

Criteria Category	Yield Testing Criteria	Yes	No
Utilizing Standard Definitions of Yield	What was the percent increase in yield? (X-X%)		
Utilizing Standard Definitions of Yield	Did they provide this percent increase in yield as a number within the paper?		
Utilizing Standard Definitions of Yield	Did they measure plot yield? (As opposed to single plant yield or grain yield per 1000 plants)		
Using Controls	Did they test for yield effects in elite germplasm?		
Using Controls	Did they backcross (or do any kind of cross) of their edits into non-transformed germplasm before yield testing?		
Using Controls	Did they test for yield effects in a non-transformed wild-type plant?		
Using Controls	Did they measure yield in a line containing a null construct?		
Using Controls	Did they measure yield in a context relevant to what is commercially expected? (i.e. did they measure hybrids vs. inbred yield; maize = hybrid, wheat = inbred, rice = inbred or hybrid)		
Matching On-Farm Practices	Did they test for yield effects in the field?		
Matching On-Farm Practices	Did they plant at a standard planting density?		
Matching On-Farm Practices	Did they mechanically harvest their fields?		
Matching On-Farm Practices	Did they directly plant the seeds in the field or transplant seedlings (YES/NO)? Is this normal (normal/not-normal)?		
Matching On-Farm Practices	Did they control for moisture before weighing (grain, leaves) to an industry standard? (maize = 15.5%, wheat = 13.5%, soy = 13-15%)		
Sufficient Replication and Experimental Design	Did they have sufficient replication of individuals within an environment (n = number of plants)? Where sufficient ≥ 20		
Sufficient Replication and Experimental Design	Did they have sufficient plot replication within an environment (p = number of plots)? Where sufficient ≥ 3 .		
Sufficient Replication and Experimental Design	Did they replicate their experiments across environments and/or years (n = number years)?		
Sufficient Replication and Experimental Design	Were enough site year locations tested? (syc = total number of site-year combinations)?		

Sufficient Replication and Experimental Design	Did they consider plant competition in their plot design?		
Sufficient Replication and Experimental Design	Did they correct for field effects (BLUPs, BLUEs, fixed effects) in their yield estimates?		
Sufficient Replication and Experimental Design	Did they avoid edge effects by only harvesting inner-rows?		