Statistics at Square One

XIII—The t tests (concluded)

T D V SWINSCOW

British Medical Journal, 1976, **2**, 408-409

(4) Difference between means of paired samples

When comparing the effects of two alternative treatments or experiments it is sometimes possible to make comparisons in pairs. For example, we may want to compare a new treatment with a traditional treatment or a new test with a standard one. Two courses are possible. Firstly, we may use the same sample twice over, so that each member of it receives both treatments, acting as his own control. Secondly, we may draw two samples and, before any treatment is started, pair each member of one with a member of the other, giving treatment A to all the members of sample 1 and treatment B to their pairs in sample 2.

The first case to consider is when each member of the sample acts as his own control. Whether treatment A or treatment B is given first or second to each member of the sample should be determined by the use of a table of random numbers. In this way any effect of one treatment on the other, even indirectly through the patient's attitude to treatment, for instance, can be minimised. Occasionally it is possible to give both treatments simultaneously, as in the treatment of a skin disease by applying a remedy to the skin on opposite sides of the body.

For example, Dr Silver is continuing his studies of bran in the treatment of diverticulosis. Having chosen the preparation that he found induced the shorter alimentary transit time (Part XII), he wonders whether this transit time would be even shorter if the bran is given in the same dosage in three meals during the day (treatment A) or in one meal (treatment B). He chooses a random sample of patients with disease of comparable severity and aged 20-44, and administers the two treatments to them on two successive occasions, the order of the treatments also being determined from the table of random numbers. The alimentary transit times and the differences for each pair of treatments are set out in table 13.1.

TABLE 13.1—*Transit times in hours of marker pellets through alimentary canal of 12 patients with diverticulosis on two types of treatment: paired comparison*

Patient	Transit times in hours: Treatment A	Transit times in hours: Treatment B	Difference A–B
1	63	55	8
2	54	62	−8
3	79	108	−29
4	68	77	−9
5	87	83	4
6	84	78	6
7	92	79	13
8	57	94	−37
9	66	69	−3
10	53	66	−13
11	76	72	4
12	63	77	−14
Total	842	920	−78
Mean	70·17	76·67	−6·5

In calculating t on the paired observations we work with the difference, d, between the members of each pair. Our first task is to find the mean of the differences between the observations and then the standard error of the mean, proceeding as follows:

Find the sum of the differences	Σd
Find the mean of the differences	$\bar{d}$
Find the sum of the squares of the differences	Σd^2 (1)
Find the square of the sum of the differences	$(\Sigma d)^2$
Divide the square of the sum of the differences by the number (n) of differences	$\frac{(\Sigma d)^2}{n}$ (2)
Subtract (2) from (1)	$\Sigma d^2 - \frac{(\Sigma d)^2}{n} = \Sigma(d-\bar{d})^2$
Divide by the number of degrees of freedom (n − 1)	$\frac{\Sigma d^2 - \frac{(\Sigma d)^2}{n}}{n-1}$
This gives the square of the standard deviation (variance)	SD^2
Divide the variance by the number of differences	$\frac{SD^2}{n}$
Take the square root	$\sqrt{\frac{SD^2}{n}}$

This is the standard error of the mean of the differences.

To calculate t, divide the mean of the differences by the standard error of the mean:

$$t = \bar{d} \div \sqrt{\frac{SD^2}{n}}.$$

The table of the t distribution is entered at n − 1 degrees of freedom (number of pairs minus 1).

Dr Silver's figures are treated as follows:

n	12
Σd	−78
$\bar{d}$	−6·5
Σd^2	3030
$\frac{(\Sigma d)^2}{n}$	507
$\Sigma d^2 - \frac{(\Sigma d)^2}{n}$	2523
Divide by n − 1 to give SD^2	229·36

$$\text{SE mean difference} = \sqrt{\frac{SD^2}{n}} = \sqrt{\frac{229\cdot36}{12}} = 4\cdot37$$

$t = -6\cdot5 \div 4\cdot37 = -1\cdot487.$

British Medical Journal
T D V SWINSCOW, MSC, MB, deputy editor

Entering the table of the *t* distribution at 11 degrees of freedom (that is, n-1) and ignoring the minus sign, we find that this value lies between 0·695 and 1·796. Reading off the probability value, we see that $0.5 > P > 0.1$.

The null hypothesis is that there is no difference between the mean transit times on these two forms of treatment. Thus it is *not* disproved. We may therefore say that there is no convincing evidence of a difference between these two methods of administering this preparation of bran.

The second case of a paired comparison to consider is when two samples are chosen, and each member of sample 1 is paired with each of sample 2. Treatment A is then applied to all members of sample 1 and treatment B to all members of sample 2. The data are analysed in the same way as above for a single sample with paired treatments, but some thought needs to be given to the composition of the pairs.

Since the aim is to test the difference, if any, between two types of treatment, the choice of members for each pair is designed to make them as alike as possible. The more alike they are, the more apparent will be any differences due to treatment, since these will not be confused with differences in the results due to disparities between each member of the pair. The likeness between the pairs applies to attributes relating to the study in question. For instance, in a test of a drug for reducing blood pressure the colour of the patients' eyes would probably be irrelevant, but their resting diastolic blood pressure could well provide one basis for selecting the pairs. Another (and perhaps related) basis is the prognosis for the disease in patients; in general, patients with a similar prognosis are best paired. Whatever criteria are chosen, it is essential that the pairs are constructed before the treatment is given, for the pairing must be uninfluenced by knowledge of the effects of treatment.

Exercise 13. A new treatment for varicose ulcer is compared with a standard treatment on 10 matched pairs of patients by measuring the number of days from start of treatment to healing of ulcer. One doctor is responsible for treatment and a second doctor assesses healing without knowing which treatment each patient had. On the standard treatment the following treatment times were recorded as numbers of days: 35, 104, 27, 53, 72, 64, 97, 121, 86, 41; and on the new treatment the following: 27, 52, 46, 33, 37, 82, 51, 92, 68, 62. What are the mean difference in the healing time, the value of *t*, the degrees of freedom, and the probability? *Answer*: 15 days, $t = 1.758$, DF = 9, $0.5 > P > 0.1$.

Letter from . . . Victoria

Specialist registers and differential rebates

JOHN KNIGHT

British Medical Journal, 1976, **2**, 409-411

The establishment of a specialist register seems to generate a lot of heat. I was reminded of this while reading M D Vickers's paper, under the heading "For Debate," condemning the setting up of a British specialist register as unnecessary and a possible danger to the high standard of British medicine.[1] Interestingly enough in that issue of the *BMJ* there was a letter from D W Sturdee *et al* objecting to the lable "Junior Hospital Doctors" being applied to those who had finished their specialist training but still held non-consultant posts.[2] There still seems to be considerable confusion about who is or is not a specialist. Perhaps your readers might find the history of the specialist register in the State of Victoria of interest.

Specialist registers

The decision to set up specialist registers in Australia was economic. Historically, specialists have charged more for their services than have general practitioners. Australian health care has been subsidised by the Commonwealth (central) Government since 1949. A subsidy was available to those who had insured themselves with a health benefit society. The Government contributed half the amount the patient got back when he presented his receipted doctor's account to the society. The contributors' funds provided the other half, which was limited to a fixed sum for each service and could not exceed 90% of the charge. By 1969 the gap between fee and rebate was quite large for GP procedures and even larger for specialist fees. In 1970 the Government introduced legislation to reduce the size of the gap. But, to achieve this and keep the cost of the scheme within bounds, there had to be changes. There were two main changes: rebates were to be higher for consultations with a specialist if the patient was referred by a GP; and some operations would attract higher rebates if the operation was performed by a specialist—but again only if the patient was referred by a GP.

When the Federal Government introduced its differential rebates there had to be a list of names of specialists whose patients, if referred by their GP, would be entitled to a higher rebate on their bill. Victoria at first had no register and so, as in other States without a specialist register, the Commonwealth Government compiled a list of specialists for medical benefit purposes.

Grey Street Anaesthetic Group, East Melbourne, Victoria, Australia
JOHN KNIGHT, FFARACS, specialist anaesthetist

Defining a specialist

The definition of a specialist caused some problems. At first the possession of a higher degree or diploma and confining one's practice to a specialty seemed reasonable. Then it was pointed out that this was unfair to those practitioners who had began