In class exercise: Establishing Detection Thresholds

Jesse Brunner

A motivating example: *Bd* and *Bsal*

We import hundreds of millions of live animals into the United States every year[1]. Even if some small fraction of those individuals harbor a disease we care about, that amounts to a fire-hose of introduced infections! In class today, let's focus on amphibian pathogens.

> The amphibian chytrid fungus, *Batrachochytrium dendrobatidis* (Bd), is the most devastating emerging pathogen on record, responsible for declines in over 500 species, including 90 that are likely extinct. It likely owes its global distribution in large measure to the international trade of African clawed frogs (*Xenopus laevis*) used for pregnancy tests and research, American bullfrogs (*Lithobates catesbeianus*) sold for food, and myriad species involved in the pet trade. More recently, the international pet trade aided the emergence of a novel chytrid fungus, *B. salamandrivorans* (Bsal), that is particularly lethal to salamanders.
>
> Bsal was introduced into Northern Europe via the pet trade from Southeast Asia. While it has not yet been detected in the wild in much of Europe, it is already prevalent in and appears to have spread among private collections in Europe. In North America, a hot-spot of salamander diversity, Bsal is apparently absent from both wild and captive amphibians. However, the risk and potentially devastating consequences of its introduction via trade led to prohibitions on the importation of 201 species of salamanders into the U.S.A. and all salamanders in Canada.
>
> — Jesse L. Brunner[2]

All of which is to say, detecting pathogens at the borders, in captive collections, and generally everywhere amphibians are held or moved is really important. There are two problems facing us. The first is the magnitude of the trade, which is, as I noted above, daunting. We will set this aside for the moment, but suffice to say there are many people working on it[3].

The second issue is that we are often trying to detect low-levels of rare infections. This is sort of the sweet-spot for confusion. Low-level infections produce weak signals, which can be hard to distinguish from background levels of signal[4]. Also, because infections are rare there are probably very few real infections to detect, making us justifiably wonder if what we are seeing is real or not. It is this second problem we are going to tackle today.

Playing with real[5] data

Imagine we have the following data from a qPCR[6] diagnostic test for *Bsal*. Remember that "Ct" is the cycle threshold, the PCR cycle where the fluorescence crosses some threshold where we can agree it is (likely) real. Each bar represents

[1] Smith, K. et al. Summarizing US wildlife trade with an eye toward assessing the risk of infectious disease introduction. EcoHealth 14, 29–39 (2017).

[2] Brunner JL (2020) Pooled samples and eDNA-based detection can facilitate the "clean trade" of aquatic animals. Scientific Reports 10:10280

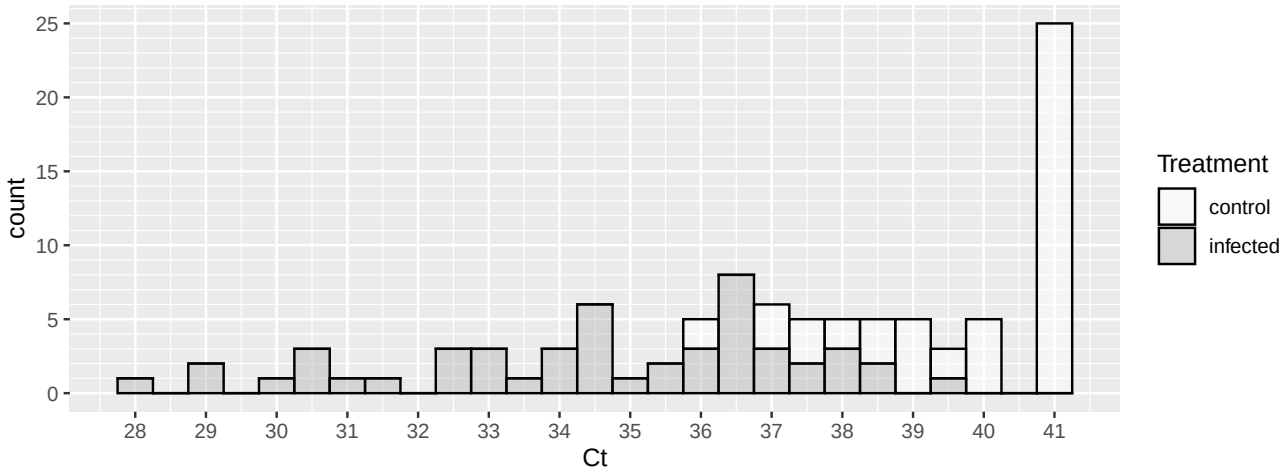
[3] Feel free to ask me about it!

[4] Is that a band on my electrophoresis gel or a shadow? Do I need to get my eyes checked?!

[5] Made up.

[6] Remember the quantitative Taqman realtime polymerase chain reaction we talked about in class? Lasers shining in to tubes with exponentially growing populations of particular DNA molecules?

samples with a given Ct value. Each sample is from an individual that was either experimentally infected (gray) or left uninfected (white). Notice that we only run this out for 40 cycles, so there is a large bar above 40, representing all the samples that never crossed the threshold. Also note that have broken out the data in half-cycle steps.



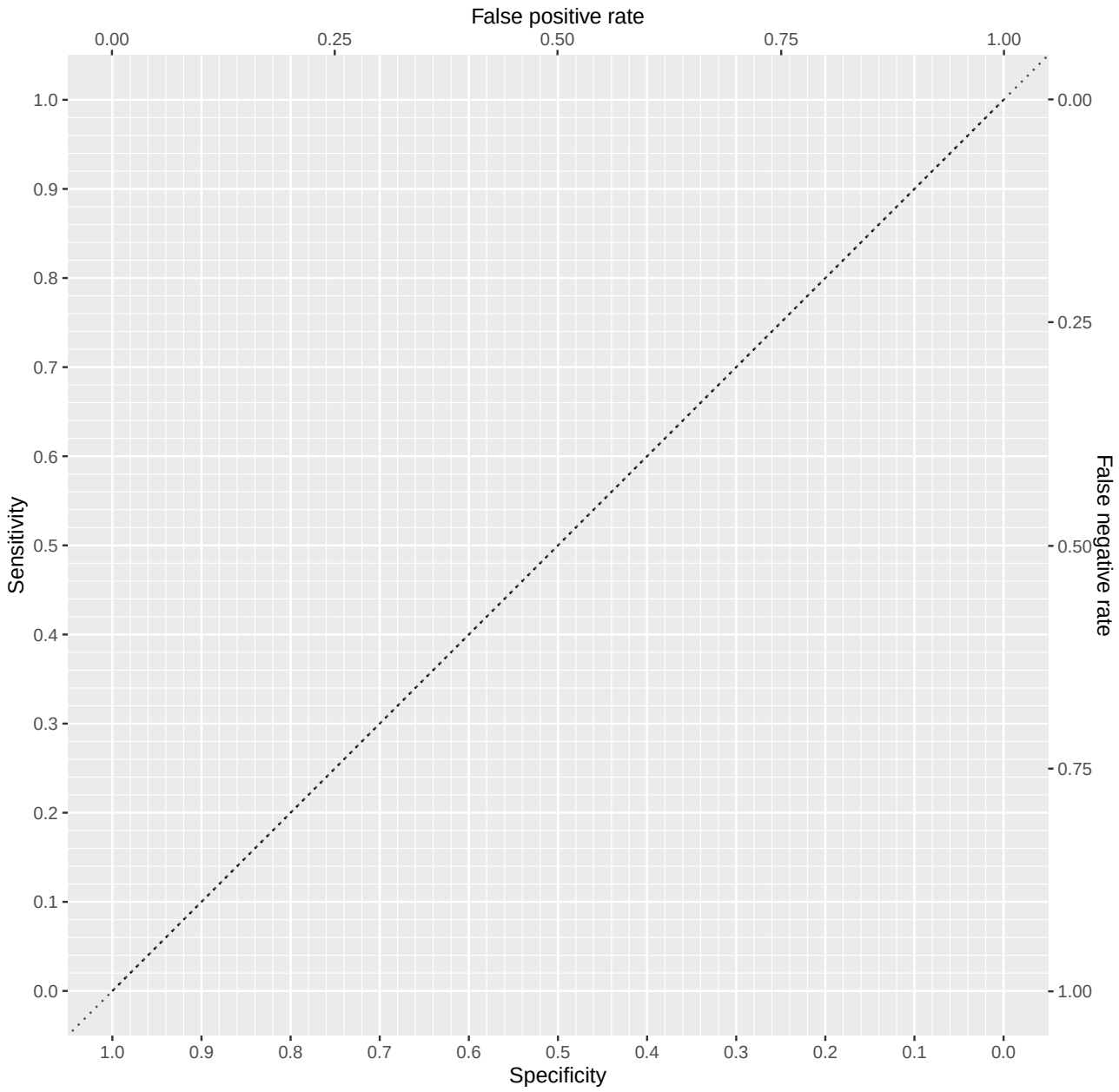
Given these data, fill in this table showing how many true positives (infected animals testing positive; TP), false positives (control animals testing positive; FP), true negatives (control animals testing negative; TN), and false negatives (infected animals testing negative; FN) we would have if we used each of the Ct values as our cutoff. Remember: anything *less than or equal to* the cutoff is considered positive.

Ct	TP	FP	TN	FN	Sensitivity	Specificity
40.0						
39.5						
39.0						
38.5						
38.0						
37.5						
37.0						
36.5						
36.0						
35.5						

Next, we need to convert these numbers into our estimates of **sensitivity**, or the proportion of infected individuals that are correctly scored as positive, and **specificity**, which is the proportion of uninfected individuals that are correctly scored as negative. Add these estimates to the table.

Finally, we have the numbers we need to construct an ROC plot. Simply place a point on the axes below corresponding to the Sensitivity and Specificity of each

potential cutoff (Ct threshold). Bonus points if you label each point with the corresponding cutoff.



After plotting your estimates, connect the dots, as it were, and extend the lines all the way to specificity of zero and all the way to sensitivity to zero. Congratulations! You have just made an ROC curve!

This should help you visualize the inherent trade-off in setting a cutoff: we can increase sensitivity, but this comes at the cost of reduced specificity, and vice versa.

What would you do?

Now, the Big Question: If you were in charge of defining the parameters of a diagnostic qPCR test that was going to be used at the borders of our country to detect *Bsal* in imported animals, what threshold would you use? What are you trying to maximize and minimize? What are you gaining? At what cost?