Impact of Naloxegol (Movantik) on Prevention of Lower Gastrointestinal Tract Paralysis in Critically Ill Adults Initiated on Scheduled IV Opioid Therapy: A Randomized, Double-Blind, Placebo-Controlled, Phase II, Single-Center, Proof of Concept Study

NCT02977286

Version Date: September 22, 2018

Approved on: October 25, 2018

12. STATISTICAL METHODS AND SAMPLE SIZE DETERMINATION

12.1 Description of analysis sets

All patients will be analyzed in the groups to which they were randomly assigned, according to the intention-to-treat principle.

12. 2 Methods of statistical analysis

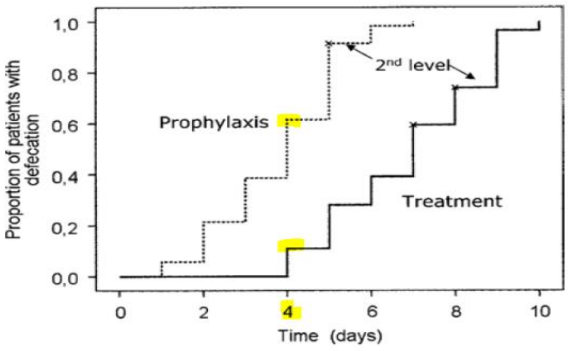
Baseline characteristics of patients will be compared using standard univariate methods (chi-square test, t-test) to confirm the randomization was successful in achieving comparable groups. To test the primary outcome (incidence of lower GI tract paralysis), given that there may be censoring (i.e., subject may die or withdraw from study before developing lower GI tract paralysis), we will present Kaplan-Meier curves for the naloxegol and placebo groups for the time to first development of lower GI tract paralysis ($\geq$ 3 days without a SBM after study randomization) during the ICU stay up to 10 days and compare the curves using a log-rank test. If all subjects complete the study (e.g., no deaths, no withdrawal or loss-to-follow-up for other reasons), then we will instead compare the proportion of patients that develop lower GI tract paralysis after study randomization during the ICU stay up to 10 days between the naloxegol and placebo groups using a chi-square test. The secondary outcome, time to first SBM will be estimated using a Kaplan Meier time to event analysis and compared between groups using a log-rank test. The secondary outcome of the number of ICU days without a SBM will be compared between the naloxegol and placebo groups using the Poisson regression for count data and adjusting for total number of days in the ICU (up to a maximum of 10) as an offset term in the Poisson model. The secondary outcome of the average daily opioid requirements will be compared between groups using either a student t-test or nonparametric Wilcoxan rank-sum test if the assumptions of the student t-test fail to be met. Length of mechanical ventilation and ICU stay will be reported as medians and interquartile ranges and will be compared using the Wilcoxon Rank Sum Test. For all statistical tests of hypotheses, a two-sided p-value of ≤ 0.05 will be considered as statistically significant. As this is a Phase II proof of concept study a p value of <0.10 may still be worth consideration for future study, however. All statistical analyses will be performed using SPSS 17.0. (SPSS, Chicago, IL) or other comparable commercially available statistical software.

12.2.1 Interim analyses

As outlined in the DSMB procedure, planned interim analyses will occur to evaluate safety after the enrollment of 10 subjects and after the enrollment of 20 subjects.

12.2.2 Determination of sample size

For the power analysis, the 4-day estimates of failure to defecate (i.e., failure to have a SBM) from the Kaplan-Meier plot presented in the Guardiola et. al. paper [4] was used as the estimate of lower GI tract paralysis. The figure below shows that the estimate of the percent of patients who *failed to defecate* (1 minus plotted proportion with defecation) by day 4 was 40% in the Prophylaxis group (i.e., the intervention group) vs. 90% in the non-prophylaxis (Treatment) group (i.e., the control group) (see yellow highlights in figure). We believe the 4-day estimate will be strongly related to the primary outcome of lower GI tract paralysis ($\geq$ 3 days without a SBM).



For this Naloxogol randomized trial, we propose enrolling 36 subjects, 18 per group. The statistical power table below shows this sample size yields over 95% power using an alpha of 0.05 if the true day 4 estimated freedom of event (defecation) rates are 90% (Placebo) vs. 40% (Prophylactic treatment), all subjects have complete follow-up, and the analysis is done using the log-rank test to compare survival curves. As sensitivity analysis, the table also shows the power of the study with this sample size under different scenarios. Because this is not the final definitive study to determine efficacy, a less stringent alpha of 0.10 may be justified for this Phase II proof-of-concept study which, if positive, would be followed by a confirmatory study [15]. While we expect most subjects to finish the study (based on the inclusion criteria), we have shown the impact on the power if approximately 20% of subjects do not finish the study. We have also varied the estimated proportion of patients failing to defecate by day 4 (i.e., not having a SBM) in the placebo group from 90% to 70% and in the prophylactic treated group from 40% to 50%.

Statistical Power of Log-Rank Test with n=18 subjects per group (36 total)

Complete study	ALPHA	Proportion Failing to Defecate at Day 4 (Placebo vs. **Prophylactic Treated**)				
		90% **vs. 40%**	90% **vs. 50%**	80% **vs. 40%**	80% **vs. 50%**	70% **vs. 40%**
100%	**0.05**	**99.8%**	**98.4%**	**92.7%**	**76.0%**	**76.0%**
	0.10	100.0%	99.4%	96.2%	84.9%	84.9%
80%	0.05	**99.7%**	**97.6%**	**90.8%**	**72.5%**	**72.5%**
	0.10	99.9%	99.0%	95.1%	82.2%	82.2%