TEMPO-2 trial: Statistical Analysis Plan (SAP)

1. Introduction

This SAP is for the TEMPO-2 study (TNK-tPA Evaluation for Minor ischemic stroke with Proven Occlusion-2).

The trial is an academic trial comparing two approved treatment approaches for acute ischemic stroke therapy. The trial is not a registration trial for the purposes of licensing a new or novel endovascular device. The trial sponsor is the "Governors of the University of Calgary". The trial is registered at Clinicaltrials.gov (NCT02398656).

Although the trial is not a registration trial, it will be conducted in Canada under a Health Canada CTA investigational drug license. Similar investigational drug licenses will be sought from international drug regulatory authorities.

2. Trial Objectives

The trial is designed to evaluate if treatment with intravenous tenecteplase (TNKase™, MetaLyse™, TNK-tPA) is superior to best standard of care in a population of minor ischemic stroke patients with proven intracranial occlusion. Patients must be treated within 12 hours of symptom onset and must not be eligible for routine treatment with intravenous alteplase (Activase™, Actilyse™).

3. Primary Outcome

Primary outcome: Return to baseline neurological functioning as measured by the mRS.

Analysis will be a responder analysis where return to baseline level of neurological functioning using a variation of the sliding dichotomy modified Rankin Scale score outcome, defined as follows:
If pre-morbid mRS is 0-1 then mRS 0-1 at 90 days is a good outcome.
If pre-morbid mRS is 2 then mRS 0-2 is a good outcome.

Pre-morbid mRS is assessed using the structured mRS prior to randomization.(1) Outcomes will be assessed by an individual blinded to the treatment assignment. The 90day mRS will be rated using the structured mRS questionnaire (see appendix 1). The 90 day mRS will be completed in person where possible and by telephone otherwise. The structured questionnaire has been showed to improve reliability in assessing the mRS both in person and by telephone.(1)

4. Sample size

A sample of 1228 patients allows us to demonstrate a 9% absolute risk difference (60% → 69% primary outcome) with 90% power between intervention and control groups.

The recent pooled thrombolysis individual patient meta-analysis showed an effect size of 10% in the subset of minor stroke patients treated with thrombolysis.(2) Enrollment in the trials included in the meta-analysis did not require patients to have an intracranial occlusion; thus it is likely that the majority of these patients did not have an intracranial occlusion. Although we expect that the effect size is higher in a population that only includes patients with intracranial occlusion we will conservatively estimate an overall 9% effect size with a change in proportion with excellent neurological outcome from 60% to 69%. The sample size for each group is 614 (1228 total).

Adjusting for alpha-spending for a single interim analysis and adding 4% loss to follow up gives a sample size estimate of 1274 patients (637 in each treatment group). There will be ongoing monitoring for safety and full details will be available in a formal safety plan. A single interim analysis for futility and efficacy will be conducted at approximately two-thirds patient enrolment (n=850). O'Brien Fleming boundaries will be used to establish the alpha spending function. Full details will be available in the DSMB charter.

It is possible that after central imaging review some patients will be enrolled in violation of the protocol or the treatment protocol may be breeched due to the dynamic nature of acute stroke. This may occur entirely in the best interests of patient care. The primary analysis population will be all patients randomized in their as-randomized assignments regardless of actual treatment – the intention to treat (ITT) population. The safety population will be defined as all patients who receive any dose of study drug. The per-protocol population will be defined as all patients who received any dose of study intervention (treatment or control) met all the inclusion and exclusion criteria and were appropriately consented.

Secondary analyses will be considered exploratory and include analysis of the pre-stated secondary outcomes and multivariable analyses of both the primary outcomes and pre-stated secondary outcomes. This Statistical Analysis Plan will be reviewed and finalized prior to breaking of the blind.

5. Interim Analyses

We will plan for a single interim analysis after two thirds patient enrolment is complete and 90-day follow-up is completed on those patients. There will be a safety analysis after 400 patients have been enrolled.

We will use O'Brien-Fleming boundaries at the interim analysis as follows:(3, 4) We will use a simple dichotomous analysis of the responder proportion (based on the mRS at 90 days from randomization). The Z-statistic for this analysis shall be derived from the normal approximation of the binomial distribution as an unadjusted two-sample test of proportions. This risk of stopping a trial early will be mitigated by having stringent futility and efficacy boundaries using O'Brien-Fleming methods (which are known to be conservative at the interim analysis stage).

O'Brien Fleming Boundary for a Binary Primary Endpoint

For a RCT comparing two treatment arms with respect to a binary outcome and one interim analysis, the binary test statistic is given as

$$Z_k = \frac{(p_{Ak} - p_{Bk})}{\left\{\sqrt{\overline{p_k}(1-\overline{p_k})/\left(\frac{1}{n_A}+\frac{1}{n_B}\right)}\right\}}, k = 1, \ldots, K = 2$$

where $\bar{p}_k = \frac{p_{Ak}+p_{Bk}}{2}$, *where* P_{Ak} *and* P_{Bk} *are the estimated response proportions in treatment arms A and B at stage k, respectively. The two-sided sequential test based on O'Brien & Fleming boundary is given as*

1. At stage 1 (interim analysis, n=850: Reject H_0 *and stop the trial at stage k if:* $|Z_k| \geq C_B(2_z0.05)\sqrt{2} = 2.834$

Else if $|Z_k| < C_B(2_z0.05)\sqrt{2} = 2.834$, *continue to stage 2.*

2. At stage 2 (final analysis): Reject H_0 *at stage k if:* $|Z_k| \geq C_B(2_z0.05) = 2.004$

Therefore, for a RCT with one interim analysis and a final analysis (i.e., K = 2), the critical boundaries at Stage 1 and Stage 2 (final analyses) are 2.834 and 2.004, respectively.

Instructions to DSMB: Stopping Rules/Guidance

- **Thus, if the Z statistic is greater than 2.834 at the interim analysis, the committee will then consider that there is statistical evidence for overwhelming efficacy.**

The committee is then entrusted with a decision to make recommendations about the continuation of the trial in the context of the data and the context of the current and known evidence about stroke treatment using their best judgment.

6. Definition of the target populations

6.1. Efficacy population

All patients enrolled in the trial randomized on an intent-to-treat basis (as randomized).

6.2. Safety population

All patients enrolled in the trial who received the intervention, any dose of study drug. All patients in the control group who received best standard of care.

6.3. Per-protocol population

All patients enrolled in the trial who received any dose of study drug and met all the inclusion and exclusion criteria and were appropriately consented.

7. Randomization

Randomization will be managed using a custom SQL server-based database that will instantly and dynamically assign treatment using the minimal sufficient balance algorithm. Randomization will therefore be conducted over the internet via a desktop computer or a web-enabled smart phone.

Randomization will be 1:1. Allocation will be 1:1 set at p(0.5) for the first 40 patients. Thereafter, a randomization minimization algorithm (minimal sufficient balance) will be utilized to ensure ongoing balance in the trial on the following 4 factors:

Age
Sex
Baseline NIHSS score
Time of onset (or last seen well) to randomization

The minimal sufficient balance (MSB) randomization is a minimization procedure that preserves balance in smaller trials, such as this one, where imbalances in important baseline prognostic variables may occur by chance and confound the primary outcome. In addition, it preserves a greater degree of randomness in patient allocation compared to permuted block designs.(5) Because of the MSB process, randomization assignments will be stochastically derived in real time using a interactive web-site and therefore concealment can never be breached. Randomization will be biased coin that will vary from fully balanced (50:50) to biased (65:35) dependent on what characteristics been previously enrolled have. The system will be enabled for smart-phone, tablet, laptop or desktop computer use. The allocation sequence will therefore be fully masked, but treatment is open-label.

Reliance on a process that requires real-time data entry makes the process susceptible to error. For example, incorrect information (eg. wrong sex or age) could be mistakenly entered into the randomization process and affect the minimization algorithm. Post-hoc, when such errors become known, the quality-controlled database entry will be considered the source of truth and the randomization database will be updated to ensure that ongoing randomized minimization utilizes the most correct data to determine balance in an ongoing way.

The randomization process is neither blocked nor stratified by site. Therefore, the number of patients enrolled into each arm of the study may not be exactly even at the time of interim analysis or when the study is completed. The proportion of patients enrolled into each arm at each site may also vary and not be equally distributed. These decisions were taken explicitly with the knowledge and belief that balance on 4 key patient characteristics in the trial overall are more important than balance by site.

8. Blinding

Treatment assignment is open-label. Blinding of the outcome assessment at 90 days will be ensured at the site by having a person who was blinded to treatment allocation and not involved in the acute treatment period conduct the assessment.

9. Missing data and imputation rules

Under the ITT principle, all patients who are randomized are included in the analysis. Therefore, missing data, especially in the outcome measures, can be problematic. Every effort will be made to keep all missing data to a minimum. We will follow a data-informed imputation process.

If a patient is known to be deceased, they will be assigned a score of 6 on the mRS, 42 on the NIHSS and 0 on the Barthel Index for all outcome time points at or following the date of death, and therefore a non-responder.

If patient is known to be alive at day 5, but has missing day 90 status, the patient will be imputed to be alive. If the patient has unknown vital status from day 5/discharge onward, the patient will be imputed to be deceased, and therefore a non-responder.

If the assessment of the primary outcome (mRS) was conducted outside of the protocol-specified time window, data obtained are still included in the analysis, with the rationale that it is a more accurate measure than those obtained by imputation. At a minimum 90-day outcome assessments will be accepted within a +/- 30-day window.

If the primary outcome (mRS at 90 days) is missing but the patient is known to be alive, the patient will be imputed to a non-responder

If the rate of missing primary outcome data is <5% no further imputation will be done. In the event that there are more than 5% missing primary outcome data, we will perform the following sensitivity analyses:

To assess the impact of those missing data by using imputation with the following methods:

- If 5-day/discharge outcome scores are available, carry forward those values to determine responder status; else, impute the patient as a non-responder.
- Assign non-responder status to all subjects with missing 3-month outcome data.
- Hot-deck or nearest neighbor method, using clinical site, age, sex, baseline NIHSS, baseline serum glucose, baseline ASPECTS, , treatment group as classification variables.

- Regression method, with age, sex, baseline NIHSS, baseline serum glucose, and treatment group as covariates.

Similar imputation methods will be employed for secondary categorical outcomes. For the raw NIHSS score, multiple imputation, regression, and mean substitution methods will be used in the sensitivity analyses. Missing covariate data, if any, will be imputed using either multiple imputation or regression method, if needed.

Finally, we will conduct a "Tipping Point" analysis to assess the influence of missing data on the primary effect size estimate and direction of effect.

10. Efficacy Analysis

10.1. Primary analysis

The primary analysis will be conducted using a two-sample test of proportions (Fisher's exact test). This will be supported by a secondary analysis will use an additive multivariable model (generalized Poisson mixed-effects model with log link) adjusting for all the minimization variables included as co-variables. Site will be considered a random effect and not pooled. Only main effects will therefore be considered in this model. The results will be expressed as a risk ratio with 95% confidence limits. Additional analyses will include a safety population analysis defined to include only those patients who received tenecteplase, a per-protocol analysis including those patients who were treated according to protocol.

The primary analysis will be unadjusted. Because the randomization is being balanced a priori according to key prognostic variables (age, sex, NIHSS, and time to treatment), we expect that the unadjusted analysis will be similar to the adjusted analysis.

A revised statistical analysis plan may be modified according to the statistical distribution of variables and finalized prior to breaking the blind.

10.2. Secondary analyses

Pre-specified secondary outcome and safety analyses of proportions will be conducted in a similar way to the primary analysis using logistic regression or using a multivariable generalized linear model with log link to derive risk ratios directly. Pre-specified secondary analyses will include the following:

10.2.1. Proportion of patients with major bleeding: This will include an analysis of symptomatic intracranial hemorrhage alone and then combined with major extracranial hemorrhage. This is the main safety outcome.

10.2.1.1. Symptomatic intracranial hemorrhage defined as new intracranial hemorrhage (ICH, SAH, IVH, SDH) associated with clinical evidence of

neurological worsening, in which, the hemorrhage is judged to be the most important cause of the neurological worsening. Clinical worsening will be guided by the NIHSS score of a minimum of 2 or more points different from baseline.

10.2.1.2. Major extracranial hemorrhage defined as life threatening, resulting in hemodynamic compromise or hypovolemic shock, requiring inotropic support or other means to maintain cardiac output, requiring blood transfusion of more than 2 units of packed red blood cells, or associated with a fall in hemoglobin greater than or equal to 5 g/L.

10.2.2. Proportion of patients with complete and partial recanalization (mAOL 2-3) post treatment. This will be assessed on CTA 4-8 hours post treatment. Recanalization will be assessed by the central core-imaging lab blinded to all clinical information.

10.2.3. Categorical shift analysis on the full range of the mRS (0-6).

10.2.4. Absence of disability defined as mRS 0-1.

10.2.5. Functional independence defined as mRS 0-2.

10.2.6. Return to exact baseline function or better. If pre-morbid mRS is 0 then mRS 0 at 90 days is a good outcome. If pre-morbid mRS is 1 then mRS 1 is a good outcome. If pre-morbid mRS is 2 then mRS 0-2 is a good outcome.

10.2.7. Comparison of the mean mRS using linear regression using the mRS as a continuous variable.

10.2.8. Lawton Instrumental Activities of Daily Living Scale (IADL)(6, 7)

10.2.9. Proportion of patients with an NIHSS 0 at day 5 (or discharge from hospital if discharged before day 5)

10.2.10. Quality of life measured on EQ5D-5L (EuroQol)(8)

10.2.11. Quality of life as measured by the “problems with usual activities” question on the EuroQol.

10.2.12. Stroke progression and recurrent stroke (separately and together).

10.2.13. All-cause mortality

10.2.14. Discharge location – home, rehab facility, long term care etc.

10.2.15. Proportion of patients getting rescue EVT

10.2.16. Economic analysis will be conducted using Canadian hospital data and quality of life measure to estimate treatment utility.

10.3. Pre-specifed subgroups of interest:

Sex
Patients treated <4.5 hours and after 4.5 hours
Outcomes in patients with recanalization vs. partial vs. no recanalization
Patients with direct evidence of occlusion on CTA vs. indirect evidence of occlusion on CTP or multiphase CTA

Occlusion location
Over age 80 vs. 80 years of age or less
Complete resolution of symptoms at randomization versus not.

References

1. Bruno A, Akinwuntan AE, Lin C, Close B, Davis K, Baute V, Aryal T, Brooks D, Hess DC, Switzer JA, Nichols FT. Simplified modified rankin scale questionnaire: reproducibility over the telephone and validation with quality of life. Stroke. 2011;42(8):2276-9.
2. Emberson J, Lees KR, Lyden P, Blackwell L, Albers GW, Bluhmki E, Brott T, Cohen G, Davis S, Donnan G, Grotta J, Howard G, Kaste M, Koga M, von Kummer R, Lansberg M, Lindley RI, Murray G, Olivot JM, Parsons M, Tilley B, Toni D, Toyoda K, Wahlgren N, Wardlaw J, Whiteley W, Del Zoppo GJ, Baigent C, Sandercock P, Hacke W. Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. Lancet. August 6th 2014.
3. De Mets DL, Furberg CD, Friedman LM. Data monitoring in clinical trials. New York: Springer; 2006.
4. Jennison C, Turnbull BW. Group sequential methods with applications to clinical trials. New York: Chapman & Hall; 2000.
5. Zhao W, Hill MD, Palesch YY. Minimal Sufficient Balance – A new strategy to balance baseline covariates and preserve randomness of treatment allocation. Stat Meth Med Res. 2012;In press.
6. Graf C. The Lawton instrumental activities of daily living scale. The American journal of nursing. 2008;108(4):52-62; quiz -3.
7. Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily
living. The Gerontologist,.9(3):179-86.
8. Dorman PJ, Waddell F, Slattery J, Dennis M, Sandercock P. Is the EuroQol a valid measure of health-related quality of life after stroke? Stroke. 1997;28(10):1876-82.