Online supplement

Efficacy and safety of NVA237 versus placebo and tiotropium in patients with moderate-to-severe COPD over 52 weeks: The GLOW2 study

E. Kerwin[*], J. Hébert[#], N. Gallagher[+], C. Martin[+], T. Overend[+], V.K.T. Alagappan[§], Y. Lu[§], D. Banerji[§]

[*]Clinical Research Institute of Southern Oregon, PC, Medford, OR, USA.

[#]Centre de Recherche Appliquée en Allergie de Québec, Canada.

[+]Novartis Horsham Research Centre, West Sussex, UK.

[§]Novartis Pharmaceuticals Corporation, East Hanover, NJ, USA.

METHODS

Study exclusion criteria

Study exclusion criteria included: patients with a lower respiratory tract infection within 6 weeks prior to screening; patients requiring long-term oxygen therapy (>15 hours a day) on a daily basis for chronic hypoxaemia or those who had been hospitalised for an exacerbation of their airways disease in the 6 weeks prior to screening; patients with a clinically relevant laboratory abnormality or clinically significant condition, such as (but not limited to) unstable ischemic heart disease, left ventricular failure, history of myocardial infarction or arrhythmia (excluding chronic stable atrial fibrillation), history of malignancy of any organ system (including lung cancer and with the exception of localised basal cell carcinoma of the skin), narrow-angle glaucoma, clinically significant prostatic hyperplasia/bladder-neck obstruction/moderate-to-severe renal impairment/urinary retention; patients with a known history and diagnosis of α-1 antitrypsin deficiency; patients with any history of asthma

indicated by (but not limited to) a blood eosinophil count >600/mm^3 at screening or onset of symptoms prior to age 40 years; patients with concomitant pulmonary disease, such as pulmonary tuberculosis (unless confirmed by x-ray to be no longer active) or clinically significant bronchiectasis; patients with a history of long QT syndrome or QTc (QT interval with Fridericia's correction) >450 ms (males) or >470 ms (females) at screening. Patients were also excluded if they were involved in the active phase of a supervised pulmonary rehabilitation programme, were contraindicated for tiotropium/ipratropium or had shown previous untoward reaction to inhaled anticholinergic agents. Also excluded were women of child-bearing potential not using an accepted form of contraception, pregnant women, and nursing mothers.

All patients gave written, informed consent to participate in the study. The study was conducted according to the principles of Good Clinical Practice and the Declaration of Helsinki [1]. The study protocol (CNVA237A2303) was reviewed and approved by institutional review boards and ethics committees at participating centres (NCT00929110).

Study design and treatment

Treatment assignment

Treatment assignment was unbiased and concealed from patients and investigating staff. After confirming a patient's eligibility, investigators contacted an Interactive Voice Response System (IVRS) that automatically and randomly assigned the patient number to a randomisation number. These randomisation numbers (not communicated to the caller) were linked to the different treatment arms, which in turn were linked to medication numbers. A separate validated system automatically and randomly assigned medication numbers to study drug packs containing each of the study drugs. Randomisation (2:1:1 ratio) was maintained at the region, and not centre level, and was stratified by smoking status (current/ex-smoker), 12/24-hour serial spirometry subgroup (Yes/No), and Holter participant (Yes/No).

Assessments

Reversibility testing

FEV1 and FVC were used to measure the reversibility response to inhaled anti-cholinergic using a standard dose of ipratopium bromide at 80 µg at Visit 2.

Spirometry

The centralised spirometric assessments taken were forced expiratory volume in one second (FEV1), forced vital capacity (FVC) and inspiratory capacity (IC). At Day 1 and Weeks 12, 26 and 52, spirometry was performed at -45 and -15 minutes pre-dose and 5, 15 and 30 min, and 1, 2, 3 and 4 h post-dose. In a subset of patients (the serial spirometry subset, n=299) additional measurements were taken at 6, 8, 10 and 12 h post-dose and then at trough on Day 1, and over 24 h post-dose (including 16 and 22 h time points) at Weeks 12 and 52.

Patients practiced the measurement of IC at screening until reproducible results could be obtained. Before undertaking an IC measurement, patients performed normal tidal breathing and then inhaled to their maximum while receiving verbal encouragement. IC measurements were performed first, before other spirometry measurements, and with approximately 3 min rest before proceeding with FEV1 and FVC.

Dyspnoea and health-related quality of life

For measuring dyspnoea, patients were interviewed by an independent, trained assessor who graded the degree of impairment due to dyspnoea at Week 1 (Baseline Dyspnoea Index; BDI) and at Weeks 12, 26 and 52 (Transition Dyspnoea Index; TDI).

Health status was assessed by the St George's Respiratory Questionnaire (SGRQ) completed by patients at baseline at Week 1 and at Weeks 12, 26 and 52, before any other assessments were made to avoid influencing the response.

Exacerbations and rescue medication use

COPD exacerbations were defined as worsening of two or more major symptoms (dyspnoea, sputum volume, sputum purulence) for at least 2 consecutive days OR worsening of one major symptom together with one minor symptom (sore throat, colds, fever without other cause, increased cough, increased wheeze) for at least 2 consecutive days [2, 3]. Exacerbations were considered to be of moderate severity if they required treatment with systemic corticosteroids and/or antibiotic without leading to hospitalisation, and were considered to be severe if they also required hospitalisation.

At Visit 2, all patients were provided with an electronic patient diary to record morning and evening daily clinical symptoms: cough, wheezing, shortness of breath, sputum volume and color, night time awakenings and rescue medication use. Designated investigator site staff determined exacerbations and rescue medication use by reviewing the dairy data with the patients at each patient visit.

Following treatment for an exacerbation, patients were expected to continue in the study if, in the opinion of the investigator, they could be safely returned to their pre-exacerbation concomitant medication. Patients who were treated with intra-muscular depot corticosteroids for an exacerbation, patients who required the addition of new concomitant COPD medications after an exacerbation and patients experiencing a further exacerbation during the treatment period were withdrawn from the study.

Adverse events

Adverse events (AEs) were coded using the Medical Dictionary of Regulatory Activities (MedDRA) and summarised by primary system organ class, preferred term, maximum severity and relationship to study drug.

For vital signs, to meet the criterion of a newly clinically notable occurrence, the patient needed to have a baseline value which did not meet the criteria for categorising the value as notable. To meet the criterion of a worsening occurrence, the patient needed to have a baseline value which was clinically notable and a subsequent worse post-baseline value.

Adjudication committee

An independent adjudication committee consisting of a pulmonologist, a cardiologist and an oncologist assessed the cause of death for patients who died from screening up to 30 days after completion of treatment. The committee was blinded with respect to the patient's study treatment, reviewed narratives, study pertinent data, discharge summaries and medical records as available to determine the most likely cause of death to either respiratory, cardiovascular, cancer, other (to be specified) or unknown.

Statistics

A hierarchical testing sequence with a Hochberg adjustment was used to control for multiplicity and type one error rate, with superiorities of NVA237 over placebo tested sequentially in the three families of end-points (primary end-point, key secondary end-points and important secondary end-points). To proceed to the next family of tests in the hierarchy all previous families of tests had to be statistically significant at a probability of the type one error rate of 0.05 after Hochberg adjustment. This testing sequence had no impact on the testing of other secondary variables.

Sample size estimation

The sample size was driven by the joint power of 0.85 for finding the significance of the primary and key secondary hypotheses. For the primary objective of trough FEV1 at Week 12, a difference of 120 mL in trough FEV1 between NVA237 and placebo, with a standard deviation of 270 mL was assumed for this study. 455 evaluable patients for NVA237 and 225 for placebo would give a two-sided test at the 5% significance level with more than 99% power. Assuming 15% of patients would drop out without providing data for the primary endpoint by Week 12, 535 patients for NVA237 and 265 for placebo were to be randomised, giving a sample size of 800.

Percentage of patients with a clinically important improvement of at least 1 in TDI focal score after 26 weeks' treatment was used here for the power estimation for the TDI following 26

weeks. Based on the review of historical data from the Phase III pivotal studies of indacaterol, percentages of patients with a clinically important improvement of at least 1 in TDI focal score after 26 weeks' treatment were estimated to be 28% for placebo and at least 42% for NVA237. Using NQuery (V6.01), assuming 20% drop out rate by Week 26, 428 evaluable patients for NVA237 and 212 patients for placebo would give a two-sided test at the 2.5% significance level with 89% power.

A difference of -4 to placebo was assumed for SGRQ mean total score after 52 weeks with a standard deviation of 13. Given a 30% drop out rate by Week 52, 374 evaluable patients for NVA237 and 186 for placebo would give a two-sided test at the 2.5% significance level with 88% power.

Analysis

For the primary efficacy analysis, values taken within 6 h of rescue medication use or 7 days of systemic corticosteroid use were excluded. For missing trough FEV1 values, the last observation of pre-dose FEV1 was carried forward (LOCF) from the last non-missing visit as long as the visit was not prior to Day 15.

A logistic regression model was used to analyse the percentage of patients achieving a minimum clinically important difference in TDI focal score (≥1 point difference) and SGRQ total score (≥4 point reduction). Odds ratio between treatment groups and the corresponding 95% confidence intervals and two-sided p-values were reported.

The total number of puffs of rescue medication per day over the 52 weeks was calculated and divided by the total number of days with non-missing rescue medication use to derive the mean daily number of puffs of rescue medication for the patient. Rescue medication data recorded during the 14-day run-in period was used to calculate the baseline, with only the last 14 days used if the run-in was longer than 14 days.

REFERENCES FOR THE ONLINE SUPPLEMENT

1 World Medical Association. Declaration of Helsinki-ethical principals for medical research involving human subjects. 2008.

2 Anthonisen NR, Manfreda J, Warren CPW, Hershfield ES, Harding GK, Nelson NA. Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Ann Intern Med* 1987; 106: 2, 196–204.

3 Seemungal TAR, Donaldson GC, Paul EA, Bestall JC, Jeffries DJ, Wedzicha JA. Effect of exacerbation on quality of life in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 157: 1418–1422.

Online supplement TABLE S1. Treatment differences in inspiratory capacity at each time-point up to 3 h 55 min post-dose and at 23 h 40 min post-dose on Day 1 and Weeks 12, 26 and 52

	Inspiratory capacity, LSM (SE) treatment differences		
	NVA237–Placebo	**Tiotropium–Placebo**	**NVA237–tiotropium**
Day 1			
25 min	0.200 (0.0225)***	0.152 (0.0261)***	0.048 (0.0226)*
1 h 55 min	0.199 (0.0233)**	0.181 (0.0269)***	0.018 (0.0233)
3 h 5 5min	0.187 (0.0251)***	0.184 (0.0290)***	0.003 (0.0248)
23 h 40 min	0.114 (0.0244)***	0.080 (0.0284)**	0.033 (0.0246)
Week 12			
-20 min	0.073 (0.0318)*	0.090 (0.0364)**	-0.017 (0.0305)
25 min	0.158 (0.0330)***	0.167 (0.0380)***	-0.009 (0.0320)
1 h 55 min	0.169 (0.0356)***	0.184 (0.0407)***	-0.015 (0.0344)
3 h 55 min	0.137 (0.0353)***	0.141 (0.0402)***	-0.004 (0.0340)
23 h 40 min	0.129 (0.0336)***	0.113 (0.0388)**	0.015 (0.0330)
Week 26			
-20 min	0.094 (0.0316)**	0.066 (0.0365)	0.028 (0.0315)
25 min	0.154 (0.0311)***	0.134 (0.0357)***	0.019 (0.0309)
1 h 55 min	0.171 (0.034)***	0.146 (0.0392)***	0.025 (0.0341)
3 h 55 min	0.191 (0.0338)***	0.14 (0.0389)***	0.051 (0.0338)
23 h 40 min	0.11 (0.0331)***	0.081 (0.0331)*	0.029 (0.0328)
Week 52			
-20 min	0.068 (0.0353)	0.096 (0.0406)*	-0.027 (0.0344)
25 min	0.152 (0.0344)***	0.116 (0.0393)**	0.036 (0.0334)
1 h 55 min	0.135 (0.0355)***	0.138 (0.0407)***	–0.003 (0.0346)
3 h 55 min	0.144 (0.0354)***	0.141 (0.0408)***	0.003 (0.0347)
23 h 40 min	0.126 (0.0346)***	0.084 (0.0398)*	0.042 (0.0339)

***p<0.001; **p<0.01; *p<0.05; LSM: Least Squares Mean; SE: Standard Error

Online supplement TABLE S2. Treatment differences in trough FVC at Day 1 and Weeks 12, 26 and 52

	Trough FVC (L), LSM (SE) treatment differences

	NVA237-Placebo	Tiotropium-Placebo	NVA237-tiotropium
Day 1	0.179 (0.0208)***	0.172 (0.0241)***	0.006 (0.0209)
Week 12	0.183 (0.0313)***	0.168 (0.0360)***	0.015 (0.0307)
Week 26	0.204 (0.0334)***	0.134 (0.0383)***	0.070 (0.0330)*
Week 52	0.179 (0.0344)***	0.180 (0.0394)***	–0.001 (0.0337)

***p<0.001; *p<0.05; FVC: forced vital capacity; LSM: Least Squares Mean; SE: Standard Error

FIGURE LEGENDS

Online Supplement FIGURE S1. FEV1 (L) at each time point up to 4 h post-dose on Day 1

Data are least squares mean ± standard error; $p<0.001$ for NVA237 and tiotropium versus placebo at all timepoints 5 min to 4 h; $p<0.01$ for NVA237 versus tiotropium at all timepoints 5 min to 4 h; FEV**1**: forced expiratory volume in 1 second.

Online supplement FIGURE S2. a) Increase from baseline in TDI focal score at Week 26 **b)** Patients achieving ≥1 point improvement in TDI focal score at Week 26

a) Data are least squares mean ± standard error; *p=0.002 versus placebo; TDI: Transition Dyspnoea Index.

b) TDI: Transition Dyspnoea Index.

Online supplement FIGURE S3. Improvement in SGRQ total score versus placebo at Weeks 12, 26 and 52

Data are least squares mean ± standard error; *$p<0.05$ versus placebo; **$p<0.01$ versus placebo; ***$p<0.001$ versus placebo; SGRQ: St George's Respiratory Questionnaire.