CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

203202Orig1s000

SUMMARY REVIEW

Office of Drug Evaluation-I: Decisional Memo

Date	February 18, 2014
From	Ellis F. Unger, M.D., Director Office of Drug Evaluation-I, Office of New Drugs, CDER
Subject	Office Director Decisional Memo
NDA #	203202
Applicant Name	Chelsea Therapeutics, Inc.
Date of Submission	August 13, 2013
PDUA Goal Date	February 14, 2014
Proprietary Name/ Established (USAN) Name	Northera droxidopa
Dosage Forms/ Strengths	Oral capsules: 100 mg, 200 mg, and 300 mg
Indication	Short-term treatment of orthostatic dizziness, lightheadedness, or the "feeling that you are about to black out" in adult patients with symptomatic neurogenic orthostatic hypotension caused by primary autonomic failure (Parkinson's disease, multiple system atrophy and pure autonomic failure), dopamine beta hydroxylase deficiency, and non-diabetic autonomic neuropathy
Action:	*Approval – Subpart H*

Material Reviewed/Consulted - Action Package, including:	
Project Manager	Anna Park
Medical Officer Clinical Review	Shari Targum, Melanie Blank
Clinical Pharmacology Review	Sreedharan Sabarinath, Rajnikanth Madabushi; Yaning Wang
Statistical Review	Jialu Zhang, James Hung
Pharmacology Toxicology	Donald Jensen, Thomas Papoian
Executive Cancer Assessment Committee	Adele Seifried, David Jacobson Kram
Chemistry Manufacturing and Controls	Lyudmila Soldatova, Kasturi Srinivasachar, Olen Stephens, Yvonne Knight
ONDQA Biopharmaceutics Review	Tien Mien (Albert) Chen, John Z. Duan, Angelica Dorantes
Method Validation	Michael Trehy, John F. Kauffman, Jamie D. Dunn, Youbang Liu
Statistical Review - Carcinogenicity Study	Steven Thomson, Karl Lin
Controlled Substance Staff	Jovita Randall-Thompson; Silvia Calderon, Michael Klein, Lori Love
Division of Good Clinical Practice Compliance, Office of Scientific Investigations	Sharon Gershon, Susan Thompson
Office of Scientific Investigation	William Taylor, Charles Bonapace, Gajendiran Mahadevan, Sripal Mada, Mehul Mehta
Division of Medication Error Prevention and Analysis	Jean Olumba, Lisa Khosla, Karen Bengston
Division of Medical Policy Programs	Sharon Mills, Barbara Fuller
Risk Evaluation and Mitigation Strategy (REMS) Review	Somya Dunn, Claudia Manzo
Office of Prescription Drug Promotion	Emily Baker
QT/IRT	Qianyu Dang, Joanne Zhang, Dhananjay Marathe, Monica Fiszman, Nitin Mehrotra
Cross-Discipline Team Leader	Shari Targum
SEALD Labeling Team	Elizabeth Donohue, Eric Brodsky
Director, Division of Cardiovascular and Renal Products	Norman Stockbridge

Regulatory Action:

The applicant is seeking approval of droxidopa for the following indication:

> "NORTHERA™ is indicated for the treatment of symptomatic neurogenic orthostatic hypotension in adult patients with primary autonomic failure (Parkinson's disease, multiple system atrophy and pure autonomic failure), dopamine beta hydroxylase deficiency, and non-diabetic autonomic neuropathy. Clinical benefit has been demonstrated in short-term trials; long-term benefits have not been verified."

I believe that the appropriate action is accelerated approval, based on new evidence of short-term effectiveness submitted August 13, 2013. Approval under Subpart H will be granted because it is critical to establish that the effect of droxidopa in NOH, a chronic illness, is maintained in at least a subset of the population treated.

The actual indication will be:

> "NORTHERA is indicated for the treatment of orthostatic dizziness, lightheadedness, or the "feeling that you are about to black out" in adult patients with symptomatic neurogenic orthostatic hypotension (NOH) caused by primary autonomic failure [Parkinson's disease (PD), multiple system atrophy and pure autonomic failure], dopamine beta-hydroxylase deficiency, and non-diabetic autonomic neuropathy. Effectiveness beyond 2 weeks of treatment has not been established. The continued effectiveness of NORTHERA should be assessed periodically."

Description/Mechanism of Action:

Droxidopa is a synthetic catecholamine acid analogue that is metabolized by dopa decarboxylase to norepinephrine, which is thought to increase blood pressure (BP) through binding and activation of adrenergic receptors. The applicant also asserts that droxidopa increases neuronal levels of NE, which could lead to sustained effects.

Disease Background:

See Office of Drug Evaluation-I Decisional Memo dated March 28, 2012.

Regulatory History:

Droxidopa is a new molecular entity that is not approved in the U.S.; however, it has been approved in Japan since 1989 for essentially the same indication now sought in the US. The drug was developed under Chelsea's IND 077248.

For a complete history of the submission and detailed reviews of the current submission, see the clinical (Shari Targum, December 5, 2013; February 5, 2014) and statistical reviews (Jialu Zhang, December 3, 2013).

For the original submission, the principal evidence of droxidopa's effectiveness was provided by Study 301. Two other studies were negative. After all primary reviews were completed and the NDA was presented at the Cardiovascular and Renal Drugs Advisory Committee Meeting, we became aware of irregularities at 2 of the clinical sites in Study 301. Data from the largest enrolling site, in particular, were highly irregular (see below). On each of 4 endpoints, there was a striking effect size, with remarkably small variances within both treatment groups. The results

from this site were highly implausible and deemed unreliable. With removal of the data from this site, the results of this single positive study were no longer statistically significant.

Of note, even if Study 301 had been clearly positive, it provided evidence of efficacy over a period of only 1 week. I believed, and most on the review team believed, that a demonstration of efficacy over a period of 1 week was not sufficient for what is unquestionably a chronic disorder. Moreover, Studies 302 and 303 were randomized withdrawal studies in droxidopa responders. The failure of those studies (although Study 302 was not entirely failed) raised further concern about the duration of effect.

After our Complete Response action on March 28, 2012, the applicant submitted a request for formal dispute resolution. The appeal was reviewed by Dr. John Jenkins, Director, Office of New Drugs, and ODE-I's Complete Response was upheld on February 8, 2013.

Dr. Jenkins' conclusion was:

> "…that at least one additional strongly positive adequate and well-controlled trial is necessary to support a demonstration of effectiveness of droxidopa in NOH. …ideally there should be evidence of the durability of the effectiveness of droxidopa since NOH is a chronic condition and it can be expected that patients will take droxidopa long-term. I note, however, that in the agreement reached between Shire and FDA on the additional trials needed to support continued marketing of midodrine the Agency agreed to accept data demonstrating a short-term benefit of midodrine as adequate evidence to support continued approval. Therefore, I believe that data strongly demonstrating a short-term clinical benefit (e.g., improvement in symptoms or ability to function) of droxidopa in patients with NOH would be adequate to support approval, with a possible requirement to verify durable clinical benefit postapproval."

In response to the Complete Response, the applicant submitted a new Study 306B, a study of subjects with Parkinson's disease that was re-engineered with a dizziness endpoint.

Significant milestones and agreements from the development program are summarized below:

- 1/2007: Orphan drug designation granted for the NOH indication
- 3/2007: Pre-IND – FDA stated that a single study could support approval if the level of significance approximates that of two studies ($p \sim 0.00125$).
- 8/2007: End-of-Phase 2 Meeting – Two phase 3 trials with a clinical endpoint, supplemented with long-term data, would be sufficient for approval. The design of Study 303, planned to have 3 months of treatment and a 2-week randomized withdrawal period, was acceptable to demonstrate durability of treatment effect.
- 9/2007: IND opened – Phase 3 protocols (Studies 302 and 303) submitted; allowed to proceed.
- 2/2008: Agreement on a Special Protocol Assessment for Study 301; agreement that the planned "long-term" exposure would be adequate for the safety evaluation. The Division stated that they expected two successful trials ($p<0.05$) to support efficacy.
- 8/2008: Fast Track designation granted
- 11/2009: Type C Meeting – discussion of sponsor's planned modifications of Study 301. The Division noted that Study 302 could not be used as one of two studies to support efficacy because it failed on its 1° outcome measure.

- 1/2010: Correspondence to sponsor – FDA agreed upon a change in the primary endpoint of Study 301 from the Orthostatic Hypotension Symptoms Assessment (OHSA) Item 1 to the Orthostatic Hypotension Questionnaire (OHQ).
- 12/2010: Pre-NDA meeting – FDA reminded Chelsea that one trial is not usually sufficient for approval. FDA requested validation data for the PRO instruments used in the studies, as well as support for the view that the observed effect size in Study 301 is clinically meaningful.
- 9/28/2011: Original submission of NDA 203202
- 2/23/2012: NDA presented at the Cardiovascular and Renal Drugs Advisory Committee Meeting. Vote: 7 to 4 in favor of approval.
- 3/28/2012: Complete Response issued
- 12/12/2012: Request for Formal Dispute Resolution of the Complete Response
- 1/10/2013: Face to face meeting with representatives of the Office of New Drugs to discuss basis for appeal
- 2/8/2013: Denial of appeal (decision written by Dr. John Jenkins, Director, Office of New Drugs)
- 7/3/13: Applicant Resubmitted application, with results of new trial, Study 306B
- 7/25/13: Incomplete Response Letter issued, based on deficiencies in data sets (inconsistencies in variable names, such that the executable programs would not function), missing narratives for subjects who discontinued 2° to adverse events, and a discrepancy in the clinical study report for 306B
- 8/13/13: Application resubmitted, and subsequently filed.
- 1/14/14: NDA presented a second time at the Cardiovascular and Renal Drugs Advisory Committee. Vote: 16 to 1 in favor of approval.

Chemistry Manufacturing and Controls:

The chemistry, manufacturing and controls (CMC) review team initially identified several deficiencies, including issues regarding testing procedures and specifications for drug substance and drug product, impurities, and questions regarding stability data for the drug product. Pending issues were resolved, except that the stability data were not sufficient for granting expiry for 300 mg capsules. Based on data subsequently submitted by the applicant, a 48-month expiration period will be granted for the 100- and 200-mg capsules packaged in high density polyethylene bottles, and a 12-month expiration period will be granted for 300 mg capsules packaged in high density polyethylene bottles. An expiration period of 36 months will be granted for 100- and 200-mg capsules packaged in aluminum foil blister packs.

Pharmacology/Toxicology:

The review team raised concerns regarding the potential neurotoxicity of DOPEGAL (3,4-dihydroxyphenylglycolaldehyde), a potentially neurotoxic metabolite to noradrenergic neurons in the locus ceruleus that may be formed from NE. Treatment with droxidopa in rats and dogs for up to 52 weeks did not produce clear evidence of neurotoxicity; however, the specific region of the brain that would be expected to be affected by DOPEGAL, the locus ceruleus, was not examined. A suggestion was made that a focused 28-day rat study could possibly examine the issue. In the final analysis, however, the pharmacology/toxicology review team opined that conduct of such an animal study would be resource intensive to perform in an interpretable manner, and they found the NDA approvable.

Carcinogenicity:

See Office of Drug Evaluation-I Decisional Memo dated March 28, 2012.

Site Inspections:

Site inspections from the original studies are outlined in the Office of Drug Evaluation-I Decisional Memo dated March 28, 2012.

The Office of Scientific Investigation (OSI) inspected 4 domestic clinical investigator sites in support of Study 306B for the NDA resubmission. Minor regulatory violations were identified during the inspection of Site 146, with issuance of one Form FDA 483 for failure to follow the investigational plan. Numerous regulatory violations were identified during inspection of Site 122, and three Form FDA 483's were issued; however, the violations were not thought to affect the primary efficacy data or subject safety. OSI recommended that the data from all sites could be considered reliable, but suggested that sensitivity analyses, omitting data from Site 122, could be worthwhile. Of note, the largest enrolling site, Site 132, was inspected, and there were no significant findings, except that subjects #04 and #10 were terminated early due to a lack of drug efficacy prior to their Visit 4, such that their data did not contribute to the primary endpoint.

Pharmacokinetics, Abuse Potential, and QT Effects:

See Office of Drug Evaluation-I Decisional Memo dated March 28, 2012.

Evidence of Effectiveness:

Studies 301, 302, and 303 were the original studies submitted to support efficacy, and they were reviewed in the first review cycle. They are summarized briefly below. See the original review memoranda and the Office of Drug Evaluation-I Decisional Memo dated March 28, 2012 for more detail. Study 306B is a new study, submitted with the applicant's complete response, and is described in more detail.

Of the 3 studies the applicant originally submitted to support efficacy, Study 301 provided the primary evidence of efficacy. Study 302 was called "supportive" by the applicant, and Study 303, an open-label extension study followed by a randomized withdrawal, was conducted to provide evidence of maintained effect and long-term safety data.

All studies enrolled adult subjects with NOH (fall in systolic BP $\geq$ 20 mmHg or diastolic BP $\geq$ 10 mmHg within 3 minutes of standing) associated with primary autonomic failure (Parkinson's disease, multiple system atrophy, or pure autonomic failure), dopamine beta-hydroxylase deficiency, or non-diabetic autonomic neuropathy. Patients with diabetes and those with significant cardiac, renal, and hepatic disease were excluded.

Use of midodrine, other vasoconstrictors, tricyclic antidepressants, and norepinephrine reuptake inhibitors was not permitted. Use of medications for Parkinson's disease was permitted.

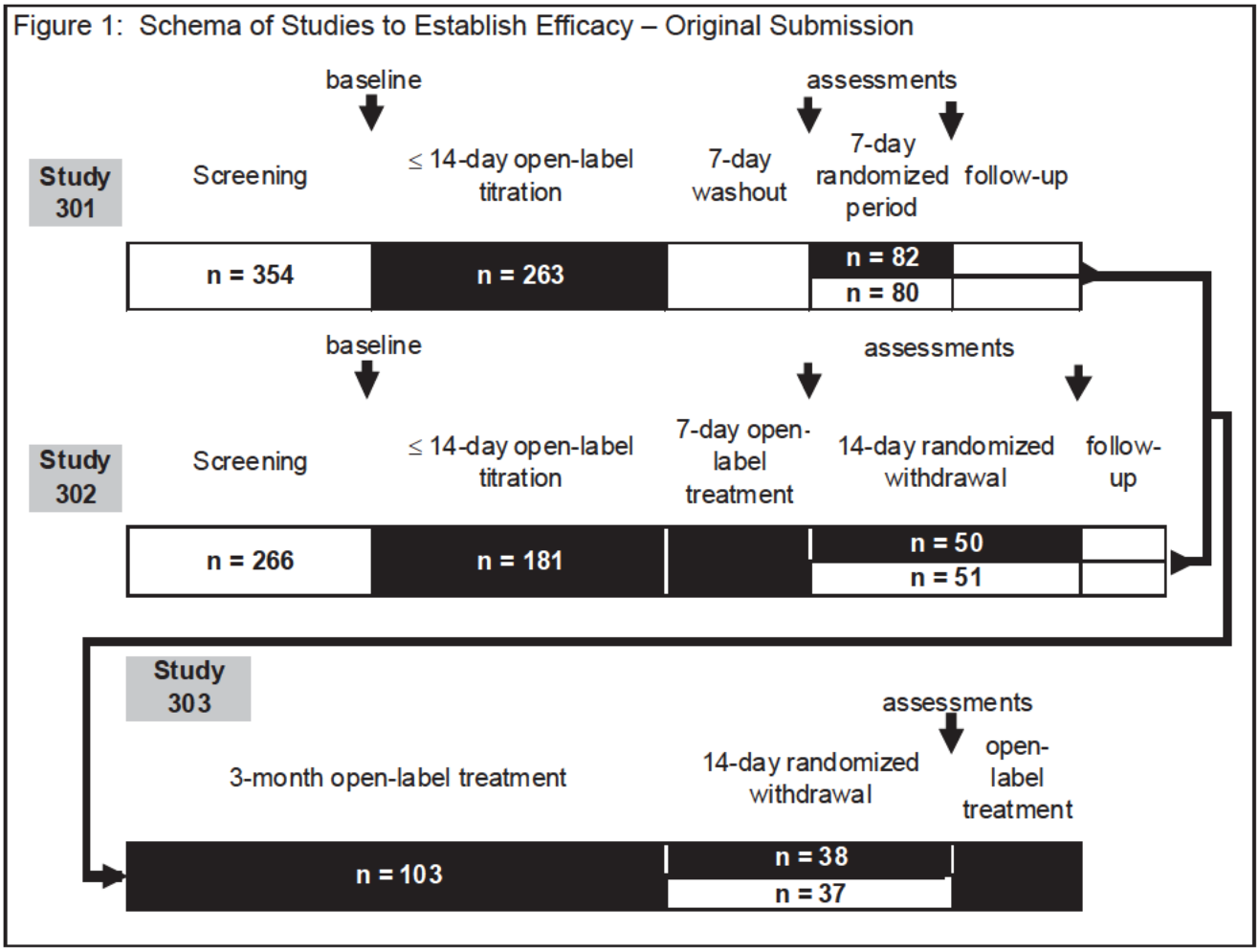
Studies 301 and 302

Studies 301 and 302 utilized *empiric predictive enrichment schemes*, whereby subjects who met general inclusion and exclusion criteria during screening were enrolled in an open-label titration period. Only subjects who tolerated droxidopa and appeared to have a favorable symptom

response during the open-label titration period were included in the randomized treatment phase, and only these subjects counted towards the assessment of endpoints:

- Subjects had to tolerate the drug's side effects and not exceed a BP of 180/110 mmHg when supine,[1] after 5 minutes of sitting, or after 3 minutes of standing.
- Subjects had to improve by ≥ 1 point on question 1 on the OHSA scale[2] and demonstrate an increase in systolic BP of ≥10 mmHg, assessed 3 minutes post-standing.

Figure 1: Schema of Studies to Establish Efficacy – Original Submission



Study 302 was conducted before Study 301.

[1] "Supine" BP was actually assessed with the head and torso elevated 30 degrees.

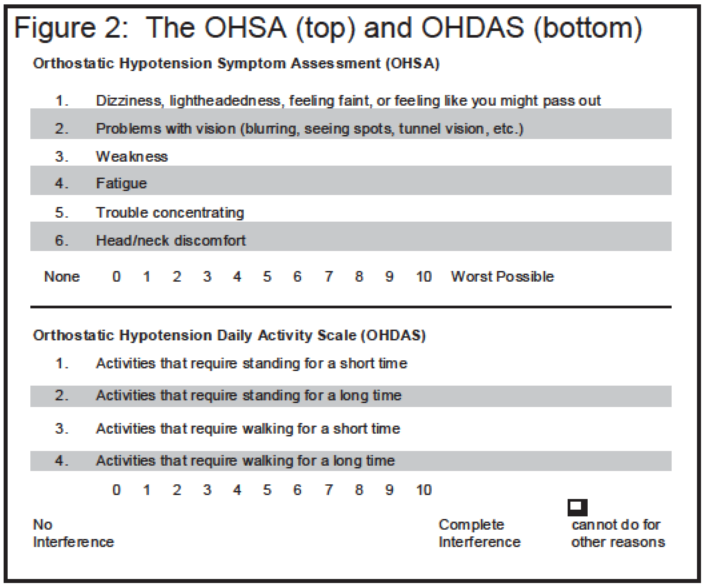
[2] Question 1 of the OHSA was assessed using an 11-point scale (0 = none; 10 = worst possible) that asked each patient to "…circle the number…that best rates how severe your symptoms from low blood pressure have been on the average over the last week….please rate the symptoms that are due only to your low blood pressure problem: Dizziness, lightheadedness, feeling faint, or feeling like you might black out."

- Study 302 was a randomized withdrawal study. See Figure 1. An optimal dose was to be found for each subject during the titration period. This period was followed by 7 days of open-label treatment and then randomized withdrawal after a total of 1 to 3 weeks treatment. The 1° endpoint of Study 302 was the change in response to Item 1 of the OSHA.[2]

- Study 301 featured an open-label titration period to find optimal dose, followed by a 7-day washout period and a 7-day randomized double-blind treatment period. See Figure 1. Study 301 is most appropriately viewed as a 1-week, randomized, double-blind, placebo-controlled study utilizing predictive enrichment (based upon an empiric response during open-label titration). The 1° endpoint of Study 301 was the change in OHQ composite score, assessed at 1 week.

The OHQ is comprised of two self-assessment questionnaires: the OHSA, which has 6 items pertaining to symptoms, and the Orthostatic Hypotension Daily Activity Scale (OHDAS), which has 4 items pertaining to ability to stand and walk (i.e., function) (*Figure 2*). The OHQ is the unweighted average of the OHSA and OHDAS, i.e., the symptom and function questionnaires are weighted equally.

Figure 2: The OHSA (top) and OHDAS (bottom)



Orthostatic Hypotension Symptom Assessment (OHSA)

1. Dizziness, lightheadedness, feeling faint, or feeling like you might pass out
2. Problems with vision (blurring, seeing spots, tunnel vision, etc.)
3. Weakness
4. Fatigue
5. Trouble concentrating
6. Head/neck discomfort

None 0 1 2 3 4 5 6 7 8 9 10 Worst Possible

Orthostatic Hypotension Daily Activity Scale (OHDAS)

1. Activities that require standing for a short time
2. Activities that require standing for a long time
3. Activities that require walking for a short time
4. Activities that require walking for a long time

0 1 2 3 4 5 6 7 8 9 10

No Interference — Complete Interference — ☐ cannot do for other reasons

The Study Endpoints and Labeling Development (SEALD) review explains that OHSA Item #1 is preferred over OHQ as a study endpoint because OHSA Item #1 represents core symptoms of NOH, whereas the OHQ includes items that have not been documented to be core disease-defining symptoms.

Studies 301 and 302 both included an open-label titration phase: droxidopa was dosed initially at 100 mg TID with incremental titration up to 600 mg TID. As explained in the reviews, the goal was to find the optimal dose for each subject based on symptom response, BP changes, and side effects.

Results:

Study 302 screened 266 patients, of whom 85 (32%) were screening failures and 80 (30% of the total, or 44% of those successfully screened) were titration failures. Thus, only 101 subjects were randomized to droxidopa (n=50) or placebo (n=51).

There was reasonable consistency of demographic and baseline characteristics between the 2 treatment groups. Approximately 44% of randomized subjects had Parkinson's disease, 30% had multiple system atrophy, 18% had pure autonomic failure, 5% had non-diabetic autonomic neuropathy, and 1 patient (1% – assigned to placebo) had dopamine beta-hydroxylase deficiency. Mean age was 65. Patients were 62% male and 96% Caucasian. Fifty-six percent of subjects were from US sites.

Study 302 did not succeed on its 1° endpoint, Item 1 of the OSHA. Both groups worsened during the 14-day randomized withdrawal period, with no statistically significant difference between groups (p=0.51). Specifically, the mean response to Item 1 of the OHSA (dizziness) worsened in the placebo group from 2.1 prior to randomized withdrawal to 4.0 after withdrawal; in the droxidopa group, the mean response worsened from 2.1 prior to randomized treatment to 3.5 after treatment. Thus, even though patients were *required* to improve by ≥ 1 point on Item 1 of the OHSA in response to open-label droxidopa to gain entry into the randomized portion of the study, there was no statistically significant difference between groups on the change in this measure following a 14-day withdrawal. Although results on Item 1 of the OHSA were not statistically significant, a *post hoc* analysis of the OHQ showed a statistically significant difference between groups.

Of note, there was no difference between groups in standing systolic BP at end-of-study.

Study 301

In light of the findings in Study 302 – the negative results on Item 1 of the OHSA and the positive *post hoc* results in the OHQ – the applicant changed the 1° endpoint of Study 301 from OHSA Item 1 to the OHQ in protocol amendment 4. The applicant also increased the size of Study 301 in order to achieve adequate power to show a difference between treatment groups. The study was ongoing when these changes were made; however, the updated statistical analytical plan was submitted to the FDA prior to performing any analyses of efficacy.

Study 301 enrolled subjects at 94 centers in 9 countries. In total, 354 patients were screened. The fractions of subjects who were successfully screened – and randomized – were similar to those in Study 302. A total of 162 subjects were in the analytic population (n=82 for droxidopa; n=80 for placebo).

Baseline characteristics were similar in the two groups. Mean age was 57. Approximately half of subjects were male. The study included subjects with Parkinson's disease (40%), pure autonomic failure (34%), multiple system atrophy (16%), and non-diabetic autonomic neuropathy (5%). No subject had dopamine beta-hydroxylase deficiency. Forty percent of subjects were from US sites.

Based on the data submitted by the applicant, there was a statistically significant difference between the droxidopa and placebo groups, favoring droxidopa, with respect to mean change in the OHQ from randomization to end of study (p=0.003). The treatment effect was 0.9 units on a scale from 0 to 10 (11 units, *Table 1*).

In contrast to Study 302, there was a nominally statistically significant increase in standing systolic BP in the droxidopa group (approximately 7 mmHg, although the results were suspect, see below). Of note however, there was no effect on the orthostatic decrease in systolic BP with standing (i.e., no effect on orthostasis), and no significant effect on diastolic BP.

Late in the review cycle, after most of the reviews were filed, the review team recognized that the applicant had not provided the 301 study results by center or by country.

Site 507 in the Ukraine enrolled the largest number of subjects (16, or 10% of the total), and demonstrated the largest effect size of any site contributing more than 3 subjects (a 3.6-unit advantage for the droxidopa group). Closer examination of the Site 507 data revealed

extraordinarily small within-group variances for all of the important end-of-study measures (composite OHQ, Item 1 of the OHSA, the CGI-I score, and standing systolic BP, Figure 3), as evidenced by the statistical certainty of the OHQ findings ($p<0.001$) (source: applicant's data files; OHQ calculated from applicant's data files; p-value: email communication from Dr. Jialu Zhang.). For each parameter, the recorded values for all subjects in the droxidopa group were markedly superior to those of all subjects in the placebo group (Figure 3). Not obvious from brief perusal of the figure, but quite improbable, was the separation of the results for GCI-I at the end of the study. For this subjective measure, every droxidopa subject had a value of 1 or 2; every placebo subject had a value of 3 or 4. There was essentially no variability within either group, and no overlap between groups, despite the lack of numerical separation.

Table 1: Primary Endpoint in Study 301: OHQ (units)

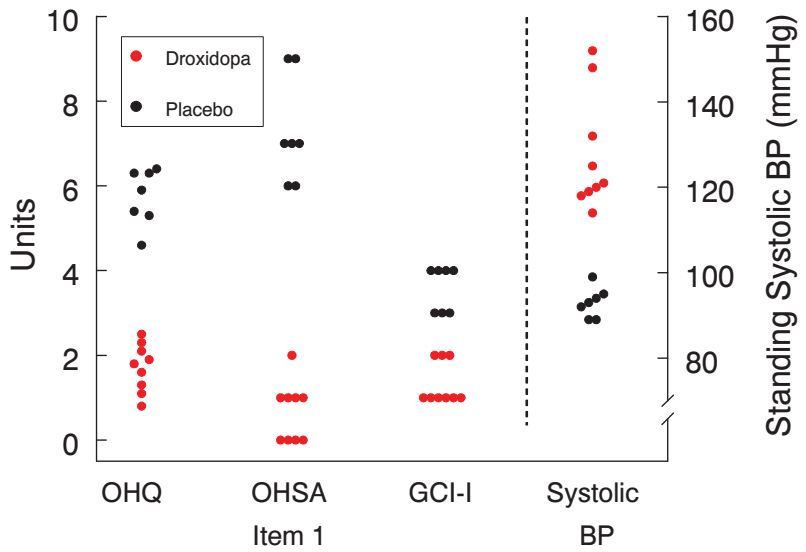
	droxidopa n=81	placebo n=79
Randomization mean (SD)	5.11 (1.96)	4.97 (2.41)
End of study mean (SD)	3.29 (2.20)	4.04 (2.61)
Change* mean (SD)	-1.83 (2.07)	-0.93 (1.69)

* p = 0.003

These results were considered to be implausible, and quite remarkable, different from all other data. Although fraud is difficult to prove in clinical trials, many on the review team were concerned about this possibility.

The results for the 1° endpoint were calculated with and without site 507, and shown in Table 2. With the removal of data from 507, the overall treatment effect on the OHQ diminished from 0.89 units to 0.56 units, and the p-value, which was persuasive for the study overall (p=0.003), was no longer statistically significant (Table 2, center). It was clear that the data contributed by site 507 were disproportionately responsible for the treatment effect.

Figure 3: End-of-Study OHQ, OHSA Item 1, CGI-I, and Standing BP at Site 507, Study 301



The effect of Site 507 on the standing systolic BP results was similar. For the entire study, there was a 7 mmHg increase in standing systolic BP in the droxidopa group, p<0.001, per the applicant's analysis. When the BP data from Site 507 were eliminated, the mean change in BP

in the droxidopa group was just 4 mmHg, and the treatment effect was not statistically significant.

It should be noted that data from Site 505 were similarly aberrant, except that the first patient randomized at that site did not appear to have anomalous results, and the overall effect size smaller.

Study 301 could be considered in two ways:

1) If data from Site 507 were deemed acceptable, that would mean that the dossier included a single positive study. But with only 1 positive study and one site disproportionately responsible for the favorable treatment effect, the data would not constitute sufficient evidence of efficacy upon which to base an approval action.

2) If data from Site 507 were deemed inadmissible, then Study 301 was not positive; none of the studies were positive.

Table 2: Site 507 in Study 301 – Effect on Primary Endpoint OHQ

	OHQ		
	randomization	**end of study**	**change**
All Sites			
Droxidopa N = 82	5.10	3.30	-1.81
Placebo N = 80	4.99	4.07	-0.92
treatment effect			-0.89
p-value			0.003
Omit Site 507			
Droxidopa N = 73	4.98	3.49	-1.49
Placebo N = 73	4.84	3.91	-0.93
treatment effect			-0.56
nominal p-value			0.07
Site 507 Only			
Droxidopa N = 9	6.07	1.70	-4.37
Placebo N = 7	6.51	5.75	-0.76
treatment effect			-3.61
nominal p-value			0.000000005

Study 303

Study 303 was a phase 3, multicenter study to examine the efficacy and safety of longer-term administration of droxidopa. The main study features included a 3-month period of open-label treatment followed by a 2-week, randomized withdrawal phase (double-blind, placebo-controlled). This was followed by a 9-month period of open-label treatment.

Subjects previously enrolled in Studies 301 or 302 could be enrolled in Study 303, including those who had at least a 1-point improvement in Item 1 of the OHSA in response to droxidopa during the open-label titration phase, but who were not randomized because they failed to demonstrate the requisite increase in BP. Thus, Study 303 was also empirically enriched with droxidopa "responders." (Moreover, these "responders" had to tolerate the drug for 3 months or longer.)

The 1° efficacy endpoint in Study 303 was the mean change in OHQ score from randomization to the end of the 2-week randomized treatment period.

Study 303, with its randomized withdrawal design, was the only trial that evaluated maintenance of efficacy after a reasonably long period of time (3 months).

Of the 102 subjects enrolled and treated in Study 303, 27 did not enter in the double-blind phase. Thus, 75 subjects were randomized 1:1 to either droxidopa or placebo for a 2-week period (droxidopa, n=38; placebo, n=37).

At the conclusion of the 2-week double-blind randomized withdrawal phase, the treatment groups did not differ with respect to total OHQ score, the 1° endpoint. The OHQ worsened in both treatment groups: by 0.90 points in the placebo group and by 0.57 points in the droxidopa group, i.e., the changes ran *counter* to a demonstration of efficacy – despite the enrichment scheme.

The applicant speculated that the pharmacodynamic effects of droxidopa might persist far beyond the drug's plasma half-life, i.e., for several days, and that such "carry over" effects, if they exist, could have obscured the difference between groups in the 14-day randomized withdrawal period. Of course, this is *possible*, but the lack of a demonstrable effect in the randomized withdrawal segment of Study 303 (and 302) is also compatible with there being no sustained drug effect.

There was no statistically significant difference in standing systolic BP between treatment groups at the end of the randomized double-blind treatment period. There was a *paradoxical trend towards decreased standing systolic BP in the droxidopa group* from randomization to the end of double-blind treatment (mean decrease of 8.4 mmHg), whereas there was no change in standing systolic BP in the placebo group

Evidence of Effectiveness – First Cycle View:

Study 302 was a negative study, based on its prespecified 1° endpoint (Item 1, OHSA), although the analysis of an exploratory *post hoc* endpoint (OHQ) was positive. Study 303, a randomized withdrawal study that was highly enriched for responders, was completely negative. On its face, Study 301 was positive, but results were driven by the largest site, Site 507. The fact that Site 507 contributed 10% of the subjects and was responsible for 37% of the treatment effect undercut the persuasiveness of the study,[3] and provided a strong rationale for not accepting Study 301 as the basis for a single-trial approval. This would have been true even if there had been no concerns about data integrity at Site 507. With data integrity issues around Site 507, and to a lesser extent around Site 505, we did not view the study as positive. In light of the concerns about Study 301 and the inconsistency of the overall findings, the collective evidence did not support approval.

In exploratory subgroup analyses, it was noted subsequently that there was a statistically significant treatment effect on OHSA Item #1 within the US (65 patients; treatment difference was 1.07 units; $p = 0.023$), and in Western countries (82 patients in the US, Italy, Germany, Canada, and Austria; mean treatment effect was 1.11 units; $p = 0.017$). As explained in Dr. Papadopoulos' Study Endpoints and Labeling Development (SEALD) review, OHSA Item #1 is preferred over OHQ as a study endpoint because OHSA Item #1 represents core symptoms of NOH, whereas the OHQ includes items that have not been documented to be core disease-defining symptoms. Thus, a nominally statistically significant finding on OHSA Item #1 within the U.S. subgroup, although clearly an exploratory analysis, provides some support of efficacy.

[3] Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products; May, 1998

The Division had repeatedly advised the applicant to provide evidence supporting the durability of droxidopa's effect; however, the study with the greatest potential to show this (Study 303) did not succeed. Similarly, Study 302 was a randomized withdrawal study that did not support durability of treatment effect after ~3 weeks of treatment.

After we issued the Complete Response, the applicant submitted a request for formal dispute resolution. The appeal was reviewed by Dr. John Jenkins, Director, Office of New Drugs, and the Complete Response was upheld. Dr. Jenkins concluded that at least one additional strongly positive adequate and well-controlled trial would be necessary to demonstrate effectiveness of droxidopa. He noted that ideally there should be evidence of durability of effectiveness; because NOH is a chronic condition and patients can be expected to take droxidopa long-term. He noted that data strongly demonstrating a short-term clinical benefit would be adequate to support approval, with a possible requirement to verify durable clinical benefit post-approval.

In response to the Complete Response, the applicant submitted a new Study 306B, a study of subjects with Parkinson's disease that was re-engineered with a dizziness endpoint.

Study 306B

The resubmission included a single phase 3 trial, Study 306B. This was an all-U.S., multicenter, randomized, parallel-group, placebo-controlled, double-blind study, to evaluate the efficacy of droxidopa in patients with symptomatic NOH associated with Parkinson's disease. For study entry, subjects had to have symptomatic NOH associated with PD, OHQ >3, clinician CGI-S >3, and a decrease of $\geq$ 20 mm Hg systolic or 10 mm Hg diastolic blood pressure within 3 minutes of standing. Following randomization (1:1) to droxidopa 100 mg or placebo TID, there was a dose titration period (up to 14 days; mock titration in the placebo arm), followed by an 8-week treatment period. The aim of dose titration was to find an optimal dose for each subject. The dose was to be increased in step-wise fashion in 100-mg increments until the subject became asymptomatic; experienced intolerable side effects; had an increase in their supine systolic or diastolic blood pressure to 180 or 110 mmHg, respectively; or reached the highest dose (600 mg TID). The primary endpoint was the OHQ composite score at Week 8.

In accordance with the study plan, an interim analysis was carried out on the first 51 patients in January, 2011. That review found a conditional power of < 0.1, the stopping rule for futility, and the Data Monitoring Committee recommended termination of the trial. Because the interim data were encouraging on the rate of falls, a 2° endpoint, the applicant opted to continue the study but split it into Study 306A (those 51 patients whose treatment codes had been unblinded for the interim analysis) and Study 306B. The primary endpoint for 306B was changed from the OHQ composite score to patient-reported falls, to be assessed at Week 8. In order to maintain study integrity, subjects in Study 306A were not to be included in the analyses of Study 306B.

After subsequent discussions with FDA, the applicant recognized the need for an additional study to support approval. In November, 2012, therefore, the 1° endpoint was changed a second time: from patient-reported falls at Week 8 to OHSA Item 1 at Week 1. At that juncture, 122 patients had been randomized in Study 306B. At one point it was discovered that an unblinded statistical team had access to the treatment codes for all Study 306 subjects, rather than the 51 subjects in the interim analysis. This is discussed below.

The population used for the analysis of the 1° endpoint was the full analysis set (FAS), comprised of all randomized subjects who received at least one dose of study drug and had an assessment of OHSA Item 1 at Week 1. According to the final statistical analysis plan, the 1°

endpoint would be tested using an analysis of covariance (ANCOVA) model adjusting for baseline OHSA Item 1 score. However, if any of the ANCOVA assumptions (including normality of the residuals) were not met, then the 1° analysis was to be changed to a non-parametric model. The violation of assumptions was to be determined by visual inspection of the plots, without a formal test of normality.

The 2° efficacy variables in Study 306B were:

- Δ OHSA Item #1 from baseline to Week 2
- Δ OHSA Item #1 from baseline to Week 4
- Δ lowest standing systolic blood pressure after standing 0 to 3 minutes, baseline to Week 1
- Δ OHSA Item #1 from baseline to Week 8
- Rate of patient reported falls from baseline to the end of the study
- Δ OHQ from baseline to Week 8

These were to be tested sequentially in the order listed if the 1° efficacy endpoint was statistically significant at $p < 0.05$.

Results:

A total of 174 subjects were randomized in Study 306B, of whom 147 had Week 1 assessments and were included in the analysis of the 1° endpoint: 69 patients in the droxidopa group and 78 patients in the placebo group. The 27 subjects who did not have Week 1 assessments were unequally distributed: there were nearly 3 times as many in the droxidopa group (20 subjects, 22%) as in the placebo group (7 subjects, 8%). Thus, almost a quarter of subjects in the droxidopa group did not undergo Week 1 assessments. Importantly, the leading reasons were consistent with dissatisfaction with drug effects: adverse events (6, half of these were hypertension), lack of efficacy/treatment failure (4), withdrawal of consent (3), and investigator decision (2). The other 5 subjects dropped out for seemingly uninformative reasons. Of the 7 subjects who dropped out in the placebo group, 4 dropped out in association with adverse events. *What is quite remarkable is that, despite not having received an active drug during the titration phase, none of the 85 placebo subjects dropped out because of treatment failure, whereas 4 subjects in the droxidopa group dropped out for treatment failure!*

Two-thirds of subjects were male; 96% were Caucasian, and mean age was 72 years. Subjects were fairly well balanced with respect to baseline characteristics, with the exception of fludrocortisone use: 20% of placebo subjects and 34% of droxidopa subjects reported fludrocortisone use at baseline.

At baseline, the mean OHSA Item 1 score was 5.1 in both groups. The applicant reported that the assumptions for the ANCOVA were not met and used a non-parametric method to conduct the analysis. The resulting treatment difference was -1.0, 95% confidence interval (CI) -2.0, 0, $p = 0.018$. Dr. Zhang, however, did not find any obvious deviation from ANCOVA assumptions, and Table 3 summarizes the results she obtained by ANCOVA, which are in agreement with those the applicant calculated using the same method. Comparing the change in OHSA Item 1 from baseline to Week 1 in the droxidopa and placebo groups, the treatment effect was -0.94, 95% CI -1.78, -0.1, $p = 0.028$.

Table 3: Primary Endpoint Results: OHSA Item 1 Score

	Droxidopa			Placebo		
	N	mean	SD	N	mean	SD
Baseline	69	5.1	2.04	78	5.1	2.33
Week 1	69	2.8	2.44	78	3.8	2.75
Least square mean difference	-0.94 with 95% CI (-1.78, -0.1)					
p-value from ANCOVA model	0.028					

Figure 4: Cumulative Distribution (left) and Histogram (right) Showing Change in OHSA Item 1: Baseline to Week 1

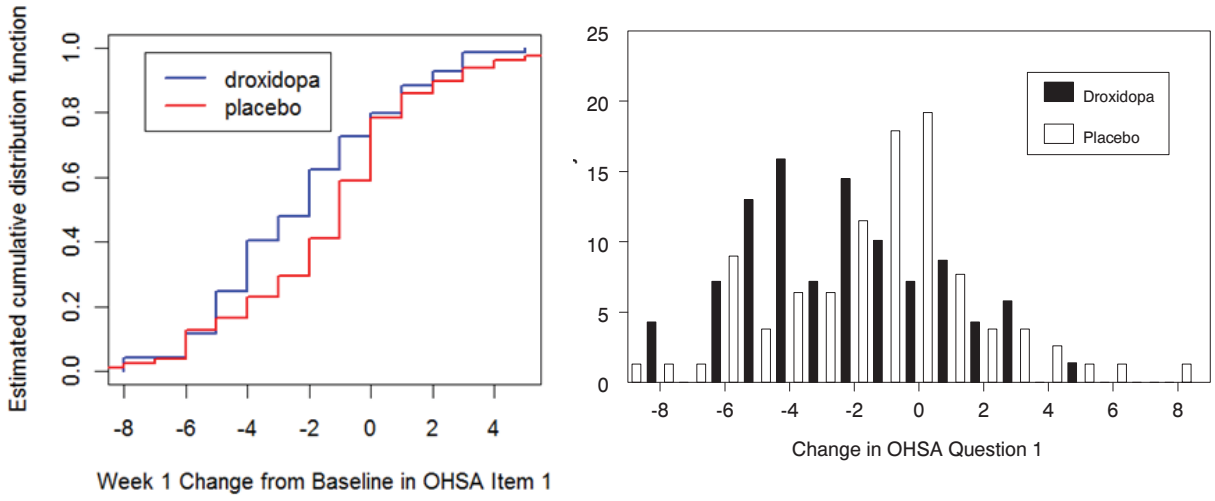
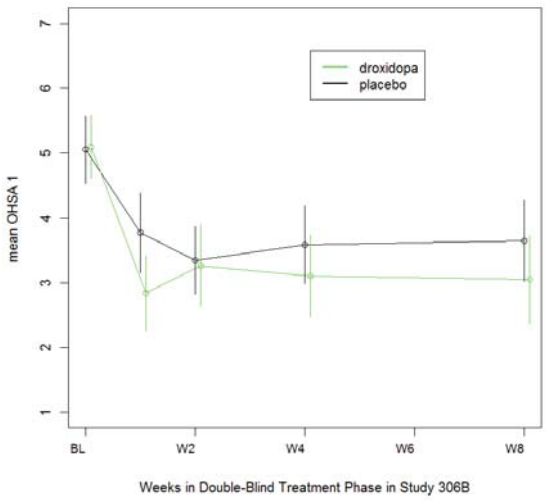


Figure 4 shows the cumulative distribution of the change in the OHSA Item 1 score from Baseline to Week 1, as calculated by Dr. Zhang (left), as well as the corresponding (non-cumulative) histogram. The histogram provides a readily interpretable snapshot of the study results, and seems appropriate for labeling. Both figures show a subset of patients with 2- to 4-unit improvements in OHSA Item #1 (more than in the placebo group).

Figure 5 shows OHSA Item #1 by study visit; the statistically significant effect at Week 1 clearly vanishes in subsequent weeks. *P*-values are 0.77, 0.26, and 0.19 for Weeks 2, 4, and 8, respectively.

Figure 5: OHSA Item #1 by Study Week



Study 306B Weaknesses:

Study integrity

The 1° endpoint of Study 306 was changed twice; and the study was split into Studies 306A and 306B. The total sample size was changed. It was discovered that an unblinded statistical team had access to the treatment codes for all Study 306 subjects, rather than just the 51 subjects in the interim analysis.

The tangible effect(s) of these issues, if any, are difficult to gauge. Sensitivity analyses did not show evidence of malfeasance, although the treatment effect tended to be diminished for subjects enrolled later in the trial. After careful consideration, the review team did not think that study integrity was a crucial issue, and members of the Advisory Committee seemed to agree.

Baseline imbalances; missing data

The review team had concerns about the baseline difference in fludrocortisone use, and particularly the differential dropout rate (20 vs. 7). The baseline difference in fludrocortisone use was difficult to interpret, but as explained in Dr. Targum's review, this imbalance did not appear to influence importantly the primary endpoint or interpretation of the study results.

The missing data, and particularly the substantial imbalance between treatment groups, was a concern to all on the review team. Subjects dropped out predominantly in the droxidopa group because of adverse events and treatment failures, strongly suggesting that missingness was not random.

Dr. Zhang imputed the missing data by carrying forward the baseline observations, and found a treatment effect of -0.45 with 95% CI -1.2, 0.3. The loss of statistical significance was predictable, given that treatment effects of zero were imputed for 20 and 7 subjects in the droxidopa and placebo groups, respectively. One might reasonably have taken a less conservative approach, to impute a treatment effect of zero for the 4 subjects in the droxidopa group who dropped out of the titration phase because of 'treatment failure' and 'lack of efficacy." In any case, no *post hoc* analysis can really compensate for the missing data. The differential dropout rate is a major limitation in the interpretation of Study 306B.

Results Driven Substantially by Largest Site

Site 132 contributed the largest number of subjects of any site, its 18 subjects representing 12% of the study. The mean treatment difference at the site was 2.6. Excluding this site would change the overall treatment effect on OHSA Item #1 from 0.94 to 0.68 units, and the *p*-value would increase from 0.028 to a non-significant value of 0.13. The site was inspected and there were no violations.

Effect size

Drs. Targum, Zhang, and Stockbridge all commented that the treatment effect size was small and/or not clinically relevant. All three took the position that the treatment effect was small compared with the intra-subject variability. The variance was calculated as 2.9 by analyzing OHSA Item 1 results at the post-baseline visits (Weeks 1, 2, 4, and Week 8). Dr. Stockbridge further noted that the distribution (Figure 4, left) shows only a shift in mean response of approximately 1 unit to the left, without the evidence of a hyper-responder subgroup that was found in Study 301 (although most of the hyper-responders were from Site 507).

I note that background variability is a feature of every biological experiment; it represents a portion of the "noise" that the treatment effect ("signal") must overcome if a study is to succeed statistically. In fact there are some who would view the demonstration of a treatment effect on top of highly variable substrate to be indicative of a robust result.

Drs. Targum, Zhang, and Stockbridge were concerned that a single patient would not perceive a treatment effect of 0.9 if their week-to-week variability was 2.9. This is a novel concept. We approve drugs to lower blood pressure, where mean effect sizes can be much smaller than the hour-to-hour or day-to-day variability in an individual patient, and the small treatment effect would be difficult to document in an individual patient in a typical outpatient setting. We have approved drugs for pulmonary hypertension where the effect on 6-minute walk distance is so small that it takes several hundred subjects to demonstrate it, and where test-to-test variability can be considerably large compared to the mean effect size. We approve drugs for depression, where the treatment effect is 2 to 3 points on the 60-point Montgomery-Asberg Depression Rating Scale. Moreover, we recognize that in most cases, not all patients will respond to a treatment. Thus, if the mean response for a population is 1 point, and 50% of patients do not respond at all, then the effect size for the 50% of the population that does respond is, on average, 2. The histogram shown in Figure 4, right, shows that some subjects had larger treatment responses than 0.9 units, the mean treatment effect.

Finally, the intra-subject variability in OSHA Item #1 in this study was calculated over a period of 8 weeks. NOH is a disease that waxes and wanes, and symptoms can be dependent on changing external factors. Thus, one should not expect symptoms to remain constant over an 8-week interval. I would argue, therefore, that the 2.9 figure does not represent true intra-subject variability, but instead reflects the variability of symptoms through a considerable period of time. But even if variability were assessed through only a few days and were found to be 2.9, I would argue that a statistically significant shift of 1 point is probably clinically meaningful. In short, the concept of comparing a mean treatment effect to typical variability within a patient population is probably unprecedented, and I do not support it.

Durability of treatment effect

The lack of treatment effect beyond 1 week is not controversial (see Figure 5). The treatment effect on OHSA Item 1 was not statistically significant after Week 1. Demonstration of durability of treatment effect has been a problem for all of the studies. The applicant has theorized that the pharmacodynamic effect of the drug persists long after it disappears from the blood. Thus, the randomized withdrawal studies that have been performed have not assessed subjects sufficiently long after discontinuing the drug, and droxidopa's carryover effects have obfuscated the studies.

Study 306B Conclusion

Overall, Study 306B, an adequate and well controlled study, was positive, but weakened by the differential dropout rate during the titration phase. It provides evidence of efficacy for the treatment of patients with NOH associated with Parkinson's disease, but as a single study, does not provide strong and robust evidence to support the efficacy of droxidopa. It provides no evidence of efficacy for long-term treatment.

Safety:

The duration of observation within the placebo-controlled segments of the original studies was brief (1 to 2 weeks). There was greater opportunity to assess safety in Study 306, which included up to 10 weeks of placebo-controlled treatment. The inadequacy of the assessment of safety in the original submission was underscored by the fact that there wasn't a single serious adverse event reported in the placebo-controlled portions of any of the studies. In Study 306, 5 subjects in the droxidopa group reported a total of 9 serious adverse events, and 4 subjects in the placebo group reported 5 serious adverse events, but no serious adverse event was reported more than once. Hypertension was reported in 1 droxidopa subject as a serious adverse event.

Ironically, reliance on the uncontrolled experience for estimation of the frequency of adverse events caused by droxidopa undoubtedly exaggerates the drug's risks, because any adverse events that are reported during uncontrolled segments of the studies have to be attributed to the drug. Given the nature of the patient population, this overestimates the risk.

Having noted the limitations of uncontrolled data, 105 of 422 patients (25%) in the long-term studies reported 224 serious adverse events, of which 20% led to study drug discontinuation and 12% resulted in death. The most commonly reported serious adverse events were syncope (14 patients, 3%), pneumonia (9 patients, 2%), dehydration (8 patients, 2%), hip fracture (6 patients, 1%), and fall and urinary tract infection (5 patients each (1%).

Many of these serious adverse events were plausibly related to droxidopa's mechanism of action, but interpretation of the data and attribution of risk is challenging. One cannot determine the extent to which droxidopa was causally related to the adverse events.

The most frequently reported common adverse events were hypertension (7% and 1% in droxidopa and placebo groups, respectively), headache, nausea, and dizziness.

In her review, Dr. Targum underscored the limited long-term exposure at the highest doses, and lack of long-term controlled studies.

Finally, Dr. Targum noted her concerns regarding cardiovascular risk. Given that droxidopa is metabolized to norepinephrine, the potential for cardiovascular risk is obvious, but difficult to characterize. Her strategy is to discourage use in patients with high baseline cardiovascular risk, and I agree with this approach for labeling.

Decision:

Dr. Targum, the CDTL, provided a superb summary of the evidence of effectiveness, and her views are summarized as follows:

The applicant submitted 4 studies in the NDA: 301, 302, 303, and 306B. Two of them, Studies 301 and 306B, were positive. Although Studies 301, 302, and 303 were enriched, Studies 302 and 303, both randomized withdrawal studies, failed to meet their 1° endpoints. Study 306A met criteria for futility.

Studies 306B and 303 both failed to show a durable treatment effect.

Considering the two positive studies, Study 301 had one site with questionable data that contributed disproportionately to the overall treatment effect. The other positive study, Study 306B, was successful at an early time point, but not at subsequent time points. The imbalance in dropouts and missing data raise problems in interpretation. The small treatment effect in the face of substantial intra- and inter-subject variability is also a problem. Collectively, therefore, Dr. Targum would consider Study 306B a positive trial, but not a "strongly positive" trial.

Dr. Zhang's opinion was that Study 306B was positive, but neither strong nor robust, and she pointed out the lack of evidence of durability of treatment effect.

Dr. Stockbridge views Study 306B as a negative trial. With regard to changes in the ongoing study, he noted that one can never be certain that decisions about trial design and analysis were not made with data in hand. More important, however, was the substantial differential in dropouts, and the difference in dropouts was not likely due to chance. HE noted that the *p*-value of 0.028 cannot stand up against any adjustments that might be made for informative censoring, i.e., the trial is negative at a *p*-value of 0.028 because of the imbalance in dropouts.

Dr. Stockbridge's view of Study 301 is also negative, and he does not believe that it would be highly supportive even if Study 306B had been positive. The only analyses that provide support for efficacy with nominal *p*-values < 0.05 are highly exploratory.

He also notes that droxidopa has not been shown to show a treatment effect beyond 2 weeks. Nevertheless, if the short-term treatment effect were persuasive, Dr. Stockbridge would grant full approval, and not accelerated approval. His reasoning is that: 1) FDA is not likely to call for removal of the drug from the market if the drug were not shown to be effective beyond a few weeks; and 2) the drug could be useful for short-term intermittent use. I disagree on point #1. We would call for removal of a drug if the verification required by Subpart H were not fulfilled. The concept of short-term intermittent use is interesting, and we suggested that the applicant study this; however, an intermittent mode of use was never studied and I believe this paradigm would need to be studied before approval for such use.

Dr. Stockbridge made strong arguments about inadequate effect size, principally that patients with NOH experience so much variation in symptoms that they could never appreciate the symptom relief afforded by a 0.9-point (mean) treatment effect, and that there is no hyper-responder patient population. As noted above, I do not accept these arguments about effect size.

<u>These are the arguments IN FAVOR OF APPROVAL of droxidopa</u>:

Study 302 was not successful on its 1° endpoint, OHSA Item #1, but it did succeed on an exploratory composite endpoint, OHQ, the endpoint used in subsequent trials. Thus, the study is supportive with a nominally positive *p*-value on an exploratory endpoint. Note also that Study 302 used a randomized withdrawal design, with randomized withdrawal following an open-label treatment period of as long as 3 weeks, with the endpoint assessed after 2 weeks. Thus, to the

extent the data show a treatment effect, the data show persistence of effect through as long as 5 weeks.

Study 301 was successful on its pre-specified 1° endpoint, OHQ, but not successful when site 507 was removed, as we believe is appropriate. Of note, however, there was still a nominal *p*-value of 0.07 after Site 507 was removed. Moreover, in exploratory subgroup analyses, the study shows a nominally statistically significant treatment effect on OHSA Item #1 within the US (65 patients, treatment difference = 1.1 units; p=0.02), and within Western countries (82 patients in the US, Italy, Germany, Canada, and Austria, mean treatment effect is 1.1 units; p=0.02).

Although concerns are recognized, Study 306B was statistically significantly positive on its pre-specified 1° endpoint, OHSA Item #1. There was an imbalance in missing data (20 vs. 7), and this missingness was informative. The study was substantially driven by the largest site where the effect size was 2.6 (overall effect size was 0.9).

But if one extends no more than "reasonable" flexibility here for an orphan disease, the 3 well-controlled studies collectively support efficacy in the short term. Study 306B is positive; Study 301 trends positively and supports approval. By no means do the data support efficacy in the long term.

<u>These are the arguments NOT TO APPROVE droxidopa</u>:

Study 302 was not successful on its pre-specified 1° endpoint, OHSA Item #1.

Study 301 was not successful once the irregular data from Site 507 were removed.

There is only one positive study, therefore, Study 306B, and 306B doesn't have the strength we typically expect to see for a single-study approval. There was an important imbalance in missing data (20 vs. 7), and missingness was informative. It was substantially driven by the largest site where the effect size was 2.6 (the overall effect size was 0.9). In the views of some, the overall effect size was not clinically important, and there was no demonstration of long-term efficacy.

These represent plausible arguments against approval, and this was the view of the review team.

<u>Conclusion</u>:

There is no doubt that the data are at the "margin" for approvability. Even if one considers 3 favorable studies, they are all imperfect in various ways.

We received a strong message from the Cardiovascular and Renal Drugs Advisory Committee to approve droxidopa. They found the positive studies credible, and noted that NOH is an orphan disease, and one that is difficult to study. The Committee's consensus was that continued benefit should be verified, such that accelerated approval was the appropriate regulatory action. There is one drug already approved for this indication (Midodrine), but it was approved under subpart H, and could be subject to withdrawal from the market. It is feasible to approve a drug under Subpart H, as long as there is no <u>fully</u> approved drug for the indication.

I believe the available data, imperfect as they are, do add up to an adequate demonstration of efficacy in the short term, and support accelerated approval. It will be critical for the applicant to ensure long-term effectiveness.

With respect to the effect size, droxidopa is to be given for its salutary effect on symptoms, and not as a disease modifying or preventive therapy. Symptoms are subjective; patients can decide to continue the drug if they feel better or discontinue the drug if they do not. As illustrated in the histogram (Figure 4, right), some subjects experienced larger treatment responses than 0.9 units, the mean treatment effect. The label will include instructions to reevaluate patients regularly to ensure that only responders receive continued treatment.

Although only patients with Parkinson's disease were included in Study 306B, the other studies evaluated subjects with a range of underlying diseases, and I see no reason to restrict the indication to patients with NOH on the basis of Parkinson's disease. This is in alignment with the advice we received from the Advisory Committee last month.

The NDA will be approved under Subpart H, where the short-term effect is construed as "reasonably likely" to predict the long-term effect that would be clinically meaningful.

The applicant will have a postmarketing requirement to conduct a study in patients with NOH to verify the durability of the treatment effect. An enriched study with a randomized, placebo-controlled withdrawal phase seems most likely to be successful; however, this trial design has been tried before with this drug, and the studies failed. If the applicant's hypothesis about pharmacodynamic carryover effects is correct, and if sufficient time is allowed for drug effects to abate after withdrawal before the 1° endpoint is assessed, the study should be a success.

One could take the position that the existence of prior studies that failed to show durability of treatment effect essentially make the short-term demonstration of efficacy reasonably unlikely to predict a long-term treatment effect here. It is possible that the applicant's theory about carryover effect is incorrect. But "reasonably likely" isn't tantamount to "substantial evidence." If the desired efficacy were verified for every Subpart H approval, we would know that we were setting the bar too high.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

ELLIS F UNGER
02/18/2014