ARE MY SEEDLINGS STUNTED? By David South

I have tested herbicides in nurseries for more than three decades, so I think I know when an herbicide treatment causes stunting. Recently, I reported results from six nursery trials, and I said no stunting of roots occurred when seedlings were treated with herbicide X. We observed no visual signs of stunting on roots and no significant decreases in root mass. In fact, at one nursery (A), the herbicide suppressed weeds and this resulted in a statistically significant (alpha = 0.05) increase in root mass.

Root weight (mg) of seedlings from six nursery trials.

Nursery	Not treated (mg root)	Herbicide (mg root)	Difference	Probability of difference occurring by chance (P>F)	My definition of "stunting" (LSD 0.05)
A	100	260	+160%	0.0001	-40%
B	160	160	0%	0.9870	-25%
C	370	350	-5%	0.5300	-13%
D	750	720	-4%	0.6292	-18%
E	510	560	+10%	0.0973	-12%
F	580	620	+7%	0.0631	-29%

However, one reviewer said my statement regarding stunting was wrong. He said stunted roots occurred at nursery C and nursery D. In both cases, the mean value for herbicide plots was 4–5 percent less than the value for untreated plots. His view of "stunting" differs from mine since he uses simple math instead of relying on common statistical tests. Since he does not use confidence intervals to determine when stunting occurs, he is willing to make a "Type 1" statistical error (i.e., calling a difference real when it really is not). For example, if WATER was substituted for the herbicide treatment in nursery D, a difference between means of up to 30 mg would occur 63 percent of the time! Therefore, the reviewer would be making a "Type 1 error" if he said the "water treatment" stunted trees (on the grounds that the roots weighed less than the average for untreated plots).

In contrast, I use a statistical test (e.g., F-test) to determine if a treatment causes stunting. If the power of the test is typical, I might require a 12–18 percent reduction in root mass before I would claim a treatment caused stunting. If seedlings in the nursery were variable, producing a test with low power, I might require a 25–40 percent reduction in root biomass before I would claim a reduction was due to the herbicide and not simply due to random chance. Even when using these values, my conclusion might be wrong (i.e. Type 1 error) in 1 out of 20 tests. If I felt this level of confidence resulted in too many "Type 2" errors (i.e., calling a real treatment effect insignificant), then I might be willing to be make a "Type 1" error 10 percent of the time (i.e., 1 out of every 10 tests).

Foresters should realize that researchers differ in their opinions regarding what constitutes "scientific proof." Some individuals do not require any replication or any confidence interval to reject a null hypothesis (e.g., herbicide X does not affect root growth of pines seedlings). Some authors now seem very willing to make claims about an uncertain future without the need for replication, confidence intervals, or rigorous hypothesis testing.