Minding your Ps and Qs: making sense of clinical trials

Susan Mayor
Medical journalist and editor

Mail Online

Home | News | U.S. | Sport | TV&Showbiz | Femail | Health | Science | Money | Video |

Health Home | Health Directory | Health Boards | Diets | MyDish Recipe Finder

Clinical trial success for new cancer drug for patients with same faulty gene as Angelina Jolie

A new cancer drug for patients with the same faulty gene as actress Angelina Jolie has shown 'impressive responses' in a clinical trial, researchers have said.

The potential drug, called BMN 673, targets DNA repair in cancer cells and is designed to attack tumours that have been left vulnerable by genetic mutations.

BRCA genes were brought into the public consciousness last month after Jolie, 37, revealed she underwent a double mastectomy when doctors told her that her faulty gene - BRCA1 - meant she had an 87 per cent risk of developing breast cancer and a 50 per cent risk of ovarian cancer.

PUBLISHED: 20:06, 4 June 2013 | **UPDATED:** 09:17, 5 June 2013

Brief history of clinical trials – from trial-by-error to randomised controlled trials

562BC: *First clinical trial recorded (Book of Daniel)*
Meat + wine vs vegetables – to maintain good health

1537: *First clinical trial of a novel therapy*
Boiling oil vs egg yolk/ oil of roses / turpentine – to heal battle wounds

1747: *First controlled clinical trial*
General diet vs general diet + oranges and lemons - in scurvy

1863: *First use of placebo in clinical trial*

1923: *First use of randomisation*

1948: *First randomised, double-blind controlled clinical trial* – MRC trial of streptomycin in pulmonary tuberculosis

1964: *Declaration of Helsinki* – set out trial ethics, including informed consent

BRITISH MEDICAL JOURNAL

LONDON SATURDAY OCTOBER 30 1948

STREPTOMYCIN TREATMENT OF PULMONARY TUBERCULOSIS

A MEDICAL RESEARCH COUNCIL INVESTIGATION

The following gives the short-term results of a controlled investigation into the effects of streptomycin on one type of pulmonary tuberculosis. The inquiry was planned and directed by the Streptomycin in Tuberculosis Trials Committee, composed of the following members: Dr. Geoffrey Marshall (chairman), Professor J. W. S. Blacklock, Professor C. Cameron, Professor N. B. Capon, Dr. R. Cruickshank, Professor J. H. Gaddum, Dr. F. R. G. Heaf, Professor A. Bradford Hill, Dr. L. E. Houghton, Dr. J. Clifford Hoyle, Professor H. Raistrick, Dr. J. G. Scadding, Professor W. H. Tytler, Professor G. S. Wilson, and Dr. P. D'Arcy Hart (secretary). The centres at which the work was carried out and the specialists in charge of patients and pathological work were as follows:

Brompton Hospital, London.—Clinician: Dr. J. W. Crofton, Streptomycin Registrar (working under the direction of the honorary staff of Brompton Hospital); Pathologists: Dr. J. W. Clegg, Dr. D. A. Mitchison.
Colindale Hospital (L.C.C.), London.—Clinicians: Dr. J. V. Hurford, Dr. B. J. Douglas Smith, Dr. W. E. Snell; Pathologists (Central Public Health Laboratory): Dr. G. B. Forbes, Dr. H. D. Holt.
Harefield Hospital (M.C.C.), Harefield, Middlesex.—Clinicians: Dr. R. H. Brent, Dr. L. E. Houghton; Pathologist: Dr. E. Nassau.
Bangour Hospital, Bangour, West Lothian.—Clinician: Dr. I. D. Ross; Pathologist: Dr. Isabella Purdie.
Killingbeck Hospital and Sanatorium, Leeds.—Clinicians: Dr. W. Santon Gilmour, Dr. A. M. Reevie; Pathologist: Professor J. W. McLeod.
Northern Hospital (L.C.C.), Winchmore Hill, London.—Clinicians: Dr. F. A. Nash, Dr. R. Shoulman; Pathologists: Dr. J. M. Alston, Dr. A. Mohun.
Sully Hospital, Sully, Glam.—Clinicians: Dr. D. M. E. Thomas, Dr. L. R. West; Pathologist: Professor W. H. Tytler.

The clinicians of the centres met periodically as a working subcommittee under the chairmanship of Dr. Geoffrey Marshall; so also did the pathologists under the chairmanship of Dr. R. Cruickshank. Dr. Marc Daniels, of the Council's scientific staff, was responsible for the clinical co-ordination of the trials, and he also prepared the report for the Committee, with assistance from Dr. D. A. Mitchison on the analysis of laboratory results. For the purpose of final analysis the radiological findings were assessed by a panel composed of Dr. L. G. Blair, Dr. Peter Kerley, and Dr. Geoffrey S. Todd.

Introduction

When a special committee of the Medical Research Council undertook in September, 1946, to plan clinical trials of streptomycin in tuberculosis the main problem faced was that of investigating the effect of the drug in pulmonary tuberculosis. This antibiotic had been discovered two years previously by Waksman (Schatz, Bugie, and Waksman, 1944); in the intervening period its power of inhibiting tubercle bacilli *in vitro*, and the results of treatment in experimental tuberculous infection in guinea-pigs, had been reported; these results were strikingly better than those with any previous chemotherapeutic agent in tuberculosis. Preliminary results of trials in clinical tuberculosis had been published (Hinshaw and Feldman, 1945; Hinshaw, Feldman, and Pfuetze, 1946; Keefer *et al.*, 1946); the clinical results in pulmonary tuberculosis were encouraging but inconclusive.

The natural course of pulmonary tuberculosis is in fact so variable and unpredictable that evidence of improvement or cure following the use of a new drug in a few cases cannot be accepted as proof of the effect of that drug. The history of chemotherapeutic trials in tuberculosis is filled with errors due to empirical evaluation of drugs (Hart, 1946); the exaggerated claims made for gold treatment, persisting over 15 years, provide a spectacular example. It had become obvious that, in future, conclusions regarding the clinical effect of a new chemotherapeutic agent in tuberculosis could be considered valid only if based on adequately controlled clinical trials (Hinshaw and Feldman, 1944). The one controlled trial of gold treatment (and the only report of an adequately controlled trial in tuberculosis we have been able to find in the literature) reported negative therapeutic results (Amberson, McMahon, and Pinner, 1931). In 1946 no controlled trial of streptomycin in pulmonary tuberculosis had been undertaken in the U.S.A. The Committee of the Medical Research Council decided then that a part of the small supply of streptomycin allocated to it for research purposes would be best employed in a rigorously planned investigation with concurrent controls.

The many difficulties of planning and conducting a trial of this nature are important enough to warrant a full description here of the methods of the investigation.

Plan and Conduct of the Trial

Type of Case

A first prerequisite was that all patients in the trial should have a similar type of disease. To avoid having to make allowances for the effect of forms of therapy other than bed-rest, the type of disease was to be one not suitable for other forms of therapy. The estimated chances of spontaneous regression must be small. On the other hand, the type of lesion should be such as to offer some prospect of action by an effective chemotherapeutic agent; for this reason old-standing disease, and disease with thick-walled

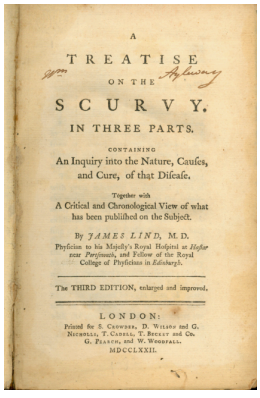
4582

How are clinical trials designed?

Clinical trials are designed to give a clear assessment of the effect of a treatment

- need to compare with control group
- the effects of chance or bias have to be removed

Controlled



- we need a control for comparison – to be sure what we see is due to the new treatment
- compare the new treatment with a control group receiving placebo or best current treatment
- increasing use of 'active comparator' studies
- superiority / non-inferiority (equivalence) trials or cross-over trials

Randomised

- fair comparison by ensuring no bias in allocating study subjects to a new treatment and control
- participants randomly allocated to different treatment groups
- randomly allocated by a process similar to flipping a coin, so which treatment they get occurs by chance
- should achieve two (or more) groups similar in every way except for the treatment they receive

Blinded

- Removes bias
- Single blind – participants don't know which treatment they are receiving
- Double blind – researchers don't know either
- Blinded assessment – researchers assessing scans etc don't know treatment allocation
- Open label – researchers and participants know which treatment is being given

Study design …

- Secondary research using clinical trials
  - Overviews – summarise primary studies – but not in a systematic way
  - Systematic reviews
    - bring together the results of previous research (usually randomised trials) about one particular treatment in a rigorous, systematic way
    - researchers try to uncover all the relevant trials and to evaluate them in a fair and objective way
  - Meta-analysis
    - the numerical results of all the trials are combined to measure how well a treatment works
    - may allow us to pick up small differences between treatments which can be very hard to spot in individual trials

What is the most reliable provider of clinical trial evidence?

- Systematic reviews / meta-analyses of randomised controlled trials
- Randomised controlled trials
- Other controlled clinical trials
- Observational studies (cohort and case-control)
- Case studies, anecdote and personal opinion

Making sense of the numbers

- Three numbers to check:
  - The size of the sample
  - The duration of follow-up
  - The completeness of follow-up

Making sense of the numbers

How are the results analysed?

- Intent to treat
- Per protocol

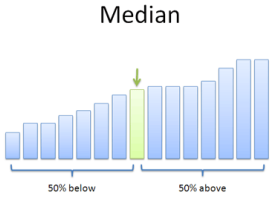
Making sense of the numbers: describing a set of data - mean and median

- Mean – the average; calculated by adding a set of numbers (for example, a set of results) and dividing by the number of values

eg The mean of 5,7,9,10,13,18,23 is 12.1

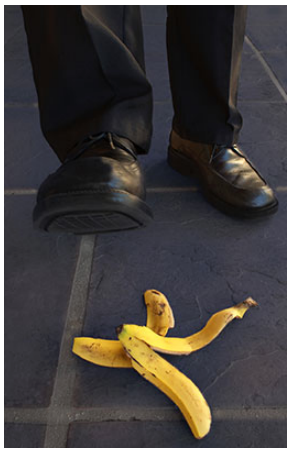
- Median – the middle value in a set of numbers when they are all placed in ascending numerical order, from smallest to largest

eg The median of 5,7,9,10,13,18,23 is 10



Making sense of the numbers: describing the effect of a treatment - hazard ratio (HR)

- The ratio of the chance of a hazard happening (such as death, heart attack or cancer recurrence) in the treatment group divided by the risk in the control group
- May be a better indication of efficacy than median as it uses data from all of the patients – not just the midway point



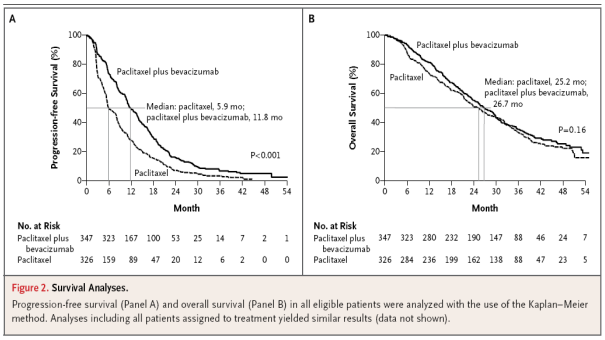


Figure 2. Survival Analyses.
Progression-free survival (Panel A) and overall survival (Panel B) in all eligible patients were analyzed with the use of the Kaplan–Meier method. Analyses including all patients assigned to treatment yielded similar results (data not shown).

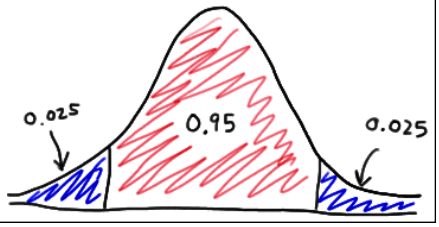
Example: phase 3 trial randomised women with metastatic breast cancer to bevacizumab plus paclitaxel or paclitaxel alone (control)

- The hazard ratio (HR) for progression was 0.6
- This means women treated with bevacizumab plus paclitaxel had 0.6 the risk of their cancer progressing compared to those treated with paclitaxel alone
- Simpler to say they had 40% less risk of their cancer progressing
- (1 - 0.6 = 0.4 = 40%)

(*NEJM 2007; 357: 2666-2676*)

Making sense of the numbers: extrapolating from the study population to all patients - confidence interval (CI)

- CI is a range in which we can be confident that the true population value lies
- studies often give the 95% confidence interval, which means 95% of the entire population will show an effect of the drug in the range given
- the size of CI is related to the sample size – larger studies usually have narrower CI



Making sense of the numbers: confidence interval

Example: The HERA trial randomised women with HER2-positive early breast cancer to Herceptin or observation after standard chemotherapy

- After one year of treatment, the hazard ratio for risk of death for the risk of death with Herceptin compared with observation was 0.66
- The 95% confidence interval for the hazard ratio for the risk of death of 0.66 was 0.47-0.91
- This means Herceptin reduces the hazard ratio for the risk of death by between 0.47 and 0.91 in 95% of the population (all women with HER2 positive early breast cancer)

(*Lancet 2007; 369: 29-36*)

Making sense of the numbers: extrapolating from the study population to all patients – p value

- The p value is the probability of an observed difference having happened by chance
- *eg* p=0.5 means that the probability of the difference having happened by chance is 0.5 in 1, or 50: 50
- If p is 0.05 or lower ($p \leq 0.05$), the finding is considered 'statistically significant'

- it means that the observed difference would occur by chance in only 1 in every 20 similar cases, or fewer

Making sense of the numbers: what does a p value mean?

Example: The HERA trial gave a hazard ratio of 0.66 for the risk of death in women treated with Herceptin compared to those randomised to observation and the p value for this was 0.0115

- this means there is a 0.0115 in 1 chance of the reduction in risk of death having happened by chance
- this equates to a 1 in 87 chance of the reduction occurring by chance
- the p value is less than 0.05, so is considered statistically significant

(Lancet 2007; 369: 29-36)

Making sense of the numbers: absolute risk reduction (ARR)

- ARR is the difference between the number of events in the intervention group and in the control group
- *eg* In a meta-analysis comparing treatment with tamoxifen plus chemotherapy to tamoxifen alone in women with estrogen-receptor positive breast cancer
- the risk of recurrence at five years was 21.6% in women treated with tamoxifen alone compared to 14.0% in those treated with tamoxifen plus chemotherapy
- ARR = 21.6% – 14.0% = 7.6%

(*Lancet 2005; 365: 1687-1717*)

Making sense of the numbers: relative risk reduction

- **Relative risk reduction (RRR)**

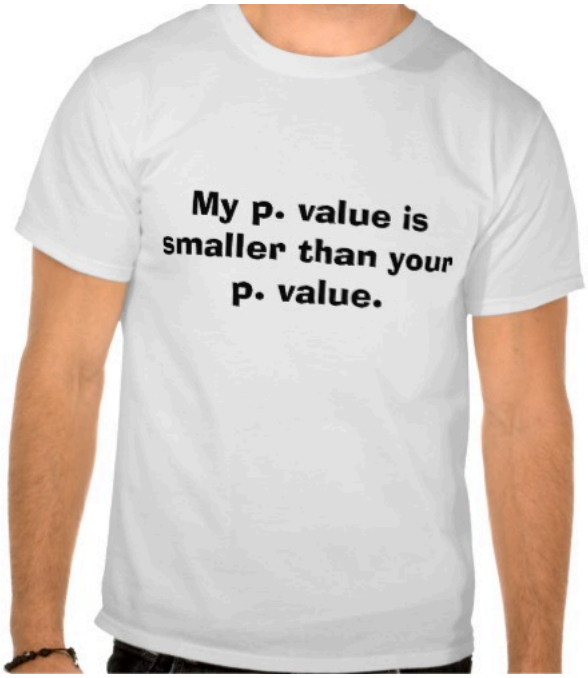
- the proportion by which the intervention reduces the event rate

- *eg* The incidence of recurrence is reduced from 21.6% in women treated with tamoxifen alone to 14.0% in those treated with chemotherapy plus tamoxifen
- RRR = (21.6-14.0) /14 = 7.6/21.6 = 35.2%

(*Lancet 2005; 365: 1687-1717*)

Summing up…

- Mean / median – the average or middle value for a group of results
- Hazard ratio – the ratio of something harmful happening in one treatment group compared to the risk in a comparison treatment group
- Confidence interval – the range in which we can be sure the true population value lies
- P value – the probability of an observed value having happened by chance



Checklist for analysing research papers

- What question is the research study asking?
- What were the main findings?
- Are the findings meaningful? – what is the 95% CI and the p value?
- ***Are the findings credible?***
  - Who carried out the study? Are they well respected?
  - Where was the study published? Is the journal peer-reviewed?
  - Who funded the research? Could this affect the interpretation of the results?
- What do the findings mean for: Healthcare professionals? Patients?

Some useful references

- Medical statistics made easy. M.Harris and G. Taylor. Martin Dunitz
  How to read a paper. Trisha Greenhalgh. BMJ Books
  Clinicaltrials.gov (US National Library of medicine)
  http://www.clinicaltrials.gov/
- Medical Research Council Clinical Trials Unit
  www.ctu.mrc.ac.uk
- Sense about science. Making Sense of Statistics
  http://www.senseaboutscience.org/resources.php/1/making-sense-of-statistics
- What is a p-value anyway? 34 stories to help you actually understand statistics. Andrew Vickers