## 1 Statistical Hypothesis and Planned Sample Size

Patients were randomised 2:1 to anastrozole + GDC-0941 and anastrozole, respectively. Patients excluded from the Per-Protocol-Population were replaced. The planned study size was 94 evaluable patients in the anastrozole plus GDC-0941 group and 47 evaluable patients in the anastrozole group, respectively, to provide 80% power to detect an effect size of 0.58 between Anastrozole and Anastrozole + GDC-0941 inhibitor at the 5% significance level. The effect size (ES) is defined as the treatment difference divided by the standard deviation, i.e. ES = [M1 - M2]/$\sigma$pooled where M1 and M2 were the mean values of the differences of proportional Ki67 changes and $\sigma$pooled = $\sqrt{[(\sigma 1^2 + \sigma 2^2)/2]}$. Taking Cohen´s standard interpretation of effect sizes into account, 0.5 is the lower limit of medium effect. An effect size of 0.5 corresponds to 33% of non-overlap between the two treatment groups. The non-centrality parameter $\delta$ is 2.83. Critical t is 1.98.

Group sample sizes of 47 patients in the Anastrozole group and 94 in the Anastrozole + GDC-0941 group also achieve 80% power to detect a difference between the group response rates of 20%. Response is defined as a 50% or higher fall in Ki67 expression. The proportion of responders in the Anastrozole + GDC-0941 group is assumed to be 60% under the null hypothesis and 80% under the alternative hypothesis. The proportion in the Anastrozole group is assumed to be 60%. The test statistic used is the one-sided Z test with pooled variance. The significance level is 5.1%. If the difference between the group response rates is 25%, then the sample size will be needed 28+56=84. On the other hand, if the difference is 30%, then only 18+36=54 patients will be sufficient to detect this difference.

Individual end-of treatment anti-proliferative response ($Response_{Ki67\text{-}Post}$), defined as the natural logarithm of percentage Ki67 positive cells of less than 1 or 1-2 at the end of study treatment, is another endpoint under which required sample size might be slightly lower. Assuming that approximately 40% of patients have activating PI3KCA mutations or PTEN deletions, the study will provide 80% power at a 5% significance level to detect an ES of 0.8.

### 1.1 Analysis Populations

All Ki-67 and Caspase3 analyses were performed on a Per-Protocol Population, defined as all randomised patients who completed 15 days (±2 days) of study treatment and for whom tumour biopsy specimens at baseline and at 15 (±2) days were available for assessment of biological response. The per-protocol population also excluded patients who had a major violation of protocol inclusion or exclusion criteria. Patients excluded from the Per-Protocol-Population were replaced.

Safety analysis was conducted on all patients who received at least one dose of the study treatment, with patients analysed according to the treatment they actually received.

### 1.2 Efficacy Analysis

The main analysis of apoptosis and proliferation were from baseline to day 15 using non-parametric statistics to compare the log (surgical/Pre-treatment) scores. Additional analyses of apoptosis and proliferation were from day 15 to definitive surgery.

Treatment comparisons were tested with and without adjustment for baseline prognostic factors. In the absence of major confounding factors the latter would be considered secondary endpoints.

On the assumption of a log normal distribution, Ki67 values were log transformed before analysis of mean ΔKi67, mean Ki67post, and ResponseKi67-Post. Ln(Ki67post) and ln(Ki67pre) were used to calculate the geometric means. 0.1 was added to every untransformed Ki67 value to avoid the mathematical anomaly that arises because the log of zero is minus infinity. As a consequence of the assumption of a lognormal distribution, ln(Ki67post)–ln(Ki67pre) was also normally distributed. This formula gave the proportional change, and as a result mean log proportional changes and CI were calculated and displayed on their original scale by back transformation. Mean ΔKi67 and mean Ki67post was compared between groups by use of the t- test, and the proportional change within groups was analysed with the paired t- test. The proportional reduction was calculated as one minus the proportional change.

Anti-proliferative response ResponseΔKi67 and end-of treatment anti-proliferative response ResponseKi67-Post were calculated in all evaluable patients. An estimate of the anti-proliferative response rates RRΔKi67 and end-of treatment anti-proliferative response rates RRKi67-Post and 95% CIs (Clopper-Pearson 1934) was calculated for each treatment arm. CIs for the difference in response rates (Satner and Snell 1980; Berger and Boos 1994) was calculated. The relative risk (treatment : control) was reported along with the associated 95% confidence interval based on logistic regression model.

A similar analyses strategy was applied for Caspase3 endpoints.

The change in tumour size was expressed as the proportional change from baseline to post-treatment. A two-sided $\chi 2$ -test was used to compare clinical or pathologic tumour response for the treatments arms. Changes in secondary outcomes from baseline to post treatment was analysed between treatment groups with the Wilcoxon Mann-Whitney test and within treatment groups with the Wilcoxon signed rank test. Associations between outcomes was investigated by use of the Spearman's rank correlation coefficient.

### 1.3 Subgroup Analysis

The effects of the study treatment was assessed separately in patients with and without PI3K mutations and/or PTEN deletions, Luminal A and B subtypes and patients with high (>14%) or low (≤14% baseline Ki67) . Additional subgroups might be defined by the exploratory biomarker analysis.

### 1.4 Safety Reporting and Analysis

Safety data was reported for all patients who received at least one dose of the study treatment. Safety data were reported separately for each treatment group. The worst toxicity during each cycle and the worst toxicity during the entire treatment will be determined separately for each patient according to the criteria specified above.

### 1.5 Exploratory Analysis

The potential relationship of exploratory biomarkers with biological response (Ki67, Caspase3) and clinical response will be explored.