the probability of people seeking care for various reasons
- Corresponds better to seasonal epidemics
  - Policy shifts could be modeled as parameter changes

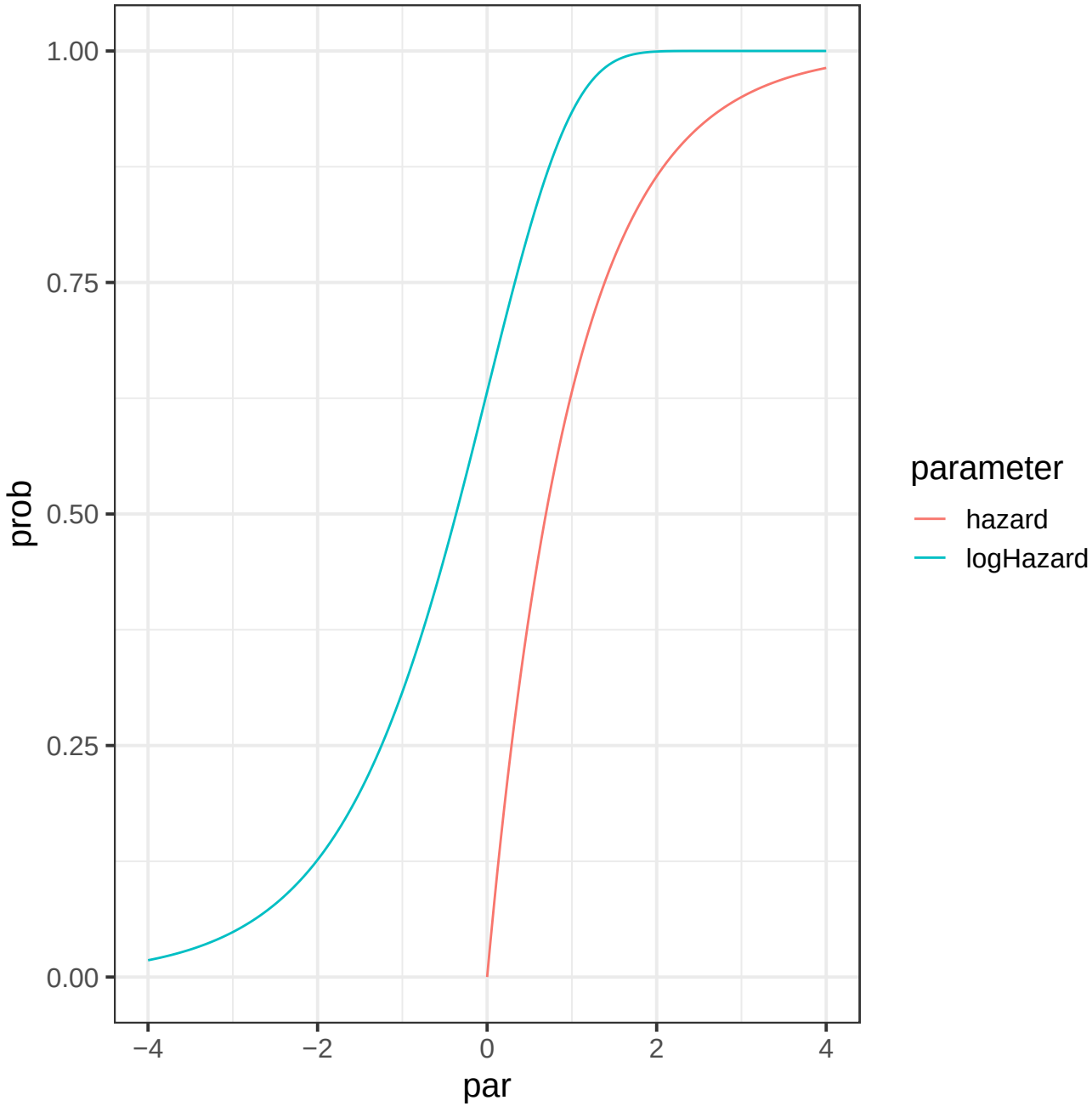
Hazard approach

- $P_{\text{missingConference}} = 1 - (1 - P_{\text{forgetting}})(1 - P_{\text{missingAirplane}})(1 - P_{\text{gettingLost}})$
- Define: $H = -\log(1 - P)$
- $H_{\text{event}} = \sum_{\text{components}} H_{\text{c}}$
- e.g., $H_{\text{test}} = H_{\text{focalSymptoms}} + H_{\text{focalContact}} + H_{\text{nonfocalSymptoms}} + H_{\text{nonfocalContact}}$

Hazard response



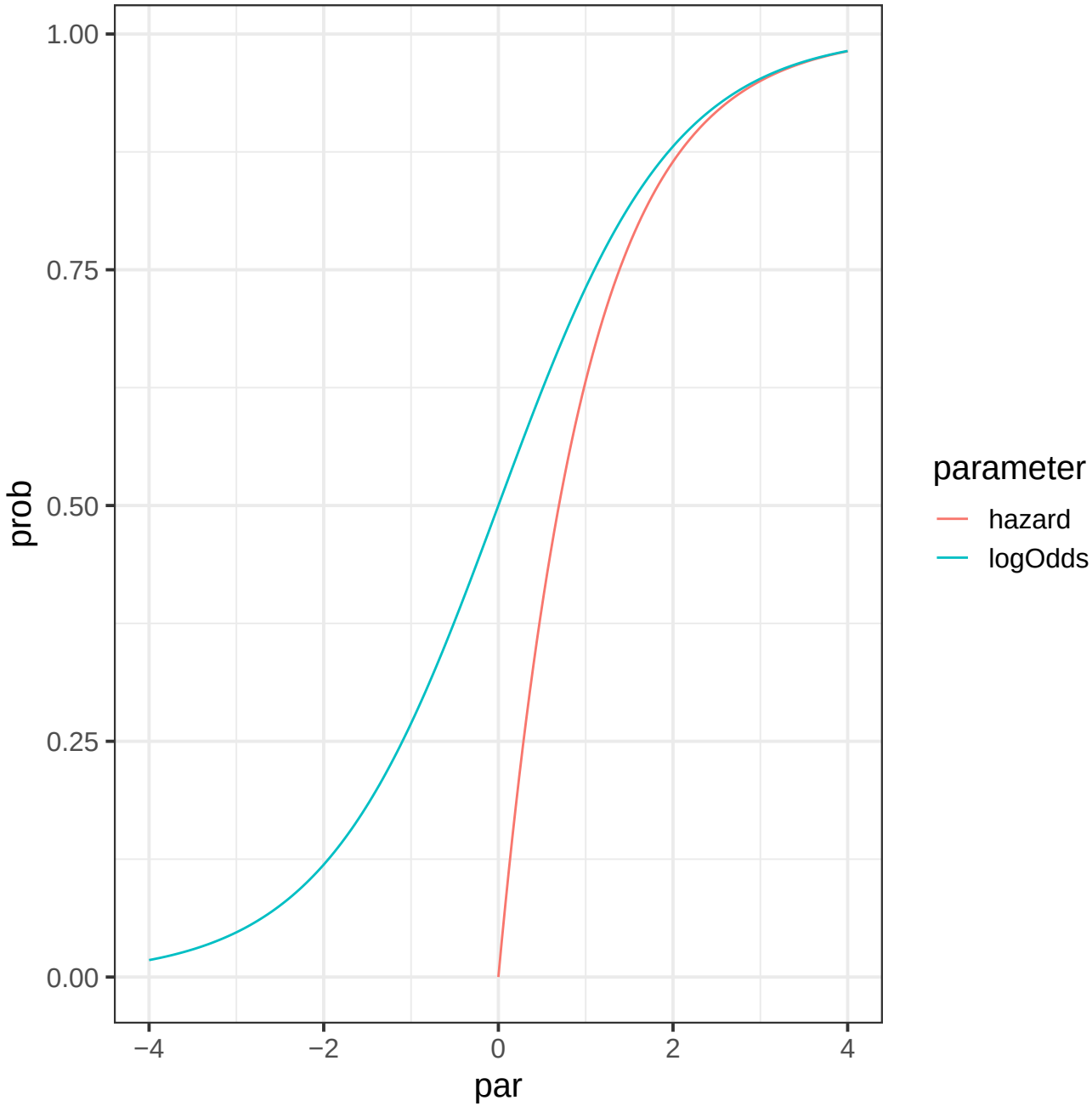
Hazard response



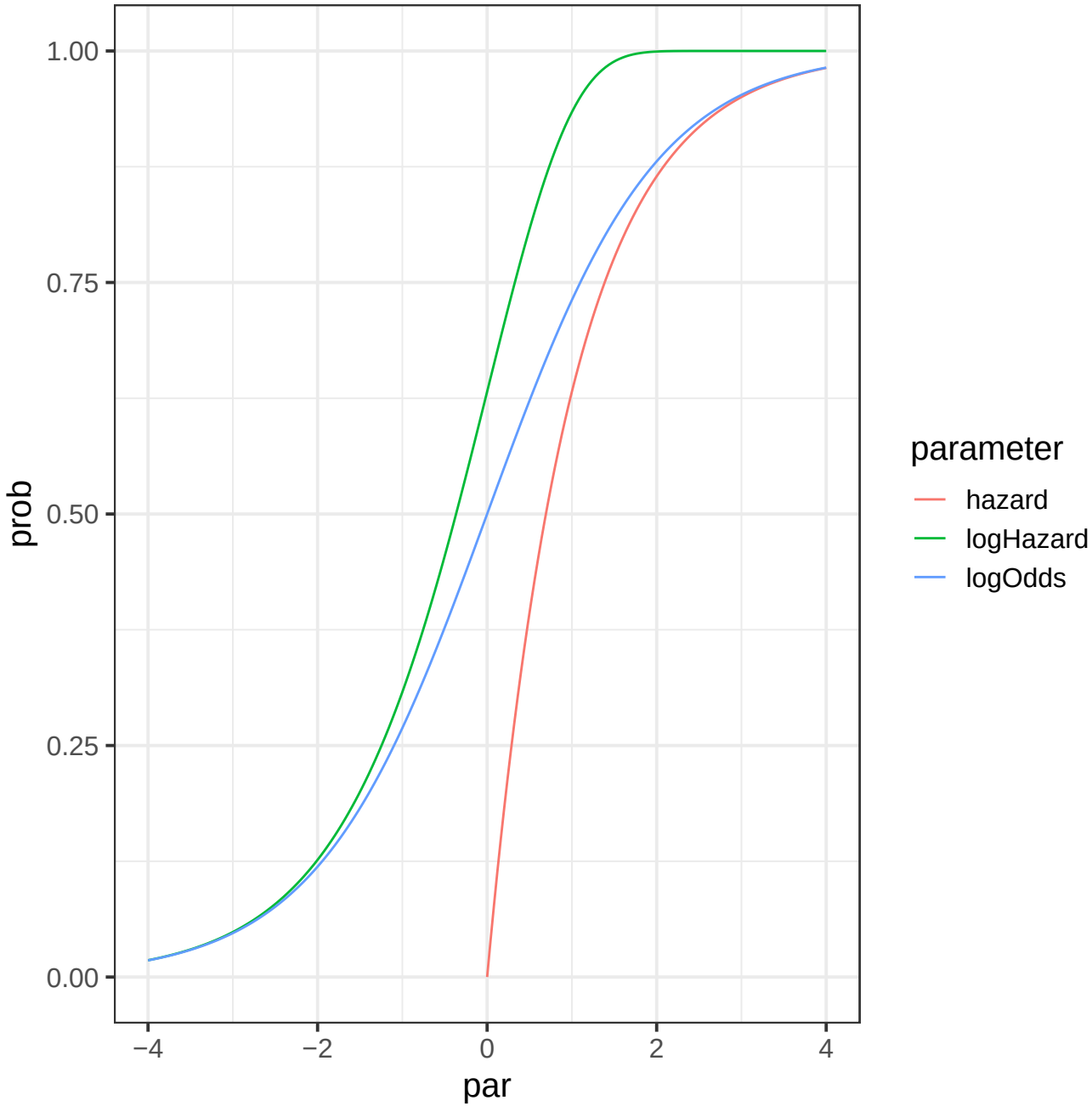
Log odds approach

- The odds corresponding a probability P is $\theta = P/(1-P)$
  - $\ell = \log(P/(1-P))$
- Principled justification for adding on the log scale in many cases
  - But not quite in this one
- e.g., $\ell_{\text{posterior}} = \ell_{\text{prior}} + \text{BayesFactor}$
  - Probability positive given positive test
  - Prop of positives among test seekers

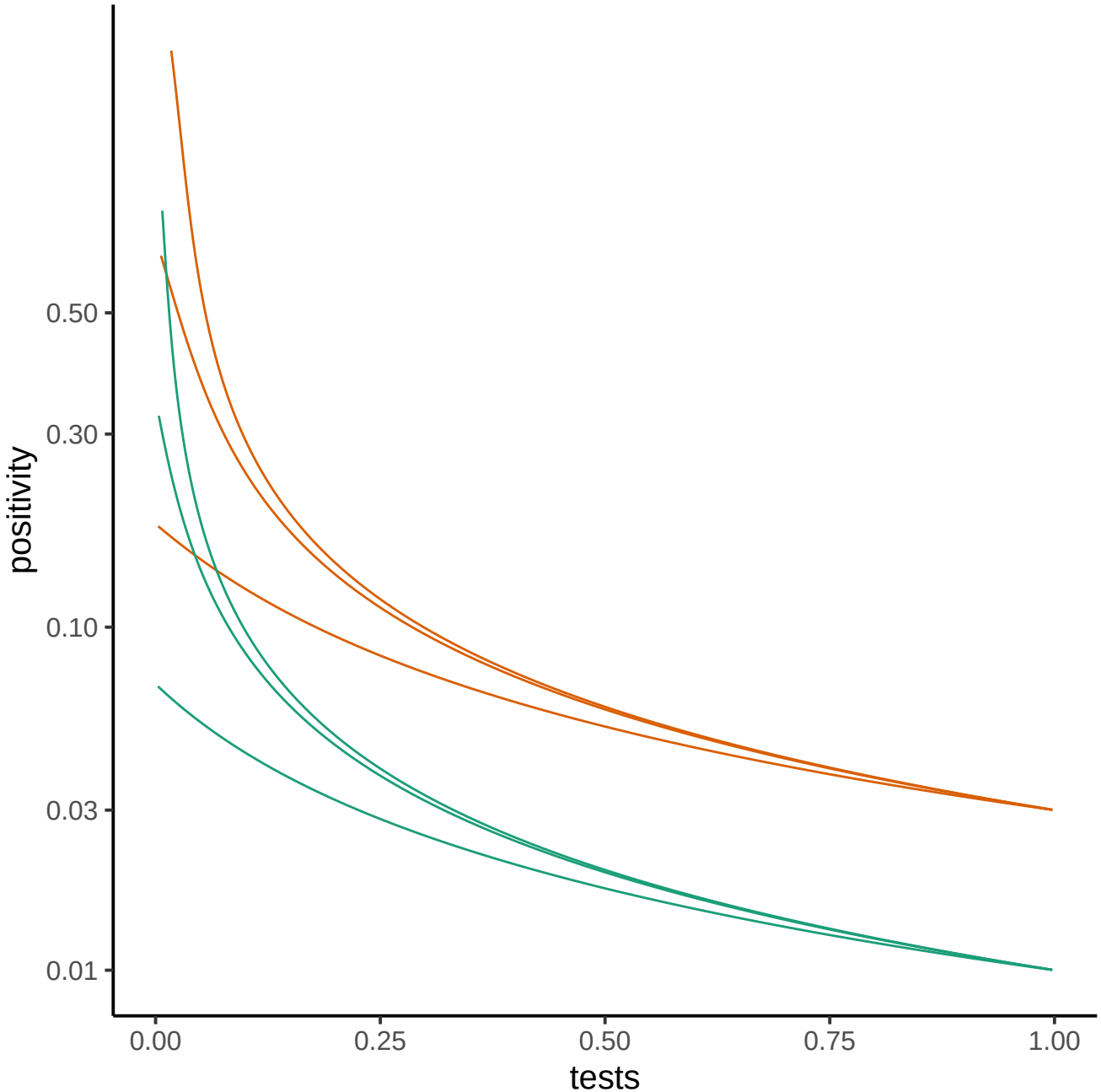
Log-odds response



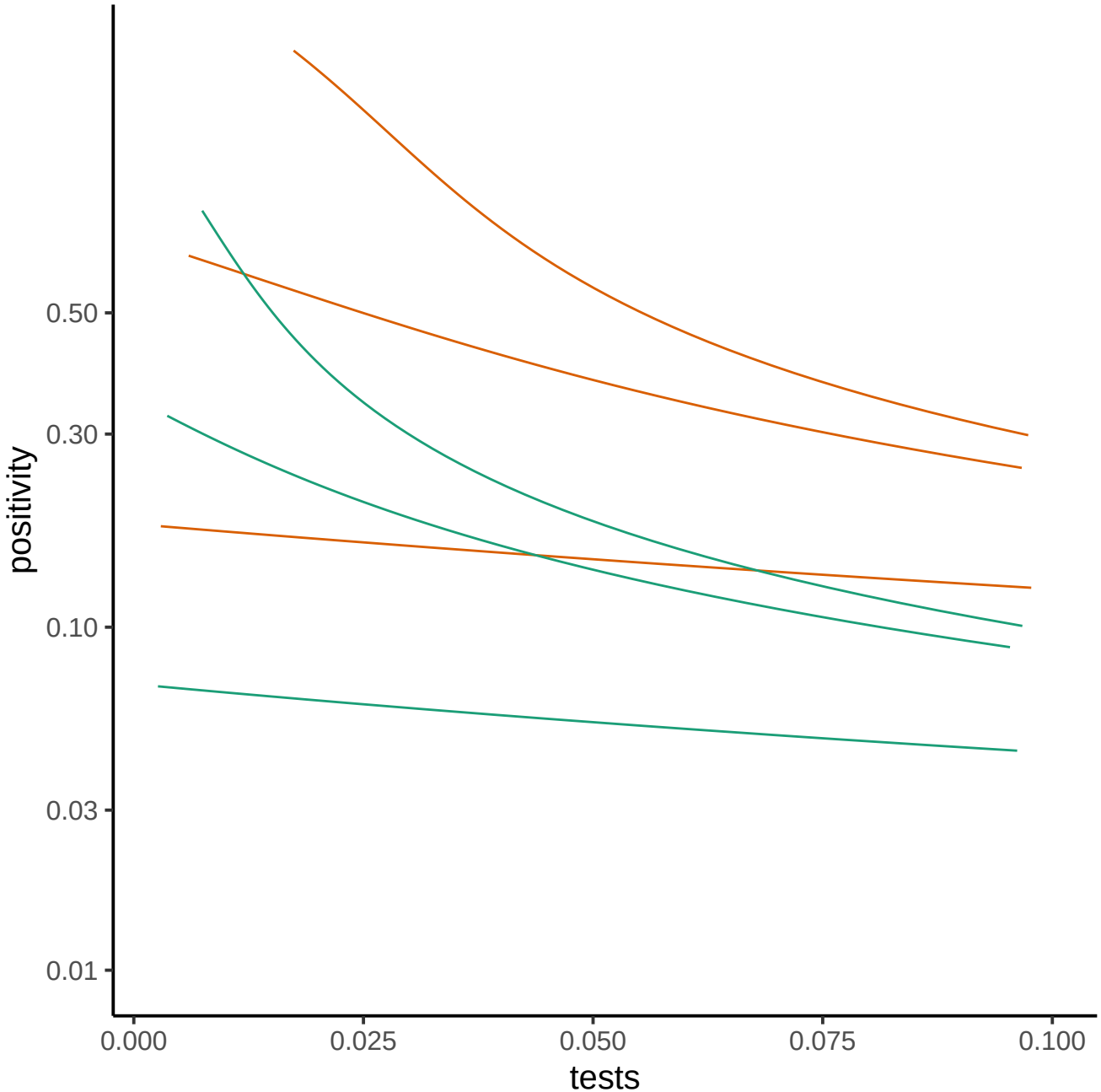
Log-odds response



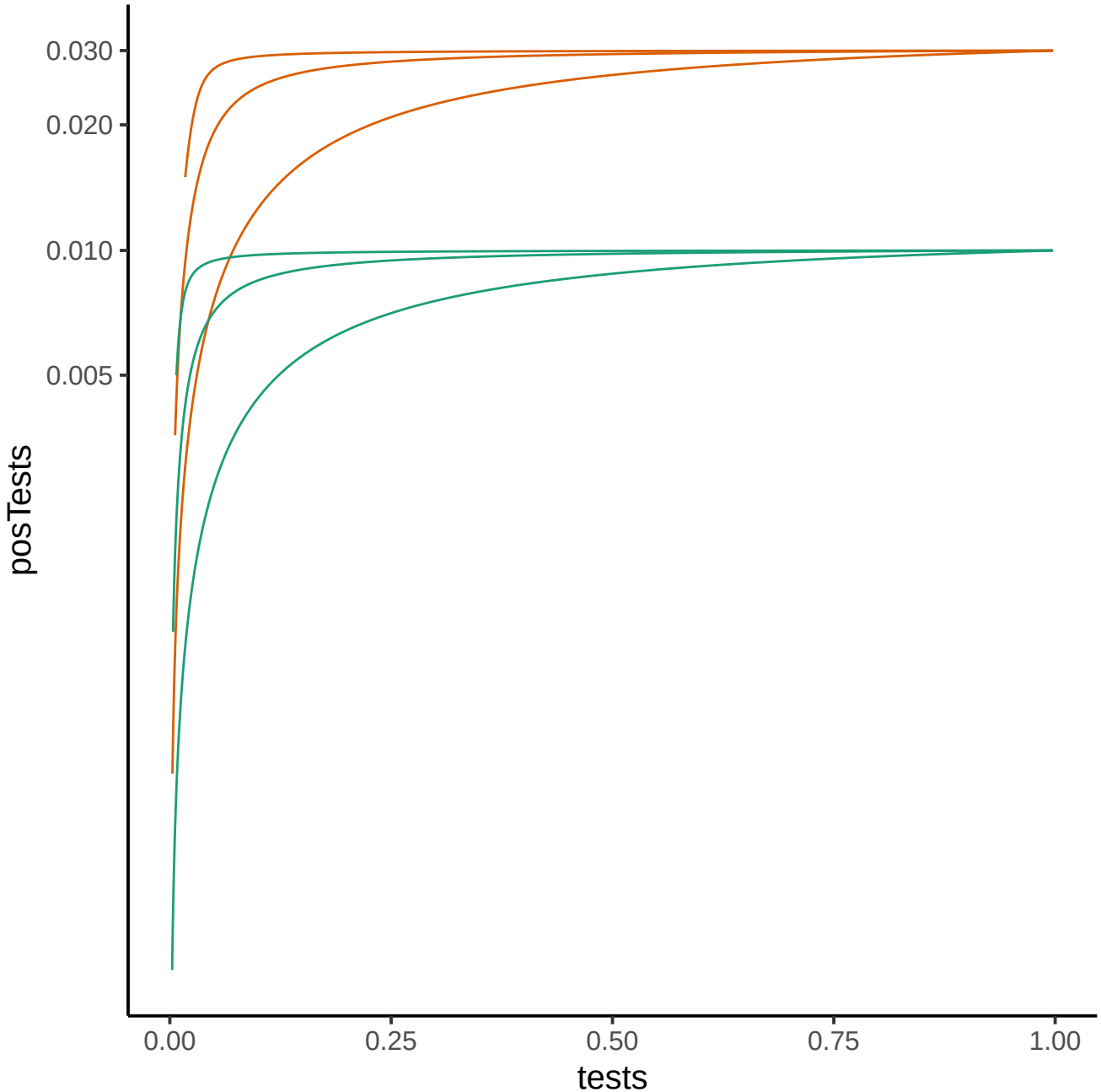
Constant odds ratio



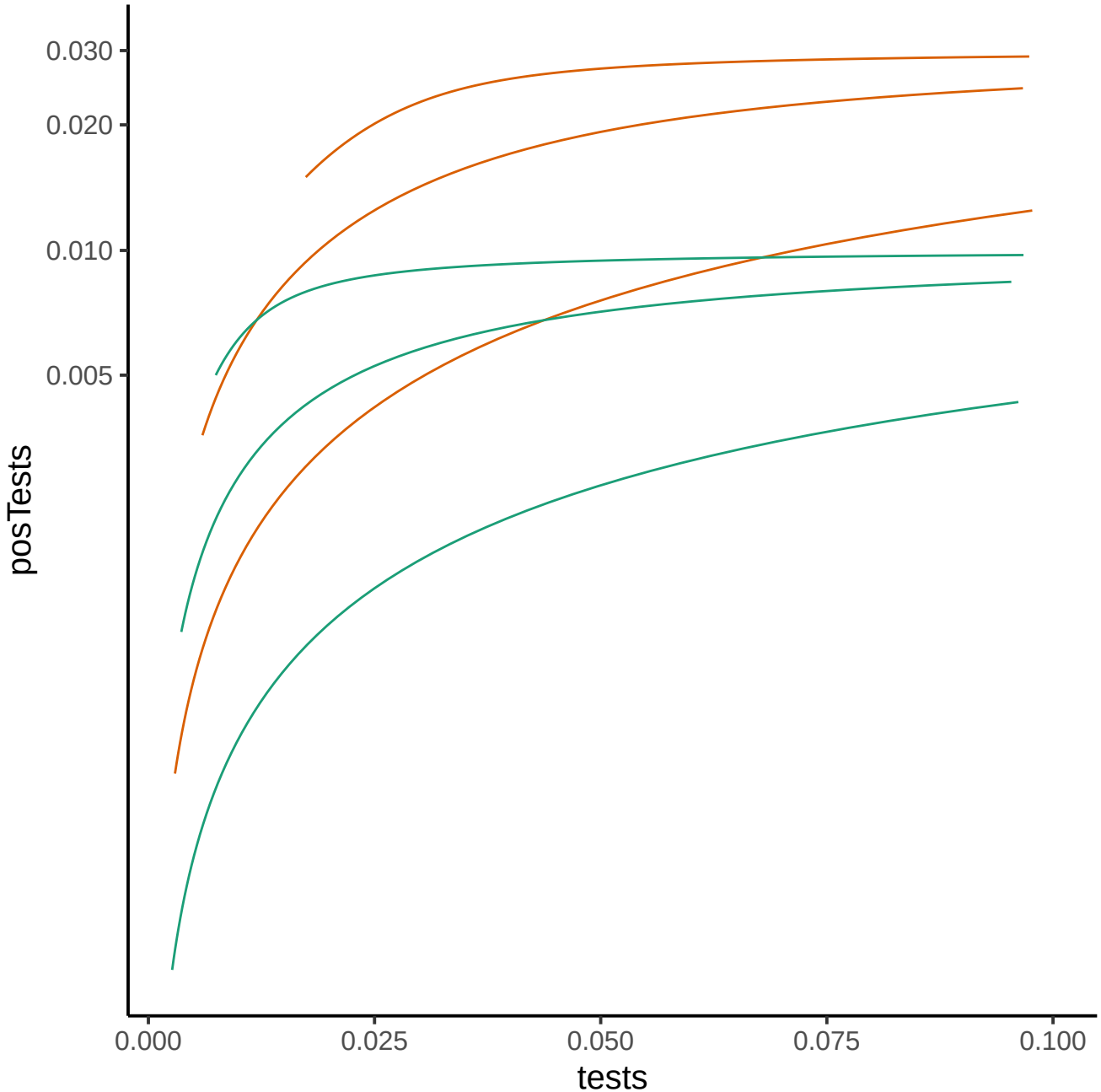
Constant odds ratio



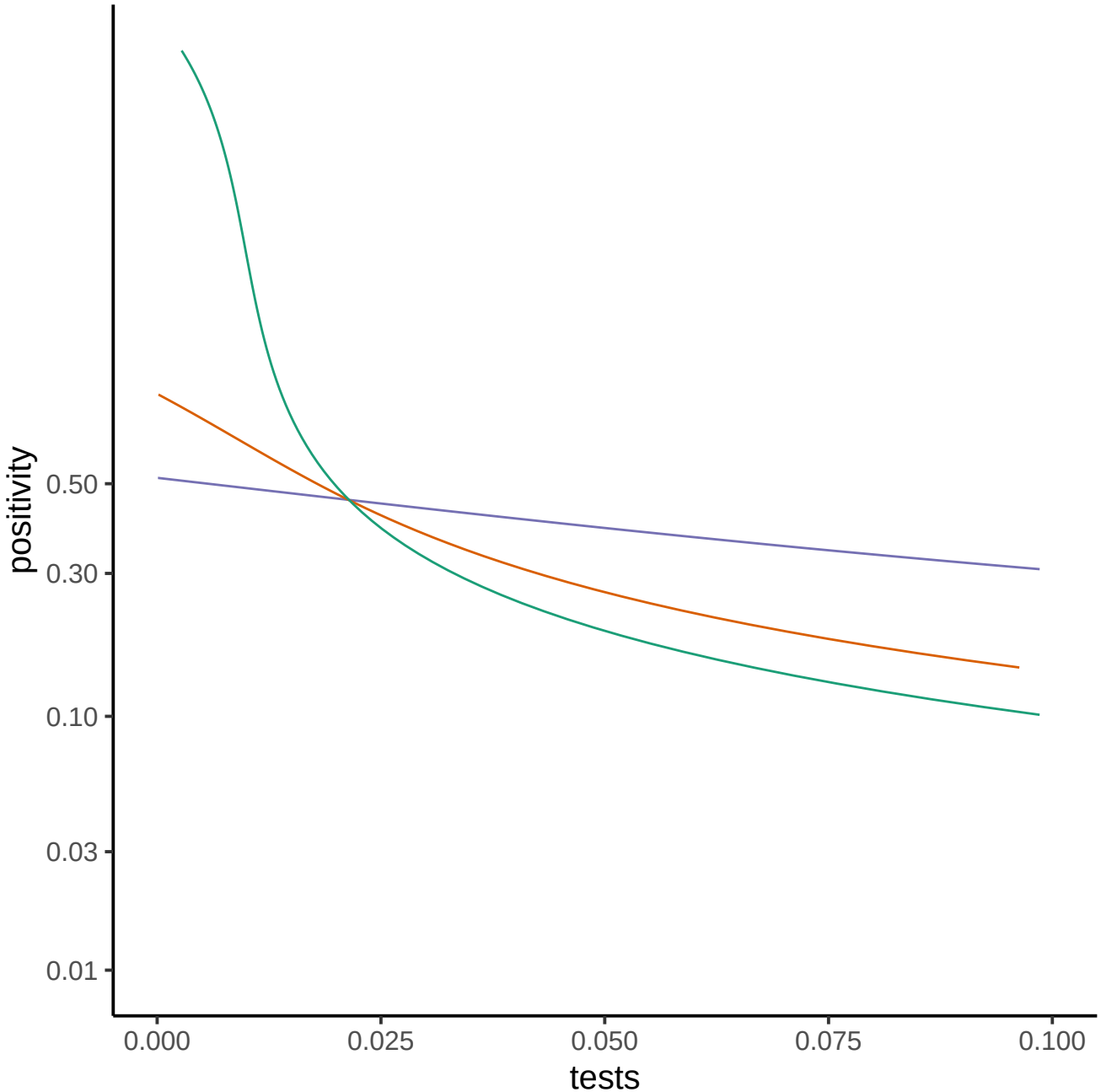
Constant odds ratio



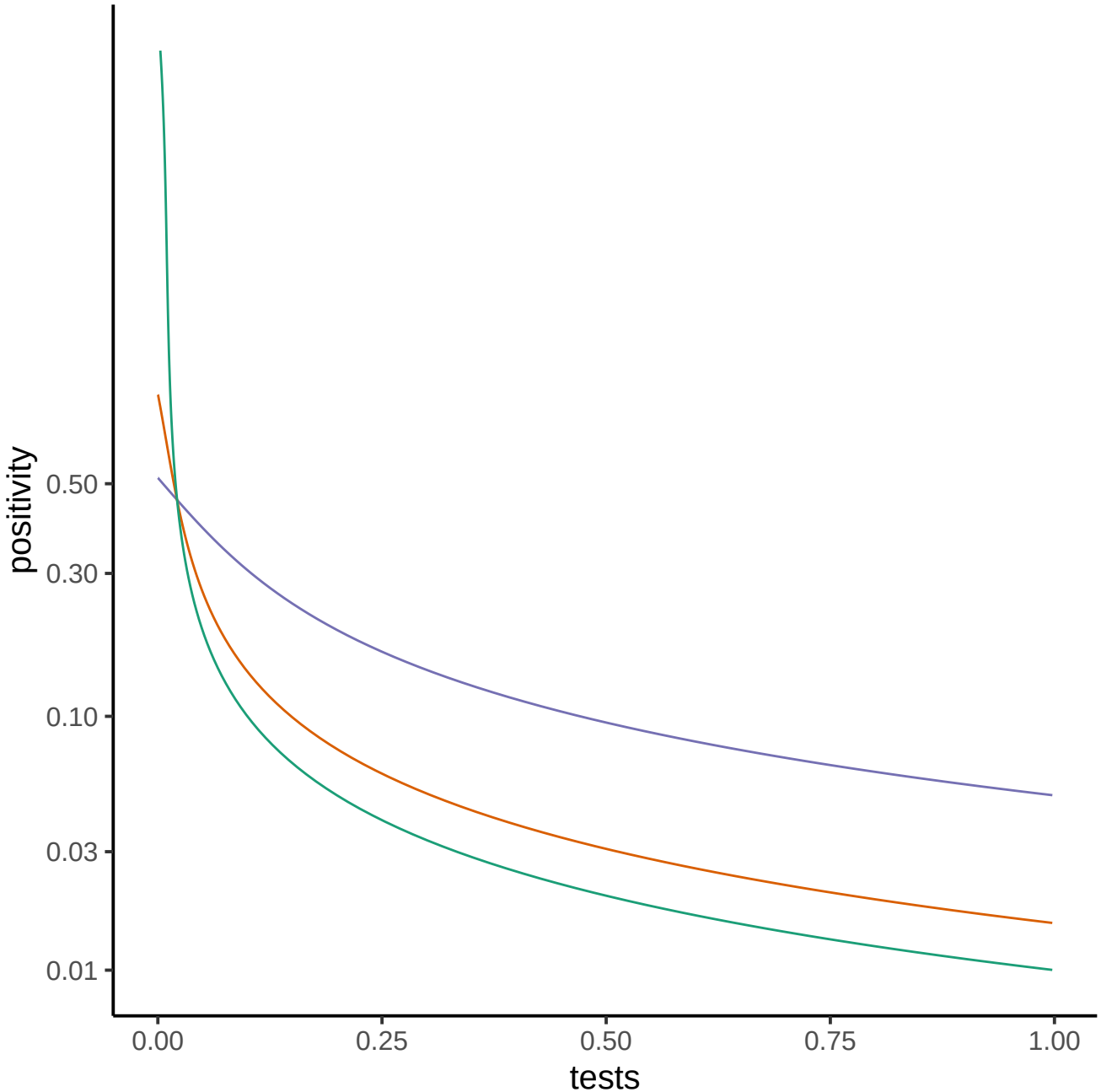
Constant odds ratio



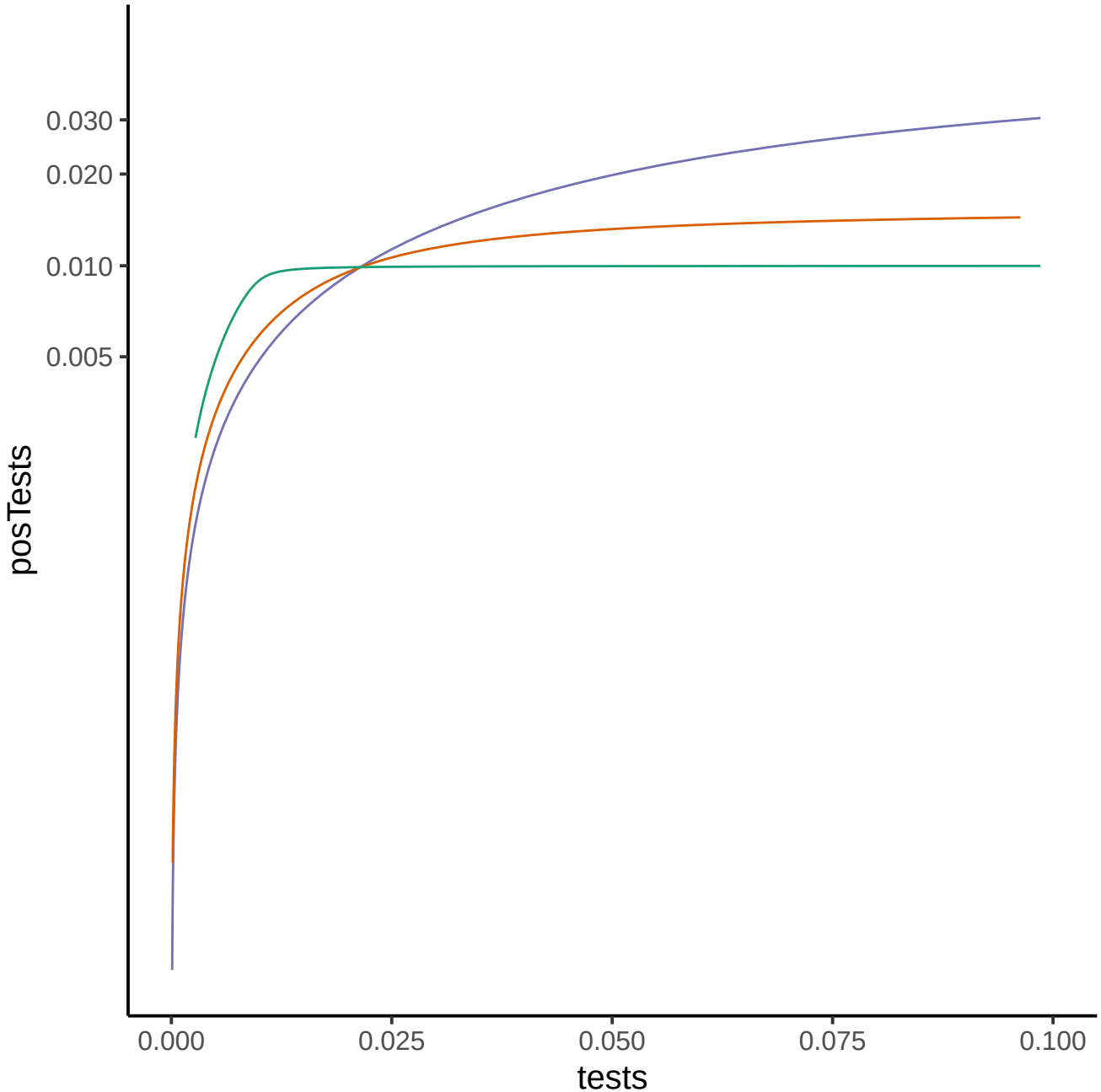
Interpreting observations



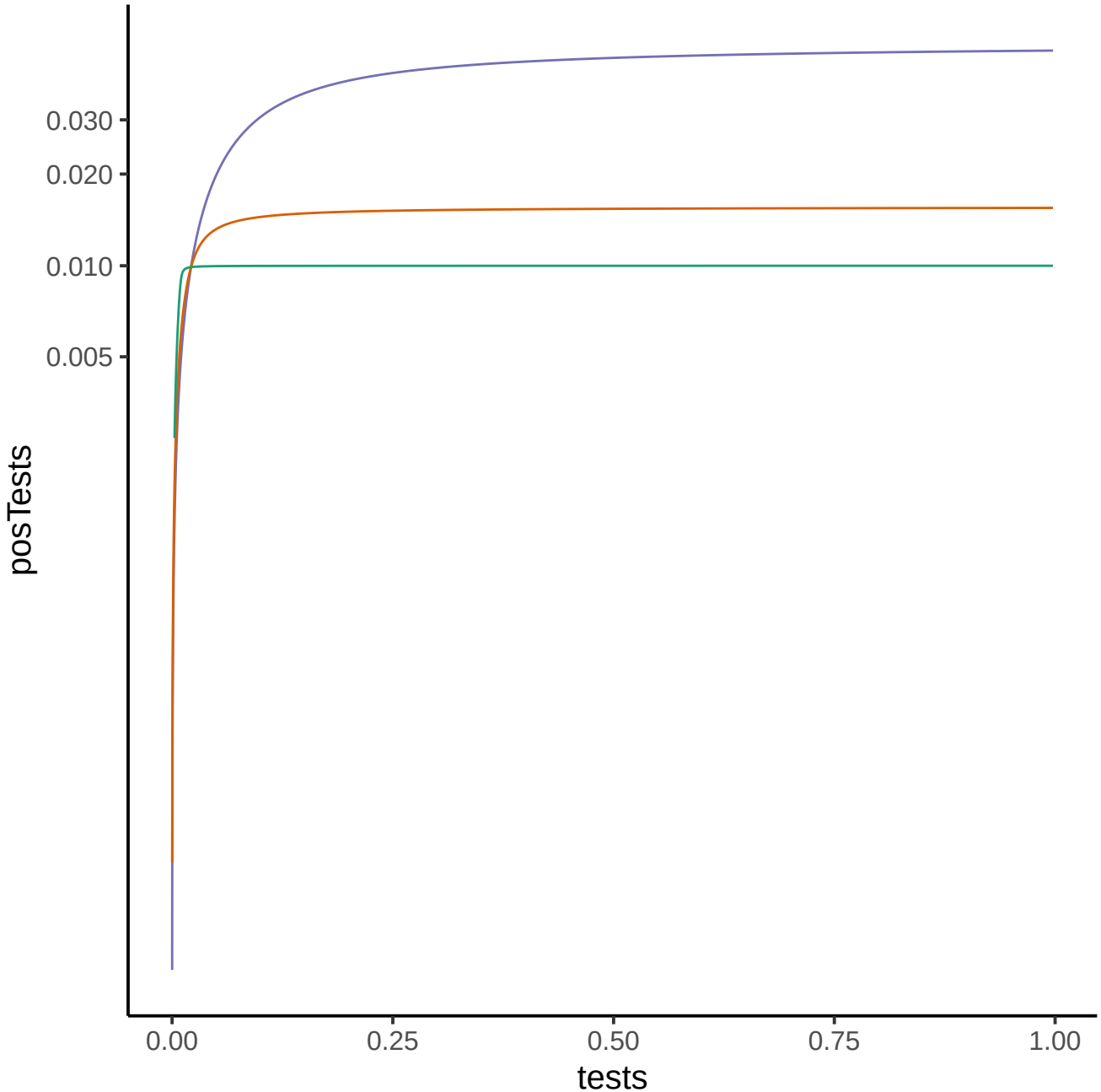
Interpreting observations



Interpreting observations



Interpreting observations



Prevalence-incidence gap

- We are thinking about the asymptotic properties of these tests as measuring prevalence when we test the whole population
- But what we're trying to measure is incidence of new cases
- Consider the population we consider eligible to take these tests
  - Does not include people with recent positives
  - Although sometimes they do take it

Modeling approaches

- Top-down models to fit a beta shape parameter together with disease dynamics
  - Could also use odds ratios and a single shape parameter to describe relationship between testing propensities in groups
- Bottom-up models to fit to likelihood of observed testing numbers and observed positives
- False-negative and false-positive results

Combine with other data streams when possible

- Medical screening, hospital discharge
- ILI surveillance reports
- Seroprevalence

Simulation-based validation

- Simulate scenarios with realistic sources of variation
- Test how well different modeling approaches can fit

Data curation

- Work with provincial and federal health agencies to improve connections between models and data
- How data are collected:
  - e.g., what multiplex tests do people take?
- How data are shared
  - Bringing models to data
  - Make shareable products as part of the research project

Thanks for your patience!

- Also:
  - Meeting organizers
  - Key collaborators: Bolker, Brown, Champredon, Li, Zhao
  - CIHR, PHAC, NSERC