A randomised controlled trial comparing the hair apposition technique with tissue glue to standard suturing in scalp lacerations (HAT study)

Ong Eng Hock M, Ooi SBS, Saw SM, Lim SH

Annals of Emergency Medicine (2002) 40, 19-26

Was the assignment of subject to treatments randomised?

Yes, Randomisation was generated using a table of random numbers and envelopes produced containing the allocation. The authors do not discuss how this was dealt with on the two sites but provision of envelopes could have been organised so that they contained the same proportion of each treatment. This is known as stratification.

Stratification is done to ensure that individual sites are not allocated, by chance, a disproportionate number of one particular treatment. If there was a difference in outcome it would be difficult to exclude some variable associated with the site.

Was the potential assignment concealed?

Yes. The envelopes containing the allocation were sealed and opaque and had to be opened in number order. This was not done until written consent had been obtained.

Were all the subjects who entered the trial properly accounted for and attributed at its conclusion? Was follow-up complete?

Yes. 189 were randomised, 188 were assessed and 1 lost to follow-up.

Did the authors justify their use of follow-up time?

Follow-up was one week at which point sutures were removed or the wound inspected. The authors felt that follow-up longer than one month would result in a large loss to follow-up although they continued to review those with problems until all had been resolved. The possibility of longer term problems such as cosmetic acceptability are probably unimportant because of the hair coverage.

Were subjects, researchers and healthcare personnel 'blind' to the intervention or treatment?

It was not possible to blind subjects, researchers or healthcare personnel and all of these could have introduced bias because of the lack of blinding. The authors mention that the primary investigators were not always around to ensure proper standardisation of wound preparation but this probably makes the assessment slightly more realistic.

A single physician assessing the outcome also could not be blinded but he or she had not been involved in the trial until that point and this may have assisted in making the assessment more objective.

Were the subjects similar at the start of the trial?

Baseline characteristics are shown in the table. No formal statistical analysis has been reported. This is reasonable since, if randomisation had been done properly then any differences seen would be due to chance – and indeed we would expect a large difference once in about every twenty characteristics looked at.

A visual comparison suggests that there is probably no difference in the groups. Although the male-female ratio seems a bit different these are relatively small numbers and a formal analysis shows that there is no statistically significant difference. A difference as great as the one seen in male-female ratio would be seen once in about every five times the experiment was repeated.

Aside from the intervention, were the subject groups treated equally?

No. Washing in the suture group was discouraged but the HAT groups were encouraged to wash after day 3.

Were data from the groups analysed according to randomisation? – i.e. intention to treat analysis?

This is stated explicitly and there were two people who should have had sutures given HAT instead. These people were both considered to have had sutures when the analysis was done.

Effect of intervention

What is the difference in measured outcome(s) between the groups?

The main outcomes were presence of complications, such as scarring, infection, bleeding, wound breakdown and healing. Secondary outcomes were length of treatment, pain perception and whether the patient would be happy to be treated again in the same way.

The results were expressed as percentages experiencing a complication and the two subtracted. A negative result implies reduced complications for the HAT group. The main differences reported were –14.2% for scarring and –4.3% for wound breakdown (-14.1% for all complications).

In addition to these differences the HAT procedure was faster (difference in median of 10 minutes), produced less reported pain (difference in median of pain score of 2 points) and was more likely to be welcomed again (difference of 52% in those who would refuse the experimental treatment in future).

The results can be interpreted in terms of risk and this is normally presented as the risk of the experimental treatment in relation to the standard treatment. Therefore the **relative risk (RR)** of complication with HAT compared with suturing is 0.074 / 0.215 = 0.34. While doubling risk is relatively easy to appreciate, understanding 0.34 of a risk is not so this is often presented as a **relative risk reduction (RRR)** instead – 1 – 0.34 = 0.66 – a 66% reduction in risk with the HAT technique.

Another way of looking at the information is to look at the **actual risk reduction (ARR)**. For complications this was 0.215 – 0.074 = 0.141 (or 14.1% less) by using HAT instead of suturing.

This number is often converted into a **Number Needed to Treat (NNT)** which mathematically is 1/ARR corrected to the next whole number. 1 / 0.141 = 7.09 converted to 8. This means that treating 8 people with the HAT procedure would, on average, benefit 1 more (have 1 less person with complications) than treating 8 with suturing.

How likely is this to be due to chance?

Chance is measured by probability – the p value. For 'any complication' the p value = 0.005. This means that the difference seen in this study (7.4% giving a complication with HAT, compared with 21.5% with suturing), or a larger difference, would only be seen by chance 5 times in 1000, or 1 in 200 or 0.5%. Any p value less than 0.05 (5%) is considered statistically significant, or beyond chance.

Is a confidence interval quoted for the intervention?

Yes.

What does this imply?

Confidence intervals are useful because they provide two pieces of information – about the statistical significance and about the clinical importance of the difference in effect. It gives you the range where the **truth** is likely to lie.

For example, for complications was a relative risk reduction of 0.66 – this is the best estimate of the truth. This would be clearly important if it were true.

The 95% confidence interval is 0.25 – 0.85. We are 95% sure that the true reduction in risk is somewhere between a 75% reduction (even more important if true) and 15% (probably still clinically important). For any complication we can be 95% sure that HAT produces a relative difference of at least 15%.

If there were in truth no difference between the treatment the relative risk would be 1.00. Because the 95% confidence interval does not include the number 1.00 within its interval we can say that there is less than 5% chance that the true result is 1.00 (no difference). This is the same as a p value of 0.05 where we said that if there was in truth no difference (RR=1.00) then seeing the difference we got in the experiment by chance is less than 5%.

Therefore the confidence interval will tell you about the clinical importance of an effect – by examining how close either end of the confidence interval is to the no difference measure (RR = 1.0). It will also tell you about the statistical significance. If the no difference measure is <u>not</u> within the confidence interval then $p<0.05$.

Note that sometimes you will not be looking at relative differences, but actual differences, where two values are subtracted from one another and the no difference measure here will be 0, so we would then be concerned where that 0 was in relation to the confidence interval derived from the experiment.

For example, looking a wound breakdown, there is an actual difference of 4.3% (95% confidence interval for the difference of 0.1% to 8.5%). Because 0 is not within the 95% confidence interval we can be sure that this is statistically significant – $p<0.05$. Interpreting the bottom end of the confidence interval we see that the difference may, in truth, only be minimal (0.1%) and we cannot be convinced of its clinical importance in relation to wound breakdown. The problem here is likely to be due to a lack of power because wound breakdown was only seen in a few people. Increasing the study size will reduce the confidence interval.

A 95% confidence interval can also be calculated for the NNT. For complications it is 5 to 24. This means that we can be 95% sure that treating 24 or less with HAT will benefit one more than suturing.

Does this paper justify the likely benefits as worth the potential harms and costs (if any)?

Costs are not formally considered but they have highlighted the reduced costs of attendance time and the need to return for suture removal and have also referred to ways of saving on use of adhesive and the reduced risk of needlestick injury to the healthcare professional.

The technique is welcomed in being less painful, more acceptable because of reduced chance of hair loss and ability to wash sooner but cannot be performed in individuals with short hair (less than 3cm length) and is limited to blunt trauma type injuries without deep wounding.

Even though the study was done in Singapore there are unlikely to be any problems into translating it into a UK environment although hair length may be too short for many in the target group.