

University of York
Department of Health Sciences
Applied Biostatistics

Exercise: Comparing the means of small samples

Question 1.

It was postulated that the increasing rate of hip fractures among elderly women may be related to women now having longer femoral necks than in previous generations. 52 x-ray pictures of hips of elderly women taken in the 1950s were compared to 52 similar pictures taken in the 1990s. All the x-rays had been taken for routine diagnostic purposes with the same equipment in the same rheumatology unit (Reid *et al.* 1994). The following table appeared :

Mean (standard deviation) dimensions of proximal femur in 52 elderly white women in New Zealand in 1950s compared with those of 52 in 1990s

Dimension (mm)	1950s	1990s	P value
Length of hip axis	124.0 (8.6)	130.5 (8.6)	0.0002
Length of femoral neck	79.4 (7.6)	84.9 (6.3)	0.0001
Width of femoral neck	38.1 (4.1)	38.6 (3.6)	0.49

- a) What method could be used to calculate P in this study?
- b) What conditions, if any, do the data have to fulfil for the method to be valid?
- c) Are they likely to be fulfilled here?
- d) From these P values, can we conclude that the length of femoral neck in elderly women has increased over time?
- e) Can we conclude that the width of femoral neck in elderly women has not increased over time?

Question 2.

Shortly after the grounding of the Braer oil tanker off the Shetland Isles, a study was conducted to ascertain whether the respiratory tracts of children were being affected by the crude oil vapour and droplet spray. Peak expiratory flow rate (PEFR) of 44 children ages 5-12 years was measured twice: three days after the shipwreck and 9 to 12 days after when the strong smell of oil had abated. Statistical analysis of the paired PEFRs (by Student's t test for difference of means for paired samples) showed no significant difference between the two sets of values ($P=0.502$) (Crum 1993).

- a) What is meant by no significant difference?
- b) What is meant by a paired t test?
- c) What assumptions are involved and are they likely to be justified here?
- d) What can we conclude about the effect of spillage on the respiratory function of children?

Question 3.

The following table shows the serum cholesterol concentrations of 20 smokers and 20 non-smokers, matched for age and sex, recruited for a study of glucose response. Statistical analysis was by two-tailed, unpaired Student's t test (Facchini *et al.* 1992).

Cholesterol concentration, mean (SEM) in mmol/l			
	Smokers	Non-smokers	t test
Total plasma	4.49 (0.19)	4.48 (0.17)	NS
Very low density lipoprotein (VLDL)	0.45 (0.06)	0.23 (0.04)	P<0.005
Intermediate density lipoprotein (IDL)	0.18 (0.03)	0.22 (0.04)	NS
Low density lipoprotein (LDL)	2.72 (0.17)	2.53 (0.13)	NS
High density lipoprotein (HDL)	1.16 (0.05)	1.51 (0.08)	P<0.005

- What is meant by a 'two-tailed, unpaired Student's t test'?
- What conditions must the data satisfy for these t tests to be valid?
- Are these likely to be satisfied here?
- What extra information could be given in the table?
- What aspect of the data has been ignored in the analysis?

References

- Crum, J.E. (1993) Peak expiratory flow rate in schoolchildren living close to Braer oil spill. *British Medical Journal* **307**, 23-4.
- Facchini, F.S., Hollenbeck, C.B., Jeppesen, J., Chen, Y.D., and Reaven, G.M. (1992) Insulin resistance and cigarette smoking. *Lancet* **339**, 1128-30.
- Reid, I.R., Chin, K., Evans, M.C., and Jones, J.G. (1994) Relation between increase in length of hip axis in older women between 1950s and 1990s and increase in age specific rates of hip fracture. *British Medical Journal* **309**, 508-9.

Questions from Martin Bland and Janet Peacock: *Statistical Questions in Evidence-based Medicine*, Oxford University Press, Oxford, 2000.