The validation of qualitative test methods

Tom Fearn

Department of Statistical Science, UCL

Types of result

- Qualitative
  - Binary (0/1) = absent/present, fail/pass
  - Ordinal, eg low, moderate, high
- Semi-quantitative
  - <1, 1-10, >1, . . .
- Quantitative
  - Measure of concentration on a continuous scale

Issues for validating qualitative tests

- The is very little information in one binary result
- Different statistical methods are needed for dealing with such data
- Study design
  - Types and numbers of samples?
- Reporting
  - What to report and how to analyse?

UCL

Types and numbers of responses

Actual	Observed		
	Negative	Positive	
Negative	TN	FP	Ntot
Positive	FN	TP	Ptot

UCL

Reliability measures

False positive rate = FP/Ntot
Specificity = TN/Ntot = 1 – False positive rate

False negative rate = FN/Ptot
Sensitivity = TP/Ptot = 1 – False negative rate

UCL

Reliability measures – some comments

The overall 'accuracy' on any population of samples

accuracy = p × sensitivity + (1 – p) × specificity

depends on the proportion p of actual positives in the population. Need to know both specificity and sensitivity in order to generalise.

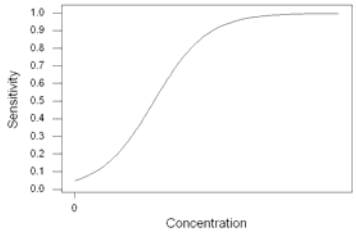
UCL

Reliability measures – some comments

- Both sensitivity and specificity may depend on sample type and the presence of interferents
- Sensitivity will usually depend on the amount of analyte present in the sample. This has implications for
  - study design
  - interpreting 'average' sensitivity from a study

UCL

Sensitivity as a function of concentration



UCL

Study design – choice of samples

- Aim to provide estimates of specificity, and sensitivity at a range of analyte concentrations
  - Desirable to include a blank sample
  - Concentration should be known for positives
  - Do not just use 'easy' samples
  - Use ongoing studies to build up a picture of the sensitivity curve

UCL

Study design – numbers of replicates

- Need substantial numbers of replicates to establish a sensitivity with and degree of precision
- For example, the 95% confidence limits for a true proportion when we observe 4 successes out of 5 are (0.41, 0.98), for 16 out of 20 they become (0.60, 0.93).

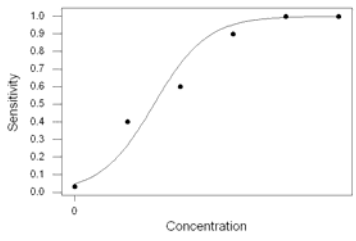
UCL

Reporting

- Report numbers of positive and negative results for each sample
- If relevant report the concentration for each positive sample
- Fine to invent scoring rules – tailor to the specific context

UCL

Modelling sensitivity as a function of concentration – logistic or probit regression



UCL

Conclusions

- The validation of qualitative test methods needs special approaches
- These need not be complicated – in fact the simpler the better
- A reference

  S.L.R. Ellison and T. Fearn, Trends in Analytical Chemistry, 24, 468-476, 2005.