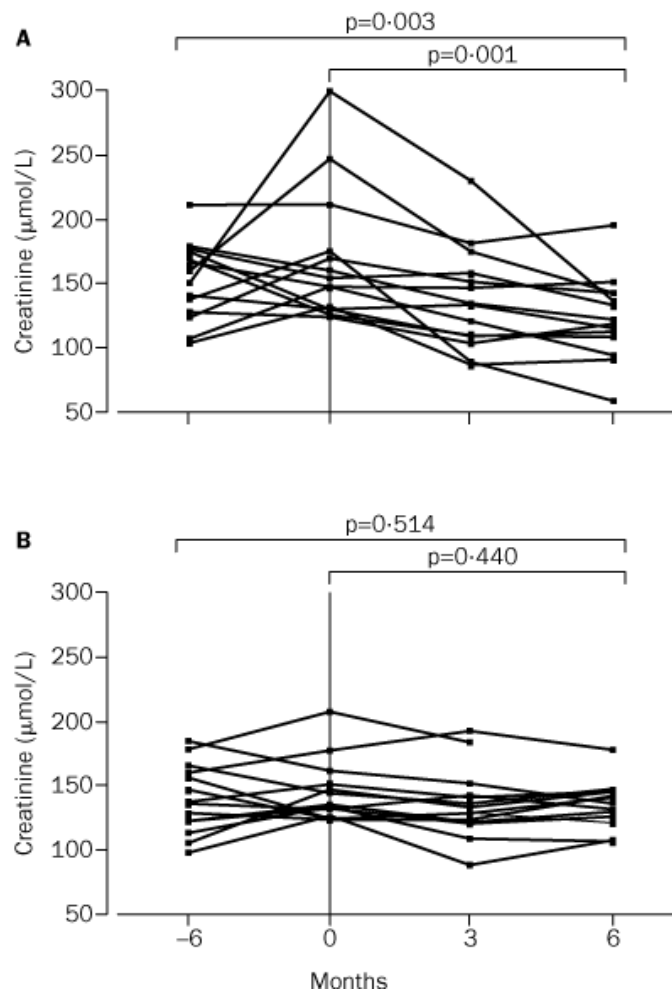


Clinical Biostatistics

Exercise: Comparing means

Liver transplant patients require drugs to suppress their immune systems and prevent rejection of the new liver. Renal dysfunction is a major complication of long-term immunosuppressive therapy with calcineurin inhibitors (CNI). In this study, 28 people who had had renal dysfunction attributable to suspected CNI toxicity were randomised to either replacement of CNI with mycophenolate mofetil (study patients, group A); or to remain on CNI immunosuppression (controls, group B). Renal function, blood pressure, uric acid, and blood lipids were measured before and 6 months after study entry (Schlitt *et al.*, 2001). The following figure was given:



- What method could be used to carry out the tests of significance shown in the figure, and why?
- What can we conclude from these tests?
- What test of significance would be better in this study?

The authors report: “At the end of the study, mean (SD) serum creatinine had fallen by 44.4 (48.7) $\mu\text{mol/L}$ in study patients compared with 3.1 (14.3) $\mu\text{mol/L}$ in controls; a mean difference of 41.3 $\mu\text{mol/L}$ (95% CI 12.4-70.2).”

- (d) What method would be used to calculate the confidence interval, and why? What condition should the data meet for this method?
- (e) The standard deviations are bigger than the means. Why should we NOT conclude that change in serum creatinine has a skew distribution?

In a study of factors predicting male fertility, conventional semen indices were measured in donors to an artificial insemination clinic. By evaluating the pregnancies that resulted, donors could be graded as to their fertility. The aim was see whether any of these factors could be used as a test for high fertility.

Two sample t tests were carried out comparing the means of the variables between the two groups. No significant differences were found in semen indices for successful and unsuccessful donors.

Semen Indices In Most And Least Fertile Donors						
	Successful donors			Unsuccessful donors		
	No.	Mean	(SD)	No.	Mean	(SD)
Volume (ml)	17	3.14	(1.28)	19	2.91	(0.91)
Semen count ($10^6/\text{ml}$)	18	146.4	(95.7)	19	124.8	(81.8)
% motility	17	60.7	(9.7)	19	58.5	(12.8)
% abnormal morphology	13	22.8	(8.4)	16	20.3	(8.5)

The study concluded that conventional semen analysis may be too insensitive an indicator of fertility potential to enable us to predict which individuals will provide useful sperm donations (Paraskevaides *et al.*, 1991).

- (f) Is there anything to suggest that the t tests may not be valid?
- (g) What are the implications of this for the t test? What could be done about it?
- (h) Are the t tests important to the conclusions of the study? In other words, if the sample were large enough for these differences to be significant with these means and standard deviations, would this change the conclusion?

References

Paraskevaides, E.C., Pennington, G.W., Naik, S., and Gibbs, A.A. (1991) Pre-freeze/post-freeze semen motility ratio. *Lancet* **337**, 366-7.

Schlitt HJ, Barkmann A, Boker KHW, Schmidt HHJ, Emmanouilidis N, Rosenau J, Bahr MJ, Tusch G, Manns MP, Nashan B, Klempnauer J. (2001) Replacement of calcineurin inhibitors with mycophenolate mofetil in liver-transplant patients with renal dysfunction: a randomised controlled study. *Lancet* **357**, 587-591.