

Clinical trial by television

Martin Bland

Emeritus Professor of Health Statistics
University of York

International Biometrics Conference, Victoria, British Columbia, July 2016

The call of television

In the spring of 2013, I received a phone call from a researcher. This was not a health or education researcher, as often happens, but a researcher from a television production company. He did, however, want to talk about randomised clinical trials.

Outline Productions (www.outlineproductions.co.uk) were thinking about doing a series of programmes about home remedies, unconventional treatments for everyday conditions. One of the examples he mentioned was to treat male pattern baldness by rubbing into the scalp bull's semen. He didn't tell me whether the users had to collect the semen themselves, nor did he explain why human semen would not have done just as well. I thought that it would be easier to source.

The idea was to carry out a series of small, short duration, randomised control trials of some of these remedies. I thought that anything which shows clinical trials in a positive light would be welcome. On television, they are usually presented in a drama where either the results are being fiddled or where terrible consequences for participants are being concealed. A few short, entertaining trials where nothing bad happens might stimulate recruitment in the future. I also fancied being on television again, though that, it emerged, was not on offer.

I explained how I would do such trials. I even devised a simple statistical test for use in the analysis of a two-group trial of 20 participants or a one-group trial with 10 participants. After several conversations, he told me that he was leaving Outline and moving to another job.

A few weeks later a new researcher called. We started again.

Eventually, a series of six programmes was broadcast by Channel 4 in UK. The original working title had been *Health Freaks on Trial*. The final series was called *Health Freaks*.

Can duct tape shrink a verruca?

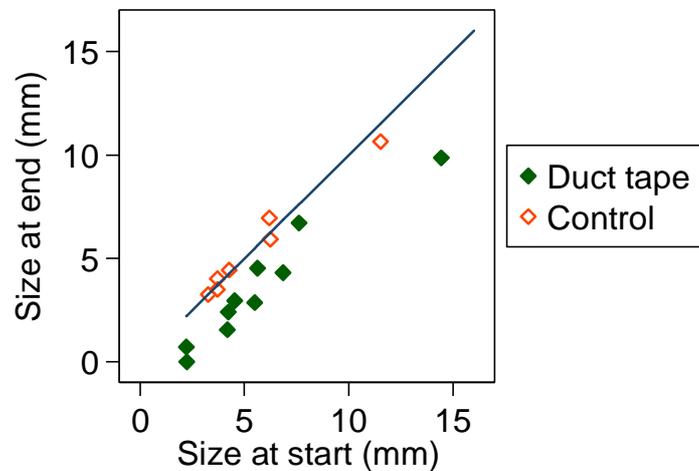
The first broadcast programme in the *Health Freaks* series described a trial of duct tape used to treat verrucae or warts. We saw an advocate of the treatment, who described how he cured a verruca on his foot by sticking a small piece of duct tape over it for six days, then changing it and repeating.

We tested this in a trial over a month. Participants were assigned to treatments by minimisation, using age group (<30 or 30+), gender, location of wart (foot or elsewhere), and duration of wart (<2 years, 2 to 5 years, 5+ years) as minimisation factors. Two groups of 11 participants took part, using surgical tape as a control treatment.

Ten duct tape and seven control participants completed the trial.

The primary outcome variable was the widest diameter of the verruca.

The figure shows the size of the wart or verruca at the end of the trial against the size at the beginning:



The diagonal line is the line of equality. Points above and to the left of this line show participants whose wart increased in size; points below and to the right of this line show participants whose wart decreased in size. Participants who were on active Treatment all had a reduction in the size of the wart, while participants on Control had very little change in wart size.

Analysis of covariance, the standard method of analysis for trials with an outcome and a baseline measurement, estimated the difference between Duct Tape and Control to be 2.1 mm, $P < 0.001$. So there was good evidence that Duct Tape was associated with smaller warts. More simply, all ten of the Duct Tape participants had reductions in size greater than any of the Control participants, $P = 0.0008$, rank sum test. I had never before been involved in a trial where the difference between treatments was so dramatic.

Participants were also asked about pain when the wart was squeezed. Results were very variable and, although the Duct Tape group appeared to do better, this was not statistically significant. One of the Duct Tape participants, and none of the Controls, reported that the verruca had disappeared, though this was not a significant difference. Significantly more Duct Tape participants than Controls reported difficulty in applying the treatment. Significantly more Duct Tape participants reported skin irritation.

My 14 seconds of fame

The original plan was that the programme's three research advisors would not appear on screen, but that we would have our names on the credits at the end. Channel 4 thought we should be seen, so we were filmed for the series. We sat at a table and chatted while a camera panned across us.

I told our University press office, they got a still of this scene and in the week of first broadcast put it on the University main webpage. Here it is:



The picture shows Rhiannon Pursall, a clinical trial coordinator, Richard Albardiaz, a medical general practitioner, and myself. I think I look like a garden gnome. In two of the programmes, this great scene was broadcast in a voiceover clip lasting 14 seconds.

Would I do it again?

Yes I would! Despite the time pressures, I really enjoyed doing it. All the people I worked with from Outline were lovely and a pleasure to work with.

I was disappointed by how little was made of the statistics, and how little time was given to the trials, but not completely surprised — I've done this before (for BBC *Horizon*).

Despite the small sample sizes, short durations, and time pressures, the trials were done to a really good standard. I thought that the programmes showed clinical trials in a positive light.

Trust Me, I'm a Doctor

Trust Me, I'm a Doctor is a series of programmes from the BBC (British Broadcasting Corporation, broadcast on BBC2). In May 2015, a researcher approached me about designing and analysing a study of duct tape and warts, inspired by *Health Freaks*.

This researcher, too, left and was replaced.

The study for *Trust Me, I'm a Doctor* had no control group. Participants were recruited via the programme and a web page. Participants were asked to get someone else to measure the diameter of the wart and to enter this on the web page. They should then apply duct tape, as in *Health Freaks*. After 28 days, the wart should be measured again. They should then submit the data on the web page.

The data

I received the results at the end of October 2015. Some things went wrong.

Participants were asked to enter their data on the website, including the diameter of the wart. They were asked after 28 days to provide the diameter of the wart again, whether it had disappeared, and, if so, how long after the start of the experiment this had happened. They could also enter the original diameter again. At follow-up, many participants entered an original diameter which was different from that entered at baseline.

The diameter of the wart was recorded by clicking a button marked 1, 2, 3, ... , 10 mm. This was done both for the first and for subsequent data entries. There was no zero button for those whose wart had disappeared. There were many non-zero diameters entered for participants who reported that the wart had disappeared within 28 days.

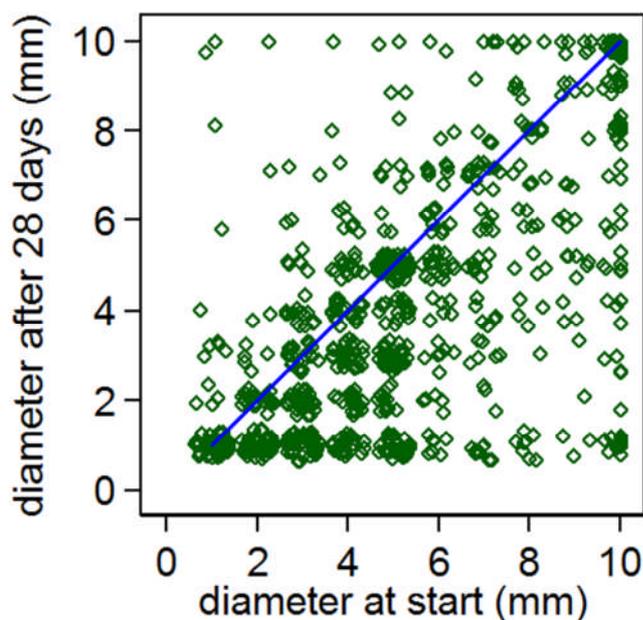
Analysis

The following choices were made for the analysis:

1. if the participant reported that the wart had disappeared by 28 days and had also recorded a diameter of 1 mm, the smallest available, or had omitted the diameter altogether, the diameter was set to zero,
2. if the participant had not entered a diameter at baseline, the original diameter as reported at follow-up was used, otherwise the first diameter entered was used.

The final amended data set included 2,780 records on 1,721 participants. 807 people provided both baseline and follow-up data. Analysis was done on these 807 people. All records without a baseline and a 28 day measurement and all records beyond 28 days, of which there were few, were dropped.

The figure below shows reported diameter after 28 days against diameter at time zero:



Again, the diagonal line is the line of equality. Points above and to the left of this line show participants whose wart increased in size; points below and to the right of this line show participants whose wart decreased in size.

The majority of participants appear to have had a reduction in reported size of the wart. The mean reduction in size was 1.33 mm (95% confidence interval or CI 1.15 to 1.50 mm).

The table shows the direction of change of the diameter of the wart from time zero to time 28 days:

	Number	Percentage
Decrease	458	56.75
No change	228	28.25
Increase	121	14.99
Total	807	100.00

Clearly there were far more decreases than increases and this was statistically highly significant, $P < 0.0001$, sign test.

In this trial, there was no control group, but there was a much larger sample than for *Health Freaks*. This enabled me to carry out further analysis.

The table below shows the duration of wart before application of duct tape was available for 716 participants:

Duration	Number	Percentage
0-3 months	58	8.10
4-6 months	40	5.59
7-9 months	41	5.73
10-12 months	42	5.87
1-2 years	144	20.11
More than 2 years	391	54.61
Total	716	100.00

The duration of the wart before the application of duct tape was available for 716 participants. Most of the participants had suffered the wart for years. As we might expect, older warts tended to be bigger:

Duration	Mean diameter at time 0
0-3 months	3.7
4-6 months	4.2
7-9 months	4.6
10-12 months	5.0
1-2 years	4.7
More than 2 years	5.8
Total	5.2

Large warts tended to shrink by more millimetres than small warts. This is not surprising, as they had further to shrink. Using regression analysis, I found that the average reduction in size went up by 0.32 mm (95% CI 0.26 to 0.38 mm) for every millimetre of diameter of the wart at time zero.

Using multiple regression, I found that the average reduction in diameter fell by 0.16 mm (95% CI 0.05 to 0.27 mm) for every category of duration. The older the wart is, the harder it is to shrink.

The *Trust Me, I'm a Doctor* broadcast

Trust Me, I'm a Doctor did not broadcast these results, but merely told people that duct appeared to work and was worth a try. The entire segment lasted for 40 seconds!

They also left my name off the credits for the programme. I was assured that this was a mistake and that my name was on the webpage for the programme.

Despite this waste, if *Trust Me, I'm a Doctor* were to call I would do it again. However, I would be wiser, next time, and ask for assurances that the data would be used in the programme and try to check their data collection.

Were the trials consistent?

We have two very different trials of duct tape for verrucae.

I thought that the *Health Freaks* trial was conducted well, with the support of an experienced trial coordinator, Rhiannon Pursall. It was a professional job. It was very small, but it convinced me.

The *Trust Me, I'm a Doctor* trial was much larger but was uncontrolled and the data were not well collected. It was subject to observer bias, because the measurements were made and submitted by the participants. It would not convince a sceptic if this were the only evidence.

A point in favour of the *Trust Me, I'm a Doctor* trial is that it enabled further analysis, because of the much larger sample size.

I compared the results for the duct tape treated group in the first trial with those for the much larger group in the second trial.

For the *Health Freaks* treated group, the mean reduction in size of the wart was 2.15 mm (95% CI 1.39 to 2.91), $n = 10$. For the *Trust Me, I'm a Doctor* participants, the mean reduction in size was 1.33 mm (95% CI 1.15 to 1.50), $n = 807$. This difference was not statistically significant, $P = 0.3$, t test. The bias, if any, is that the uncontrolled measurements showed a smaller decrease. This was exactly the opposite of what I expected.

For the *Health Freaks* treated group, the proportion of participants for whom the wart disappeared was 10% (95% CI 0.3% to 44.5%), $n = 10$. For the *Trust Me, I'm a Doctor* participants, the wart was reported have disappeared by 24.4% (95% CI 21.4% to 27.4%), $n = 807$. This difference was not statistically significant, $P = 0.5$, Fisher's exact test. The uncontrolled measurements showed more disappearances, but the numbers are very small for *Health Freaks*.

Did either programme help public understanding of clinical trials?

Health Freaks did, I think, though the programme spent a lot of time rubbishing strange ideas and warning of some dangerous ones. *Trust Me, I'm a Doctor* does so in general, I think, as it does quite a lot of simple, well-explained trials. But not this one, which was uncontrolled, had data collection problems, and for which the results were not broadcast.

For more about my experiences with *Health Freaks*, see Bland, M.: Health freaks on trial: duct tape, bull semen and the call of television (*Significance* 2014; **11**(2): 32-35). For my earlier experience with television, see Bland, M.: The *Horizon* homeopathic dilution experiment (*Significance* 2005; **2**, 106-109).

Thanks

I thank all those at Outline Productions and *Trust Me, I'm a Doctor* for involving me in these interesting projects.

I thank the National Institute for Health Research for the award which funds my conference visit.

I thank the University of York for providing me with a base and an office.