Why we measure E. coli and a note on units

Since *E. coli* (which should always be italics, although formatting limitations prevent me from italicizing this genus and species in this post’s title) has such a big scary reputation in the popular press, it’s important to remember why we measure *E. coli*, and know how the Texas Commission on Environmental Quality (TCEQ) regulates *E. coli* in public waters. As it turns out, neither why we measure *E. coli* nor how we regulate it is totally straightforward.

First, *E. coli* is a species of bacteria. It is a species that includes many sub-species, some of which are human pathogens (most commonly *E. coli* associated with food poisoning in the US) and many of which aren’t. *E. coli* are relatively easy to measure, live in the guts of mammals and birds, and have some other nice characteristics that make them useful as a proxy for mammalian or avian fecal contamination. But that’s it! We use them as a proxy for other pathogenic organisms potentially transmitted via feces in water, and not so much because we are worried about the *E. coli* themselves. Recent debates in Rio de Janeiro over microbiological water quality for Olympic events highlight the importance of this point and the potential implications of building regulations on indicator organisms and not specific human pathogens (see [http://bigstory.ap.org/article/d92f6af5121f49d982601a657d745e95/ap-investigation-rios-olympic-water-rife-sewage-virus](http://bigstory.ap.org/article/d92f6af5121f49d982601a657d745e95/ap-investigation-rios-olympic-water-rife-sewage-virus)). *E. coli* is not a perfect indicator of fecal contamination, and finding better (and still practical!) indicators is an active area of research. If you’re interested, Google “microbial source tracking”.

Second, the TCEQ regulates *E. coli* levels by setting a mean value over time that a water body must not exceed (the mean must be computed from at least 10 samples over two years), as well as a maximum single sample value that should never be exceeded. This is pretty standard, since the EPA recommends this. But it means that after a single sampling event such as ours, it is possible to say that a water body is violating the standard, but it is impossible to say that a water body is compliant. Also, you have to follow specific sampling and analysis protocols in order for your *E. coli* measurements to be truly comparable with the standards. (We did not follow this protocol.) This promotes quality control and also eliminates potential differences produced by different methodologies, even if they are both done correctly!

Another important caveat with regard to comparing our *E. coli* measurements with the TCEQ standards is that we used a technique which gives us units of MPN/100 ml, while the standard is specified in CFU/100 ml. MPN is “most probable number” of viable *E. coli*, and CFU is “colony forming units”, a proxy for individual viable *E. coli*. A recent paper¹ found that regardless of how careful you are in executing your lab analyses, MPN values will tend to be higher and show greater variability than CFU measurements. These differences are produced by the statistical assumptions built into MPN estimates, not just lab procedure variability, so they will always be there!