Statistical Plan for TEAM Study

Official title: Safety and Efficacy of Endovascular Treatment of Unruptured Intracranial Aneurysms in the Prevention of Aneurysmal Haemorrhages: A Randomized Comparison With Indefinite Deferral of Treatment in 2002 Patients Followed for 10 Years

Brief title: Trial on Endovascular Aneurysm Management

Acronym: TEAM

NCT number: NCT00537134

Protocol ID: MCT-80799

Secondary ID: ISRCTN62758344

Document Date (Protocol final revision date): March 3, 2007

This statistical plan is taken from the Protocol (pp 21-26)

Number of patients for TEAM

In accordance with the philosophy of care trials, the number of patients that need to be recruited in order to lift the uncertainty provides prudent boundaries without which unproven care would be practiced without limits. This number, however, cannot be used to qualify the trial as 'unfeasible'.

Although more sensitive tests will be performed to assess some hypotheses (Kaplan-Meyer methods for hemorrhagic events and mortality), we used two-sided log rank tests to estimate sample size. We took into account patients lost for analyses, and adjusted the alpha for interim analyses. The total size of the population would be approximately 2002 (taking into account 12% (or 240) patients being lost for analysis (8-10% of unrelated deaths and 1-2% lost to follow-up), based on the following hypotheses: The overall benefit of endovascular management could be demonstrated with a total of 1688 patients, which is the number to achieve 80% power at a 0.0167 significance level (to account for 1-5-year interim analyses) to detect a difference of 0.04 (hazard ratio: 0.4896) between the null hypothesis that both proportions are 0.08 (the disease/treatment-related morbidity/mortality reaches 8% at 10 years) and the alternate hypothesis that the proportion of the endovascular group is 0.04. Increasing the size to 2002 would permit:

1) to verify a significant difference under the same conditions between 7-9% in the conservative group to 3-5% in the endovascular group (Table 1);

2) to achieve 80% power at a 0.0167 significance level to detect a difference of 0.05 (between 0.84 and 0.89) in the proportions surviving in groups 1 and 2 (hazard ratio: 0.6684; proportion lost to follow-up 0.02 (Table 2). A significant difference in the overall mortality, an outcome resistant to bias, but not chosen as primary, could offer another convincing evidence for a benefit. These numbers are felt to be conservative. The incidence of hemorrhagic events was 3.75%/5 years (poor outcome in 83%; mortality in 65%=2.44%) in ISUIA, while most other series are in the 1-2% range for annual rates. Most endovascular series include an over-representation (30-45%) of posterior circulation aneurysms, as well as of patients with a previous history of SAH (30-50%), which are important risk factors for haemorrhage. For these reasons a 10-year incidence of haemorrhage in the conservative group would range from 7.5 to 20%, with a poor outcome in 6 (ISUIA) to 17% (meta-analyses).

This will leave some alpha for interim analyses: 0.001 when 1000 patients have reached a mean follow-up of 5 years, 0.002 when 2000 patients have reached a mean of 5 years, 0.002 when all 2000 patients have reached 5 years and some alpha (0.005) to compare overall mortality at 10 years.

Planned analyses

As in any trials of invasive therapy, early risks may be followed by later benefit. Therefore the hazard ratio will be unfavourable during the recruitment years, while interventions are being performed.

Initially, the DSMC will assure that treatment-related complications and hemorrhagic events are within confidence intervals compatible with a safe and meaningful trial. In order to describe how and when hemorrhagic events occur, analyses will include Kaplan-Meyer life-table methods to assess the 5-, and 10-year survival without neurological dependency or mortality, and survival without haemorrhage, among all those allocated immediate treatment (including the few who did

not undergo it) and all those allocated deferral of any intervention (including the few who will eventually be treated). The 'survival' functions will be compared graphically and using a log-rank statistic.

The main statistical tests will involve comparisons between the probabilities of mortality:

1/ from haemorrhage, excluding perioperative complications,

2/ from haemorrhage or from complications of treatment, or

3/ comparisons of the 5 and 10-year probabilities of combined disease/treatment-related mortality/morbidity, in the absence of other causes of death or disability.

Descriptive statistics will be done on demographic variables and potential risk factors to compare the two groups at baseline. Means, standard deviations and range will be presented for quantitative variables such as size of aneurysms and frequency tables for categorical variables (such as number of patients with a previous history of SAH, or multiple aneurysms).

Statistics will be broken down by center and by treatment arm. Comparability of the groups will be assessed through independent ANOVAs (quantitative data) or Mantel-Haentzel and X^2 tests (categorical data).

Assuming comparability of groups across centers, the primary outcome, neurological dependency or death (for both intent-to-treat and per-protocol populations) will be compared between groups through a z-test for independent proportions at 5 and 10 years. Similar analyses will be done for disease or treatment-related mortality combined mortality-morbidity.

Secondary outcomes and overall morbidity will be compared between groups through independent t-tests (quantitative variables) or X^2 statistics (categorical data). The analyses of neurological data at follow-up will control for baseline data using logistic regression, ANOVA or Cox regression multivariate models.

All tests will be interpreted with adjustment for multiplicity to have the 0.05 level of confidence at 10 years only. Finally, a logistic regression will be used to find variables capable of predicting haemorrhages or complications in both groups. The method planned is a stepwise forward with alpha <0.05 to enter a predictor. Possible predictors include the status of the aneurysm (previous history of SAH vs. unruptured only), size of the aneurysm ($\geq$7mm vs $<$7mm), location (posterior vs. anterior circulation aneurysms), as well as other baseline characteristics.

Frequency of analyses

Initial assessment of the rate of hemorrhagic events in the control group, the morbidity of active treatment, the frequency of cross-over, will be performed after 300 patients to ensure that events are within confidence intervals compatible with a meaningful trial (judged by DSMC). Interim analysis on efficacy in the prevention of hemorrhagic events will be conducted after 300 patients have completed 1 year of follow-up, to ensure that patients are not unnecessarily exposed to the intervention or to the risk of haemorrhage. Safety data will be reviewed periodically by the DSMC that will meet on a 3-month basis initially. This committee will provide periodic reports and decide to conduct comparative analysis whenever they are felt necessary. The primary outcome will be assessed at 5 and 10 years.

Subgroup analyses

It is possible that treatment morbidity, or incidence of haemorrhages, will be affected by some characteristics of patient/aneurysm, such as size (large vs. small) or location (anterior or posterior) or previous history of SAH. These factors will be analyzed by introducing an interaction term in the logistic model.

TABLE 1

Sample sizes to detect a significant difference in disease or treatment-related combined morbidity/mortality.

Numeric Results with Proportion Lost to Follow Up = 0.1100

Power	N	N1	N2	S1	S2	Hazard Ratio	Two-Sided Alpha	Beta
0.8007	1317	659	658	0.9000	0.9500	0.4868	0.0167	0.1993
0.8006	862	431	431	0.9000	0.9600	0.3875	0.0167	0.1994
0.8006	595	298	297	0.9000	0.9700	0.2891	0.0167	0.1994
0.8004	1925	963	962	0.9100	0.9500	0.5439	0.0167	0.1996
0.8002	1155	578	577	0.9100	0.9600	0.4328	0.0167	0.1998
0.8007	749	375	374	0.9100	0.9700	0.3230	0.0167	0.1993
0.8002	3187	1594	1593	0.9200	0.9500	0.6152	0.0167	0.1998
0.8002	1669	835	834	0.9200	0.9600	0.4896	0.0167	0.1998
0.8008	990	495	495	0.9200	0.9700	0.3653	0.0167	0.1992
0.8000	6641	3321	3320	0.9300	0.9500	0.7068	0.0167	0.2000
0.8002	2727	1364	1363	0.9300	0.9600	0.5625	0.0167	0.1998
0.8004	1408	704	704	0.9300	0.9700	0.4197	0.0167	0.1996

Numeric Results with Proportion Lost to Follow Up = 0.1200

Power	N	N1	N2	S1	S2	Hazard Ratio	Two-Sided Alpha	Beta
0.8007	1332	666	666	0.9000	0.9500	0.4868	0.0167	0.1993
0.8006	872	436	436	0.9000	0.9600	0.3875	0.0167	0.1994
0.8006	602	301	301	0.9000	0.9700	0.2891	0.0167	0.1994
0.8004	1947	974	973	0.9100	0.9500	0.5439	0.0167	0.1996
0.8002	1168	584	584	0.9100	0.9600	0.4328	0.0167	0.1998
0.8007	757	379	378	0.9100	0.9700	0.3230	0.0167	0.1993
0.8002	3223	1612	1611	0.9200	0.9500	0.6152	0.0167	0.1998
0.8002	1688	844	844	0.9200	0.9600	0.4896	0.0167	0.1998
0.8003	1001	501	500	0.9200	0.9700	0.3653	0.0167	0.1997
0.8000	6717	3359	3358	0.9300	0.9500	0.7068	0.0167	0.2000
0.8002	2758	1379	1379	0.9300	0.9600	0.5625	0.0167	0.1998
0.8004	1424	712	712	0.9300	0.9700	0.4197	0.0167	0.1996

Report Definitions

Power is the probability of rejecting a false null hypothesis.
N is the combined sample size.
N1 sample size in group 1.
N2 sample size in group 2.
S1 is the proportion surviving in group 1.
S2 is the proportion surviving in group 2.
The Hazard Ratio is the ratio of hazard2 and hazard1. It is Log(S2)/Log(S1).
Alpha is the probability of rejecting a true null hypothesis.
Beta is the probability of accepting a false null hypothesis.

Summary Statements

A two-sided log rank test with an overall sample size of 1688 subjects (of which 844 are in group 1 and 844 are in group 2) achieves 80% power at a 0.0167 significance level to detect a difference of 0.0400 between 0.9200 and 0.9600--the proportions surviving in groups 1 and 2, respectively. This corresponds to a hazard ratio of 0.4896. The proportion of patients lost during follow up was 0.1200.

TABLE 2

Sample sizes to detect a significant difference in overall mortality

Numeric Results with Proportion Lost to Follow Up = 0.0100

Power	N	N1	N2	S1	S2	Hazard Ratio	Two-Sided Alpha	Beta
0.8002	1505	753	752	0.8200	0.8800	0.6442	0.0167	0.1998
0.8004	1078	539	539	0.8200	0.8900	0.5872	0.0167	0.1996
0.8001	2103	1052	1051	0.8300	0.8800	0.6861	0.0167	0.1999
0.8002	1422	711	711	0.8300	0.8900	0.6254	0.0167	0.1998
0.8001	3186	1593	1593	0.8400	0.8800	0.7332	0.0167	0.1999
0.8003	1982	991	991	0.8400	0.8900	0.6684	0.0167	0.1997
0.8001	5487	2744	2743	0.8500	0.8800	0.7866	0.0167	0.1999
0.8003	2995	1498	1497	0.8500	0.8900	0.7170	0.0167	0.1997
0.8000	11946	5973	5973	0.8600	0.8800	0.8476	0.0167	0.2000
0.8000	5141	2571	2570	0.8600	0.8900	0.7727	0.0167	0.2000
0.8000	46186	23093	23093	0.8700	0.8800	0.9179	0.0167	0.2000
0.8001	11158	5579	5579	0.8700	0.8900	0.8368	0.0167	0.1999

Numeric Results with Proportion Lost to Follow Up = 0.0200

Power	N	N1	N2	S1	S2	Hazard Ratio	Two-Sided Alpha	Beta
0.8002	1520	760	760	0.8200	0.8800	0.6442	0.0167	0.1998
0.8004	1089	545	544	0.8200	0.8900	0.5872	0.0167	0.1996
0.8003	2125	1063	1062	0.8300	0.8800	0.6861	0.0167	0.1997
0.8005	1437	719	718	0.8300	0.8900	0.6254	0.0167	0.1995
0.8001	3219	1610	1609	0.8400	0.8800	0.7332	0.0167	0.1999
0.8000	2002	1001	1001	0.8400	0.8900	0.6684	0.0167	0.2000
0.8001	5543	2772	2771	0.8500	0.8800	0.7866	0.0167	0.1999
0.8001	3025	1513	1512	0.8500	0.8900	0.7170	0.0167	0.1999
0.8000	12068	6034	6034	0.8600	0.8800	0.8476	0.0167	0.2000
0.8001	5194	2597	2597	0.8600	0.8900	0.7727	0.0167	0.1999
0.8000	46657	23329	23328	0.8700	0.8800	0.9179	0.0167	0.2000
0.8000	11271	5636	5635	0.8700	0.8900	0.8368	0.0167	0.2000

Report Definitions

Power is the probability of rejecting a false null hypothesis.
N is the combined sample size.
N1 sample size in group 1.
N2 sample size in group 2.
S1 is the proportion surviving in group 1.
S2 is the proportion surviving in group 2.
The Hazard Ratio is the ratio of hazard2 and hazard1. It is Log(S2)/Log(S1).
Alpha is the probability of rejecting a true null hypothesis.
Beta is the probability of accepting a false null hypothesis.

Summary Statements

A two-sided log rank test with an overall sample size of 2002 subjects (of which 1001 are in group 1 and 1001 are in group 2) achieves 80% power at a 0.0167 significance level to detect a difference of 0.0500 between 0.8400 and 0.8900--the proportions surviving in groups 1 and 2, respectively. This corresponds to a hazard ratio of 0.6684. The proportion of patients lost during follow up was 0.0200.

Software : PASS 2000, Power Analysis and Sample Size for Windows; NCSS, Kaysville, Utah