

IFM Quality Services TUTORIAL for Participant Laboratories

This tutorial explains about participation in proficiency testing and how to get the most from the programs.



© Copyright IFM Quality Services Pty Ltd 2004

What is Proficiency testing?

- Scheduled part of laboratory quality assurance where a group of laboratories compare their results of a test with many others who have tested the same material(s).
- Mechanism by which a laboratory can demonstrate competence based on practical evaluation.



Why participate?

- Requirement of certification and accreditation bodies
- Proof to management of competence
- Proof to higher authorities and/or clients of competence



- Opportunity for increasing understanding of quality issues in a test
- Opportunity for comparison of methodologies with other labs
- Opportunity to learn and teach
- Increase confidence of laboratory (feel good J)



What makes a “good” PT Program ?

- Definite aim
- Design to meet that aim
- Technically appropriate
- Accredited (for acceptance with ISO 17025)
- Provides information that participants need/desire



PT Aims

- Usually, PT programs aim to compare results and to find possible reasons for variation
- Specific aims are sometimes covered at special request

Design program to meet these aims

- **Compare results**
 - Have a measurable way of making a comparison
- **Identify possible reasons for variation**
 - Questionnaire
 - Detailed data analysis looking at factors within methodology

Technically appropriate

- Programs are designed in conjunction with technical advisers
- Design against a standard (usually the base standard for the method/test)
- Samples/methodology should as far as possible mimic what is done in the lab



Accredited

- ILACG13:2000
- This entails:
 - Maintaining a quality system
 - Designing programs appropriately
 - Ensuring samples meets specified criteria for performance homogeneity and stability



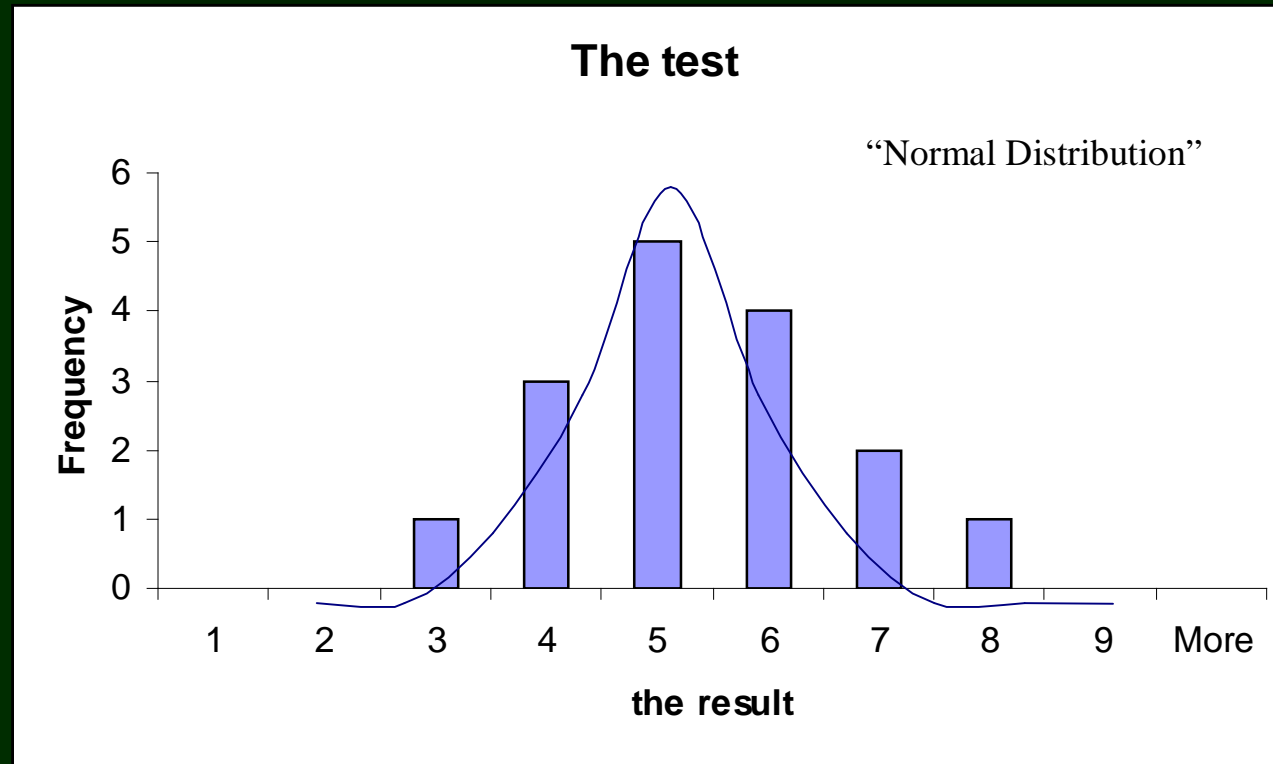
- Reporting according to the requirements
- Report all data
- Reports accurately
- Report reasons for variation
- Provide means of assessing performance



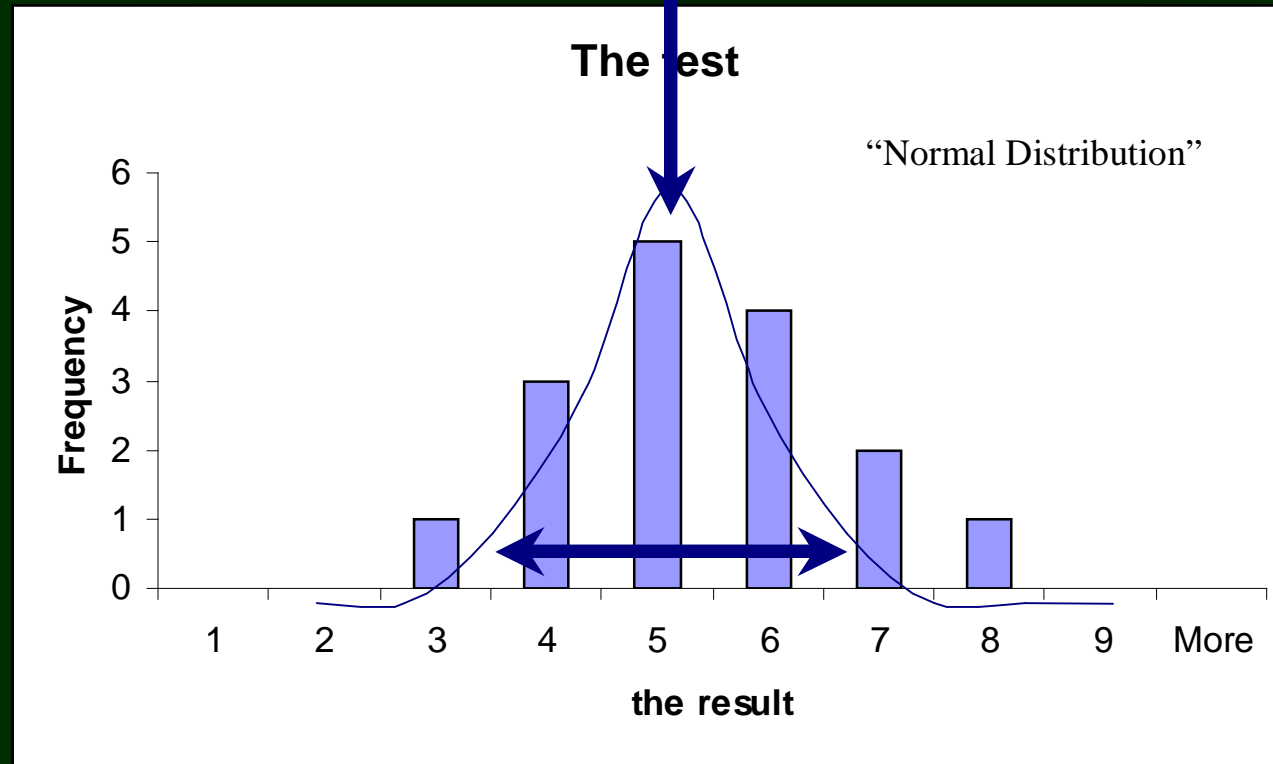
Assessing the result

- It is preferred if the results fall into a normal distribution (Bell shaped curve).
- When this occurs, standard assessing practices can be applied.
- And now..... a little background about robust statistics

Assessing the Results

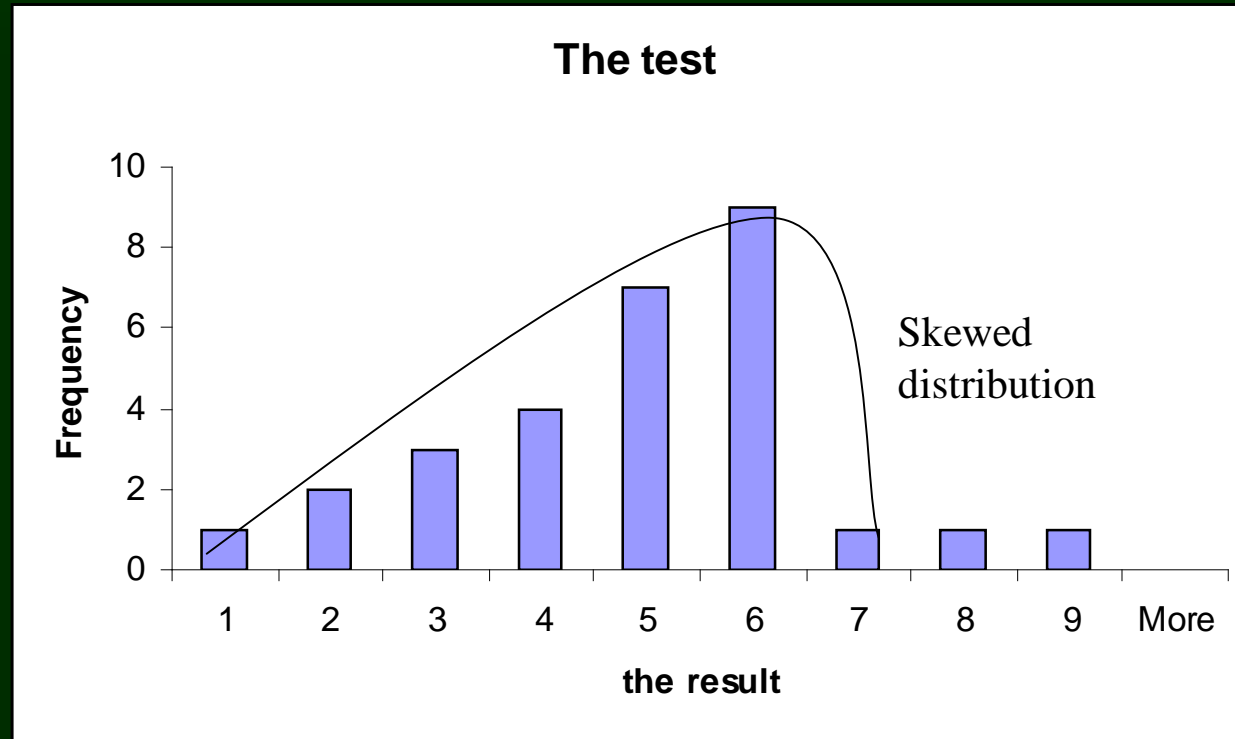


- All PT stats search for the consensus (or true) result and the distance a “correct” result can be from the consensus



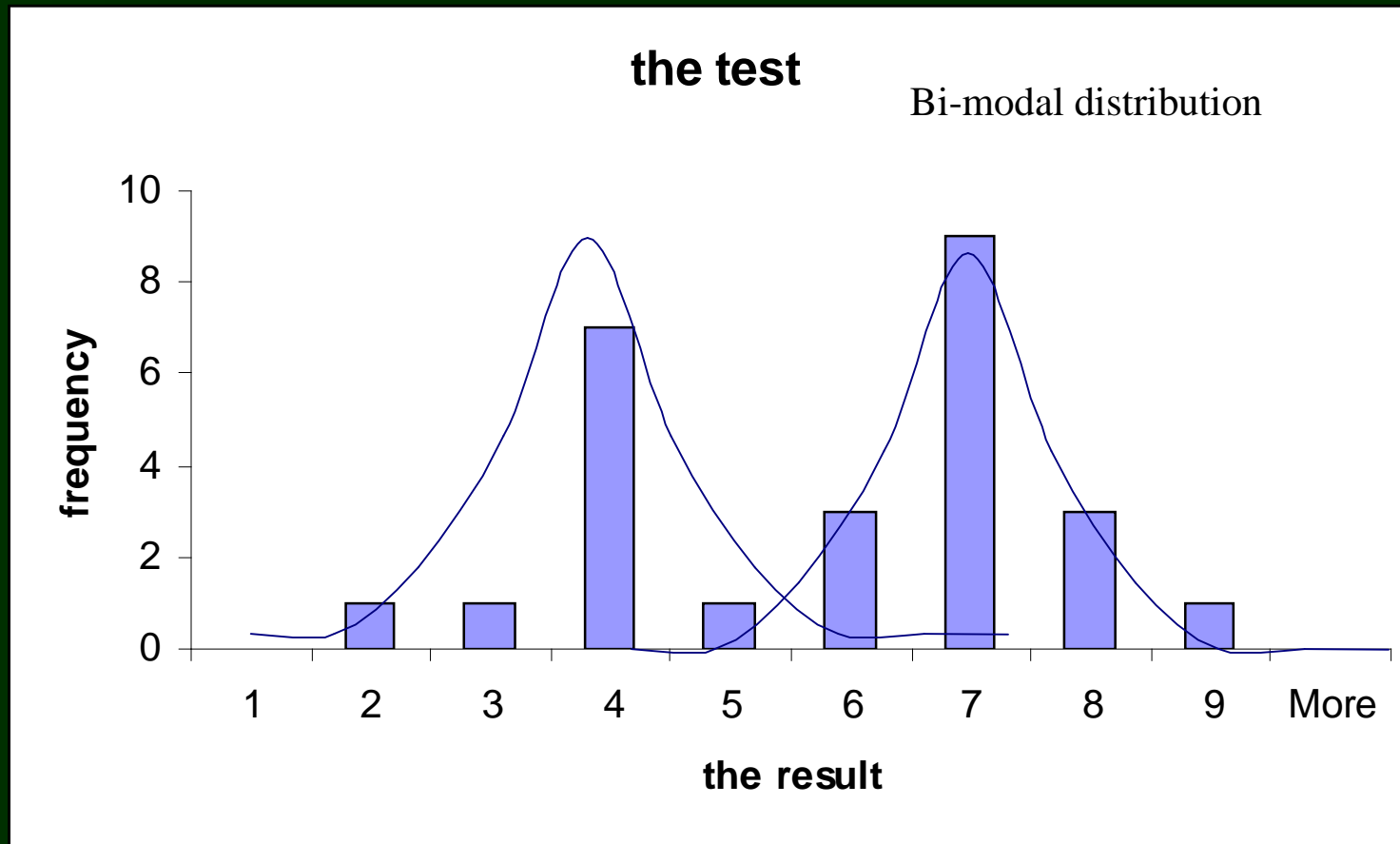
- Traditional stats use the average and standard deviation
- Robust statistics use the median and the inter-quartile range.
- When the distribution of results is not “normal”, then care must be applied before applying certain statistical rules, as the rules applied may have limited value.



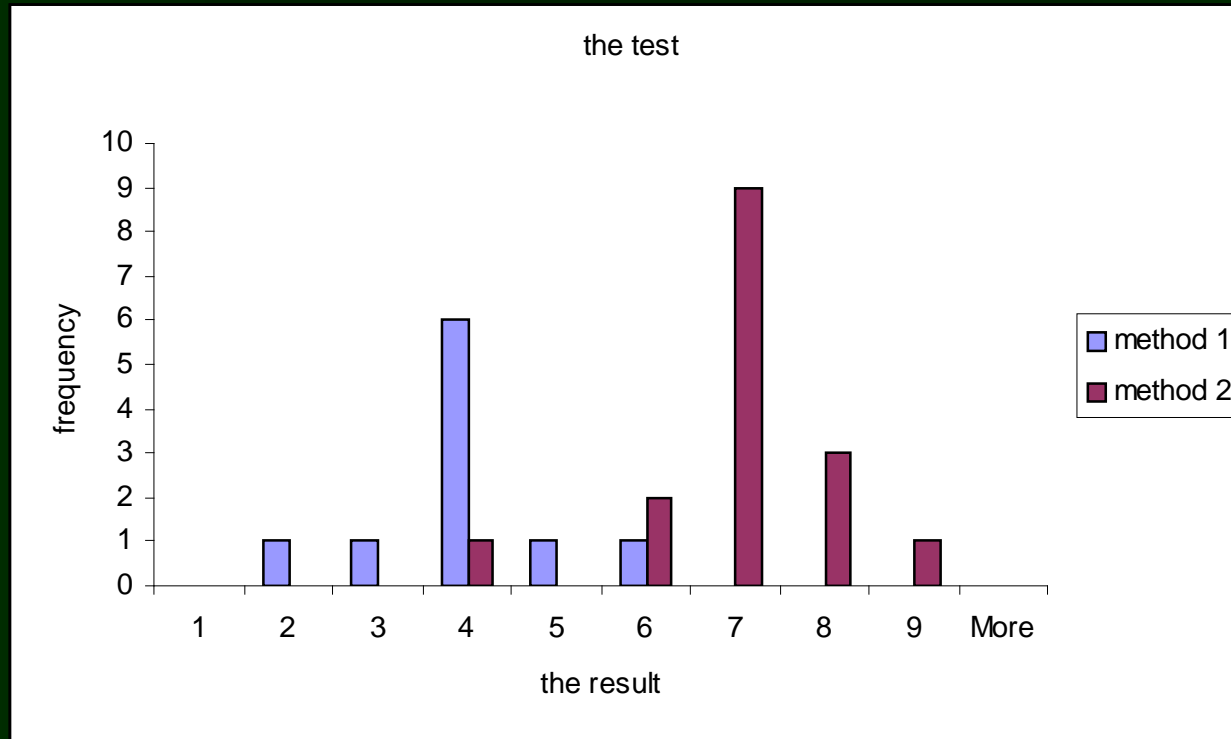


If the distribution is skewed, perhaps some factor has limited the evaluation of the result (for example, a an aspect of test methodology may not be able to detect readings higher than a certain value

- This is another example of when care should be taken for evaluating results



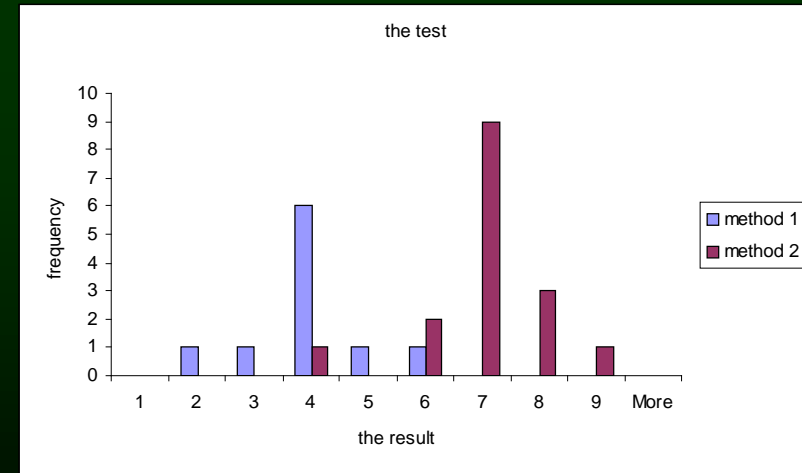
Perhaps two methods are involved here?



- 1. Maybe both methods are valid
 - Split each method and perform evaluation on both



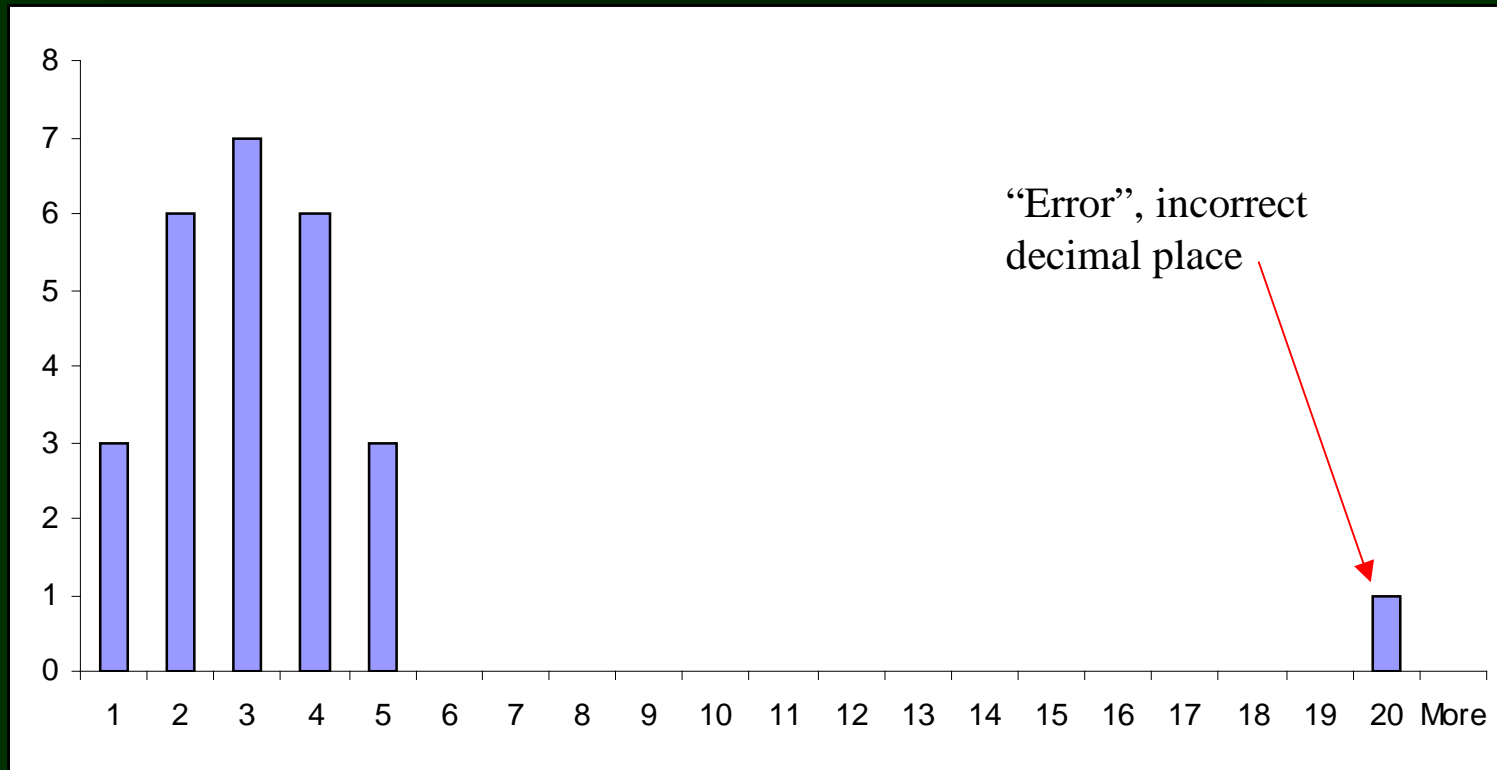
© Copyright IFM Quality Services Pty Ltd 2004



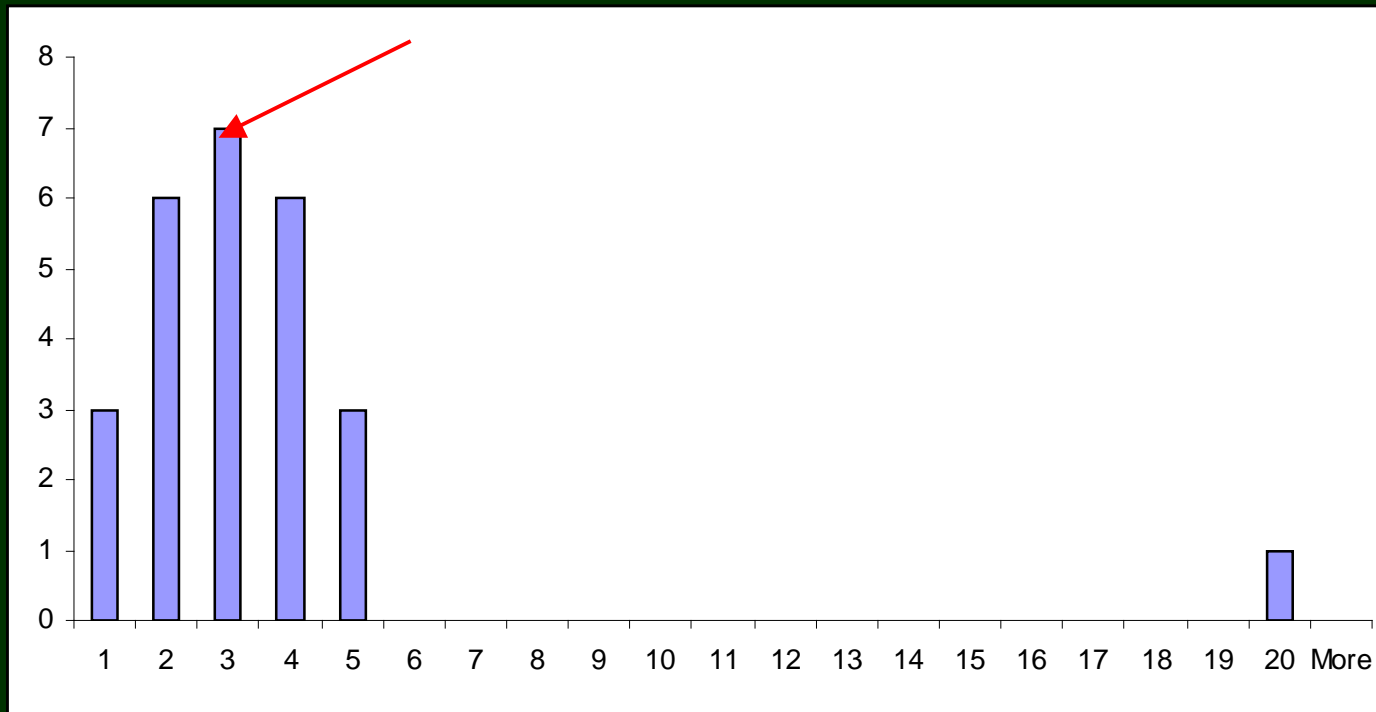
- **2. Maybe only method 2 is valid.**

- Perform stats on method 2 and apply these evaluation criteria to method 1 (some method 1 participants might “pass”)
- OR method one cannot be evaluated (leave up to participant to decide if correct)
- OR participants using method 1 fail

Erroneous reporting



		with outlier	without outlier
traditional	mean	3.65	3.00
robust	median	3.00	3.00
traditional	SD	3.54	1.22
robust	NIQR	1.48	1.48



		with outlier	without outlier
traditional	mean	3.65	3.00
robust	median	3.00	3.00
traditional	SD	3.54	1.22
robust	NIQR	1.48	1.48

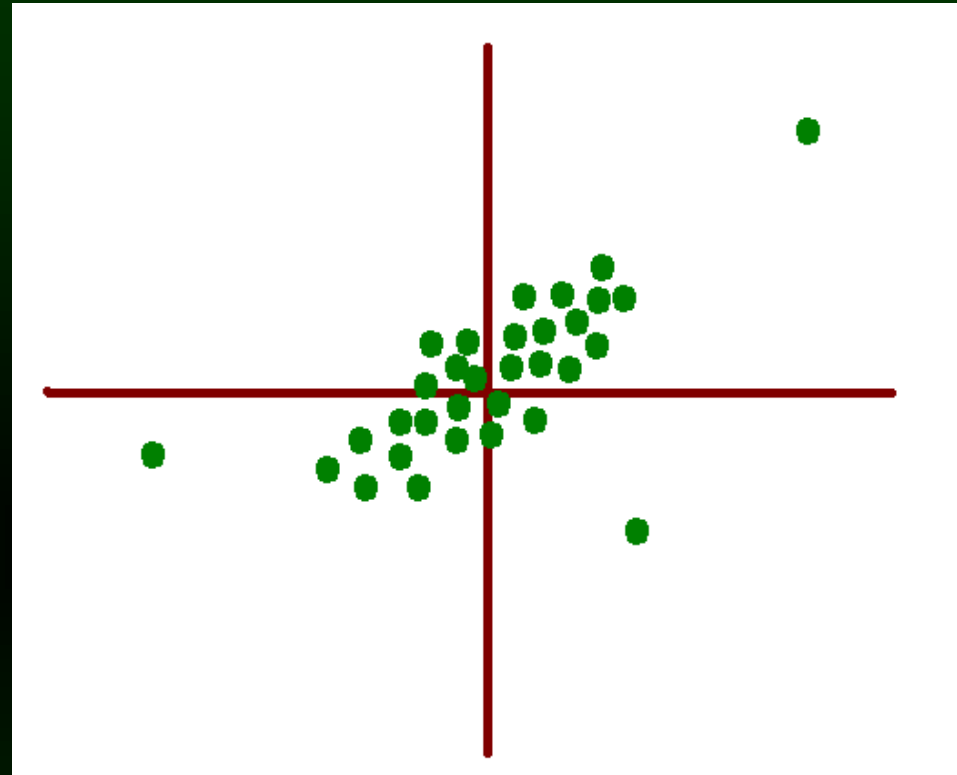


Within and Between Analyses

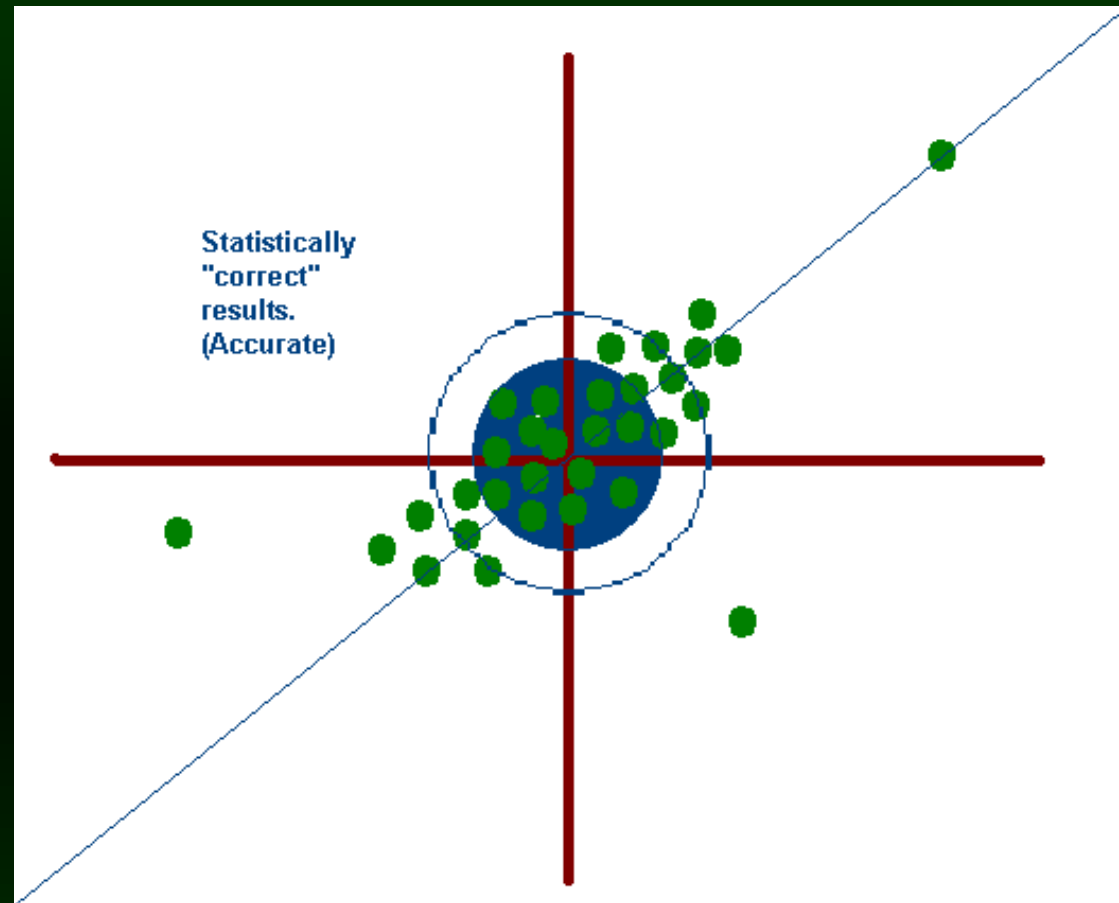
- Within and between scores are approximations to accuracy and precision.
- For these purposes,
 - Accuracy means how “correct” a result is
 - Precision reflects the ability with which we can produce exactly the same reading for the same determination



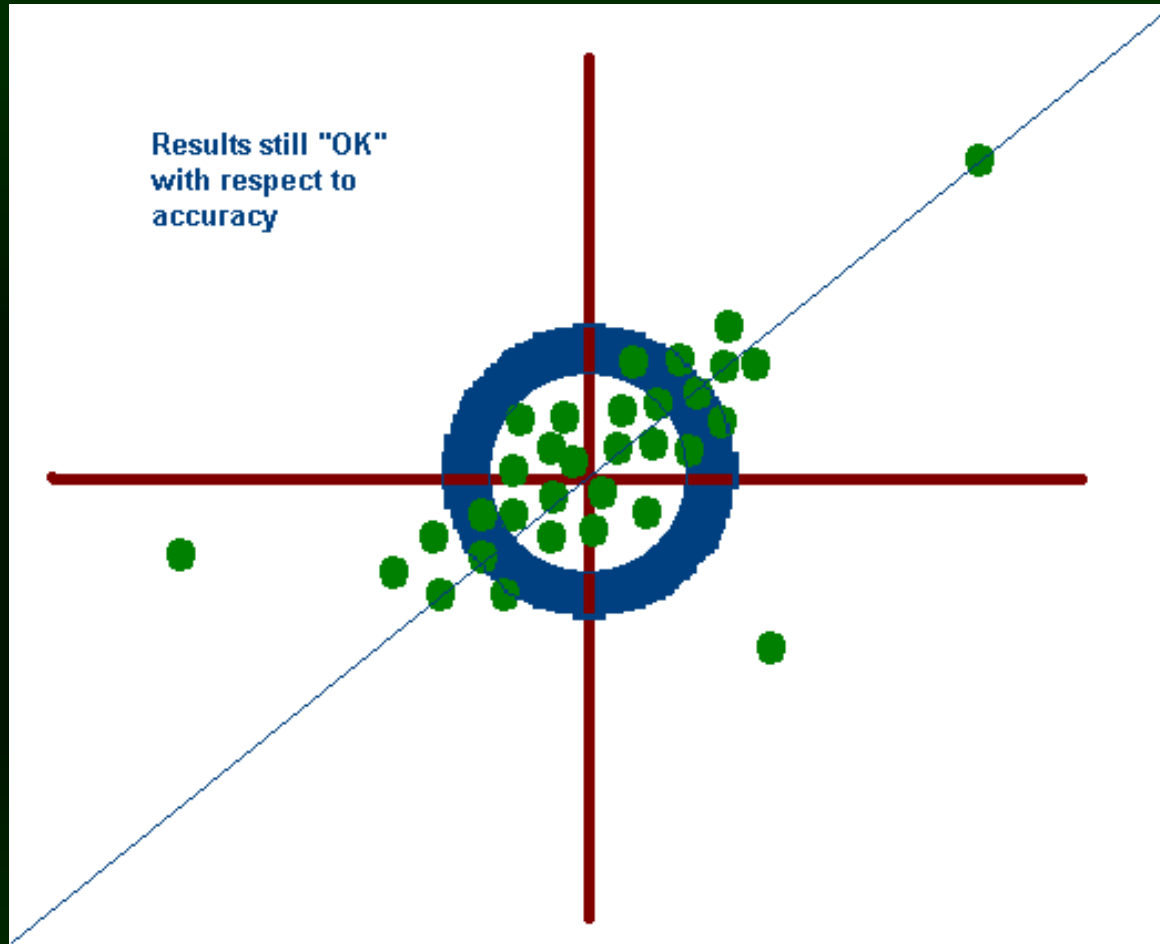
- Let the two lines represent the results for two samples
- The dots represent the results submitted by individual participants for these two samples
- The “correct” result falls at the intersection of the two lines



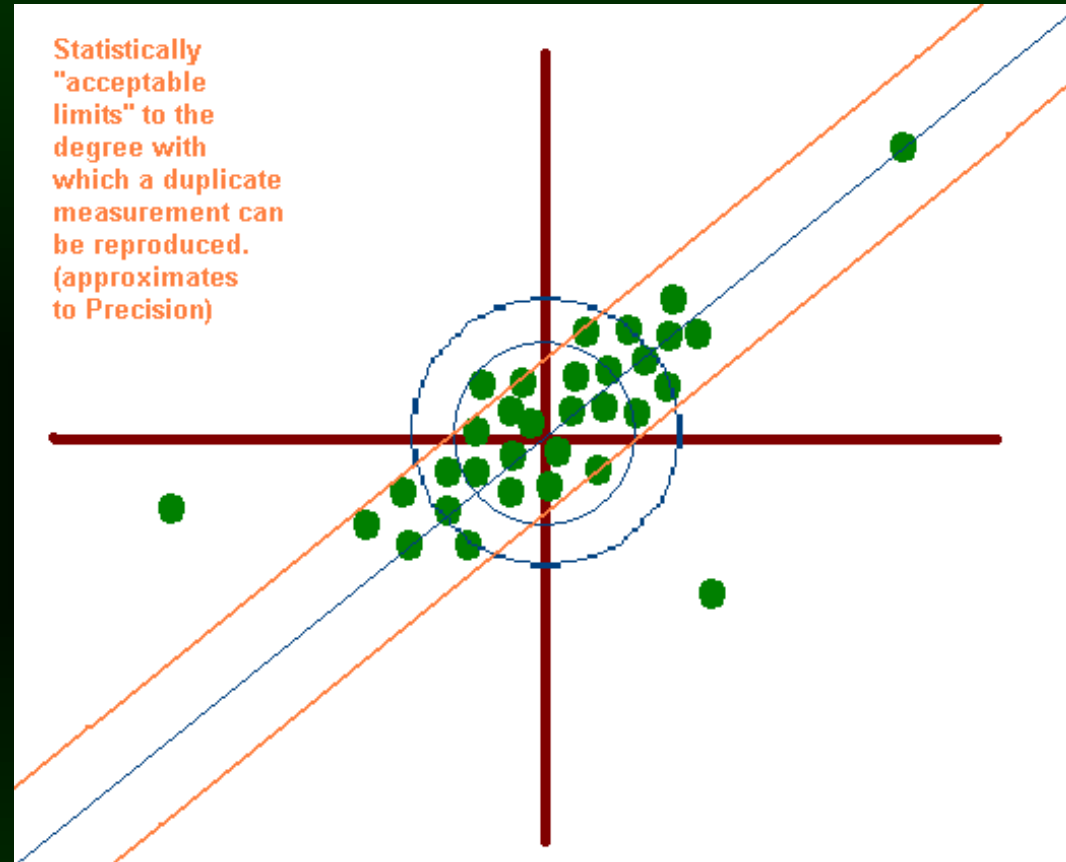
- The statistically “correct” results will fall in a kind of target board around the intersection of the two lines



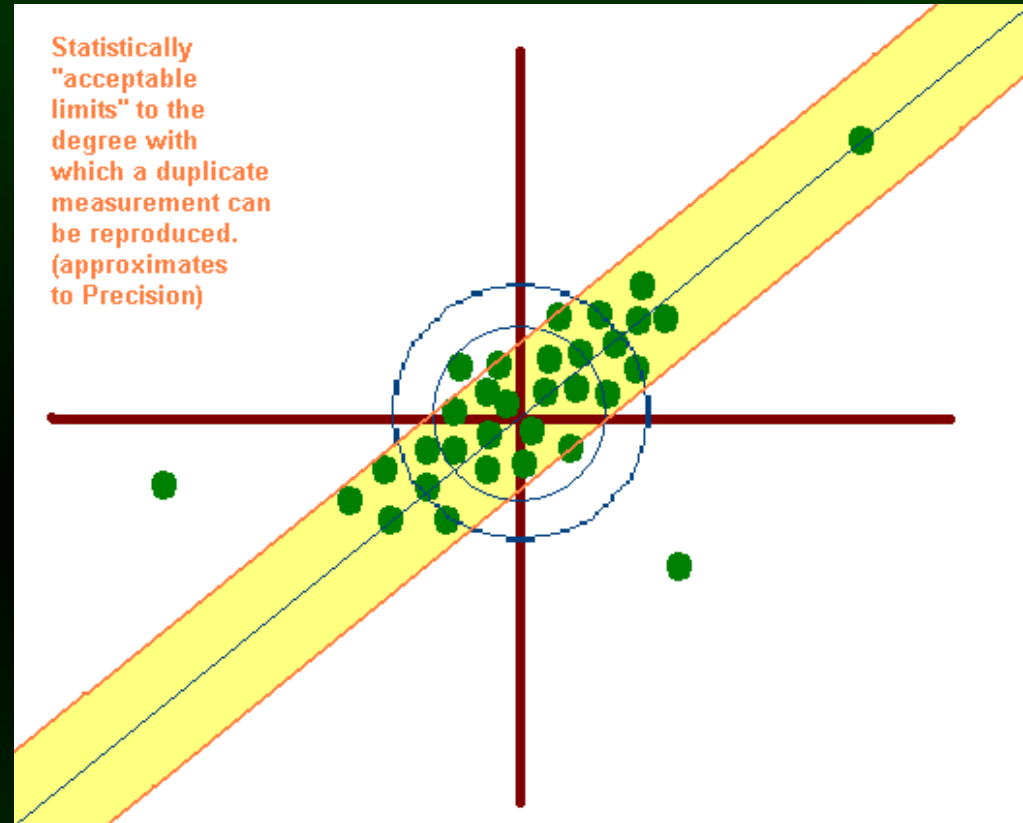
- According to the confidence limit set, there may be further “allowances” for accuracy.

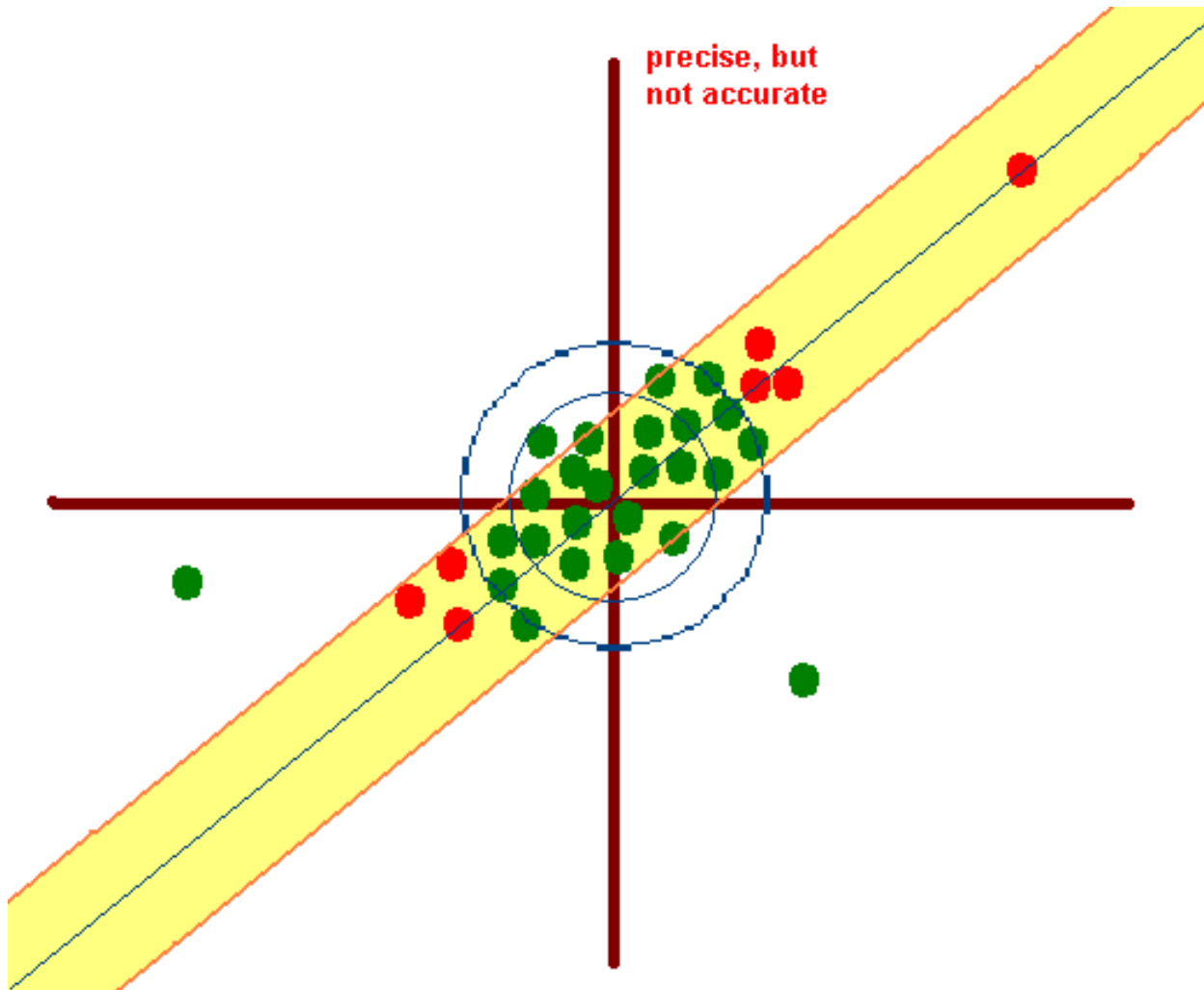


- An instrument or operator may be able to reproduce a result with very close precision



- The results within the yellow band are precise, but not necessarily accurate





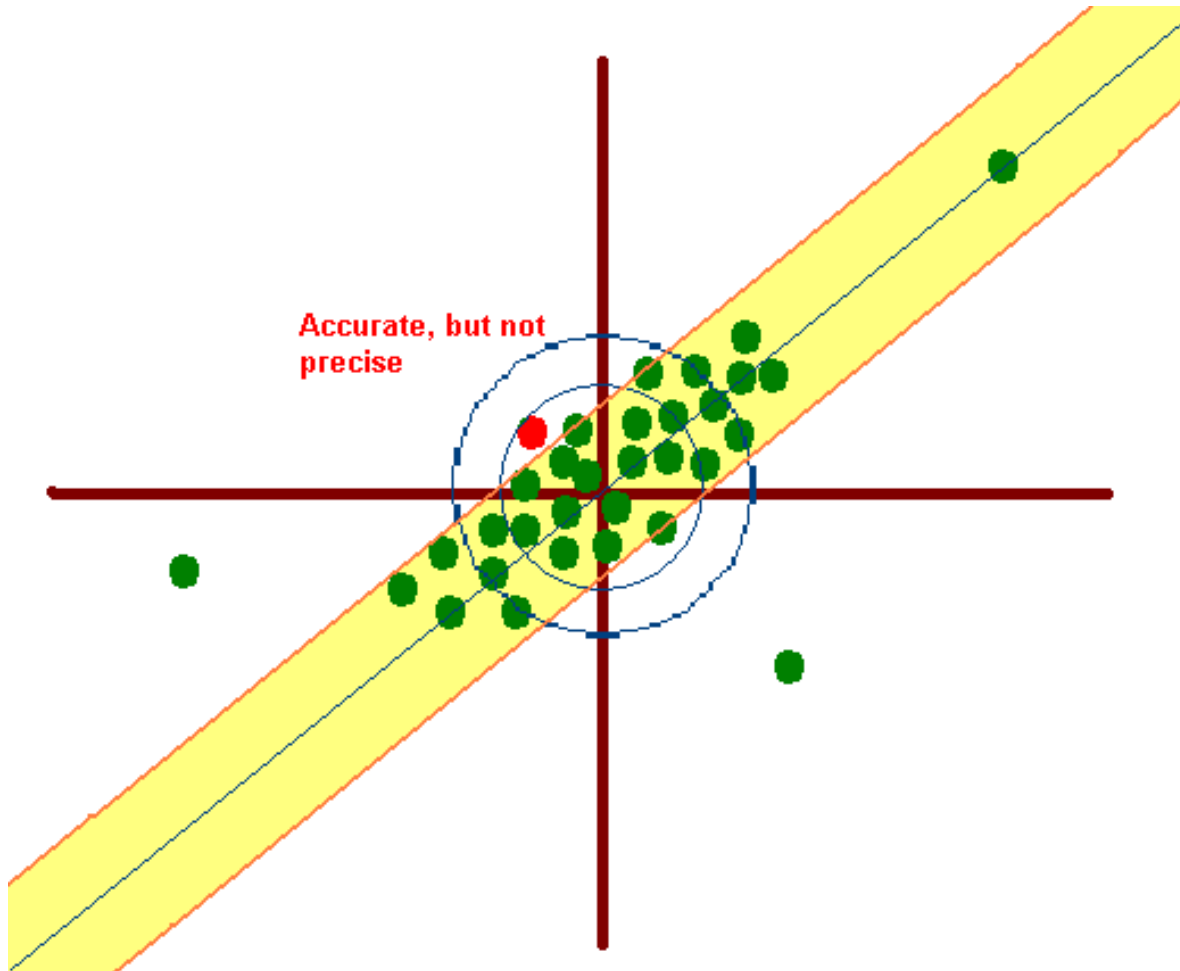
- Acceptable “within” laboratory Z score, but unacceptable “between” laboratory Z score.

Possible Examples for “good” within but “bad” between Z scores

- Instruments are OK, but approach/method used is not in line with other labs
- Instrument may be very good, but out of calibration (slope good, but amplitude incorrect – eg tare of balance)
- There is a bias (equipment, environment or operator related) which causes a consistent trend in one direction

- Errors occurring in this pattern are usually systematic.
- (There is usually a pattern or a bias)
- These types of errors are often more easily explained (not necessarily more easily fixed!)



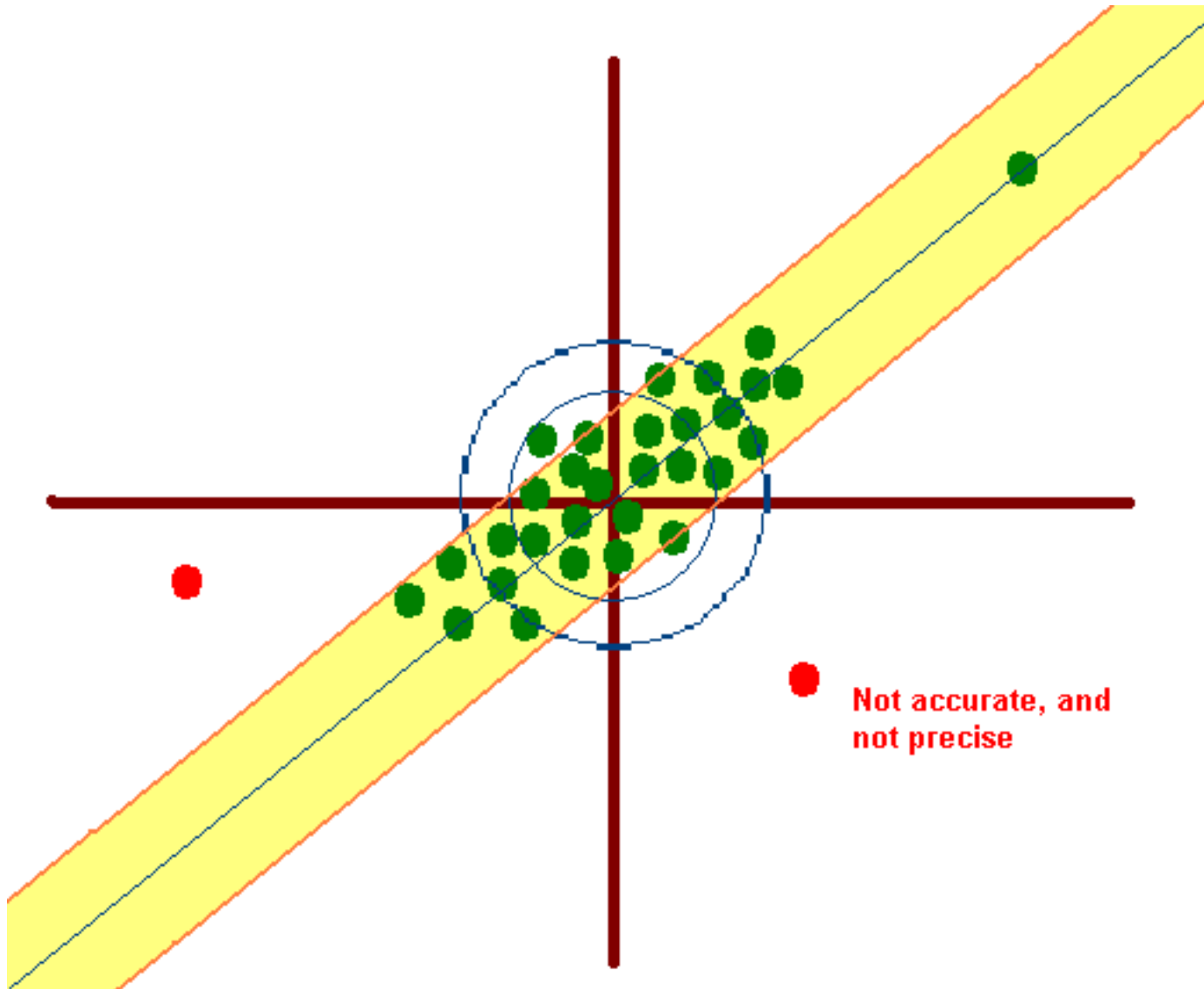


- Between laboratory Z score is acceptable, but within laboratory Z score not acceptable

Possible examples for “good” between but “bad” within Z scores

- Instrument may not be sensitive enough for the determination

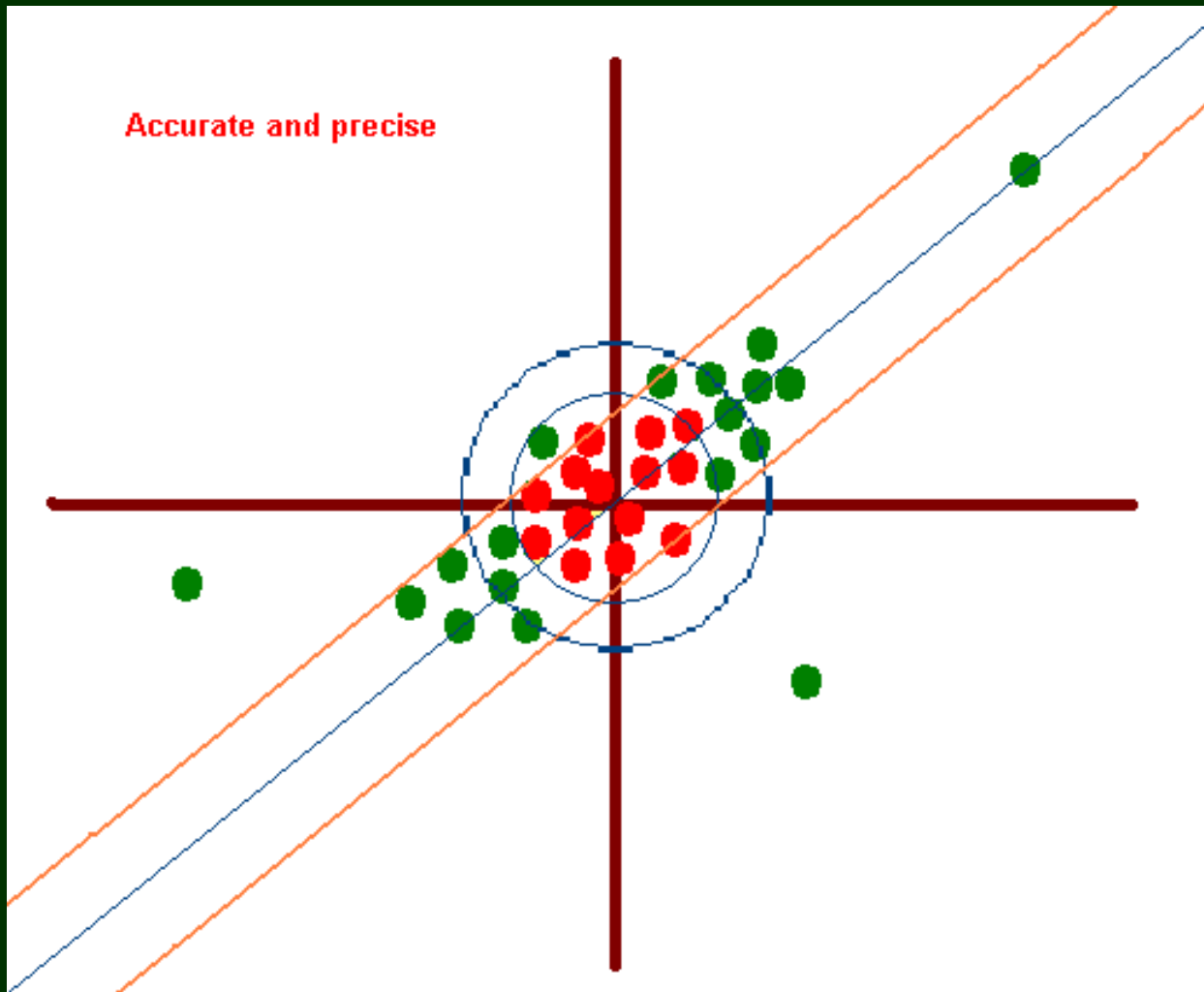
- Both within and between laboratory Z scores will be unacceptable



Both within and between Z scores “bad”

- In this situation everything should be thoroughly reviewed.





- Both within and between laboratory Z scores are acceptable

• J

Random
errors

Systematic
errors

Systematic
errors

Random
errors

When you receive the report.....

- Read it
- Compare your laboratory's performance with the evaluation criteria
- Plot your results on a graph to determine whether your laboratory is more prone to random or systematic errors
- Report to your staff



Do you have to perform “follow-up”?

- Don't panic
- The usual quality principles apply:
 - Evaluation of the nature of the problem
 - Cause analysis
 - Solution building
 - Measure success of solution implementation



Check that the reported result is actually what was obtained

- Most 'errors' are in reporting such as:
 - sample mix-up,
 - decimal places,
 - reading scales on meters,
 - number transposition (3.52 vs 5.32)
 - Reading graphs

Check equipment and methodology

- Was the correct method used?
- Were instructions followed?
- Was equipment in calibration (and did controls perform appropriately?)
- Were staff trained?
- Can your questionnaire response assist in identification of possible problems?



Try again

- Irrespective of whether a possible cause for error was found, repeat the test.
- A cause of error is not always found.
- Repeating the test can restore confidence
 - OR (if still unsuccessful), can confirm that more investigation is needed
 - (If possible, test a sample where the result is not known by the laboratory.)