

October 28, 2003

RE: Final Statistical Review for PMA P030032, Hylaform® (Hylan B Gel) for Correction of Soft Tissue Contour Deficiencies, Genzyme Corporation (August 1, 2003)

Introduction and Summary of Clinical Studies

Hylaform® (Hylan B Gel) is a sterile, nonpyrogenic, viscoelastic, clear, colorless gel implant composed of cross-linked molecules of hyaluronan. Hyaluronan is a naturally occurring polysaccharide of the intercellular matrix in human tissues, including skin. Hylaform is injected into the dermal tissue for correction of skin contour deficiencies caused by wrinkles and scars.

The sponsor conducted a prospective, multi-center, randomized, double-blind safety and efficacy trial. Subjects were randomized (1:1) to either Hylaform or a commercially available collagen implant (Zyplast). A touch-up was allowed 2 weeks after initial treatment if necessary. The primary efficacy objective was to establish either the superiority or non-inferiority of Hylaform to Zyplast for the correction of nasolabial folds, using serial photographic documentation and blinded Independent Panel Review (IPR) photographic evaluation. Efficacy was based on the blinded IPR Wrinkle Assessment Scores of the Week 12 or 14 photographs (12 weeks following last device implantation), although there were earlier follow-ups at 3 days, 2, 4, and 8 weeks post-implantation. Safety was determined by the incidence of adverse events up to 12 week following last implantation. Secondary objectives included a physician live assessment and patient self-assessment. One hundred thirty-three subjects were injected with Hylaform and 128 with the control at 10 U.S. centers. The sponsor demonstrated non-inferiority by showing that the lower bound of the one-sided 97.5% confidence interval for the difference in mean IPR median scores (Zyplast – Hylaform) was -0.38, larger than the prespecified maximum tolerable difference of -0.5 (delta). The superiority hypothesis was not met. Adverse events were comparable between the two treatment groups.

In order to evaluate the effects of repeat treatment with Hylaform in addition to the safety and efficacy of Hylaform Plus, the protocol was amended to include a “repeat treatment phase” using a split-face design of Hylaform and Hylaform Plus. Each of the 133 Hylaform treated patients was offered a repeat treatment and 96 accepted. Safety data through 4 weeks is presented in the PMA, although claims for efficacy of repeat treatments or of Hylaform Plus have been deferred. Therefore, this review will focus on the initial phase only.

Reviewer’s Comments

The sponsor's clinical study was well designed and the PMA clearly written and comprehensive. The study was designed as both a non-inferiority and superiority study, with separate criteria for each one. The 6-point Genzyme grading scale was specifically created and validated for this study, utilizing photographs that could easily maintain the blind and be put in random order. Patient masking was also successful. In the Hylaform group, 36 (27.1%) believe they received Hylaform; 18 (13.5%) believed they received Zyplast; and 76 (57.1%) did not know. In the Zyplast group, 31 (24.2%) believed they received Zyplast; 25 (19.2%) believe they received Hylaform; and 69 (53.9%) did not know. The two treatment groups were well balanced on all demographic features, which included age, weight, height, gender, ethnicity, smoking history, and sun exposure. Therefore, I am satisfied that the randomization and the masking were successful.

One area of concern, however, is the fact that 94% of the subjects were women and 80% of the subjects were Caucasian. Although this distribution is likely representative of the target population of people who would seek this treatment in real life, the 26 non-Caucasian subjects may not have been sufficient to assess how Hylaform performs in all skin types, which can differ with respect to scarring, for example. Therefore, this issue should be evaluated from a clinical perspective.

Patient accountability was excellent. Of the 261 patients randomized and treated, 255 completed the 12 weeks of the initial study phase (130/133 Hylaform and 125/128 Zyplast). Three patients in each treatment group withdrew from the study. The primary efficacy and all safety analyses were done on the ITT population. However, 10 Hylaform and 11 Zyplast patients who did not have 12-week IPR scores, (including the 6 who discontinued) were excluded from the primary efficacy ITT analysis. Missing values were not estimated and there was no data imputation. The per-protocol subgroup consisted of 115 Hylaform and 109 Zyplast subjects.

The study was sized to test both non-inferiority and superiority. The sponsor confirmed their estimate the standard deviation (1.3 pts) from an external data collection on 32 Zyplast treated patients over 12 weeks. This estimate was used to calculate sample size. It was calculated that 125 patients per group would afford more than 80% power for a non-inferiority test with a delta of 0.5 points on a 6-point scale, and 95% power for a superiority test of 1 full point, allowing 15% for potential dropouts. Thus, the study was sized adequately for effectiveness. For safety, however, in order to detect statistical differences between Hylaform and Zyplast, the adverse event rates would have to be further apart than what might be considered the minimally clinically significant increase. This issue will be further discussed below with the safety assessment.

The primary efficacy variable, which was the mean of the median IPR scores (medium of 3 evaluators) at 12 weeks, was 2.3 for Hylaform patients and 2.2 for Zyplast. The lower bound of the one-sided 97.5% confidence interval for the

difference was -0.38, indicating that Hylaform was less than 1/2 point inferior, thus satisfying the non-inferiority criteria. The analysis was done on the Intent-to-Treat population and on a per-wrinkle basis. To construct the confidence intervals, the sponsor used a repeated measures analysis of covariance model, with the covariates being treatment group, site, patient (within treatment group), and baseline IPR score. The patient's scores for the right and left nasolabial folds served as the repeated measure. This analysis was appropriate, because it adjusted for any differences in baseline values, and accounted for the within-patient correlation between the two sides of the face. Adjustment for baseline differences was particularly important because patients were entered into the study based on a "live" wrinkle score of 3 or 4, but the analysis was based on more variable scores derived from photographs. Because it was believed that 1 point on the Wrinkle Severity Scale was distinguishable, the sponsor was held to a rather strict criteria of 1/2 point for non-inferiority. As mentioned earlier, the superiority claim, held to a higher threshold, was not met.

The sponsor performed subgroup analyses for mean IPR median scores at 12 weeks for subgroups defined by touchup-status (yes/no), smoking history (3 categories), and sun exposure (3 categories). Although touch-up patients generally had higher median IPR scores than those not requiring touch-up, the scores were similar between Hylaform and Zyplast within a stratum. No significant differences in median IPR scores between treatment groups were noted based on smoking history or sun exposure, using a confidence interval analysis. Thus, it is acceptable to pool these subgroups for the analysis. As for pooling across sites, the sponsor provided a site by site analysis of the primary efficacy endpoint, but there was insufficient power (i.e., sample size) to meet the non-inferiority criteria for each site individually. What is important here is that there was no treatment by site interaction. The direction of the difference (Zyplast – Hylaform) was always negative, indicating Zyplast did a little better at each center.

The sponsor also performed a categorical analysis comparing the proportion of patients who experienced at least a 1 point improvement in both nasolabial folds at 12 weeks. This proportion was higher for Zyplast (9.5%) than Hylaform (4.1%), although not statistically significant. Other secondary endpoints of investigator live (severity scale) and investigator and patient global assessment showed were extremely close (~1/10 point).

The analysis of the primary endpoint (median IPR scores at 12 weeks) was also performed on the per-protocol population. Eighteen Hylaform and 19 Zyplast patients were excluded from the per protocol analysis due to major protocol deviations, leaving 115 Hylaform and 109 Zyplast subjects. The mean of the median IPR scores were identical to the ITT population, and the lower bound of the confidence interval was -0.36, a hair tighter than the ITT analysis. Thus, the non-inferiority criteria were met for the per-protocol population.

It is interesting to note that the trend of Zyplast doing just a little better holds up if one examines the results by individual visit (3 days, 2, 4, 8, 12 weeks). This is true for all 3 analyses: the blinded IPR scores, the investigator live assessment, and the patient assessment. I couldn't find one comparison (other than at baseline) where Hylaform received higher mean scores. However the differences were often just 1/10 of a point.

Although the sponsor does not make a big point of it, I think it is important to point out that 56.1% of folds returned to baseline by 4 weeks, 68.9% by 8 weeks, and 73.3% by 12 weeks. Zyplast wrinkles did somewhat better as shown in the table below. However, neither product provides any long-lasting correction.

Number (%) of Folds Returned to Baseline		
	Hylaform*	Zyplast*
2 Weeks	95 (38.2%)	54 (21.9%)
4 Weeks	138 (56.1%)	63 (26.3%)
8 Weeks	175 (68.9%)	112 (46.7%)
12 Weeks	178 (73.3%)	151 (65.1%)
*denominators vary slightly depending on visit		

As mentioned earlier, patients were entered into the study based on a "live" wrinkle score of 3 or 4, but the analysis was based on the IPR blinded assessments, which showed more variability. The following table shows the percentage of folds that were rated less than 3 at baseline by the IPR.

Percent Folds Rated <3 at Baseline by IPR Assessment*		
	Hylaform	Zyplast
With rounding	147/256 (57.4%)	124/252 (49.2%)
Without rounding	188/256 (73.4%)	169/252 (67.0%)
*All folds rated at least a 3 on the "live" assessment for study entry		

This table illustrates the gap between the "live" ratings and the masked IPR scores, which was approximately a 1-point mean difference. However, because there was no difference between the two treatment groups with regards to distribution of baseline scores, and because baseline score was used as a covariate in the non-inferiority analysis, I don't think the fact that many IPR scores were less than 3 is a problem. As an aside, the analyses were performed on raw IPR scores with no rounding.

As far as safety is concerned, the incidence of adverse events was identical in each treatment group (88%) and most were mild. The occurrence of individual events appears to be comparable in both groups. Given the sample size, huge differences would have had to have been observed for rates to be statistically different (e.g. 2% vs 13%, or 10% vs 26%). Therefore, statistical testing on the

adverse event rates have little meaning and the safety profile should be evaluated from a clinical perspective.

Conclusion

In summary, I find that the sponsor conducted a well-designed clinical study with excellent follow-up and the data support the non-inferiority of Hylaform with respect to safety and effectiveness. However, there are some clinical issues (including adequate representation of skin types) that must be considered in the evaluation of this product. Since the correction does not last more than a few months at best, patients choosing this treatment will likely seek repeat treatments. The long-term safety and efficacy of that have not been evaluated.

Phyllis M. Silverman, M.S.