DEPARTMENT OF HEALTH AND HUMAN SERVICES Public Health Service

Food and Drug Administration
Center for Biologics Evaluation and Research
Office of Biostatistics and Epidemiology
Division of Biostatistics

STATISTICAL REVIEW AND EVALUATION

BLA

BLA/Supplement Number:	125400/0
Product Name:	Apligraf (oral)
Indication(s):	Surgically created gingival and alveolar mucosal surface defects
Applicant:	Organogenesis
Date submitted:	May 13, 2011
Review Priority:	Standard
Statistical Branch:	Therapeutics Evaluation
Primary Statistical Reviewer:	John Scott, Ph.D. 2/3/2012
Concurring Reviewer (1):	Shiowjen Lee, Ph.D. 2/3/2012
Concurring Reviewer (2):	Boguang Zhen, Ph.D. 2/3/2012
Medical Office/Division:	OCTGT / DCEPT
Clinical Reviewer(s):	CBER Clinical reviewer: Agnes Lim, M.D. CDRH Clinical consult: Bob Betz, D.D.S.
Project Manager:	Terrolyn Thomas

Table of Contents

1. EXECUTIVE SUMMARY

This is an original biologic license application (BLA) intended to support a license for Apligraf (oral) for the treatment of surgically created gingival and alveolar mucosal surface defects in adults. Apligraf (oral) is a living, bilayered tissue construct consisting of neonatal foreskin-derived cells and structural proteins. Apligraf (oral) is the same final product as the sponsor's current commercially-available Apligraf product, which was approved by the Center for Devices and Radiological Health (CDRH) under pre-market approval (PMA) P950032 on November 10, 1998. The original indication was for the treatment of non-infected partial and full-thickness skin ulcers due to venous insufficiency of greater than 1 month duration and which have not adequately responded to conventional ulcer therapy.

1.1 Conclusions and Recommendations

In the pivotal study, 006-PER-002, Apligraf (oral) met its primary efficacy endpoint of demonstrating that the proportion of Apligraf-treated sites with 6 month keratinized tissue (KT) > 2 mm exceeds 50%. The product also met four of six secondary efficacy endpoints: superiority in color matching and texture matching relative to control (free gingival graft), KT > 1 mm success rate in excess of 80%, and superiority in patient preference to control. I did not identify any major statistical issues with the study.

This application was the subject of a Cellular, Tissue and Gene Therapies Advisory Committee meeting held on November 17, 2011 in Silver Spring, MD. There were voting questions on effectiveness and safety. All fifteen voting members of the committee voted "Yes" on the question, "Based on the data provided, is Apligraf effective for the treatment of surgically created gingival surface defects in adults?". Fourteen voting members voted "Yes" and one member voted "No," on the question, "Do the data presented demonstrate the safety of Apligraf for the proposed indication?"

1.2 Brief Overview of Clinical Studies

The sponsor has submitted the results of two completed clinical studies of Apligraf in the oral indication, a pilot study (05-PER-001) and a pivotal study (06-PER-002-CTX). Study 006-PER-002-CTX was a within-subject controlled trial involving 96 subjects with 85 subjects included in efficacy analyses. This study met its primary efficacy endpoint of demonstrating that the proportion of Apligraf-treated sites with 6 month KT > 2 mm exceeds 50%.

Apligraf also met four of six secondary efficacy endpoints in study 006-PER-002-CTX: superiority in color matching and texture matching relative to control, KT > 1 mm success rate in excess of 80%, and superiority in patient preference to control. Apligraf

was not superior to control on sensitivity at week 1 or on pain at day 3. As these endpoints were tested in fixed sequence, there are no multiplicity issues.

The pilot study, 005-PER-001, did not meet its primary efficacy endpoint of demonstrating that Apligraf was non-inferior (1 mm margin) to control on change in width of attached gingiva from baseline to month 6. In fact, both in study 005-PER-001 and study 006-PER-002, control was significantly superior to Apligraf on 6 month KT width and attached gingival width.

1.3 Major Statistical Issues and Findings

I did not identify any major statistical issues that would affect the interpretation of the results of the pivotal trial, 006-PER-002-CTX. The statistical analyses were appropriately chosen and correctly performed as prespecified in the study protocol. I was able to reproduce the primary analyses of all primary and secondary efficacy endpoints using the SAS datasets submitted by the sponsor.

2. INTRODUCTION

2.1 Overview

Apligraf is a living, bilayered tissue construct consisting of neonatal foreskin-derived cells and structural proteins. The fibroblast layer consists of bovine Type 1 collagen and human fibroblasts which produce additional human matrix proteins. The keratinocyte layer is formed by human keratinocytes (epithelial cells) which first multiply and then differentiate into a cornified epithelial layer. The layers adhere as one unit to form the final product.

The proposed indication is for the treatment of surgically created gingival and alveolar mucosal surface defects in adults. The product is applied over a vascular wound bed to regenerate site-appropriate oral mucosal tissues. Apligraf has not been marketed in the United States or any foreign country for the oral indication. However, Apligraf (oral) is the same final product as the sponsor's current commercially-available Apligraf product, which was approved by CDRH under PMA P950032 on November 10, 1998. The original indication was for the treatment of non-infected partial and full-thickness skin ulcers due to venous insufficiency of greater than 1 month duration and which have not adequately responded to conventional ulcer therapy.

The sponsor originally intended to submit a PMA to CDRH for the use of Apligraf in the oral indication. However, the product has now been designated as a biologic and, following formal and informal pre-BLA discussion between the sponsor and a Center for Biologics Evaluation and Research (CBER) review team, the planned PMA was converted to a BLA and submitted for review to CBER.

Since 1990, there have been approximately 787 Apligraf-treated patients in 19 completed Organogenesis-sponsored clinical trials and an estimated –b(4)---- commercial units shipped for patient treatment.

The sponsor has submitted the results of two completed clinical studies of Apligraf in the proposed oral indication: a pilot study (05-PER-001) and a pivotal study (06-PER-002-CTX). The sponsor also submitted the results of an additional pilot study (07-PER-004-CTX) which was performed in a different indication from that being sought in the BLA (submerged under a surgical flap rather than non-submerged as in studies 05-PER-001 and 06-PER-002), and is only intended to provide supportive evidence of safety.

The sponsor has also submitted study reports for twelve clinical studies performed in non-oral indications. These will not be reviewed below.

2.2 Data Sources

This is an electronic submission in eCTD format. This memo is focused on the review of efficacy results generated from studies 05-PER-001 and 06-PER-002-CTX. This statistical review is based on a clinical overview (module 2.5), study protocols and clinical study reports for studies 05-PER-001 and 06-PER-002-CTX (module 5.3.5).

The sponsor has submitted data files in SAS Transport format, including analysis datasets for 06-PER-002-CTX (module 5.3.5.1.25.3) and pooled safety datasets for studies 05-PER-001 and 06-PER-002-CTX (module 5.3.5.3.25.3). I have verified all critical efficacy findings reported in the clinical study report for 06-PER-002-CTX, including any efficacy claim included in draft labeling, against the data provided by the sponsor.

No electronic efficacy datasets were submitted for pilot study 05-PER-001. The sponsor included printouts of data line listings and output from analyses performed with the SPSS software package as appendices to the clinical study report for study 05-PER-001. I have verified the primary analysis of the primary efficacy endpoint for study 05-PER-001 using these printouts.

3. STATISTICAL EVALUATION

My statistical evaluation is based on a brief overview of the pilot study, 05-PER-001, and a more detailed examination of the design and findings of the pivotal study, 06-PER-002-CTX.

3.1 Evaluation of Efficacy

3.1.1. Study 05-PER-001

Study Design and Endpoints

Study 05-PER-001 was a pilot study designed to assess the safety and efficacy of Apligraf in establishing a functional zone of attached gingiva. The primary objective was to determine if Apligraf can provide a functional zone of attached gingiva comparable to a free autogenous palate graft. The secondary objectives were to evaluate:

1. Inflammation
2. Color and texture match of the graft to the adjacent tissue
3. Resistance to oral muscle pull
4. Probing depth
5. Clinical attachment level
6. Subject preference or satisfaction
7. Change in recession depth
8. Width of keratinized tissue

This was a prospective, randomized, within subject controlled (matched for teeth and gingival condition), single center, pilot study. The first three patients were used to help determine surgical and material handling techniques, and were not included in the statistical analysis. Subsequent patients were enrolled when the first three patients completed the first four weeks of follow up. Treatment site and order of treatment were randomized. Following screening and randomization, subjects received both a palate graft and Apligraf, with primary endpoint evaluations at Month 6. There were patient follow-up visits at Week 1, Month 1, Month 3, and Month 6.

The primary efficacy variable was the change in the amount of attached gingiva at Month 6 compared between treatments. The secondary efficacy variables were the change from baseline in width of keratinized tissue, recession depth, inflammation score, color and texture match of the grafted tissue to the adjacent tissue, resistance to oral muscle pull, clinical attachment level, probing depth, subject discomfort and subject preference or satisfaction at Month 6, compared between treatments.

The primary efficacy endpoint and most secondary efficacy endpoints (probing depth, recession, clinical attachment, attached gingiva, keratinized tissue width, plaque index) were analyzed with a repeated measures analysis of covariance (ANCOVA), with within-subject terms for treatment and controlling for baseline values of the efficacy variables. The sponsor used Friedman's test for related outcomes to test for differences in patient perceptions of duration of pain, bleeding, swelling, and sensitivity among surgical sites. The Wilcoxon signed rank test was used to compare other secondary variable scores at each time point postoperatively (inflammation scores, change in bleeding on probing, muscle pull, patient preference and tissue color and texture).

The trial was designed to demonstrate non-inferiority between treatment and control in the change in amount of attached gingiva over the 6-month observation period. The primary hypothesis tested was H_0: D <= -1 mm vs. Ha: D > -1 mm, where D = mean difference within subjects (treatment change over 6 months – control change over 6 months).

Note that no multiplicity adjustment was used for the secondary efficacy endpoints. Consequently, any p-values should be interpreted with caution.

Results and Conclusions

Twenty-five patients were screened and enrolled. The first three patients were considered training patients, and the remaining 22 were included in analyses. The first patient was enrolled in September, 2005 and the final patient visit was conducted in May, 2006.

Table 4 from the clinical study report, reproduced below, summarizes the efficacy results from the trial. Apligraf treatment showed an average increase of attached gingiva at Month 6 compared to baseline from 0.30 mm to 1.14 mm, which the sponsor described as an adequate functional zone. Apligraf treatment established at least 2 mm of keratinized tissue width in 81.8 % of the cases. Apligraf sites exhibited an average increase in keratinized tissue width from 1.13 mm at baseline to 2.50 mm at 6 months. The confidence intervals reported in Table 4 may be misleading. These were calculated based on estimated standard errors from the repeated measures ANCOVA models the sponsor used to compare Apligraf to control treatment. The models, however, treated time as a within-subject factor but treated site as a between-subject factor, even though the treatment comparison was also within-subject. However, this does not affect the qualitative conclusions of the study.

When Apligraf was compared to the control treatment, the control sites showed a greater change from baseline in the amount of attached gingiva compared to Apligraf sites ($p<0.001$). There was also a significantly larger change from baseline to 6 months in width of keratinized tissue in the control site compared to the Apligraf site ($p<0.001$). The sponsor did not report the results of the primary non-inferiority hypothesis test for this trial in the clinical study report. However, based on the fact that the point estimate for the mean difference between Apligraf and control in change in amount of attached gingiva from baseline to Month 6 was in the null region ($\hat{D} = -1.58$), non-inferiority could not be demonstrated.

Section 3 Table 4: Clinical Variables Over Time and Change in Clinical Variables from Baseline to 6 Months by Treatment Group (n=22)

	Baseline Mean (95% CI)	3 Months Mean (95% CI)	6 Months Mean (95% CI)	$p^{\S}$	Change Mean (95% CI)	$p^{\S}$
Probing Depth						
Apligraf	1.37 (1.18,1.56)		1.41 (1.24,1.58)		-0.04 (-0.26,0.18)	
Control	1.36 (1.17,1.55)	-	1.68 (1.51,1.85)	0.211	-0.32 (-0.54,-0.11)	0.071
Recession						
Apligraf	2.42 (2.17,2.66)	2.41 (2.14,2.68)	2.16 (1.89,2.44)		0.25 (0.09,0.42)	
Control	2.36 (2.11,2.60)	2.02 (1.75,2.29)	2.02 (1.74,2.30)	0.250	0.34 (0.18,0.50)	0.453
Clinical Attachment						
Apligraf	3.79 (3.49,4.08)		3.58 (3.32,3.83)		0.21 (-0.05,0.47)	
Control	3.71 (3.42,4.01)	-	3.70 (3.45,3.95)	0.888	0.02 (-0.25,0.28)	0.278
Attachment Creep						
Apligraf					1.57 (1.21,1.93)	
Control	-	-	-		1.43 (1.07,1.79)	0.572
Attached Gingiva						
Apligraf	0.30 (0.13,0.46)		1.14 (0.77,1.50)		0.85 (0.48,1.21)	
Control	0.27 (0.11,0.44)	-	2.71 (2.34,3.07)	<0.001	2.43 (2.06,2.79)	<0.001
Keratinized Tissue Width						
Apligraf	1.13 (0.92,1.33)	2.61 (2.36,2.87)	2.50 (2.18,2.82)		1.37 (0.97,1.77)	
Control	1.24 (1.03,1.44)	4.50 (4.24,4.76)	4.57 (4.25,4.89)	<0.001	3.33 (2.93,3.74)	<0.001
Plaque Index						
Apligraf	0.21 (0.13,0.28)	0.30 (0.22,0.37)	0.27 (0.20,0.35)		0.07 (-0.02,0.16)	
Control	0.30 (0.22,0.37)	0.27 (0.20,0.35)	0.30 (0.22,0.37)	0.427	-0.00 (-0.09,0.09)	0.278

The sponsor noted that when only sites with positive attached gingiva were evaluated, no significant differences were detected between groups for attached gingiva ($p=0.184$), and keratinized tissue width ($p=0.057$). However, this appears to have been an entirely post hoc comparison and, furthermore, marginally significant differences between groups in a small pilot study cannot be taken as assurance that no true differences exist.

When assessing periodontal health around test and control teeth, the sponsor found no statistically significant differences between Apligraf and control in the change from baseline to 6 months in probing depth, recession and clinical attachment. There was also no statistically significant difference in resistance to muscle pull, inflammation, or bleeding on probing within the two groups. Apligraf treatment showed a significantly higher subject satisfaction/preference score and significantly better tissue color (more equally red compared to surrounding tissue than control) and tissue texture (more equally firm compared to surrounding tissue than control).

Since the control site treatment necessitates harvesting an autologous palatal graft, patient pain perception was compared between the combined control graft site and control donor sites vs. the Apligraf site. At Week 1 at the Apligraf site, 13.6% of subjects reported no pain, 22.7% reported mild pain, 40.9% reported moderate pain and 22.7% reported severe pain. For the control treatment, where pain was scored as the greater of the graft and donor site pain scores, 4.5% of subjects reported no pain, 36.4% reported mild pain, 50.0% reported moderate pain and only 9.1% reported severe pain. The sponsor also performed post hoc testing of duration of pain, but these results are not readily interpretable. Surgical site biopsy specimens at 6 months showed intact Stratified Squamous Epithelium and tissue of normal thickness and architecture.

In summary, the study demonstrated some efficacy of Apligraf in the proposed indication, but failed in its primary hypothesis test of demonstrating non-inferiority in change in gingival attachment relative to control. In fact, Apligraf was found to be significantly inferior to control in gingival attachment and width of keratinized tissue. The sponsor's conclusions from the study were as follows:

> The results indicate that Apligraf is a safe and effective alternative to harvesting tissue from the palate. Apligraf was unable to produce as much keratinized tissue and attached gingiva, but resulted in equivalent periodontal health (as measured by clinical attachment levels, recession, muscle pull, and inflammation) and a shorter duration of pain and sensitivity. Apligraf showed a significantly higher subject satisfaction/preference score and a significant difference between tissue color (more equally red compared to surrounding tissue) and in tissue texture (more equally firm compared to surrounding tissue).
>
> Based on the results of this study, future studies should consider focusing more attention on patient comfort (i.e. a more tolerable subject pain experience), esthetics (i.e. integration with surrounding tissues in terms of tone and texture) and clinical measures associated with maintenance of periodontal health (i.e. pocket depth and clinical attachment level).

These conclusions seemingly helped guide the design of the pivotal study, 006-PER-002-CTX, including the fact that there was no direct superiority or non-inferiority comparison between Apligraf and control on the primary endpoint in that study.

3.1.2 Study 006-PER-002-CTX

Study Design and Endpoints

Study 006-PER-002-CTX was a prospective, randomized, within-subject controlled, multicenter trial intended to serve as a pivotal trial for establishing the safety and efficacy of Apligraf in the oral indication. Up to 96 subjects with recession-type defects aged 18 – 70 were to be enrolled, including up to 14 training subjects. Training subjects were included in safety analyses and supportive analyses of efficacy only.

Study procedures and follow-up

Eligible subjects had two study teeth identified and randomized to treatment with Apligraf or Free Gingival Graft (Day 0); up to three adjacent teeth were treated per quadrant. However, only one tooth in each quadrant was identified as a study tooth. Alveolar bone level, surgical position margin and reference point (REF)-graft base measurements were obtained. Graft bed preparation, graft placement and post-surgical dressing placement were the same for both treatment sites. Photos were taken before, during and after the surgical procedure. Any changes in medications or adverse events were noted. Subjects were prescribed antibiotics and provided with a 0.12% chlorhexidine mouth rinse and oral hygiene instructions.

Subjects were given a paper diary to be completed daily through day 14. Subjects were asked if the surgical dressing stayed on at each of the surgical sites and the palatal donation site. Subjects recorded pain associated with each surgical site and the palatal graft donation site by assessing the pain for each as none, mild, moderate and severe. Additionally, this diary was used to record any medications the subject has taken for mouth-related pain. The diary was not completed on day 0 because subjects may still have been under the effects of anesthesia. In addition, study personnel placed telephone calls to subjects within 48 hours post-surgery, 72 hours post-surgery, and at 2-weeks post-surgery to perform a well-being check.

The first follow-up visit occurred one-week post-surgery. Any changes in medications or adverse events were documented. Photos of the test sites and palatal graft donation site were taken and clinical measurements and assessments of bleeding and swelling were obtained. The investigator removed any dressing that had not fallen off by this visit. Sensitivity was assessed with a puff of air. The subject's surgical procedure preference was recorded and oral hygiene instructions reviewed.

Further follow-up evaluations occurred at 4 weeks, 3 months and 6 months post-surgery. Changes in medications and adverse events were documented at each visit. Photos of the test sites and clinical measurements were obtained and texture and color of the test areas

were evaluated. At 4 weeks, pain and sensitivity were also assessed. An oral exam was performed at 4 weeks and 6 months. At 3 and 6 months post-surgery a dental cleaning was performed. At 6 months post-surgery a Subject Preference Questionnaire was completed, radiographs of the study teeth obtained and a pregnancy test administered to females of childbearing potential.

Objectives and endpoints

The primary objective of the trial was to evaluate the ability of Apligraf to achieve a clinically acceptable threshold for keratinized tissue (KT) at 6 months (KT > 2 mm).

The secondary objectives were to determine if Apligraf is superior to control or to a fixed standard for:

1. Color same as adjacent tissues after 6 months (superiority to control)
2. Texture same as adjacent tissues after 6 months (superiority to control)
3. KT > 1 mm for CelTx™ after 6 months (superiority vs. a 80% success standard)
4. Patient preference after 6 months (superiority to control)
5. Surgical site sensitivity mild or absent after 1 week (superiority to control)
6. Pain absent after 3 days (superiority to control)

The endpoint corresponding to the primary efficacy objective was proportion of patients with KT > 2 mm at 6 months for the Apligraf-treated site. The secondary efficacy endpoints were:

1. Color: The proportion of patients having the same color as adjacent tissues after six months.
2. Texture: The proportion of patients having the same texture as adjacent tissues after six months.
3. KT ≥ 1 mm: The proportion of patients achieving KT ≥ 1 mm after six months.
4. Patient Preference: The proportion of patients that stated a preference for Apligraf after six months.
5. Sensitivity: The proportion of patients with surgical site sensitivity mild or absent after one week.
6. Pain: The proportion of patients with surgical site pain absent after 3 days.

The sponsor also assessed six additional efficacy endpoints at 6 months:

1. Width of attached gingiva
2. Width of keratinized gingiva
3. Resistance to muscle pull
4. Clinical attachment level
5. Recession
6. Inflammation
7. Bleeding on probing

The safety endpoints were to evaluate changes in soft tissue healing, postoperative infections, excessive postsurgical bleeding and/or swelling and to assess treatment-specific and systemic Adverse Events (AEs) observed and/or reported.

Patient Disposition, Demographic and Baseline Characteristics

The sponsor screened 119 subjects, 23 of whom were not enrolled due to violations of inclusion or exclusion criteria. Of the 96 subjects enrolled, 11 (the first two for each surgeon) were considered training subjects. The remaining 85 subjects were considered 'pivotal' subjects. Training subjects were not included in the primary efficacy analyses, but were analyzed separately and also pooled with pivotal subjects in supportive analyses.

Subjects were enrolled at 4 centers and treated by 6 different surgeons. At one of the centers, a total of three subjects were treated by two surgeon, so all three of the subjects from this center were considered training subjects. Thus, only three centers contributed pivotal subjects to the efficacy analysis. Sample sizes were approximately evenly distributed across these three centers.

The sponsor defined intent-to-treat (ITT), modified intent-to-treat (mITT), per protocol (PP) and safety analysis populations in the protocol, and planned on performing the primary efficacy analysis in the mITT population with a supportive analysis in the PP population. The mITT population was defined as all subjects who received the treatment as randomized to each side of the mouth and were followed for at least one week. However, all 96 enrolled patients were treated as randomized and completed all required follow-up visits, so that the ITT, mITT, PP and safety analysis sets turned out to be identical.

Table 10-3 of the clinical study report summarizes baseline demographics for the study population and baseline dental problems are summarized in Table 10-4.

Table 10-3 Selected Demographics (mITT Population)

Parameter	Statistics	Pivotal	Training	All
Gender				
Female	n (%)	46 (54.1%)	6 (54.5%)	52 (54.2%)
Male	n (%)	39 (45.9%)	5 (45.5%)	44 (45.8%)
Age (years)	n	85	11	96
	Mean (SD)	46.85 (12.700)	49.40 (16.662)	47.14 (13.134)
	Median	48.33	53.28	48.75
	Min, Max	18.0, 70.8	21.2, 70.3	18.0, 70.8
Race				
White	n (%)	77 (90.6%)	10 (90.9%)	87 (90.6%)
Black of African American	n (%)	1 (1.2%)	0 (0.0%)	1 (1.0%)
Native Hawaiian or Other Pacific Islander	n (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Asian	n (%)	4 (4.7%)	1 (9.1%)	5 (5.2%)
American Indian or Alaska Native	n (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Other	n (%)	3 (3.5%)	0 (0.0%)	3 (3.1%)
Site				
Site 10	n (%)	30 (35.3%)	4 (36.4%)	34 (35.4%)
Site 15	n (%)	0 (0.0%)	3 (27.3%)	3 (3.1%)
Site 16	n (%)	27 (31.8%)	2 (18.2%)	29 (30.2%)
Site 17	n (%)	28 (32.9%)	2 (18.2%)	30 (31.3%)
Previous Tobacco Use				
Yes	n (%)	34 (40.0%)	4 (36.4%)	38 (39.6%)
No	n (%)	51 (60.0%)	7 (63.6%)	58 (60.4%)

Table 10-4 Dental History (mITT Population)

Dental Problems	Statistics	Pivotal	Training	All
Dental Problem				
None	n (%)	12 (14.1%)	0 (0.0%)	12 (12.5%)
Bad breath	n (%)	16 (18.8%)	4 (36.4%)	20 (20.8%)
Loose teeth	n (%)	14 (16.5%)	4 (36.4%)	18 (18.8%)
Bleeding gums	n (%)	29 (34.1%)	3 (27.3%)	32 (33.3%)
Food collection in teeth	n (%)	21 (24.7%)	7 (63.6%)	28 (29.2%)
Grinding or clenching	n (%)	42 (49.4%)	5 (45.5%)	47 (49.0%)
Clicking or popping jaw	n (%)	21 (24.7%)	2 (18.2%)	23 (24.0%)
Sensitivity to cold	n (%)	41 (48.2%)	5 (45.5%)	46 (47.9%)
Sensitivity to biting	n (%)	8 (9.4%)	0 (0.0%)	8 (8.3%)
Sensitivity to heat	n (%)	24 (28.2%)	1 (9.1%)	25 (26.0%)
Sensitivity to sweets	n (%)	21 (24.7%)	1 (9.1%)	22 (22.9%)
Sores in mouth	n (%)	17 (20.0%)	2 (18.2%)	19 (19.8%)
Other	n (%)	6 (7.1%)	0 (0.0%)	6 (6.3%)

Table 10-5 contains the treatment area characteristics for the Apligraf and control sites for each patient; there are no apparent strong imbalances in area characteristics between treatments.

Table 10-5 Treatment Area Characteristics (mITT Population)

Parameter	Statistics	Pivotal		Training		All	
		CelTx	Control	CelTx	Control	CelTx	Control
Treated Side							
Left	n (%)	39 (45.9%)	46 (54.1%)	7 (63.6%)	4 (36.4%)	46 (47.9%)	50 (52.1%)
Right	n (%)	46 (54.1%)	39 (45.9%)	4 (36.4%)	7 (63.6%)	50 (52.1%)	46 (47.9%)
Treatment Jaw							
Mandible	n (%)	80 (94.1%)	80 (94.1%)	9 (81.8%)	9 (81.8%)	89 (92.7%)	89 (92.7%)
Maxilla	n (%)	5 (5.9%)	5 (5.9%)	2 (18.2%)	2 (18.2%)	7 (7.3%)	7 (7.3%)
Additional Treated Teeth							
0	n (%)	51 (60.0%)	47 (55.3%)	5 (45.5%)	5 (45.5%)	56 (58.3%)	52 (54.2%)
1	n (%)	21 (24.7%)	25 (29.4%)	6 (54.5%)	6 (54.5%)	27 (28.1%)	31 (32.3%)
2	n (%)	13 (15.3%)	13 (15.3%)	0 (0.0%)	0 (0.0%)	13 (13.5%)	13 (13.5%)

There were a number of protocol deviations, but none were considered by the sponsor to be major. There was one minor inclusion/exclusion deviation. In one subject, the identified study teeth at baseline were not the teeth treated on Day 0. This subject was retained in analyses and the treated teeth were used for 6-month efficacy assessments. All other protocol deviations were procedure-related or visit schedule-related.

Statistical Methodologies

The primary hypothesis tested for the primary endpoint of KT > 2 mm with Apligraf was H_0: $\pi_c \leq 0.50$ against H_A: $\pi_c > 0.50$ where π_c is the proportion of Apligraf-treated sites with KT > 2 mm at 6 months. This hypothesis was tested by comparing the lower one-sided 95% exact binomial confidence limit for proportion of Apligraf-treated sites with KT > 2 mm to 0.50. It is not clear why 50% was chosen as the benchmark success rate, nor is it clear why the sponsor chose to use a one-sided $\alpha = .05$ instead of the more conventional one-sided $\alpha = .025$ for this analysis. However, as discussed below, the success rate observed was well above 50% which resulted in the rejection of the null hypothesis even at $\alpha = 0.025$.

For each of the secondary endpoints of color, texture, sensitivity and pain, the null hypothesis tested was $H_0: \pi_{10} - \pi_{01} \leq 0$ against H_A: $\pi_{10} - \pi_{01} > 0$ where π_{10} is the proportion of subjects for whom the Apligraf-treated site achieves success while the control site does not, and π_{01} is the proportion of subjects for whom the control site

achieves success while the Apligraf-treated site does not. These hypotheses were tested using McNemar's marginal homogeneity test for paired samples.

For the secondary endpoint of KT > 1 mm, the hypothesis tested was H_0: $\pi_c \leq 0.80$ against H_A: $\pi_c > 0.80$ where π_c is the proportion of Apligraf-treated sites with KT > 1 mm at 6 months. And for the secondary endpoint of patient preference, the hypothesis tested was H_0: $\pi_c \leq 0.50$ against H_A: $\pi_c > 0.50$ where π_c is the proportion of patients preferring Apligraf to control treatment. The sponsor used an exact binomial test for these endpoints.

Each secondary endpoint was tested at a one-sided $\alpha = .025$, using a fixed sequence closed testing procedure to account for multiplicity. The sequence of testing was the same as the order of secondary endpoints given above.

In addition to the primary analysis for each efficacy endpoint, the sponsor also planned on performing supportive analyses controlling for the following covariates:

1. Baseline clinical attachment level,
2. Baseline recession depth,
3. Baseline alveolar bone level, and
4. History of tobacco use.

These analyses were performed using logistic regression models for the endpoints compared to a fixed standard, and generalized linear mixed models with probit link functions for the endpoints compared between Apligraf and control sites.

The sponsor also planned on testing for a center*treatment interaction at $\alpha = .05$ which, if significant, would have led to further examinations of treatment effect by center.

Based on the prospective statistical analysis plan, all other efficacy endpoints and safety endpoints were to be analyzed descriptively only. However, according to the sponsor, at a pre-PMA meeting held in 2009, a statistical reviewer at CDRH requested that the sponsor include post hoc inferential testing of the seven "other effectiveness endpoints" in the planned PMA submission. These post hoc analyses have been included in this BLA submission.

The sponsor developed a missing data plan but, as no subjects were lost to follow-up, the missing data analyses were not performed. There were no interim analyses planned or performed.

Results and Conclusions

Eighty-one of the 85 pivotal subjects (95.3%) met the primary endpoint of month 6 KT > 2 mm at the Apligraf-treated site, which yielded an exact binomial 95% CI of (88.4%, 98.7%). (The sponsor reported a two-sided 95% CI despite planning on performing a one-sided test at $\alpha = .05$.) The null hypothesis that the success rate is no greater than

50% was rejected (p < .0001). All 11 training subjects also met the primary endpoint at the Apligraf-treated site, and all 96 subjects met the primary endpoint at the control site.

The sponsor also performed post hoc analyses of the success rate for KT > 2.5 mm and KT > 3 mm at the Apligraf site. These analyses are summarized in Table 11-4, but should be interpreted with caution.

Table 11-4 *Post hoc Analysis*: Keratinized Tissue ≥ 2.5 mm and ≥ 3.0 mm at 6 Months (mITT Population)

	Statistics	Pivotal CelTx
KT ≥ 2.5 mm	n (%)	60 (70.6%)
Test if Pc2 = 50%	p-value	<0.001
	95% CI	(59.7, 80.0)
KT ≥ 3.0 mm	n (%)	59 (69.4%)
Test if Pc2 = 50%	p-value	<0.001
	95% CI	(58.5, 79.0)

The sponsor also performed a post hoc subgroup analysis of the primary endpoint based on whether Apligraf was used to treat a defect spanning a single tooth (60% of pivotal subjects) or spanning multiple teeth (40% of pivotal subjects). The overall efficacy results were qualitatively consistent between these subgroups.

With respect to the secondary efficacy endpoints, Apligraf was statistically significantly superior to control treatment for color matching, texture matching, and patient preference, and had a KT > 1 mm success rate that was significantly greater than 80%. There was no significant difference between Apligraf and control on sensitivity. Because of the pre-planned fixed sequence testing of endpoints, no statistical significance testing could be performed to assess differences in pain between Apligraf and control sites. (Note that, despite this, the sponsor did calculate a p-value for the pain comparison, and found p = .56.) Tables 11-5, 7, 9, 10, 11 and 13, below, summarize the results for these secondary endpoints.

Table 11-5 Color: Test for Discordance at Month 6 (mITT Population)

		Pivotal CelTx	
		Equally Red	Not Equally Red
Month 6			
Pivotal Control	Equally Red	23	0
	Not Equally Red	56	6
		p-value < 0.0001	

Table 11-7 Texture: Test for Discordance at Month 6 (mITT Population)

		Pivotal CelTx	
		Equally Firm	Not Equally Firm
Month 6			
Pivotal Control	Equally Firm	46	0
	Not Equally Firm	35	4
		p-value < 0.0001	

Table 11-9 Keratinized Tissue ≥ 1 mm at Month 6 (mITT Population)

	Statistics	Pivotal CelTx	Pivotal Control
Month 6	n (%)	85 (100.0%)	85 (100.0%)
Test if Pc1 = 80%	p-value	<0.0001	
	95% CI	(95.8, 100.0)	

Table 11-10 Subject Preference after Month 6/Early Termination (mITT Population)

	Statistics	Pivotal CelTx	Pivotal Control
Overall preference	n (%)	61 (71.8%)	24 (28.2%)
Test if Pt = 50%	p-value	<0.0001	
	95% CI	(61.0, 81.0)	
Preference based on appearance only	n (%)	65 (76.5%)	20 (23.5%)
Test if Pa = 50%	p-value	<0.0001	
	95% CI	(66.0, 85.0)	

Table 11-11 Sensitivity: Test for Discordance at Week 1 (mITT Population)

		Pivotal CelTx	
		Not Sensitive	Sensitive
Week 1			
Pivotal Control	Not Sensitive	67	3
	Sensitive	1	0
		p-value = 0.3173	

Table 11-13 Pain Assessment: Test for Discordance at Day 3 (mITT Population)

		Pivotal CelTx	
		No Pain	Pain
Day 3			
Pivotal Control	No Pain	54	7
	Pain	5	18
		p-value = 0.5637	

The sponsor also provided summaries and, at the request of a CDRH review team, post-hoc statistical analyses of the "other effectiveness endpoints" of Recession Depth, Recession (%), Probing Pocket Depth (including mesial and distal), Clinical Attachment Level (CAL), Keratinized Tissue Width (including mesial and distal), Attached Gingiva, Bleeding on Angulated Probing, Muscle Pull Resistance, Plaque scores (buccal and lingual), Inflammation Score, Bleeding, and Swelling. The p-values from these post hoc analyses are not readily interpretable, but the major qualitative results are summarized briefly below.

The sponsor assessed changes from baseline to month 6 in each of the endpoints listed above, using the Wilcoxon signed-rank test for ordinal variables and McNemar's test for dichotomous variables. For the Apligraf-treated site, there were significant ($p < .05$) improvements from baseline to month 6 in KT width, attached gingiva width and recession depth. No other endpoints yielded significant changes from baseline to month 6. For the control site, there were significant improvements from baseline to month 6 in KT width and attached gingival width. The improvement in recession depth at the control site did not quite reach the level of statistical significance ($p = .077$).

The sponsor also compared the Apligraf and control sites on each of these endpoints at month 6, again using the Wilcoxon signed-rank test for ordinal variables and McNemar's test for dichotomous variables. The control site had significantly ($p < .0001$) greater KT

width and attached gingival width at month 6 than the Apligraf site. The mean (SD) KT width at month 6 was 3.21 (1.14) mm for the Apligraf site compared to 4.57 (1.00) mm for the control site. The mean (SD) attached gingival width at month 6 was 1.77 (1.32) mm for the Apligraf site compared to 3.17 (1.17) mm for the control site. There were no other significant differences between sites on 'other effectiveness endpoints' at month 6.

The sponsor performed supportive analyses of primary and secondary site-specific efficacy endpoints controlling for four baseline covariates. The results of these analyses were consistent with the primary, uncontrolled analyses. The sponsor also performed supportive analyses of training subjects only and of training subjects pooled with pivotal subjects. These results were also qualitatively consistent with the analyses of pivotal subjects only.

3.2 Evaluation of Safety

In the pilot study 05-PER-001, investigators reported no signs of local or systemic reaction to Apligraf. Twenty-two adverse events were reported among 17 patients. All were judged unrelated to treatment. None were severe. There were no serious adverse events or deaths. Polymerase chain reaction (PCR) studies at 6 months showed no sign of Apligraf DNA persistence in the patients' graft sites.

In the pivotal study 06-PER-002-CTX, 25% of the subjects experienced an adverse event (AE) during the study and a total of 43 events were reported in the study. Fifteen AEs reported by 6 subjects were assessed by the investigator to be related (unlikely, possible, probably, definite) to study treatment. Of these 15, 2 AEs reported by two subjects occurred at the Apligraf site (gingival injury and gingival pain) and were assessed as possibly or probably related, respectively. Three serious adverse events (SAEs) were reported during the study and all three events were assessed by the investigator to be either not related (pneumonia and chest pain) or of unlikely relationship (metastatic malignant fibrous histiocytoma). An additional non-oral cavity malignancy was reported during the study, which in the assessment of the investigator was not related (follicular thyroid cancer).

The three study locations were the Apligraf-treated site, control-treated site and the palatal harvest site. Adverse events occurring at these locations were:

- Two subjects experienced AEs occurring at the palatal harvest site (post-procedural hemorrhage and thrombosis). Palatal harvest morbidity is a well-established risk with harvest of a free gingival graft (FGG).
- Two subjects experienced AEs occurring at the control-treated site (gingivitis and skin exfoliation).
- Three subjects experienced AEs occurring at the Apligraf-treated site. Two subjects, both in the training cohort, had inadvertent placement of the Apligraf transwell membrane to the oral mucosal defect at the time of Apligraf placement which resulted in AEs of gingival injury and gingival pain. For both of these

subjects the membrane was able to be removed without sequelae and ≥ 2 mm KT was regenerated at the treatment site. This potential safety issue was able to be addressed through training of the investigators and did not occur in subjects beyond the training cohort. The third subject experienced an AE of mouth ulceration.

A detailed review of the safety data is deferred to the clinical review team.

3.3 Gender, Race, Age and Other Special/Subgroup Populations

Of the 85 pivotal subjects in study 06-PER-002, 39 (46%) were men and 46 (54%) were women. All 39 men met the primary efficacy endpoint of KT > 2 mm at Month 6, while 42 / 46 (91%) of the women met the primary efficacy endpoint. In terms of race, 77 subjects (91%) were white and 8 (9%) were non-white. The primary efficacy endpoint was met by 74/77 (96%) of white subjects and by 7/8 (88%) of nonwhite subjects. For age, 79 subjects (93%) were between the ages of 18 and 65 and 6 (7%) were 65 or older. The primary efficacy endpoint was met by 75/79 subjects (95%) between the ages of 18 and 65, and by all 6 subjects aged 65 or older.

There were no subjects studied under the age of 18. At the 11/17/2011 Advisory Committee meeting, committee members noted that this product might be used in patients under the age of 18, particularly in the context of orthodontic treatment. Consequently, the sponsor is currently planning on performing a pediatric study, possibly as part of a Post-Marketing Requirement (PMR).

By study center, the number (percent) of pivotal subjects in study 06-PER-002 who met the primary endpoint of KT > 2 mm for the Apligraf-treated site at 6 months was 26/30 (87%) at center "10," 27/27 (100%) at center "16" and 28/28 (100%) at center "17."

4. SUMMARY AND CONCLUSIONS

4.1 Statistical Issues and Collective Evidence

I did not identify any major statistical issues that would affect the interpretation of the results of the pivotal trial, 006-PER-002-CTX. The statistical analyses were appropriately chosen and correctly performed as prespecified in the study protocol. I was able to reproduce the primary analyses of all primary and secondary efficacy endpoints using the SAS datasets submitted by the sponsor.

In terms of collective evidence, the sponsor's claim of effectiveness for Apligraf in the oral indication rests almost entirely on the results of 006-PER-002-CTX, a within-subject

controlled trial involving 96 subjects with 85 subjects included in efficacy analyses. This study met its primary efficacy endpoint of demonstrating that the proportion of Apligraf-treated sites with 6 month KT > 2 mm exceeds 50%.

Apligraf also met four of six secondary efficacy endpoints: superiority in color matching and texture matching relative to control, KT > 1 mm success rate in excess of 80%, and superiority in patient preference to control. Apligraf was not superior to control on sensitivity at week 1 or on pain at day 3. As these endpoints were tested in fixed sequence, there are no multiplicity issues.

The pilot study, 005-PER-001, did not meet its primary efficacy endpoint of demonstrating that Apligraf was non-inferior (1 mm margin) to control on change in width of attached gingiva from baseline to month 6. In fact, both in study 005-PER-001 and study 006-PER-002, control was significantly superior to Apligraf on month 6 KT width and attached gingival width.

4.2 Conclusions and Recommendations

In the pivotal study, 006-PER-002, Apligraf (oral) met its primary efficacy endpoint of demonstrating that the proportion of Apligraf-treated sites with 6 month KT > 2 mm exceeds 50%. The product also met four of six secondary efficacy endpoints: superiority in color matching and texture matching relative to control, KT > 1 mm success rate in excess of 80%, and superiority in patient preference to control. I did not identify any major statistical issues with the study.

This application was the subject of a Cellular, Tissue and Gene Therapies Advisory Committee meeting held on November 17, 2011 in Silver Spring, MD. There were voting questions on effectiveness and safety. All fifteen voting members of the committee voted “Yes” on the question, “Based on the data provided, is Apligraf effective for the treatment of surgically created gingival surface defects in adults?”. Fourteen voting members voted “Yes” and one member voted “No,” on the question, “Do the data presented demonstrate the safety of Apligraf for the proposed indication?”