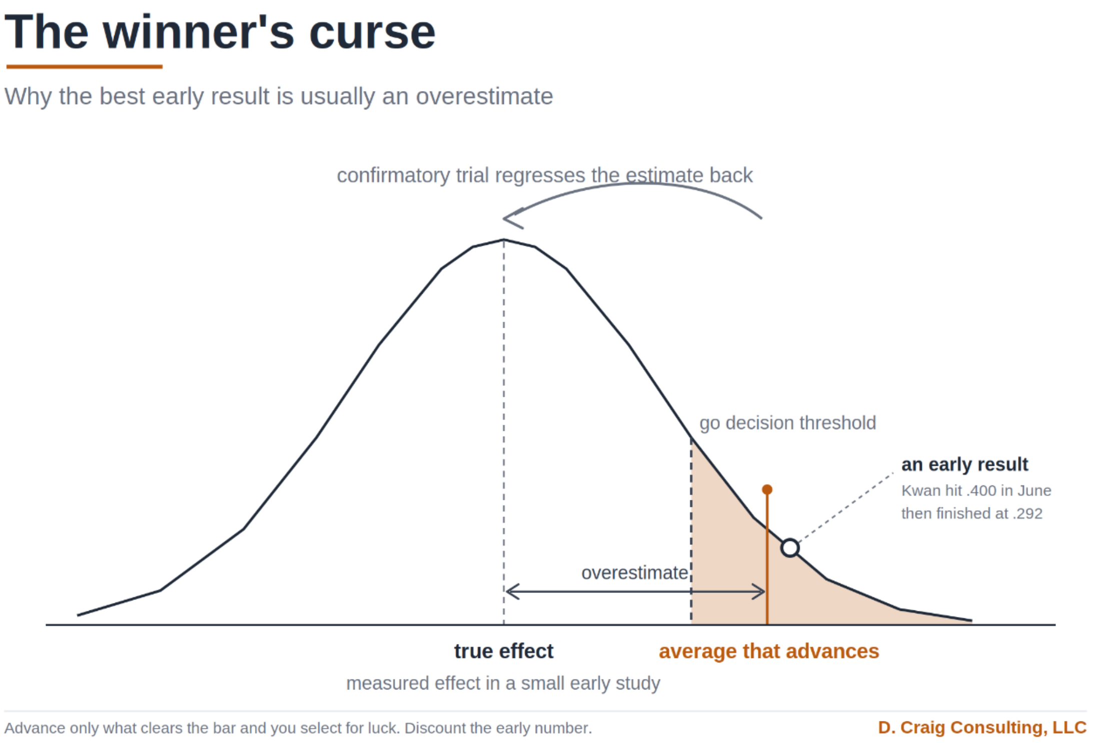
The winners curse



The winner's curse

Why the best early result is usually an overestimate

confirmatory trial regresses the estimate back
go decision threshold
an early result
Kwan hit .400 in June
then finished at .292
overestimate
true effect
average that advances
measured effect in a small early study

Advance only what clears the bar and you select for luck. Discount the early number.

D. Craig Consulting, LLC

Early in a program, a small study reads out better than you dared hope. The effect size is big. The p-value cooperates. Everyone in the room exhales. I have been in that room, and I have learned to distrust it.

The reason is a piece of math that does not care how good your science is. Call it the winner's curse.

Every effect you measure is the real effect plus noise. When you advance only the molecules, the endpoints, the subgroups that clear a bar, you are not simply selecting for truth. You are also selecting for the times the noise happened to run your way. The compounds that win an early look are disproportionately the ones that got lucky, which means the headline number is, on average, too high. And the smaller the study, the more room luck had to work.

Then you run the real trial, the one powered to settle the question, and the effect drifts back toward its honest value. Sometimes it holds. Often it shrinks. Occasionally it disappears, and a program that looked like a winner in Phase II becomes another line in the Phase III attrition tables. That is not a failure of nerve or talent. It is arithmetic doing what arithmetic does.

Baseball makes the same point every summer. In 2024, Steven Kwan of the Guardians was still hitting .400 as late as June, far enough into the season that people started invoking Ted Williams. Nobody who follows the game booked him for a .400 season, because .400 is not what a Kwan at-bat is worth. It was a real number sitting on a stretch of at-bats that happened to fall his way. He finished the year at .292, back around his true level for a career .276 hitter, once the season ran long enough to tell the truth. The hot streak was not a lie. It was an estimate taken before the noise had time to average out. Your Phase II readout is a hot streak in June, not the final line.

The idea also has an older name in auction theory, where the bidder who wins is usually the one who overestimated the prize and therefore overpaid. If you have ever won a competitive licensing deal against a room full of smart people, sit with that for a second. The winner was, statistically, the most optimistic party at the table. Now picture that asset de-risked mostly on a thin early dataset. You have paid a premium for someone else's lucky read.

None of this is an argument for pessimism. It is an argument for a discount. Write down your endpoints before you look at the data. Assume the early estimate is inflated, and size the next study off a haircut rather than the headline. Ask for the effect to show up twice before you put real money behind it. Be most careful precisely when you are most excited, because the excitement is the tell that the number has already done its work on you.

The best operators I know are not the ones who find the biggest early signals. They are the ones who know how much to believe them.

I write more about building capital-efficient biotechs in The New Rules of Biotech, available on Amazon (https://www.amazon.com/dp/B0H4WZV6CW). All proceeds are donated to Destroy Duchenne and other patient support groups.

#Biotech #DrugDevelopment #ClinicalTrials #Biostatistics #VentureCapital