October 7, 2024

Dr. Nikolay Nikolov, M.D.
Office Director, Office of Immunology and Inflammation
Center for Drug Evaluation and Research
Food and Drug Administration
10903 New Hampshire Avenue
Building 22, Suite 4177
Silver Spring, MD 20993

NDA 218489
VLY-686/Tradipitant Capsules for the Treatment of Symptoms in Gastroparesis
Original New Drug Application

Dear Dr. Nikolov:

I write concerning the "Complete Response Letter" dated September 18, 2024, regarding our new drug application for tradipitant in the treatment of patients with gastroparesis.

At the outset, we wish to be clear: the complete response letter is unlawful. The agency has disregarded the clear statutory mandate of 21 U.S.C. § 355(c)(1), which obligates the agency to approve a New Drug Application or provide a Notice of Opportunity for a Hearing. We are sending this letter because we were shocked by the sheer disregard for the facts, evidence, and basic scientific principles contained in the complete response. This letter should not be construed as waiving any of our rights or as seeking anything other than the Agency's reflection upon its approach toward this review, the serious harms to the American public and to the regulated public by the Agency's disregard of facts, evidence, and scientific principles. For the avoidance of all doubt, this is not a request for a formal dispute resolution nor is it a request for a Notice of an Opportunity for a Hearing.

In the complete response letter, you offer conclusions on the inadequacy of the evidence of efficacy. Despite the fact that this document was supposed to encapsulate the conclusion of a *12-month* scientific review, nowhere is there any inclusion of even a data point that would allow any scientist to critically review this letter. This letter is essentially useless for guiding further discussion with the agency and instead is a heavily biased conclusory critique with flawed scientific logic. While it is hard to decipher the origins of the summarily stated conclusions, we have nonetheless tried to piece together the review logic, and we are shocked by the gross distortion of facts and misuse of scientific methods. This cannot possibly reflect a legitimate scientific review: it does not adhere to basic principles of a scholarly and balanced examination of evidence and application of analytic methods. It also cannot possibly reflect a legitimate regulatory review: it does not provide reasoned explanations or engage with the evidence we presented on its merit—including voluminous evidence from experts. The complete response

letter offers not a single piece of data but instead relies on vague references to parts of the evidence and then arrives at equally vague conclusions. The letter is replete with misinformation and distortion of facts and selectively censors scientific data to achieve its ends.

Misrepresenting or cherry-picking facts is unacceptable, especially when it leads to conclusions and actions that are harmful to the American public that the Agency is supposed to protect. We do not know why the review division or you as a signatory would be motivated to conduct a biased review with willful misrepresentation of facts.

We also do not know what level of evidentiary review that Dr. Donohue and you personally performed before finalizing and signing this complete response letter. Unfortunately, as reflected in agency internal communications obtained through the Freedom of Information Act, it is evident that Division personnel, past and present, have participated in systematic retaliatory behavior towards our company. This retaliatory behavior was motivated by Vanda's exercise of its basic First Amendment rights to challenge unlawful agency action and to disagree with the agency on scientific and legal grounds. This response letter confirms the biased and retaliatory behavior by weaponizing pseudoscientific methods and censoring scientific data.

Vanda has repeatedly requested from you and others that the Agency convene an independent advisory committee of experts to evaluate our application, but the Agency has refused. We are confident that no group of experts would agree with your conclusions or the methods you used to arrive at your conclusions. The Federal Food, Drug, and Cosmetic Act instructs that a new drug application demonstrates substantial evidence of efficacy where "qualified experts" "could" "fairly and responsibly" "conclude[]" that the drug will have the effect. The review undertaken here is not what the statute contemplates. It is instead biased, unfair and irresponsible.

Unlike the absence of specificity in your critique of the application, we provide below a detailed account of several of the gross misrepresentations and manifestations of bias reflected in the letter that you endorsed. We have focused on the discussion of the evidence from Trial 2301 for now; we stand ready to provide a similar discussion with respect to other features of your letter if you are interested in having a legitimate discussion.

You wrote the following about Trial 2301 in the complete response letter of September 18, 2024:

> *"Although the phase 2 trial, Trial 2301, achieved a statistically significant difference in the primary endpoint of change from baseline to Day 28 (Week 4) in the weekly average of individual daily nausea severity scores, the results were not statistically persuasive because there were methodological shortcomings with the analysis that could bias results (e.g., exclusion of randomized and treated subjects with post-baseline data for reasons not specified in the statistical analysis plan [SAP], use of a manually derived pooled site variable as a covariate in the analysis that is not in accordance with the SAP algorithm for deriving the pooled site variable). Statistical analyses for the primary endpoint that addressed these*

concerns by including all randomized subjects in the efficacy analysis and removing the manually derived pooled site variable were not robust with respect to missing data assumptions (as demonstrated in the tipping point analysis provided in response to our February 1, 2024, information request"

Trial 2301 did demonstrate a statistically significant difference in the primary endpoint of change from baseline to Day 28 (Week 4) in the weekly average of individual daily nausea severity scores. The word "Although" at the beginning of the sentence, however, preconditions the reader of the complete response letter to think that the positive outcome of Study 2301 may actually not be true. The sentence goes on to say that the outcome is "not statistically persuasive" without defining how to assess what is and is not "statistically persuasive." You then go on and explain that this conclusion is reached because there are "methodological shortcomings with the analysis that could bias results."

You then give examples of two such "shortcomings." First, you refer to the "exclusion of randomized and treated subjects with post-baseline data for reasons not specified in the statistical analysis plan." You do not explain what this is referring to, but prior communications with the review division inform our understanding. You omit that this is referring to a total of only 11 randomized individuals that were not included in the primary endpoint analysis. The impression your unqualified and imprecise statement leaves is that Vanda randomly and for unknown and impermissible reasons excluded an unspecified number of subjects from the analysis. But that is not true, and the Division was already provided with a clear and unrefuted explanation on December 21, 2023. Vanda explained that the statistical analysis plan specified that "a patient must have at least half non-missing scores during each time interval in order to have an average non-missing average score". That is, the statistical analysis plan prespecified that if an individual was missing more than half of their daily diary entry data then the patient will be excluded. This is a standard approach for handling individuals that have a large volume of missing data in order to avoid inaccurate imputations. In the same response we also provided substantial additional detail on these 11 individuals. Of these, 7 were randomized on placebo and 4 were randomized on drug. Ten of the eleven had data entries for fewer than 14 of the 28 study days. One of these 10 patients had only 2 post baseline days, one had 4, three had 5, one had 6, one had 7, one had 8, one had 9 and one had 12. All of these 10 subjects therefore were missing more than half of the postbaseline days as prespecified in the analysis plan. The remaining one of the 11 patients was missing more than half of the baseline days and was therefore excluded as well. Contrary to your statement, which implies that the exclusion of these patients was arbitrary and not prespecified, these 11 patients were properly excluded from the analysis due to excessively missing data. The Division knew this directly as of December 21, 2023. Even further, we performed a sensitivity analysis to determine what would happen to the primary endpoint of the study if these 11 patients were included in the analysis anyway. The results of the analysis showed that inclusion of these 11 patients in the analysis did not change the outcome of the study; instead, it showed a very similar effect with the primary analysis with a drug placebo difference of 0.48 and a p-value = 0.0155. This was shared with the Division on December 21, 2023. In short, the first "shortcoming" of including or excluding these 11 individuals was

irrelevant and did not affect the outcome of the study. Your conclusion that the exclusion of these individuals "could bias results" of the study cannot be true, as it made no difference if they were included or excluded.

Second, you mention the "use of a manually derived pooled site variable as a covariate in the analysis that is not in accordance with the SAP algorithm for deriving the pooled site variable." It is true that the SAP specified an algorithm for pooling sites, a common approach to minimize variability from site to site, which results in pooling smaller sites to (a) larger site(s). As we explained to the Division on February 12, 2024, the prespecified algorithm resulted in an unbalanced pooling with just one large pooled site of 14 sites and an additional 5 pooled sites of 3-6 sites each. A manual pooling produced a more balanced 6 sites of 4-8 sites each. For sensitivity analysis we provided the Division with the results of the analysis using the Program Pooling versus the optimized manual Pooling. The results were very similar with a treatment difference of the manual pooling of about 0.47 and a p-value of 0.0193 versus the original report of 0.5 and p-value of 0.0099. Here again, the pooling method did not have a material effect on the outcome and therefore it could not have biased the results.

Additionally, the Division asked us to impute the data of missing values for all randomized subjects for all post baseline values as well as baseline values. Again here the primary result for all randomized subjects showed a similar effect with or without the inclusion of pooled site in the model with an effect of 0.46 and p-value=0.0195 or 0.41 and p-value=0.0381 respectively. In the analysis where the actual site was used instead and despite the significant reduction of statistical power, still the difference favored the drug with an effect size of 0.35 and a p-value=0.0899. The combination of the two suspected "shortcomings" when included in the model with the addition of the 11 patients and without a poolsite in the model *still* resulted in the similar statistical answer favoring the drug. The Division was provided this information on February 12, 2024.

The Division also asked for a sensitivity analysis of the primary endpoint statistical outcome with the method of the tipping point analysis. We had previously provided a tipping point analysis on the original endpoint on December 21, 2023, that showed that the tipping point analysis confirmed "robustness" when a range of 0-5 was used. A range of 0-5 would be highly implausible given that the effect size tested is approximately 0.5, the baseline about 3.5, and the maximum severity scale value is 5. The Division then asked us to perform additional tipping point analysis for the significant findings when all randomized patients (including the 11), with the inclusion or not of the poolsite parameter. The Division did not specify or justify any range that the tipping point analysis should be tested. We provided tipping point analysis on an arbitrary scale of 0-2.0. This range again for the same reasons as explained above is much larger than a plausible range. Nonetheless the results did demonstrate robustness more so with the pooled site in the model versus the less powered no pooled site analysis. Nonetheless, you concluded that "statistical analyses for the primary endpoint that addressed these concerns by including all randomized subjects in the efficacy analysis and removing the manually derived

pooled site variable were not robust with respect to missing data assumptions (as demonstrated in the tipping point analysis provided in response to our February 1, 2024 information request)."

This account shows that in the CRL there are significantly factual misrepresentations that lead to an incorrect impression that Study 2301 "was not statistically persuasive," when it actually was very persuasive. Your conclusion is weighted almost entirely on the single tipping point analysis. A tipping point analysis is a sensitivity type analysis that examines the stability of the statistically positive effect observed in the primary analysis. In the tipping point analysis missing data are imputed and a penalty is applied so that the drug effect is decreased by a certain magnitude and placebo is increased by that same magnitude. This is performed in a step wise fashion and the tipping point is the magnitude of penalty at which the original result is no longer nominally significant. As such, the tipping point analysis will be heavily influenced by the amount of missing data to be imputed and penalty applied and by the range within which the tipping point analysis will be performed. Starting with the missing data, the post hoc inclusion of 11 patients almost tripled the size of missing data at week 4 from 6 patients in ITT and as such heavily influenced negatively the tipping point analysis. Furthermore, the Division never discussed or justified a range for this tipping point analysis. The range is extremely important because if set at a large value it may be implausible. For example, our original tipping point of 0-5 would suggest that drug and placebo will effectively have a difference of 10, which is twice the value of the actual severity scale and over 10-fold the observed effect size that is being tested. Without a range definition and the clinical understanding of implausibility, the tipping point analysis is uninterpretable. Yet you concluded that this same tipping point analysis showed that the primary results were not robust. That is factually incorrect. If a tipping analysis was performed at an appropriate range of twice the size of the effect size, for example 0-0.8, it could be concluded that the results are robust.

The faulty logic of the review's analysis is evident. According to the Division, you must first include 11 patients that were properly excluded for missing more than half their data, then remove the pooled site from the model, impute missing baseline data form postbaseline data, then perform a tipping point analysis and then, without any prespecified range, conclude that Study 2301 does not provide efficacy of evidence. These are extraordinary errors of judgment, and it is critical that they are immediately corrected. Under this review logic, *any* positive study can be shown to be not robust by improperly including patients with large volumes of missing data, removing the reasonable inclusion of a pooled site assumption, and then ignoring the statistical model's results to instead perform a tipping point analysis where the range is not defined. The result of such a process will render many studies not robust and will lead to incorrect negative conclusions. The review goal is not to minimize false positives without regard to the inflation of false negatives. Of course, the practical impact will be that fewer useful drugs can ever be approved and will remain unavailable to the public to use under the guidance and supervision of a licensed physician. Congress set the approval standard at "substantial evidence" and not some higher standard because it wanted to achieve an appropriate balance between minimizing false positives and minimizing false negatives.

The faulty review conducted here threatens the integrity of the entire drug approval process and must be corrected. Congress carefully calibrated the amount of evidence required for determining drug efficacy; deviations from that evidentiary standard have a major impact on the health of Americans and the economy.

Sincerely,

_____

Mihael H. Polymeropoulos, MD
Vanda Pharmaceuticals Inc.
2200 Pennsylvania Avenue, N.W., Suite 300-E
Washington, DC 20037
(202) 734-3400
mihael.polymeropoulos@vandapharma.com