Addendum to the EAC Assessment Report

EAC review of power calculation in TROOPER

1. Introduction

At the consultation of the draft guidance, the company disagreed with the statement that the TROOPER trial had a small sample size, and noted that the TROOPER study was fully powered as a non-inferiority study. The consultee from the company stated that the trial was the correct size to address the study question and it was designed and analysed by statistical experts from the Imperial College Clinical Trials Unit.

2. Power calculation review

The EAC reviewed the TROOPER trial (Bourne et al. 2017) and its power calculation and found limited information about the statistical calculation in the paper. It contacted the study authors for more details about the power calculation.

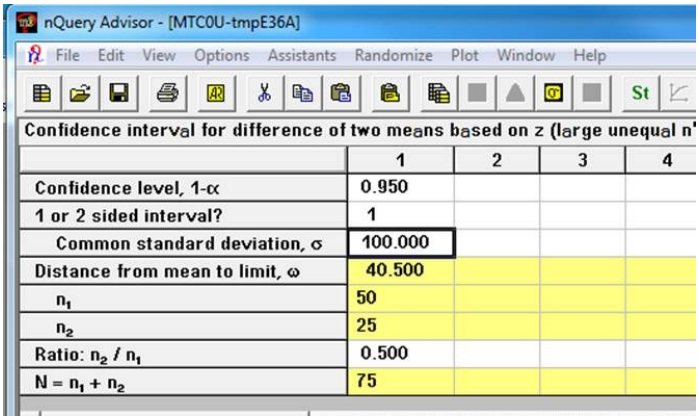
The study authors explained that the study was 'powered' to assess that the new intervention was not inferior to the current standard of care. Their sample size calculation was based on estimated effect size and standard deviation from previous pulmonary rehabilitation studies. The authors felt that the non-inferiority margin chosen of 40.5 meters for the 6-minute walk test (one of the primary outcomes of the trial) was suitable because it is smaller than the minimally clinically important difference and aligns with established literature on this outcome. The study authors considered that 40 meters difference represented a reasonable minimum loss of efficacy that was acceptable. If there was no difference between the intervention arms and a standard deviation of 100 (reported in the publication), the calculation found that 75 participants were needed (2:1 ratio) to estimate the lower 95% CI bound for the mean difference to be no more than the non-inferiority threshold for 6-minute walk test (40.5 meters).

The study authors shared the sample size calculations done by the statisticians at the Imperial College clinical trials unit who did the analysis (see notes in the next section), noting that the analysis was originally done using NQUERY (but using PASS when responding to the EAC request).

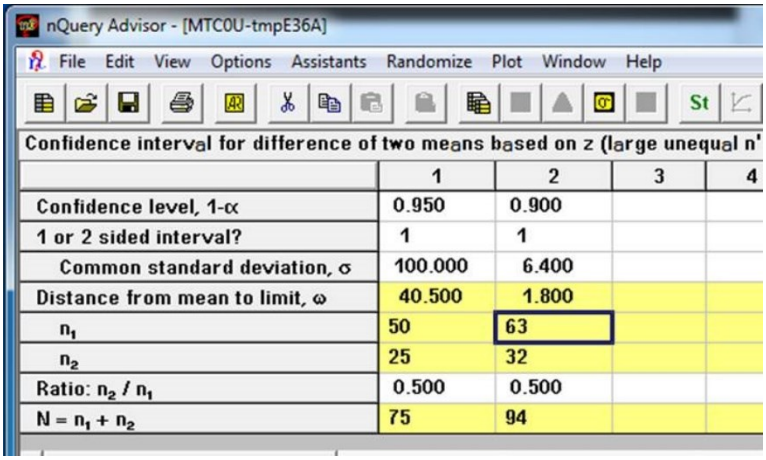
3. Notes on sample size calculations from TROOPER study statisticians

6-minute walk test as a primary outcome: 54 meters (Redelmeier et al 1997) is the minimum value of a significant increase in patients perception of exercise performance. Therefore taking 40 meters to be the margin of non-inferiority.

A standard deviation of 100 seems reasonable to assume based on the published literature. Sample size was calculated in NQuery using a one-sided 95% CI and ratio of 2:1.



COPD assessment test score (CAT) as a primary outcome: Kon et al (2014) has the minimally important difference for a CAT score as 1.8. This gives the CAT at baseline with 95% CIs (20.8 to 22.0) estimated from a sample size of 565 which enables the standard deviation to be estimated as 6.4 (one sided) (SD= SQRT(N)x(UCI-LCI)/ 3.92).



4. Conclusions

After receiving further information, the EAC discussed the evaluation of the TROOPER trial sample size calculation with statisticians at YHEC and an external statistician. The EAC commented that

1. The choice of non-inferiority limit depends on clinical judgement but the non-inferiority limit should certainly always be less than the clinically significant difference. It should normally also be set at a level where both groups also show this magnitude of change in outcome.

In the TROOPER trial, the use of 1.8 for the clinically meaningful difference in CAT score appears to be in line with the current opinion (2 is considered a significant change). However, for the 6MWT, the clinically meaningful difference was taken as 40.5m and this is greater than the 30m stated in the ERS/ATS 2014 guidance.

2. In terms of the sample size calculation, the EAC consulted with a statistician who confirmed that the calculations were based on a lower power than would normally be expected (50% instead of the normal 80-90%) and the 90% significance level used for the calculation based on CAT score was not justified. Using 80% power, 95% significance level and 30m and 1.8 as the non-inferiority thresholds, the EAC calculated sample sizes of 300-350. The statistician's thoughts were that the trial was promising in terms of its results but that a bigger trial would probably be needed to confirm findings.