

**Food for thought:  
Statistical limitations on the  
power of validation**

# Sensitivity and specificity

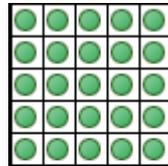
- ▶ Can only be measured relative to a chosen 'gold standard' that defines what is known to be present
  - Widely considered to be sequencing but should more properly include a range of techniques with different capabilities.
  - NB it is likely that the gold standard in itself is flawed (i.e. <100% sensitive)

|                     |   | Gold standard |    |
|---------------------|---|---------------|----|
|                     |   | +             | -  |
| Experimental result | + | TP            | FP |
|                     | - | FN            | TN |

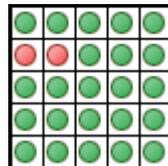
TP = True positive  
FP = False positive  
FN = False negative  
TN = True negative

- **SENSITIVITY** – Proportion of gold standard +ves correctly identified  $TP/(TP+FN)$
- **SPECIFICITY** – Proportion of gold standard -ves correctly identified  $TN/(TN+FP)$

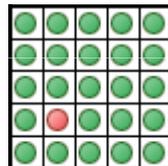
# Example



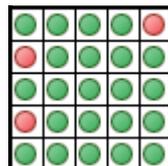
Sensitivity =  $25/25 = 100\%$



Sensitivity =  $23/25 = 92\%$



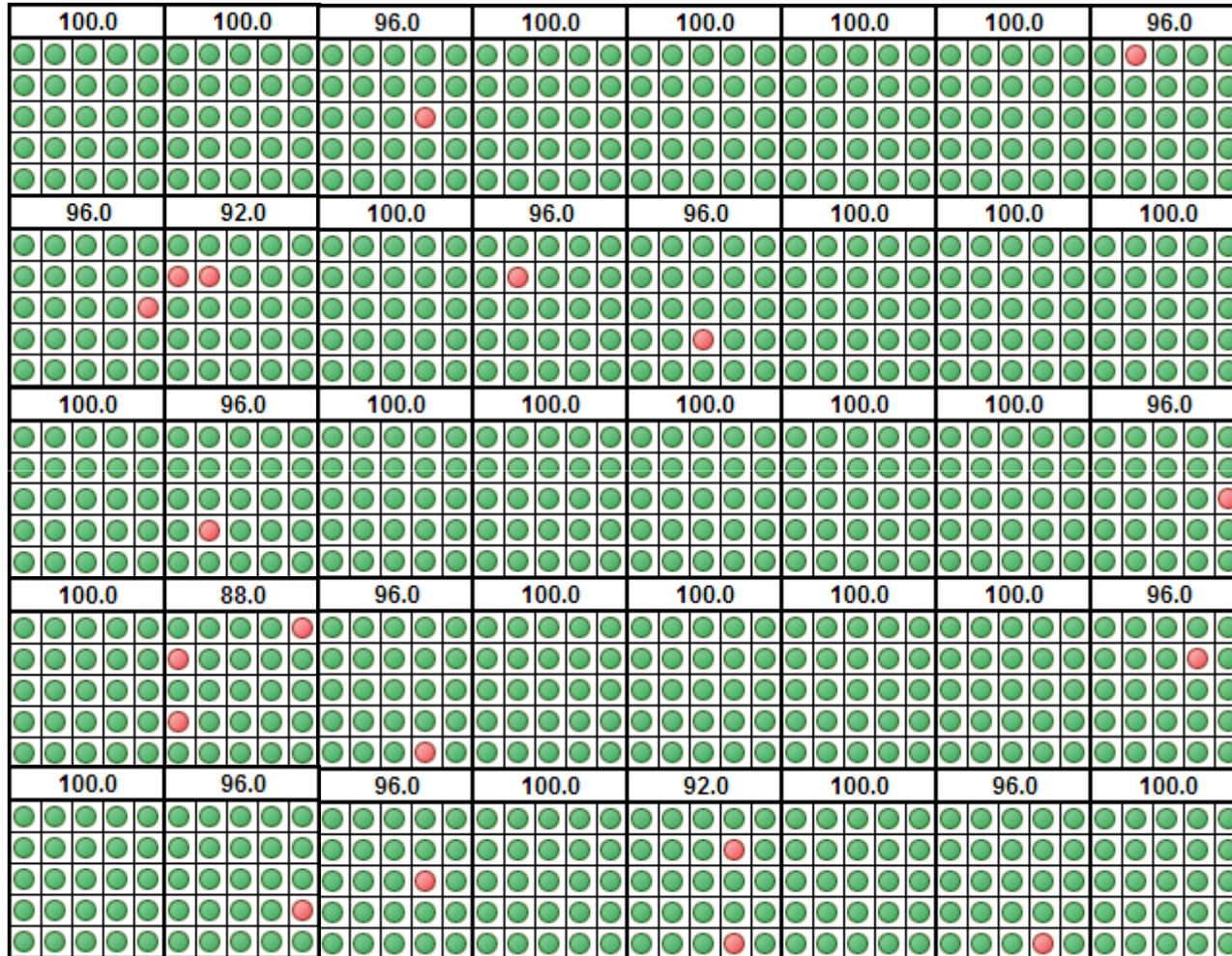
Sensitivity =  $24/25 = 96\%$



Sensitivity =  $22/25 = 88\%$

Overall sensitivity =  $984/1000 = 98.4\%$

# Example



Overall sensitivity =  $984/1000 = 98.4\%$

# Sensitivity

|   | Sensitivity |
|---|-------------|
| <p>More than 500 mutations have been identified in the CFTR gene, making it an excellent system for testing mutation scanning techniques. To assess the sensitivity of denaturing gradient gel electrophoresis (DGGE), we collected a representative group of 202 CFTR mutations. All mutations analyzed were detected by scanning methods other than the DGGE approach evaluated in this study. DGGE analysis was performed on 24 of the 27 exons and their flanking splice site sequences. <b>After optimization, 201 of the 202 control samples produced an altered migration pattern in the region in which an alteration occurred. The remaining sample was sequenced and found not to have the reported mutation.</b> The ability of DGGE to identify novel mutations was evaluated in three Asian CF patients with four unknown CF alleles. Three novel Asian mutations were detected-K166E, L568X, and 3121-2 A--&gt;G (in homozygosity)-accounting for all CF alleles. <b>These results indicate that an optimized DGGE scanning strategy is highly sensitive and specific and can detect 100% of mutations.</b></p> | <b>100%</b> |
| <p>...<b>A larger set of 32 mutant DNA specimens was then analyzed using these optimized tandem CAE-SSCP/HA protocols and materials and yielded 100% sensitivity of mutation detection...</b></p>   | <b>100%</b> |

# Sensitivity

- ▶ The two papers looked at different numbers of mutations (201 and 32 respectively) and are therefore not comparable – they have different ‘power’.
- ▶ Although all mutations in the studies were found this does not mean all possible mutations would be.

## Rule of three (approximation but accurate)

We can say with 95% confidence that the probability of a false negative, given a study of **n** samples with no false negatives, is **3/n**.

∴ for n = 201 (with no false negatives)

$$\begin{aligned} \text{probability of a false negative} &= 3/201 \\ &= 0.015 \\ &= 1.5\% \end{aligned}$$

∴ With 95% confidence sensitivity  $\geq 98.5\%$

- ▶ If you want to be 99% confident the rule is **4.6/n**
- ▶ THESE RULES APPLY TO ANY PROPORTIONAL MEASUREMENTS (I.E. SPECIFICITY)

# Sensitivity

|  | Sensitivity with 95% confidence |
|--|---------------------------------|
| <p>More than 500 mutations have been identified in the CFTR gene, making it an excellent system for testing mutation scanning techniques. To assess the sensitivity of denaturing gradient gel electrophoresis (DGGE), we collected a representative group of 202 CFTR mutations. All mutations analyzed were detected by scanning methods other than the DGGE approach evaluated in this study. DGGE analysis was performed on 24 of the 27 exons and their flanking splice site sequences. <b>After optimization, 201 of the 202 control samples produced an altered migration pattern in the region in which an alteration occurred. The remaining sample was sequenced and found not to have the reported mutation.</b></p> <p>The ability of DGGE to identify novel mutations was evaluated in three Asian CF patients with four unknown CF alleles. Three novel Asian mutations were detected-K166E, L568X, and 3121-2 A--&gt;G (in homozygosity)-accounting for all CF alleles. <b>These results indicate that an optimized DGGE scanning strategy is highly sensitive and specific and can detect 100% of mutations.</b></p> | <b>≥98.5%</b>                   |
| <p>... <b>A larger set of 32 mutant DNA specimens was then analyzed using these optimized tandem CAE-SSCP/HA protocols and materials and yielded 100% sensitivity of mutation detection...</b></p>   | <b>≥91.6%</b>                   |

# Considerations

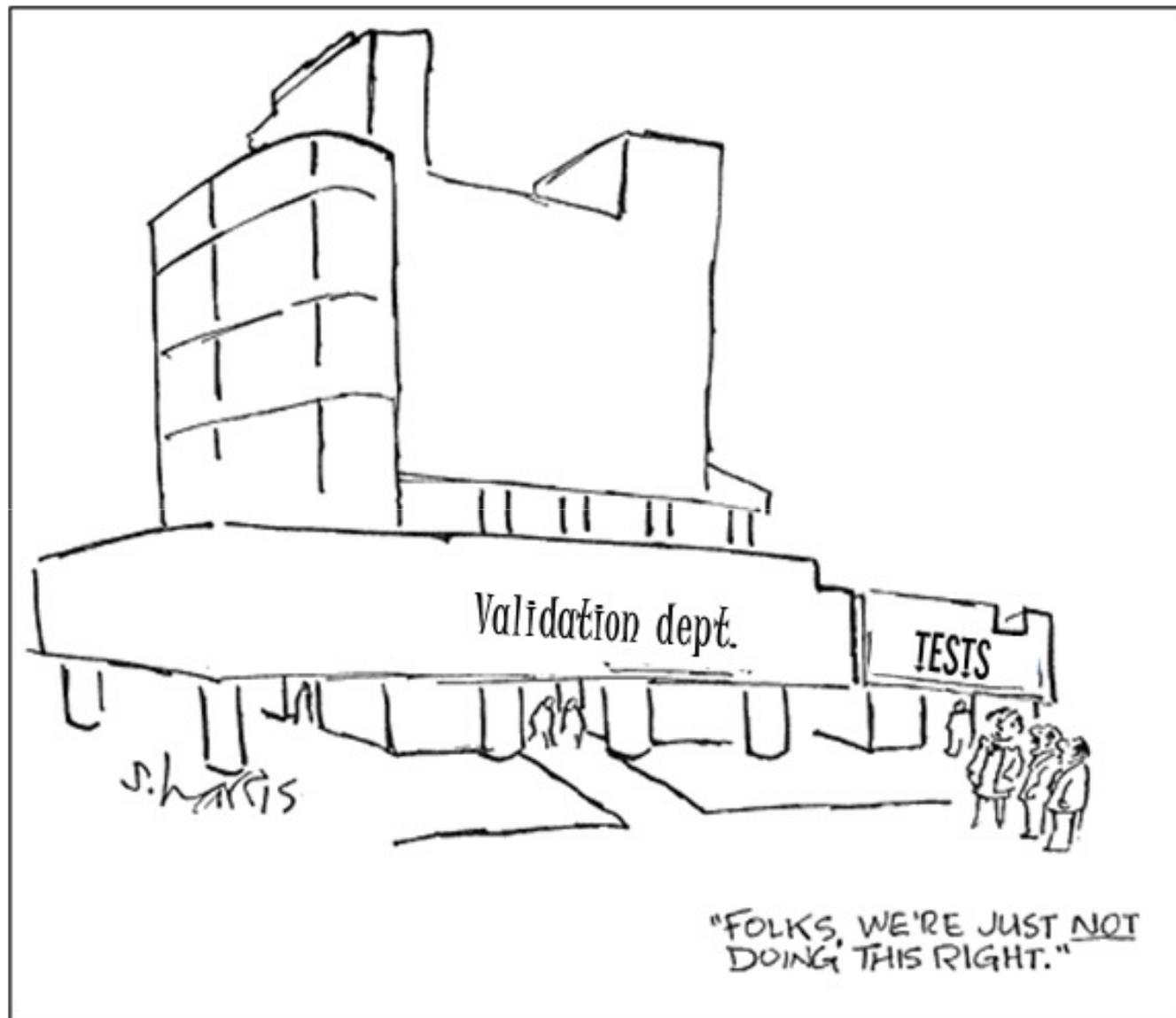
- ▶ Ideally validation studies should be carried out with enough samples to show the required sensitivity (and specificity)
- ▶ Normally studies should be balanced  
i.e. similar numbers of positive and negatives give equal power to estimate  $Se$  and  $Sp$
- ▶ Analyses should be blinded
- ▶ The positives used should not have been used to optimise the assay
- ▶ Positives should be representative of what is expected in practice
- ▶ If any positives are missed more complex statistics are required to calculate estimates of accuracy

**However it is often the case that insufficient positive controls are available to show the desired sensitivity.**

- ▶ In such cases it may be necessary to perform inter-laboratory study to gain sufficient positive controls
- ▶ Most importantly results should be reported honestly: This means results should always be given as a range of values with a confidence level

e.g. **≥98% (95% confidence)** or **85 to 95% (95% confidence)** if that is the case

**Validation must be practical!**



"FOLKS, WE'RE JUST NOT  
DOING THIS RIGHT."