Dockets Management Branch
Food and Drug Administration
Department of Health and Human Services
5630 Fishers Lane
Room 1061
Rockville, Maryland 20852

5559 '03 MAR 28 A11:03

Re: Supplement to Docket No.: 02P-0029/CP 1
Citizen Petition Requesting FDA Not Recommend the Mylan Estradiol Transdermal System (ETS) as a Generic Substitute to the Climara® Once-A-Week Estradiol (TDS) Transdermal System

Dear Sir or Madam:

Berlex Laboratories, Inc. (Berlex) and 3M Drug Delivery Systems, a division of the 3M Company (3M, formerly Minnesota Mining & Manufacturing), submit this supplement to their January 16, 2002 citizen petition (Docket No. 02P-0029/CP 1). Berlex originally submitted this citizen petition jointly with 3M Pharmaceuticals, but product responsibility within 3M has been recently transferred to the Drug Delivery Systems Division.

This supplement is in response to comments submitted by Mylan Technologies, Inc. (Mylan) related to citizen petition (Docket No. 02P-0029/CP 1) and the November 27, 2002 supplement to this citizen petition. Berlex is the holder of the NDA for Climara®, a 7-day estradiol transdermal delivery system.

In a previous Citizen Petition (Docket No. 98P-0434), Berlex and 3M have expressed theoretical concerns that the Mylan Estradiol Transdermal System (ETS) product is indeed not bioequivalent to Climara. In order to substantiate these theoretical concerns, Berlex conducted a bioequivalence study of the Mylan and Climara transdermal systems in healthy postmenopausal women using the buttock as the application site.[1] That study demonstrated a lack of bioequivalence of the Mylan ETS and Berlex's Climara once-a-week estradiol transdermal system at the buttock application site.

On April 17, 2003, the law firm of Rothwell, Figg, Ernst & Manbeck submitted comments in opposition to our Citizen Petition (Docket No. 02P-0029) on behalf of Mylan. The goal of these comments was to invalidate the results of Berlex's bioequivalence study at the buttock application site. In this supplement to our Citizen Petition (Docket No. 02P-0029), Berlex intends to demonstrate that the allegations of Rothwell, Figg, Ernst & Manbeck on behalf of Mylan (hence referred to as Mylan) are not true and that the bioequivalence study conducted by Berlex will stand on its own merit.

[1] The full study report was submitted to the FDA's Office of Generic Drugs on December 3, 2001 and was included in the initial submission to Docket No. 02P-0029.

02P-0029 SUP 3

Key Mylan statements have been taken from their commentary of April 17, 2003 and have been numbered below, along with our responses. The first set of Mylan statements relates to the introductory part of Mylan's comments, appearing on pages 1 to 3. The page and paragraph numbers of the Mylan document have been included for ease of reference.

(1) *The Mylan commentary states that granting any of the four Berlex requests in their Citizen Petition (Docket No. 02P-0029) "would have the net effect of removing from the market the only generic alternative to Climara" (pg. 2, ¶ 1).*

There is a process for establishing bioequivalence in the USA. The four Berlex requests for FDA action in their Citizen Petition (Docket No. 02P-0029) were made because this process has not been satisfied for the Mylan applications. In the Citizen Petition (Docket No. 02P-0029), Berlex and 3M submitted the results of their bioequivalence study that showed that the Mylan ETS patches were not bioequivalent to Climara patches on the buttock application site. It is in the interests of patients and prescribers to have generic products that comply with FDA regulations and are bioequivalent to the innovator product. This was shown not to be the case with the Mylan ETS product. The FDA must take appropriate action to prevent a product from being marketed as a generic if the product is not bioequivalent to the innovator product.

Additionally, given that FDA Commissioner Dr. McClellan has stated, "there are a lot of things the FDA can do within its mandate to help consumers get access to the effective treatments quickly at the lowest possible cost,"[2] the Agency may wish to consider that removing the A/B rating for the Mylan product will not result in higher costs to the consumer. In fact, a recently published analyses shows that the costs associated with using Climara may be less than that for the Mylan patch, due to the latter's much greater likelihood to lift and detach during normal use.[3]

(2) *Mylan states that "ANDA Nos. 75-181 and 75-233 have been reviewed and approved by the FDA, demonstrating that they are fully compliant with all current statutory and regulatory requirements designed to ensure the safety and efficacy of generic drug products, as well as current FDA practices and recommendations for demonstrating bioequivalence specifically of transdermal products" (pg. 2, ¶ 2).*

The Mylan ANDAs were approved in the year 2000, and were compliant with regulatory requirements and reflected the assumptions for transdermal drug delivery that were in

[2] http://www.fda.gov/fdac/features/2003/203_mcclellan.html
[3] Jones, J.P., Rowe, M.M., and Harrison, L.I. 2001. Replacing branded estradiol transdermal systems with generic alternatives does not result in a cost savings. Abstract presented at the Academy of Managed Care Pharmacy, 14th Annual Meeting, Salt Lake City, UT, April 3-6, 2002.

place at that time.[4] However, even at that time, Berlex and 3M opined that the Agency made a mistake by accepting the clinical data in these ANDAs as evidence of bioequivalence. Berlex and 3M have extensive expertise in transdermal dosage forms and hormone replacement therapy, and it was discussed in Docket No. 98P-0434 that the bioequivalence assessment of a 7-day estradiol patch was a complex scientific issue. The study presented by Mylan raised serious concerns that should have precluded the acceptance of these data as evidence of bioequivalence (as discussed in Docket No. 98P-0434). It was therefore suggested by Berlex and 3M that the Mylan data be presented to an Advisory Committee panel of experts for resolution of these issues. Nevertheless, the Agency accepted the Mylan data without presentation to an Advisory Committee.

The theoretical concerns expressed by Berlex and 3M at the time of the Mylan ANDAs have indeed become concrete. In our Citizen Petition (Docket No. 02P-0029), we presented new clinical data that demonstrate that one of the key regulatory assumptions of Mylan's approval was incorrect – i.e., bioequivalence at one application site was incorrectly assumed to assure bioequivalence at other application sites. Thus, the Mylan ANDA products approved in 2000 as being fully substitutable for the reference product, Climara, may in fact not be as safe and efficacious as Climara when applied to the buttock.

The FDA has the right and the duty to reconsider any Center for Drug Evaluation and Research (CDER) approved drug application as new data become available, especially if the new data relate to safety and efficacy. The criteria and evidence specified for assessing actual or potential bioequivalence problems in the regulations [21 CFR § 320.33(b)] are stated to include: "(b) Evidence from well-controlled bioequivalence studies that such products are not bioequivalent drug products." Thus, Mylan's contention that their ANDA approval is sufficient despite changes in scientific knowledge is not justifiable. It is the responsibility of the Agency to assess bioequivalence problems based on new scientific data from a well-designed and appropriately conducted clinical study that demonstrate that these products are not bioequivalent.

(3) *Mylan states that "the 'new data' that Berlex presents with its petition is derived from a flawed, under-powered study, the design of which encourages a finding that the two products are not bioequivalent. . ." (pg. 2, ¶ 3).*

The Berlex study used a well-established crossover design to measure bioequivalence. This design is recommended in FDA guidance[5] for all bioequivalence studies. Mylan, in

[4] Guidance for Industry. Statistical Procedures for Bioequivalence Studies Using a Standard Two-Treatment Crossover Design. Center for Drug and Evaluation and Research, Food and Drug Administration, July, 1992.

[5] Guidance for Industry. Bioavailability and Bioequivalence Studies for Orally Administered Drug Products – General Considerations. Center for Drug and Evaluation and Research, Food and Drug Administration, draft dated October, 2000, final issued March, 2003.

fact, utilized a crossover design in their ANDA bioequivalence study (also recommended in the guidance documents from that time).[6]

The point of contention between Mylan and Berlex regarding the comment about the *"flawed, under-powered study"* relates to the number of subjects included in the study. The number of subjects is calculated based on the expected intrasubject coefficient of variability (CV), the desired power of the statistical analysis, and the acceptable ratio of the pharmacokinetic parameters for the test and reference products.

Mylan contends that *"The Berlex study was based on an assumption of the ratio between test and reference that was too low, resulting in the use of too few test subjects to provide any meaningful results" (pg. 2, ¶ 4).* However, we note that the Berlex study was appropriately powered to meet recognized criteria. The ideal bioequivalence ratio is 1, which gives the greatest assurance of equivalent performance of the test and reference products. A deviation of ± 5% from 1 is considered an acceptable range by the pharmaceutical industry for bioequivalence testing. It is also accepted by both the scientific and regulatory community that if a product is in the range of 95% to 105% of the reference product, and meets the bioequivalence requirements, that it will give the same therapeutic performance as the reference. The Berlex study was appropriately powered to meet these criteria.

It is probable that a product that is 90% to 110% of the reference product, which meets the bioequivalence requirements, will be therapeutically equivalent to the reference product. Although the Berlex study was not intended to be powered for this eventuality, inclusion of 39 subjects rather than the required 34 subjects per the protocol meant that the study was actually powered for bioequivalence testing of a product (at 80% power and a CV of 25%) that could be 92.6 % to 108% of the reference product. Thus, the Berlex study with 39 subjects was actually sufficiently powered to test a larger difference of approximately ± 8%, which implies that even with this larger ratio, the Mylan product would not have been shown to be bioequivalent.

(4) *Mylan states that "the procedures used for base-line correction . . . were not in accordance with specific recommendations provided to Berlex by the FDA for the conduct of such studies with estradiol patches" (pg. 2, ¶ 3).*

Berlex used a single blood sample collected immediately prior to study initiation as the estradiol baseline measurement, rather than two blood samples collected prior to study initiation as recommended by the FDA on page 10 of their March 17, 2000 response (Docket No. 98P-0434). While we respectfully consider FDA opinion regarding their suggested method of baseline calculation, we recognize and consider other methods that could be more appropriate in special circumstances. In this case, consider the magnitude of the values: the average estradiol baseline concentration for the reference product was

[6] Guidance for Industry. Statistical Procedures for Bioequivalence Studies Using a Standard Two-Treatment Crossover Design. Center for Drug and Evaluation and Research, Food and Drug Administration, July, 1992.

3.4 pg/mL and the peak or maximum blood or plasma concentration (Cmax) was 259.7 pg/mL. Thus, the baseline was only 1.3% of the Cmax value, which is insignificant when one considers the CV of the Cmax parameter of 38.3%.

Although in the above statement, Mylan raises the FDA's suggestion of calculating the baseline from multiple concentrations in their response, Mylan omitted the fact that the FDA has additional comments in that same response letter that support the Berlex decision: "*Assuming his [Dr. Cabana's] assertions regarding baseline variability of as much as 154 percent for estradiol are correct, this represents a very small magnitude of variability when plasma levels resulting from transdermal estradiol are analyzed . . . The magnitude of baseline variation of ± 0.6 pg/mL when added to concentrations of 165 pg/mL yields an increase in variability of only ± 0.4 percent, a negligible increase*" *(pg. 7, ¶ 1 of the March 17, 2000 FDA response).* Our observation is in concert with the FDA's opinion, showing a negligible contribution of the baseline to the overall post-application concentrations. Thus, almost identical test/reference ratios in the Berlex study were obtained for the Cmax parameter, with and without baseline subtraction, and almost identical confidence intervals, indicating that there was no bias in the baseline subtraction method.

(5) *Mylan also states that "The procedures used . . . for re-applying patches that fall off were not in accordance with specific recommendations provided to Berlex by the FDA for the conduct of such studies with estradiol patches" (pg. 2, ¶ 3).*

The discussion regarding patch re-application was provided on pages 16 and 17 of the March 17, 2000 FDA response. The Agency suggested that if a patch falls off during a bioequivalence study, that subject should be withdrawn from the study. The FDA explained their rationale on page 17 of the document as follows: *"Taking a subject out of the study increases the chance of failure of both the bioequivalence and adhesion studies because of the lower statistical power in the bioequivalence evaluation and a negative result in the adhesion analysis."* It is this bias that we wished to avoid in our bioequivalence assessment. As was explained in the study report, after careful consideration, the data from all subjects who had a patch fall off during a treatment period were included in the pharmacokinetic analyses since they were not different from the remaining subjects. This procedure gave the study its highest power to determine bioequivalence.

It should be noted that on Page 16, the Agency states that *"equivalent skin adhesion is a criterion for approval of a generic transdermal product, and FDA will not approve a generic transdermal product that fails this comparison."* The Mylan product clearly fails this comparison at the buttock site. The Mylan ETS patch had a high rate and early occurrence of patch lift off at the buttock site, as demonstrated by data collected in the recent Berlex bioequivalence study. Specifically, patch lift or fall-off occurred in 59% of Mylan applications compared to 18% of Climara applications, and the median lift time for the Mylan product was 35.5 hours (less than 2 days of the 7-day intended patch dosing) after application (23 occurrences) while the median lift time for Climara was

119 hours (6 occurrences).[7] Although this study was not solely designed to measure different degrees of patch adhesion, the measure of the time of first occurrence of patch lift/fall-off was a valid, appropriately documented metric.

Furthermore, examining the Mylan ANDA bioequivalence study from the study report submitted to the New Jersey Formulary shows the same tendency for inequivalent adhesion at the abdomen site. Lifting was observed beginning at 48 hours post-application with the Mylan patch, whereas no lifting was observed throughout the entire 7-day dosing period with the Climara patch.

(6) *Mylan states that "While the under-powered Berlex study fails to demonstrate that the Mylan ETS meets the bioequivalence standards of the FFDCA when compared to Climara, it does not demonstrate that the two products are inequivalent" (pg. 2, ¶ 3).*

The Berlex study was appropriately designed and powered to test the hypothesis of bioequivalence. The study found statistical differences between the products and demonstrated that the Mylan ETS patches were not bioequivalent to Climara patches on the buttock application site.

(7) *Mylan states that these issues have previously been raised by Berlex in the first Citizen Petition and were denied by the Agency. They conclude that "the present petition adds nothing to support the arguments that the FDA has already rejected" (pg. 3, ¶ 1).*

The FDA's response to our Citizen Petition (Docket No. 98P-0434) was based on a theoretical regulatory assumption at that time - that bioequivalence at one application site assures bioequivalence at other application site. Citizen Petition (Docket No. 02P-0029) added new and relevant scientific data that showed that bioequivalence at one site did not necessarily result in bioequivalence at a second site. Our findings have been made public and were properly submitted in Docket No.02P-0029. The FDA has the right and the duty to consider any and all new sets of data that are relevant as per 21 CFR § 320.33(b).

(8) *Mylan summarizes that there is "nothing new in the present petition, and the data presented as allegedly demonstrating non-bioequivalence is from a study containing fundamental flaws in its design that render its results essentially meaningless" (pg. 3, ¶ 2).*

Berlex and 3M have demonstrated that the Mylan allegations about the bioequivalence study at the buttock application site are not accurate, and that the study results indeed support the lack of bioequivalence.

---

[7] Jones, J.P., Rowe, M.M., and Harrison, L.I. 2001. Replacing branded estradiol transdermal systems with generic alternatives does not result in a cost savings. Abstract presented at the Academy of Managed Care Pharmacy, 14th Annual Meeting, Salt Lake City, UT, April 3-6, 2002.

(9) "This second attempt by Berlex to eliminate generic competition for its Climara product should be rejected, and the Berlex Petition, including the request in the November Supplement, denied" (pg 3, ¶ 2).

It is in the interests of patients and prescribers to have generic products that comply with FDA regulations and are bioequivalent to the innovator product. The data in Citizen Petition (Docket No. 02P-0029) show that the Mylan ETS product is not bioequivalent to the Climara reference product at the buttock application site. The FDA must take appropriate action to assure that product substitutions do not happen between products that are not bioequivalent.

The next Mylan statement relates to their item I on page 3.

I. THE MYLAN ANDAs DO NOT MEET ALL STATUTORY AND REGULATORY REQUIREMENTS

(10) Mylan states that FDA followed a "rigorous, well-established procedure that applies proven, scientifically sound principles to the determination as to whether or not a proposed generic product can be considered an appropriate substitute for the branded reference product" (pg. 3, ¶ 3).

It should be pointed out that the Mylan ANDAs were the first and are still the only ANDAs for a 7-day transdermal estradiol product. All other approved ETS products utilized the NDA process for approval. Therefore, the process used by FDA for the review and approval of the Mylan ANDA was not a "well-established procedure" for assessing transdermal bioequivalence.

Mylan also cites the regulatory response of March 17, 2000 to our Citizen Petition (Docket No. 98P-0434) as evidence that our concerns about the ANDA approval process for a transdermal estradiol product have been suitably considered and answered by the FDA. In Mylan's comment, it is not recognized that the FDA opinion regarding the application site issue reflected the scientific understanding from the year 2000, prior to the presentation of data to the Agency that demonstrated that bioequivalence at one application site does not necessarily mean bioequivalence at all application sites.

The next set of Mylan statements relates to their item II on pages 4-8.

II. THE NEW STUDY SHOWS NON-BIOEQUIVALENCE

(11) Mylan states that "in fact, even a cursory look at the data presented in the published study shows that the mean serum concentration profiles of the two products are virtually identical" (pg. 4, ¶ 1).

It is not sufficient for two products to have similar mean serum concentration profiles to be determined to be bioequivalent. The FDA purposefully stipulates (in guidance documents that were available at the time of the Mylan study) that generic applicants must analyze the area under the curve (AUC) and Cmax data statistically to determine if the products are bioequivalent.[8] Mylan also appears to support this opinion when it contradicts itself and states that: *". . .similarity between the plasma concentration profiles, per se, is not, and never has been, a requirement for bioequivalence"(pg. 8, ¶ 3).*

(12) Mylan states that the buttock bioequivalence study was "flawed in such a way to make it unsuited for FDA submission, even if bioequivalence had been concluded" (pg. 4, bottom, pg. 5, top).

This statement is unsubstantiated and we disagree with the premise that the study is flawed. Berlex considered and incorporated all relevant FDA guidance and opinion when the Berlex bioequivalence study was planned. For this reason, partial AUC analyses were included in the data analysis as recommended in the FDA bioequivalence guidance published in October, 2000. This guidance also recommended that any subject whose baseline was greater than 5% of Cmax be removed from all pharmacokinetic analyses.[9] One subject in one period was in this category, and we removed this subject from all pharmacokinetic analyses. The statistical plan was also modified to incorporate FDA guidelines that recommend a 90% confidence interval analysis for bioequivalence assessments.[10] Thus, as an additional analysis for the Cmax and AUC variables for estradiol, a 90% confidence interval was constructed for each of these assessments without multiplicity adjustment (Appendix 1 provides detailed discussion of statistical issues).

Furthermore, the FDA's Division of Scientific Investigation (DSI) has inspected both the study site and the bioanalytical laboratory and no untoward observations were found. (Appendix 2 memo from Michael Skelly [DSI] to Dale Conner)

(13) Mylan states that "the data presented in the Berlex study has not proven bioinequivalence, and is at best inconclusive" (pg. 5, top).

The Berlex study was appropriately designed and powered, in accordance with FDA guidances,[9,10] and was able to detect differences between the test products. The study

---

[8] Guidance for Industry. Statistical Procedures for Bioequivalence Studies Using a Standard Two-Treatment Crossover Design. Center for Drug and Evaluation and Research, Food and Drug Administration, July, 1992.

[9] Guidance for Industry. Bioavailability and Bioequivalence Studies for Orally Administered Drug Products – General Considerations. Center for Drug and Evaluation and Research, Food and Drug Administration, October, 2000, see page 24, Appendix 2.

[10] Guidance for Industry. In Vivo Bioequivalence Studies Based on Population and Individual Bioequivalence Approaches. Center for Drug and Evaluation and Research, Food and Drug Administration, October, 1997; see page 11.

demonstrated that the Mylan ETS and Climara are not bioequivalent at the buttock application site.

(14) Mylan states that the Berlex study "deviates significantly from the Mylan study showing bioequivalence between the ETS and Climara at the abdomen application site" (pg. 5, ¶ 1).

It is true that the Berlex study design is different than the Mylan bioequivalence study in the calculation of the confidence intervals; this was purposeful.

The need for controlling an overall level of significance in making multiple inferences based on the same data is well recognized. When two or more test formulations are compared with a reference formulation, the problem of multiple inferences arises as this comparison involves two or more confidence intervals. The overall coverage probability of such confidence intervals would be lower than the nominal coverage probability of 90% if the size of each confidence interval were kept at 90%. In other words, the probability of falsely declaring at least one test formulation as being bioequivalent to the reference formulation would be inflated. In order to alleviate this problem, multiple comparison methods are used. The Bonferroni procedure is one of them. Although conservative, the Bonferroni procedure is commonly used when the problem does not involve more than two inferences.

When the Berlex statisticians designed the protocol for their bioequivalence study, involving two test formulations and a reference formulation, the principle of multiple inferences (simultaneous confidence intervals in this case) was applied. The sample size was calculated using the Bonferroni procedure, leading to a 95% confidence interval for inferring bioequivalence of each of the pairs: (1) modified and Climara patches, and (2) generic and Climara patches. It was appropriate to apply the same principle while analyzing the data from this study and was so stated in the study protocol. Therefore, a 95% confidence interval, as well as the traditional 90% confidence interval, was computed for each of the product comparisons.

(15) Mylan states that the Berlex study "deviates from both industry and FDA-accepted guidelines for conducting and powering such a study" (pg. 5, ¶ 1).

There is in fact agreement by industry and FDA on suitable methods to calculate sample size, such as the work of Hauscke et al. (1992)[11] cited in the Mylan commentary, but there are no recommendations by the FDA on the appropriate ratio to use. Both the industry and the FDA favor a ratio that is close to 1. The ratio of ± 5% is the one most often suggested. We are not aware of *"numerous examples (pg. 5, ¶ 2)"* where sample size was calculated for the bioequivalence testing of ratios much greater than ± 5% for

[11] Hauschke, D., Steinijans, V.W., Diletti, E., and Burke, M. 1992. Sample size determination for bioequivalence assessment using a multiplicative model. J. Pharmacokinetics Biopharmaceutics 20:557-561.

any transdermal product, and certainly not ± 15%. If one were to use the Mylan study results as suggested in their April 17, 2003 commentary to help calculate the appropriate sample size for the Berlex study, one would have selected about 32 subjects. This in fact is in excellent agreement with the sample size calculation for the Berlex study.

(16) Mylan also suggests that "it is quite possible, for example, for two bioequivalent products to have a point-estimate mean ratio ranging from approximately 0.85 to 1.18 (or more), such that the associated 90% confidence intervals would fall within 0.8 to 1.25" (pg. 6, ¶ 1).

We do not accept this contention, that simply passing the confidence interval criteria, confers therapeutic equivalence on the product. True, it is theoretically possible, as Mylan contends, to have confidence limits within the (0.8, 1.25) interval even if the test/reference product ratio is as large as 1.15, 1.20, or even 1.24, if one includes enough subjects to reduce the variability accordingly. For example, one could theoretically have an AUC ratio of 1.20 with a confidence interval of (1.150, 1.250) if one included 450 subjects in the bioequivalence study. From a public health point-of-view, this is not reasonable and is counter to the intention of bioequivalence. Bioequivalence testing is used as a primary measure of therapeutic equivalence and as such, bioequivalence testing must be sensitive to avoid clinically meaningful differences between the test and reference products. It is unlikely that hormone-replacement therapy products that differ by more than 15% on average could be considered therapeutically equivalent. The FDA has an obligation to assure public health and prevent clinically meaningful differences between the test and reference products (the intended purpose of bioequivalence), as cited in 21 CFR § 320.33(b).

(17) Mylan states that "Based on the assumed values for the mean ratios (i.e. 1.05) and intra-subject variability (25%), Berlex selected sample size of 42 test subjects (n=42) of which only 40 completed the study and of which only 39 were used in the final data analysis" (pg. 7, ¶ 1).

This is an incorrect and misleading statement. It is clearly stated in section 9.8.5 of the Berlex study report entitled Determination of Sample Size that a sample size of 34 subjects was calculated as appropriate to test a ratio of 1.05. Berlex initiated the study with a greater number of subjects, 42 subjects to be exact, for two reasons. First, to have a fair, unbiased evaluation, Berlex desired to overpower the study. Second, to have the required number of subjects complete the study, sample size allowed for a loss of as many as 20% of the subjects. Actually, Mylan's comments on pg. 7, *pg. ¶ 1*, rather than being a criticism of the Berlex study, serve to reinforce the appropriateness of the Berlex study design: *"Based on the industry standard method of Hauscke et al., it would normally be reasonable to power a pivotal study to a degree somewhat larger than that used in the Berlex study. Also, it is common for subjects to drop out of a study such as that executed by Berlex (and in fact, two did), thus one would normally recruit even higher numbers of subjects to account for that."*

(18) "Had the study completed 48 subjects rather than 39, it is probable that the 0.8 to 1.25 bioequivalence criteria would have been met"(pg. 7, ¶ 1).

A valid and adequately powered study, conducted according to well-designed protocol, produces data that is reliable and scientifically sound. The Berlex study met this scientific criteria. The Berlex study was powered to detect a difference of 5% between the test and reference products, which was a reasonable assumption when planning the study. Based on that assumption, the number of subjects included was sufficient to provide adequate power for that study. It is inappropriate to raise hypothetical propositions, such as possible results with greater numbers of subjects, after the study results are known. For a valid, well conducted, and powerful study, any hypothetical propositions raised retrospectively are unscientific.

(19) "B. The Berlex Study Was Not Properly Base-Line Corrected . . . the use of a 'single base-line' value instead of an average base-line value introduces an unacceptable degree of inaccuracy into the results" (pg. 8, top).

Berlex used a scientifically acceptable method to calculate the estradiol baseline concentration, as already discussed under statement *(4)* above. Furthermore, Berlex's method of handling the baseline data was appropriate and consistent with good statistical procedures. When an individual plasma concentration was below the limit of quantification (BLQ), it was assigned a value of 0. Because of this, the CV for baseline measurements appeared relatively high. The BLQ was 5 pg/mL. So when a measurement was BLQ, the true, but unknown, value would have been between 0 and 5 pg/mL. However, this missing value needed to be imputed. The choice of 0 (the lowest possible value of plasma concentration) as an imputed value causes the largest spread among the baseline measurements. Hence the CV for baseline values was relatively large. However, we should not over-interpret the CV, because the width of a confidence interval depends on the variance and not on the CV. For baseline corrected Cmax, the variance was: var(Cmax – baseline) = var(Cmax) +var(baseline) – 2cov(Cmax, baseline), where var stands for the variance of a variable and cov for the covariance between the Cmax and baseline. The variance of baseline was about 15 pg/mL^2, whereas that of the Cmax is about 12,000 pg/mL^2. Thus, var(Cmax – baseline) depends mostly on var(Cmax). Furthermore, the baseline means are almost equal for all three formulations. Hence the relative contribution of baseline to a confidence interval for the baseline-corrected Cmax is very small.

This is the reason why with- and without-baseline corrected methods give almost identical confidence intervals. Almost identical means and standard deviations, along with high pairwise correlations for 3 baselines obtained in the Berlex study, imply that the within-subject variability is very small. In this situation, multiple baselines for each formulation will not have any appreciable advantage over a single baseline prior to each period.

The next Mylan statement relates to their item III on pages 8-9.

III. BERLEX'S SUGGESTION THAT THE SAMPLING INTERVAL USED IN THE ORIGINAL MYLAN STUDY MASKED NON-BIOEQUIVALENCE AT THE ABDOMEN SITE IS SOUND

(20) *Mylan states that "contrary to Berlex's assertion, the times at which peak mean serum estradiol concentrations are reached with both Mylan's ETS and Climara is at approximately 18 hours when applied to the buttocks . . . This finding is consistent with Mylan's own bioequivalence study, which showed that approximately 94% of the observations on subjects administered Climara to the abdomen showed peak plasma estradiol concentrations at ≤ 24 hours and 73% showed peak levels at ≤ 18 hours. In addition, in discussing differences in rate of absorption, Harrison and Harari comment that absorption is expected to be faster (and hence Tmax occurs earlier) when Climara is applied to the buttocks compared to the abdomen" (pg. 8, ¶ 2).*

The comment of Harrison and Harari above clearly states that the Tmax for buttock application is faster than for abdomen application; thus, the abdomen Tmax should be later than 18 hours. This point was missed by Mylan. Mylan's observation that *73% showed peak levels at ≤ 18 hours* for the abdomen application should have alerted Mylan to a flaw in the blood-sampling schedule of their bioequivalence study. The expected Tmax for the Climara reference product is actually around 36 hours when applied to the abdomen. This Tmax information was available to Mylan before they designed their study protocol, both from the open scientific literature and in the Climara package insert. Mylan should have consulted the pharmacokinetic information on Climara that was available and utilized a more appropriate blood sampling schedule.

The next set of Mylan statements relates to their item IV on page 9.

IV. BERLEX'S ASSERTION THAT ADHESION IS WORSE WITH MYLAN'S ETS IS UNSUPPORTED

(21) *Mylan observes that Berlex reached conclusions about adhesion even though the study was not designed to measure adhesion (pg. 9, ¶ 1).*

While it is true that the Berlex study was not designed to assess the comparability of lift between Mylan's ETS and Climara on a numeric scale, the study was adequately designed to asses the time to the first event of adhesive failure (patch lift or patch fall, whichever occurred first). This measure of adhesion was planned per protocol, was subjected to the appropriate statistical tests, and the results are valid and interpretable.

(22) *Mylan contends that the Berlex conclusions about adhesions were determined from anecdotal information (pg. 9, ¶ 2).*

In the Berlex study, patch adhesion was in fact monitored by a trained member of the investigator's staff per the protocol at every blood sampling time (at least once daily). If

any lifting was observed, tape was to be applied. The adhesion data collected in this manner reached statistical significance. This was a planned study endpoint, and was appropriately measured and analyzed; therefore, it was a valid observation. Clinically important differences between the products were observed for the number of patches lifting, the number of patches falling-off, and times for the occurrence of patch lift and/or fall-off, all of which could affect therapy of a patient.

Clinically relevant observations are always useful, even if they were not the main objective of the study. In fact, it is incumbent upon the investigator to point out such findings. Thus, our analysis of the adhesion data was appropriate.

(23) *Mylan contends that "There were no criteria whatsoever provided to the subjects for assessing what constituted a loss of adhesion, or for assessing the degree of loss of adhesion . . . Furthermore, subjects were even instructed to re-apply a patch that had fallen off, in contradiction to customary and accepted procedures for either an adhesion or a bioequivalence study" (pg. 9, ¶ 3).*

As discussed above under Mylan comment 22, there were criteria in the protocol that specified when patch taping would occur. Furthermore, the protocol specified that a member of the investigator's staff would assess patch adhesion at every blood sampling time (at least once daily) and document any observations.

One can debate the inclusion of subjects whose patches were reapplied. Berlex chose not to remove these subjects to give the study more power to conclude bioequivalence, a point that actually favors the Mylan product.

The last Mylan statement relates to their CONCLUSION on pages 10-11.

CONCLUSION

(24) *Mylan believes that since the argument in the Berlex Citizen Petition have already been rejected by the Agency, and since the buttock bioequivalence study was flawed, that these data are inconclusive and the Berlex Petition should be denied (pg. 9, bottom, pg. 10, top).*

Berlex and 3M strongly disagree with this conclusion. Berlex has conducted a well-designed bioequivalence study to provide concrete data that establishes that the Mylan ETS and Climara are <u>not</u> bioequivalent when applied to the buttock. Concerns about the study design have been demonstrated to be unwarranted. FDA audits of the clinical site and the bioanalytical facility have confirmed that the study was conducted appropriately and per the protocol.

The Agency cannot overlook scientific data from an appropriately conducted study (as per 21 CFR § 320.33(b)), and must rule that the two products are not bioequivalent when dosed according to the labeled directions, which includes application of the transdermal system either on the abdomen or buttock.

Respectively Submitted,

Sharon W. Brown
Director, Drug Regulatory Affairs
Berlex Laboratories, Inc.
P. O. Box 1000
Montville, NJ 07045
(973) 487-2162

Marie Kuker, RPh., RAC
Manager, Regulatory Affairs
US and International
3M Pharmaceuticals SBC
Bldg. 270-3A-08, 3M Center
St. Paul, MN 55144-1000

Copies:
Mr. G. Buehler, Director, Office of Generics
Dr. D. Conner, Supervisor, Division of Bioequivalence
Dr. D. Greibel, Deputy Director, Division of Reproductive and Urologic Drug Products
Ms. M. Kober, Chief, Regulatory Project Management, Division of Reproductive and Urologic Drug Products
Ms. C. Rogers, Regulatory Counsel, Office of Regulatory Policy
Dr. D. Shames, Director, Division of Reproductive and Urologic Drug Products
Dr. S. Wood, Director, Division of Policy & Program Development
Dr. J. Woodcock, Director, Center of Drug Evaluation & Research