Outliers in graded exercise testing

A potential issue with outliers in graded exercise testing is that they tend to focus exclusively on the VO2 vs. time relationship. This likely makes sense in many analyses that focus on VO2, such as VO2max tests and VO2 kinetics tests. However, breakpoint methods use additional variables such as VCO2, VE, petO2, and petCO2, which can have their own outliers. The relationship with VO2 and time in an incremental graded exercise test is usually quite linear, but that is not necessarily the case for the other variables. Generally, it seems like when any of VO2, VCO2, or VE is an outlier, the other two are also outliers, so it may be fine to identify outliers with one relationship alone. Given that VO2 is mostly linear, it would make the most sense to use that one. That aside, I think a distance metric that uses at least time, VO2, VCO2, and VE would be worth exploring because it assesses outliers from all three measures. Although all three values are generally outliers together, it is possible that this may not always be the case.

Another issue with relying on the VO2 vs. time relationship to find outliers is that the linearity of that relationship breaks down if someone demonstrates a plateau or even a drop in VO2 at the end of a test. If such an event occurs, then one can only reasonably fit a linear relationship up to VO2max, but they must exclude the data points after it. It’s possible that this could be fixed with a loess, spline, or a non-linear model, but then how do you choose the specifications for those models? Another potential correction is to fit two linear models in the event of a plateau or decrease in VO2 at the end of the test. However, the difficult part about that is you need to determine where the plateau starts. It is easier to determine the presence of a plateau than to determine the precise starting location of the plateau. This may be accomplished with a breakpoint method.