Sleep Medicine 38 (2017) 108–112

Contents lists available at ScienceDirect

Sleep Medicine

journal homepage: www.elsevier.com/locate/sleep

Original Article

Identifying clinically important difference on the Epworth Sleepiness Scale: results from a narcolepsy clinical trial of JZP-110

Lawrence Scrima [a, 1], Helene A. Emsellem [b], Philip M. Becker [c], Chad Ruoff [d], Alan Lankford [e], Gary Bream [f, 2], Moise Khayrallah [f, 3], Yuan Lu [g], Jed Black [d, g, *]

[a] Sleep-Alertness Disorders Center, Inc., Aurora, CO, USA
[b] The Center for Sleep & Wake Disorders, Chevy Chase, MD, USA
[c] Sleep Medicine Associates of Texas, Dallas, TX, USA
[d] Stanford Sleep Medicine Center, Redwood City, CA, USA
[e] Sleep Disorders Center of Georgia, Atlanta, GA, USA
[f] Aerial BioPharma, LLC, Morrisville, NC, USA
[g] Jazz Pharmaceuticals, Inc., Palo Alto, CA, USA

ARTICLE INFO

Article history:
Received 13 February 2017
Received in revised form
10 July 2017
Accepted 12 July 2017
Available online 22 July 2017

Keywords:
Narcolepsy
Excessive sleepiness
Hypersomnolence
JZP-110
ADX-N05
Epworth Sleepiness Scale

ABSTRACT

Background: While scores ≤10 on the Epworth Sleepiness Scale (ESS) are within the normal range, the reduction in elevated ESS score that is clinically meaningful in patients with narcolepsy has not been established.

Methods: This post hoc analysis of a clinical trial of patients with narcolepsy evaluated correlations between Patient Global Impression of Change (PGI-C) and ESS. Data of adult patients with narcolepsy from a double-blind, 12-week placebo-controlled study of JZP-110, a wake-promoting agent, were used in this analysis. Descriptive statistics and receiver operating characteristic (ROC) analysis compared PGI-C (anchor measure) to percent change from baseline in ESS to establish the responder criterion from patients taking either placebo or JZP-110 (treatments).

Results: At week 12, patients ($n = 10$) who reported being "very much improved" on the PGI-C had a mean 76.7% reduction in ESS score, and patients ($n = 33$) who reported being "much improved" on the PGI-C had a mean 49.1% reduction in ESS score. ROC analysis showed that patients who improved were almost exclusively from JZP-110 treatment group, with an area-under-the-curve of 0.9, and revealed that a 25% reduction in ESS (sensitivity, 81.4%; specificity, 80.9%) may be an appropriate threshold for defining a meaningful patient response to JZP-110 and placebo.

Conclusions: A ≥25% reduction in patients' subjective ESS score may be useful as a threshold to identify patients with narcolepsy who respond to JZP-110 treatment.

1. Introduction

The chronic sleep disorder of narcolepsy is clinically characterized by a symptom pentad that includes excessive sleepiness (ES), cataplexy, disrupted nighttime sleep, hypnagogic/hypnopompic hallucinations, and sleep paralysis. Among this pentad, ES, which can be severe, is the only symptom that is present in all patients with narcolepsy and is the only symptom that is the essential component of narcolepsy diagnostic criteria [1,2]. ES is generally the first symptom to appear and has a substantial effect on daily function, also resulting in increased risk of accidents that is associated with narcolepsy [3,4].

JZP-110 (formerly known as ADX-N05) is a selective dopamine-norepinephrine reuptake inhibitor [5] that is being developed as a wake-promoting agent to treat ES and impaired wakefulness associated with narcolepsy and obstructive sleep apnea. Results of a phase 2a placebo-controlled crossover study in patients with narcolepsy suggested that a two-week treatment with JZP-110 significantly increased wakefulness, which was objectively assessed on

* Corresponding author. Jazz Pharmaceuticals, Inc., 3180 Porter Drive, Palo Alto, CA 94304, USA.
E-mail address: Jed.Black@jazzpharma.com (J. Black).
[1] Current affiliation: Sleep Expert Consultants, LLC, and Rocky Mountain Health Diagnostics, LLC, Aurora, CO, USA.
[2] No current affiliation.
[3] Current affiliation: QuatroBio, Raleigh, NC, USA.

http://dx.doi.org/10.1016/j.sleep.2017.07.006

the Maintenance of Wakefulness Test (MWT), and decreased patient-reported sleepiness, which was assessed using the Epworth Sleepiness Scale (ESS) [6] relative to placebo [7]. A subsequent 12-week phase 2b randomized controlled trial in 93 patients with narcolepsy provided further support for the efficacy of JZP-110 [8]. Relative to placebo, JZP-110 significantly increased wakefulness and decreased ES. In addition, a significantly greater percentage of JZP-110-treated patients improved, as rated by the clinicians on the Clinical Global Impression of Change (CGI-C) scale and by the patients on the Patient Global Impression of Change scale (PGI-C) scale.

While the decrease in ES with JZP-110 was statistically significant relative to placebo, it is also important to establish whether the observed changes are clinically relevant from the patient's perspective, ie, a statistically significant difference may not necessarily reflect patient-perceived tangible benefits. The normal range of the ESS has been clearly established, as have been thresholds representing mild (11–14), moderate (15–17), and severe (≥18) ES [6]; however, the minimal clinically important difference, defined as the smallest difference in score that patients perceive as providing benefit according to the definition by Jaeschke et al. [9], has not been previously characterized for ES using the ESS, although it is an old concept first utilized in the 1830s by Ernst Heinrich Weber and referred to as a "just noticeable difference" in experimental perception psychology research [10]. An analysis that incorporates a patient-reported measure of improvement as an anchor can be used to determine what represents a clinically meaningful change from the patient's perspective on the measure of interest, in this case, ES. Such an approach has previously been used for demonstrating clinically meaningful changes in pain [11] and is recommended by the US Food and Drug Administration for assessing treatment benefits [12].

A recent analysis of clinically meaningful difference with sodium oxybate treatment for narcolepsy using the ESS was conducted and revealed support for a 20% reduction in ESS as a potentially useful cut-off for defining a clinically meaningful response to treatment [13]. However, that analysis used a clinician-reported assessment as the anchor, ie, the CGI-C. This post hoc analysis was conducted to help evaluate the percentage change in ESS score that could represent a clinically relevant responder criterion for patients taking either placebo or JZP-110 (treatments), using the PGI-C as the anchor.

2. Methods

2.1. Study design

The methodology and the results of the primary analysis of this study have been previously described [8]. Briefly, the study was a double-blind, placebo-controlled study, in which patients were randomized to receive placebo ($n = 49$) or JZP-110 ($n = 44$) 150 mg/day for the first four weeks that was increased to 300 mg/day for weeks 5–12. Eligible patients were between 18 and 70 years of age, with an ICSD-2 diagnosis of narcolepsy and a baseline score ≥10 on the subjective ESS for ES and a mean baseline MWT sleep latency ≤10 min based on the average of the first four trials of a five-trial objective MWT for ES. The CGI-C [14] and the change from baseline in the MWT were co-primary efficacy endpoints at week 12 or the last assessment in the study [8]. Moreover, the change from baseline on each of the five individual MWT sessions and the ESS score for ES, along with the percentage of patients who reported any improvement on the PGI-C, were secondary endpoints. The PGI-C is rated by the patient using a seven-point Likert-type scale that ranges from "very much improved" to "very much worse" to evaluate the patient's perspective of global change in health [14].

2.2. Post hoc analyses

Descriptive statistics were used in this post hoc analysis, comparing treatment groups, using a two-sided *t*-test (changes from baseline) or Fisher's exact test (percentage of patients with PGI-C improvements). Receiver operating characteristic (ROC) curve analysis was used to compare the anchor measure, PGI-C, to the percent change from baseline on the ESS to establish the patient-reported responder criterion for this sample. In this analysis, a "true" response was defined, regardless of treatment allocation, as patients who reported PGI-C ratings of "very much improved" or "much improved." The accuracy of the patient outcomes for predicting a "true" response was evaluated using the area under the curve (AUC) of the ROC curve. A determination of an appropriate threshold for a responder definition was derived from the optimal balance between sensitivity (true positive) and specificity (true negative) from the ROC curve.

Spearman correlation based on a last-observation-carried-forward imputation for the entire population was used to evaluate the relationship between the PGI-C and the CGI-C at week 12.

3. Results

3.1. Demographic and clinical characteristics

Patients were predominantly female (64.5%) and white (74.2%), with a mean (standard deviation [SD]) age of 38.7 (12.1) years, and the majority of patients did not have cataplexy (64.5%) [8]. The mean (SD) MWT sleep onset latency was 5.7 (4.5) and baseline ESS score was 17.3 (3.3), indicating moderate to severe ES [8].

3.2. Primary analysis

The change from baseline at week 12 in ESS score with JZP-110 was significantly greater than that with placebo (−8.5 vs. −2.5; $P < 0.0001$) (Fig. 1A), resulting in a mean ESS score of 8.8, which is within the normal range. On the global impression scales (Fig. 1B), significantly more patients treated with JZP-110 than with placebo reported improvement on the PGI-C (93.0% vs. 38.3%; $P < 0.0001$) at week 12 and were also rated by the investigator as improved on the CGI-C (86.0% vs. 38.3%; $P < 0.0001$).

3.3. Post hoc analysis

The percent change in ESS showed positive values among patients who reported "very much worse" and "much worse," with higher values among the former's PGI-C category, reflecting worsening of sleepiness (increase in ESS score). In contrast, negative percent changes, indicating decreases in ESS scores (improvement in ES), were observed regardless of treatment allocation for patients who reported PGI-C improvement (Fig. 2); greater decreases in ESS were associated with greater magnitudes of the self-reported PGI-C improvement.

The AUC of the ROC curve (Fig. 3) was 0.9, indicating good accuracy for predicting a "true" response. In determining an appropriate trade-off between sensitivity and specificity, given that 80% sensitivity is commonly accepted, a 25% reduction in ESS, the subjective measure of ES, was identified as the optimal threshold on the basis of the ROC analysis and corresponded to an 81.4% true-positive rate (sensitivity) and 19.1% false-positive rate (80.9% specificity; Fig. 3).