CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

21-201s000

STATISTICAL REVIEW(S)

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmacoepidemiology and Statistical Science
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/Serial Number:	21-201
Drug Name:	Aethoxysklerol (Polidocanol) 0.5%, 1%, [redacted]
Indication(s):	Sclerotherapy of varicose veins
Applicant:	Chemische Fabric Kreussler & Co, GmbH
Date(s):	Received 12/01/03, user fee (10 months) 8/02/04
Review Priority:	Standard
Biometrics Division:	Division 3, HFD-725
Statistical Reviewer:	Steve Thomson, HFD-725
Concurring Reviewer:	Division Director: Mohammed Huque, Ph.D.,HFD-725
Medical Division:	Dermatological and Dental Drug Products, HFD-540
Clinical Team:	Efficacy Reviewer: Brenda Vaughan, M.D.,HFD-540 Safety Reviewer: Brenda Carr, M.D.,HFD-540 Team Leader: M. Luke, M.D., Ph.D.,HFD-540
Project Manager:	Lea Carrington, HFD-540
Keywords:	Active control/noninferiority, Bayesian analysis, clinical studies, Cochran-Mantel-Haenszel, multi-center, NDA review, multiple Comparisons, post hoc/prospective analyses

Table of Contents

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

The objective of this study was to compare the efficacy and safety of Aethoxysklerol to Sotradecol in the treatment of varicose veins in the lower extremities. Complete disappearance of varicosities was specified as the primary endpoint. Although the comparator drug, Sotradecol, was active, the dose levels of the comparator were below the labeled doses. Therefore the Medical team specified that the comparisons between Aethoxysklerol and Sotradecol should be superiority comparisons. The analysis was conducted as one two center study and one single center study, following the a single protocol. The analysis was stratified on three vein sizes. The primary endpoint was whether or not there was complete disappearance of varicosities in the target vein. Even without an adjustment for the multiplicity of tests, in both studies no overall or within vein size treatment comparisons were statistically significant (over studies and vein sizes all $p \geq 0.1758$). Further, in neither study did any of these comparisons of this dichotomous endpoint meet the requirements to show non-inferiority (see Appendix 1). Similar results from a preliminary Bayesian analysis of the pooled studies are given in Appendix 6.

1.2 Brief Overview of Clinical Studies

This submission has a long and varied history. Although the Sponsor seems to have originally considered it as a multi-center trial comparing Aethoxyslkerol versus Sotradecol, the powering was done within center by vein size. So the original Agency statistical reviewers considered it as three separate single center studies following a common protocol. Doses were stratified by vein size as follows:

1) Veins under 1 mm were to receive 0.5% Aethoxysklerol or 0.25% Sotradecol.
2) Veins 1 - 3 mm were to receive 1.0% Aethoxysklerol or 0.5% Sotradecol.
3) Veins 3 - 6 mm were to receive 3.0% Aethoxysklerol or 1.5% Sotradecol.

A post hoc adjustment separated the results from the three centers into two nominal studies, one with the California and Michigan centers, the other based on the single Ohio center. The primary endpoint was disappearance of the varicosities as determined by a panel of three vascular surgeons, by week 16 after last treatment. The Medical team specified that was to be interpreted as complete disappearance and that Aethoxysklerol should be required to show superiority over Sotradecol.

Table 1 below provides the number of enrolled patients per center (actually defined as the number of patients with demographic information) and the number of subjects who dropped out/were lost to follow-up. Note that no detailed information on dropouts was provided. These are just the patients for whom no complete follow-up information was included.

Table 1. Number Randomized (Number of drop outs, if greater than 0)

Vein size	California		Michigan		Ohio		Total
	Sotra-decol	Aethoxy-sklerol	Sotra-Decol	Aethoxy-Sklerol	Sotra-decol	Aethoxy-sklerol	
Overall	70 (1)	63 (3)	21 (3)	20	75 (6)	75 (2)	324 (15)
< 1 mm	23	21 (2)	10 (1)	7	25 (4)	25	111 (7)
1 mm - 3 mm	23 (1)	20 (1)	8 (2)	8	25 (2)	25 (2)	109 (8)
3 mm - 6 mm	24	22	3	5	25	25	104

Note the relatively small number of dropouts (4% in the pooled MICA study and 5.3% in the Ohio study). It seems quite likely that there was a restriction on randomization, at least in terms of blocksize, particularly in the Ohio study. However no such details of the randomization are provided in either the reports or the protocol.

Further details on the design are provided in Section 3.1.

1.3 Statistical Issues and Findings

Issues

The actual method of sclerotherapy was left to the investigator, so we could expect investigator differences. The use of only three investigators does not provide much information on potential between method/center effects, and hence these are modeled as fixed effects. Finally the patient could return for multiple treatments (actually up to 3, though the protocol specified more). Only success was to count.

From a statistical point of view the primary difficulty with this submission comes from the mostly post hoc, sometimes not necessarily consistent recommendations for analysis. For example:

1. The decision to split the studies into two separate studies was a post hoc suggestion that would not be consistent with current practice. Note that although this would have power implications, the original three center studies were powered for a different endpoint, anyway.

2. In the third and last protocol, submitted after completion of the studies, the Sponsor proposed a 5-point scale, labeled "Disappearance of Varicosities", as the primary endpoint scored 1=worse, 2=same, up to 5=complete disappearance. At various meetings the Sponsor was apparently reminded that this should be analyzed as a dichotomous variable, but at other meetings the original 5-point scale seems to have been implicitly accepted.

3. The Sponsor was informed that since Sotradecol was to be used at below labeled dose, at least for vein sizes below 3mm, it was to be treated as a placebo, and superiority comparisons should be used. However, at several other times the Sponsor was also told that these could be treated as non-inferiority comparisons, and could be tested using the procedure described in Appendix 1. The Sponsor's analysis seems to conclude that since there were no statistically

significant differences using the Disappearance of Varicosities scale, Aethoxysklerol was shown to be as effective as Sotradecol. (Note this violates the adage: "Absence of proof is not proof of absence.")

4. No details of the randomization were provided.

5. Due to apparent difficulties in data collection, blinding, etc., cited by the Sponsor, the Michigan center was closed early. The Sponsor's comments do seem to challenge the quality of the data from this center. Further, the FDA DSI investigation has found problems with the Ohio study.

6. Note the concentration of each drug depended upon the target vein size. That is, the actual treatment was different for the different vein sizes. The protocol analysis treated treatment as crossed with vein size. However, this reviewer prefers to treat treatment effects as nested within vein sizes. Further details of the current analyses are given in section 3.1.
.

On February 26, 2004, prior to any FDA analysis of the data, at an internal clinical /statistics meeting the Clinical Team determined that the primary analysis would be a superiority comparison of the dichotomized complete disappearance of varicosities. That determination was used to guide the analyses presented here.

At the April 15, 1996, guidance meeting the Division Director recommended that for comparison purposes Sotradecol should be considered as a placebo and tests should be cast as tests of superiority. However, as was noted above, at several later meetings the Sponsor was encouraged to also analyze these as non-inferiority trials. On March 25, 1997, in a memo to the Sponsor's statistician from the FDA statistician, the Sponsor was told that bioequivalence could be tested by the following procedure:
1) Calculate the 95% CI for the difference test mean - active control mean, giving bounds (LL,UL).
2) LL should not be more than 20% worse than the mean for the active control drug.
3) The test and active control are determined to be equivalent if the 95% CI includes 0 and the lower limit is not less than -0.2 times the active control mean.

Such an analysis was done for the primary dichotomous endpoint, and original 5 point disappearance of variscosities scale. Details are given in Appendix 1.

Findings

A post hoc adjustment separated the results from the three single center studies into two nominal studies, one with the California and Michigan centers, the other based on the Ohio center. The primary endpoint was success on complete disappearance of the varicosities by week 16 after treatment, as determined by a panel of three vascular surgeons. Again, the Medical team specified that Aethoxysklerol should be required to show superiority over Sotradecol (Sodium Tetradecyl Sulfate). The observed success rates in each center were as follows:

Table 2. Complete Disappearance of Varicosities Rates

Vein size	MICA California		MICA Michigan		Ohio	
	Sotradecol	Aethoxy-sklerol	Sotradecol	Aethoxy-sklerol	Sotradecol	Aethoxy-Sklerol
Overall	23.2 %	31.7%	16.7%	20.0%	17.4%	24.7%
< 1 mm	17.4 %	36.8%	11.1%	14.3%	23.5%	16.0%
1 mm - 3 mm	22.7 %	21.1%	33.3%	12.5%	13.0%	26.1%
3 mm - 6 mm	(b) (4)					

As noted above, even without an adjustment for the multiplicity of tests, in both nominal studies, no overall or within vein size comparisons between treatment were statistically significant (all p≥ 0.1758). These results were supported by a preliminary Bayesian analysis (see Appendix 6).

From Appendix 1, in both studies we see that noninferiority was not shown for the primary endpoint.

2. INTRODUCTION

2.1 Overview

According to the Sponsor, Aethoxysklerol is the most extensively used drug worldwide for the treatment of varicose veins. Although it is currently marketed in six concentrations (0.25%, 0.5%, 1%, 2%, 3%, and 4%), the Sponsor indicates that concentrations of 0.5%, 1.0%, and 2.0% are indicated for the treatment of varicose veins of the legs. The Sponsor submitted the summaries or published reports of a number of small or open label results, plus the results from two (originally one pooled) pivotal trials, as given below:

Table 3. Referenced Studies

Study	Location	# centers	Start date - Finish date	Description	Aethoxysklerol/ # subj entered	Sotradecol / # subj entered
Controlled Phase 3 Studies						
MICA	US	2	3 March 93 - 19 Feb. 96	Randomized, Double-blind, Active controlled	0.5% / 29 1.0% / 31 3.0% / 27	0.25%/ 33 0.5% / 32 1.5% / 27
OHIO	US	1	6 Jan. 93 - 26 July 95	Randomized, Double-blind, Active controlled	0.5% / 25 1.0% / 25 3.0% / 25	0.25%/ 25 0.5% / 25 1.5% / 25

Table 3. (cont.) Referenced Studies

Study	Location	# centers	Start date - Finish date	Description	Aethoxysklerol/ # subj entered
Open-label Studies					
ASK 94-002	Japan	1	3 Feb. 95 - 11 Jan. 96	Randomized, Open-label, Phase 2	0.5% / 22 1.0% / 50 2.0% / 63 3.0% / 26
ASK 96-001	Japan	1	27 Aug. 96 - 17 May 97	Randomized, Open-label, Phase 2	0.25%/ 20 0.5% / 51 1.0% / 29
AET-AS25	Germany	1	Dec 1995 Completed	Open-label with different formulation	2.7 mL / 40 1.7 mL / 39
AET-P2	US	1	April 1991 Completed	Open-label, Phase 1	2 mL / 10

In addition, several published reports were included.
In this review only the two active controlled Phase 3 studies are reviewed. Note that the regulatory history of the pivotal studies is summarized below.

2.1.1 Regulatory History

IND 35,139 for Aethoxysklerol® (0.5%, 1%,) for the treatment of varicose veins was submitted to the Division of Medical Imaging, Surgical and Dental Drug Products on July 2, 1990. The Sponsor indicates that a protocol for phase 3 studies to evaluate the efficacy of a large clinical trial was submitted to this Division in December, 1992. According to the Sponsor's Subsection 8B, by January 1994 this had been assigned to the Division of Dermatology. Later it was returned to the Division of Medical Imaging and was again returned to the Division of Dermatology in August 1995.

Note that in the review dated December 15, 1994, the Sponsor was warned that since three different dosages were being used the analysis would require a correction for multiplicity. This was confirmed in a review dated August 29, 1994. In this review the Sponsor was also recommended to consider ANOVA instead of several t-tests as originally proposed in the protocol.

From the FDA Medical Officer's review dated December 30, 1993, the original protocol(s) for this study were submitted to the Division on February 22, 1993. It is not clear, but it seems likely, that this was the same protocol the Sponsor indicates was submitted to the Agency in December, 1992. The original proposal was powered for three single center studies in Michigan, California, and Ohio. The Sponsor requested to pool the Michigan and California (MICA) studies. The re-labeled MICA study was initiated on March 3, 1993 and completed by February 19, 1996. The Ohio study was started on January 6, 1993 and completed by July 26,

1995. Due to concerns about the recruitment and disorganized data collection, the Sponsor closed down of the Michigan center early.

On July 1995, a revised protocol for the NDA was submitted. Note that this was near the completion of the studies. The study was originally powered to test for treatment differences in a score defined by the sum of a 5-point disappearance of varicose veins scale, plus a 4 point pigmentation sale, and a 3 point neovascularization scale. The 1-5 disappearance scale was coded with 1=worse than before treatment, 2=same as before, 3=the minority disappeared, 4=the majority disappeared, 5=complete disappearance. That is, level 5 was labeled as "Complete disappearance." In the July 1995 protocol the primary endpoint was specified as clinical improvement (see section 3.1). In a letter dated April 1, 1996, the Agency told the Sponsor that only "Disappearance of Varicose Veins" would be recognized as the primary efficacy endpoint. That letter also indicated refinements to several of the Sponsor's rating scales. A pre-NDA meeting was held on April 15, 1996. According to the FDA minutes of this meeting the Sponsor agreed to Disappearance of Varicose Veins as the primary endpoint. On January 29, 1997, after completion of the studies, the Sponsor submitted a revised protocol for the completed studies using this endpoint. For the analysis, in each study the results over all vein sizes were to be combined, and then stratified by vein size.

At a later guidance meeting on September 23, 1998, the FDA recommended a dichotomized version of the disappearance of varicosities as the sole primary efficacy variable. There is a reference from an e-mail (Jan. 31, 1996) to a meeting (Dec. 4, 1995) where complete disappearance (presumably dichotomized disappearance) was specified as the primary endpoint. However, this reviewer has not found any other documentation of this specification. It should be noted that Project management sent an e-mail to the Sponsor on January 31, 1996, indicating that since there had been no formal end of phase 2 meeting, there were no commitments from the Agency.

According to the Sponsor's minutes, at the meeting on April 15, 1996 the Sponsor was also told that the proposed concentrations of the comparator drug Sotradecol were off label, particularly for vein sizes below 3mm. Hence, at this meeting the Division Director recommended that for comparison purposes Sotradecol should be considered as a placebo and tests should be cast as tests of superiority. The Sponsor did respond with an argument that in practice Sotradecol was used at these lower concentrations. However, from the Sponsor's minutes, the Sponsor was informed that unless the manufacturer of Sotradecol would change the current labeling, the Sponsor would be required to show superiority of Aethoxysklerol for these vein sizes, i.e. (from FDA minutes) Sotradecol would be treated as a placebo. However, there was some discrepancy between the Medical Officer's comments and the Division Directors comments.

At the April 15, 1998 guidance meeting, the trial was described as a non-inferiority /equivalence trial. At the June 15, 1998 guidance meeting the Sponsor was told that "If the test drug is claimed to be as effective as the comparator, this claim should be tested using a 95% confidence interval for both the mean score and the proportion of patients with complete

disappearance." Furthermore, at the September 23, 1998, guidance meeting the Sponsor was told again that this was a non-inferiority/equivalence trial and that if superiority was demonstrated then no adjustment was needed for multiple comparisons. Although this likely referred to testing non-inferiority followed by superiority, that was not made explicit in the comments.

On March 25, 1997, in a memo to the Sponsor's statistician from the FDA statistician, the Sponsor was told that bioequivalence could be tested by the following procedure:
1) Calculate the 95% CI for the difference test mean - active control mean, giving bounds (LL,UL).
2) LL should not be more than 20% worse than the mean for the active control drug.
3) The test and active control are determined to be equivalent if the 95% CI includes 0 and the lower limit is not less than -0.2 times the active control mean.

On November 29, 1999, the Agency recommended that the Sponsor withdraw the submission to avoid a refuse-to-file decision, based on concerns about human pharmacokinetic data. On October 21, 2002, there was a further guidance meeting on outstanding issues related to the re-submission of this NDA. Only CMC issues were discussed in detail. The revised submission was received on October 2, 2003.

On February 26, 2004, prior to any detailed FDA analysis of the data, at an internal clinical/statistics meeting the Clinical Team determined that the primary analysis would be a superiority comparison of the dichotomized complete disappearance of varicosities.

2.2 Data Sources

The Sponsor submitted a compact disk with SAS data sets for general demographic data and the final endpoint variables in each of the Ohio and MICA studies. Upon request the Sponsor provided the randomization list for the California study.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

The following was the title of the protocol of the Sponsor's pivotal trials:

TITLE: Double Blind, Prospective, Randomize Comparative, Multicenter Trial between Aethoxysklerol (Polidoconal) and Sotradecol (Sodium Tetradecyl Sulphate) in the Management of Varicose Veins of the Lower Extremities.

The objective of these studies was to compare the efficacy and safety of the sclerosing agent, Aethysklerol, to Sotradecol, at that time an approved sclerosing drug, in the treatment of varicose veins in the lower extremities.

Originally this study was to have been done in three centers, one center in each of Michigan, California and Ohio. The California and Michigan centers were combined in the so-called MICA study. The Ohio center was treated as a separate study. Due to apparent difficulties in data collection, blinding, etc., cited by the Sponsor, the Michigan center was closed early.

The protocols for the study specified three different dosages depending upon vein sizes:
1) Veins under 1 mm were to receive 0.5% Aethoxysklerol or 0.25% Sotradecol.
2) Veins 1 - 3 mm were to receive 1.0% Aethoxysklerol or 0.5% Sotradecol.
3) Veins 3 - 6 mm were to receive 3.0% Aethoxysklerol or 1.5% Sotradecol.

Note that no details of the randomization were supplied, except that a revised randomization list was sent to California. Treatment assignment was from a randomized list within each center, but no details were provided. In the Ohio study, all six drug by vein size combinations were equal sized, which suggests a restriction on the randomization.

The final version of the protocols indicates that the primary efficacy variable is the "disappearance of varicosities." This was to be independently judged and scored by three vascular surgeons based upon comparison of a set of pre-injection photographs with a set taken 16 weeks after the last treatment. According to the protocol, the extent of disappearances in the target vein was to be evaluated by each vascular surgeon on a five point scale:

1 = worse than before treatment
2 = same as before
3 = the minority disappeared
4 = the majority disappeared
5 = complete disappearance of varicosities

A secondary endpoint was the "overall clinical improvement." This was an assessment by the same vascular surgeons of the vein disappearance, pigmentation, and neovascularization seen in the final set of photographs. The degree of improvement is supposed to be measured on an 11 point scale:

0 = no improvement or worse
0-2 = poor
2-4 = fair
4-6 = moderate
6-8 = good
8-10 = excellent
10 = perfect cosmetic result

Note that the data set provided for analysis includes only the mean scores of the vascular surgeons for these endpoints. But that was judged to be sufficient for this report.

A final secondary endpoint was the overall patient satisfaction, on a four point scale:

1 = unsatisfied
2 = moderately satisfied
3 = satisfied
4 = very satisfied

As noted above, on February 26, 2004, prior to any FDA analysis of the data, at an internal clinical/statistics meeting the Clinical Team determined that the primary analysis would be a superiority comparison of the dichotomized complete disappearance of varicosities. This

analysis, with supporting analyses of the 5-point disappearance scale, is presented in sections 3.1.1-3.1.3 below.

Demographic information on patients is given in Appendix 5.

3.1.1 MICA Study

Superiority Comparisons for Primary Endpoints

As was discussed in section 2.1, Overview, the primary endpoint specified by the Medical team is a dichotomization of the Sponsor's Disappearance of Varicosities scale. This was to be tested using a superiority comparison. Note that the dichotomized complete disappearance response (i.e., "success") was defined as all three evaluators giving an assessment of 5 (Complete Disappearance) on the Disappearance of Varicosities scale. Otherwise it was not scored as complete disappearance (i.e., as a "failure").

Table 4. Complete Disappearance of Varicosities (Relative Success Rate & Percentage)

Vein sizes	California		Michigan		
	Sotradecol	Aethoxysklerol	Sotradecol	Aethoxysklerol	p-values
Overall	16/69 (23.2%)	19/60 (31.7%)	3/18 (16.7%)	4/20 (20.0%)	0.3127[1]
< 1 mm	4/23 (17.4%)	7/19 (36.8%)	1/9 (11.1%)	1/7 (14.3%)	0.1758[2]
1 mm - 3 mm	5/22 (22.7%)	4/19 (21.1%)	2/6 (33.3%)	1/8 (12.5%)	0.5688[2]
3 mm - 6 mm	[redacted]	[redacted]	[redacted]	[redacted]	(b) (4)

[1] CMH test stratified on center x vein size
[2] CMH test stratified on center

Thus, either overall or for each vein size individually, even without adjusting for the multiplicity of tests, there is no statistically significant difference between treatment groups ($p \leq 0.3127$ and all $p \geq 0.1758$ respectively). Adjusting for multiplicity over all four comparisons using Holm's method leads to a minimum significance level of 0.7032.

At some points in the history of this submission, an analysis using an ANOVA of the mean of the 5-point varicosities scale was recommended. This could be accomplished with the following model:

Expected Score = $\mu + vein_i + center_j + vein*center_{ij} + treat_k(vein_i) + center*treat_{jk}(vein_i)$,
for i=1,2,3, j=1,2, and k=1,2.

Table 5. Mean Disappearance on 5-point Varicosities Scale - Mean (Std. Dev.)

Vein sizes	California		Michigan		
	Sotradecol	Aethoxysklerol	Sotradecol	Aethoxysklerol	p-values
Overall	4.4 (0.6)	4.5 (0.5)	4.0 (1.0)	4.4 (0.5)	0.2698[1]
< 1 mm	4.3 (0.5)	4.6 (0.4)	3.8 (1.2)	4.2 (0.6)	0.1793[2]
1 mm - 3 mm	4.3 (0.8)	4.3 (0.7)	4.2 (0.9)	4.4 (0.5)	0.1892[2]
3 mm - 6 mm	(b) (4)				

[1] Overall test of treatment differences (H_0: difference=0)
[2] Contrast on within vein size treatment differences

Thus, again, either overall or for each vein size individually, there is no statistically significant difference between treatment groups ($p \leq 0.2698$ overall and over vein sizes all $p \geq 0.1793$). Appropriate adjustment for multiplicity would these significance levels even further.

The original protocol specified a randomized complete block design to compare the mean results across the centers for each category of vein size. That is, within each vein size:

Expected Score = $\mu + center_j + treat_k + center*treat_{jk}$, for j=1,2 and k=1,2.

Within a given center a two sample t-test is to be used in each of the three categories of vein sizes.

Table 6. Mean Disappearance on 5-point Varicosities Scale (Protocol Analysis)

Vein sizes	California			Michigan			Overall
	Sotradecol	Aethoxy-sklerol	p-values[2]	Sotradecol	Aethoxy-sklerol	p-values[2]	p-values[1]
< 1 mm	4.3 (0.5)	4.6 (0.4)	0.0646	3.8 (1.2)	4.2 (0.6)	0.4104	0.0697
1 mm - 3 mm	4.3 (0.8)	4.3 (0.7)	0.8058	4.2 (0.9)	4.4 (0.5)	0.5464	0.9327
3 mm - 6 mm	(b) (4)						

[1] Tests of treatment differences within each vein size using randomized block design above.
[2] Paired Sample t-test within vein size

Thus, within each vein size no differences were statistically significant (all $p \geq 0.0646$). Note that with an adjustment for multiplicity (Holm's method) the significance level of the most statistically significant comparison would be 0.2091.

3.1.2 Ohio Study

Superiority Comparisons for Primary Endpoints

For the primary endpoint, the dichotomized endpoint based on the mean disappearance on the 5-point varicosities scale, we get the table of success rates:

Table 7. Complete Disappearance of Varicosities

Vein sizes	Ohio		
	Sotradecol	Aethoxysklerol	p-values
Overall	12/69 (17.4%)	18/73 (24.7%)	0.2908[1]
< 1 mm	5/21 (23.5%)	4/25 (16.0%)	0.5060[2]
1 mm - 3 mm	3/23 (13.0%)	6/23 (26.1%)	0.2648[2]
3 mm - 6 mm			(b) (4)

[1] MH test stratified on vein size.
[2] Chi-Square test.

Thus the overall success rate was 17.4% on Sotradecol and 24.7% on Aethysklerol. However, either overall or for each vein size individually, even without an adjustment for multiplicity, there is no statistically significant difference between treatment groups ($p \leq 0.2908$ and all $p \geq 0.1853$ respectively).

As noted above, at several meetings, the Sponsor was encouraged to analyze the mean disappearance scale. This could be done using a GLM nested Model:

Expected Score = $\mu + vein_i + treat_k(vein_i)$ for i=1,2,3 and k=1,2

Table 8. Mean Disappearance on 5-point Varicosities Scale- Mean (Std. Dev.)

Vein sizes	Ohio		
	Sotradecol	Aethoxysklerol	p-values
Overall	4.2 (0.7)	4.2 (0.8)	0.1203[1]
< 1 mm	4.3 (0.5)	4.0 (0.8)	0.1039[2]
1 mm - 3 mm	4.0 (0.8)	4.3 (0.9)	0.1909[2]
3 mm - 6 mm			(b) (4)

[1] Overall test of treatment differences.
[2] Contrast on within vein size treatment differences.

Thus again, either overall or for each vein size individually, there is no statistically significant difference between treatment groups ($p \leq 0.1203$ and all $p \geq 0.1039$ respectively).

Again, the original protocol specified that within each center a two sample t-test is to be used to compare treatments in each of the three categories of vein sizes. Using this analysis we get the following:

Table 9. Mean Disappearance on 5-point Varicosities Scale (Protocol analysis)

Vein sizes	Ohio		
	Sotradecol	Aethoxysklerol	p-values[1]
< 1 mm	4.3 (0.5)	4.0 (0.8)	0.1055
1 mm - 3 mm	4.0 (0.8)	4.3 (0.9)	0.2818
3 mm - 6 mm	(b) (4)		

[1] Paired sample t-test within vein size.

Thus, using the two-sample t-test within each vein size, there is no statistically significant difference between treatment groups (all p≥ 0.1055 respectively). As above, an appropriate adjustment for multiplicity would only exaggerate the non-significance.

3.1.3. Pooled MICA and Ohio Studies

Originally the MICA and Ohio Studies were merely separate subsets of centers in a single study. The separation into two studies was a post-hoc adjustment, supposedly for replication. The p-values presented below correspond to the results of CMH tests on the dichotomized endpoint (which was specified as the primary endpoint by the Medical team), and two additional ANOVA analyses treating the 5-point ordinal scale as a continuous variable. The actual proportions and mean scores being compared in the tests are given in the tables above. Note that the analysis specified in the protocol was to have been an ANOVA with randomized blocks stratified on vein size.

Table 10. P-values for Complete Disappearance of Varicosities in Pooled Study

Vein sizes	Dichotomized Endpoint	Mean Disappearance (5 point scale)	
	CMH tests	Nested Model[3]	Randomized Block[6]
Overall	0.1443[1]	0.4395[4]	
< 1 mm	0.5468[2]	0.4404[5]	0.8892
1 mm - 3 mm	0.7658[2]	0.3228[5]	0.3884
3 mm - 6 mm	(b) (4)		

[1] CMH test stratified on center x vein size.
[2] CMH test stratified on center.
[3] Nested Model: Expected Score = $\mu + vein_i + center_j + vein*center_{ij} + treat_k(vein_i) + center*treat_{jk}(vein_i)$, for i=1,2,3, j=1,2,3, and k=1,2.
[4] Overall test of treatment differences (H_0: difference=0)
[5] Contrast on within vein size treatment differences.
[6] Randomized Block: Within each vein size, Expected Score = $\mu + center_j + treat_k + center*treat_{jk}$, for j=1,2 and k=1,2.

Even without adjusting for the multiplicity of measures, none of the three analyses of the disappearance of varicosities show statistically significant differences (all p≥ 0.0976 unadjusted). Note the Sponsor's actual analysis seemed to be consistent with results above. However, the Sponsor concluded that these results show that Aethoxysklerol is as good or better than Sotradecol.

3.2 Evaluation of Safety

Please see Medical Officer's safety review.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race and Age

Note that no race data was included and that almost all subjects were female, so only a breakdown by age is appropriate:

Table 11. Complete Disappearance of Varicosities By Age Group

		California		Ohio		Michigan	
Vein Size	Age Group	Sotradecol	Aethoxy-sklerol	Sotradecol	Aethoxy-sklerol	Sotradecol	Aethoxy-sklerol
Overall	21-35	2/11 (18%)	8/18 (44%)	5/32 (16%)	6/15 (40 %)	2/6 (33%)	3/9 (33%)
	36-50	14/42(33%)	8/23 (35%)	4/25 (16%)	10/42 (24%)	0/9	0/10
	51-65	0/16	3/19 (16%)	3/12 (25%)	2/16(12.5%)	1/3 (33%)	1/1 (100%)
< 1 mm	24-35	1/4 (25%)	1/5 (20%)	3/10 (30%)	1/3 (33%)	1/4 (25%)	1/4 (25%)
	37-50	3/13 (23%)	5/9 (56%)	2/8 (25%)	2/15 (13%)	0/4	0/3
	51-65	0/6	1/5 (20%)	0/3	1/7 (14%)	0/1	.
1-3 mm	21-35	1/3 (33%)	3/4 (75%)	1/9 (11%)	3/8 (37.5%)	1/1(100%)	1/3 (33%)
	36-50	4/14 (29%)	0/5	1/9 (11%)	3/13 (23%)	0/3	0/5
	51-65	0/5	1/10 (10%)	1/5 (20%)	0/2	1/2 (50%)	.
3-6 mm	24-35	0/4	4/9 (44%)	1/13 (8%)	2/4 (50%)	0/1	1/2 (50%)
	36-50	7/15 (47%)	3/9 (33%)	1/8 (12.5%)	5/14 (36%)	0/2	0/2
	51-64	0/5	1/4 (25%)	2/4 (50 %)	1/7 (14%)	.	1/1 (100%)

With the possible exception of the youngest patients in the California and Ohio studies, success rates in disappearance of varicosities are quite uniform across the treatment groups for each age group in each study. The protocol specified that each principal investigator was to use the technique for sclerotherapy of their choice. The apparently somewhat discrepant results across centers may reflect such differences in treatment. Of course they could also be just an artifactual result of the study.

When further broken down by vein size, the same pattern seems to hold, but only for veins 1mm or greater. Of course here, again, the study was not powered to detect this type of comparison, so this just an observation, and is especially prone to perturbation due to the small cell size in the vein x age group x center x treatment table.

4.2 Other Special/Subgroup Populations

For results on marginal vein subgroups see the efficacy tables above.

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

Issues

The actual method of sclerotherapy was left to the investigator, thus one would expect investigator differences. However, the use of only three investigators does not provide much information on these potential between method effects. Dosing is described as being repeated once a week for up to 6 or more weeks until clearance, though apparently only three treatments were the maximum. This makes it difficult to determine the actual dosage used.

From a statistical point of view the primary difficulty with this submission comes from the mostly post hoc, sometimes not necessarily consistent recommendations for analysis. For example:

1. The decision to split the single three center study into two separate studies would not be consistent with current practice. The original powering of the study was for three separate single center studies. However, the original study was powered for a different endpoint anyway.

2. In the third and last protocol, submitted after completion of the studies, the Sponsor proposed a 5-point scale, labeled "Disappearance of Varicosities", as the primary endpoint. At various meetings the Sponsor was apparently reminded that this should be analyzed as a dichotomous variable, but at other meetings the 5-point scale seems to have been implicitly accepted.

3. The Sponsor was also informed that since Sotradecol was to be used at below labeled dose, at least for vein sizes below 3mm, it was to be treated as a placebo, and superiority comparisons should be used. However, at several other times the Sponsor also was told that these should be treated as noninferiority comparisons, and could be tested using the procedure described in Appendix 1.

4. No details of the randomization were provided.

5. Due to apparent difficulties in data collection, blinding, etc., cited by the Sponsor, the Michigan center was closed early. The Sponsor's comments do seem to challenge the quality of the data from this center. Furthermore, the FDA DSI investigation has found problems with the California study. Whether they are substantial or not is not yet apparent.

6. Because of the change in concentrations of the treatment drugs for different vein sizes, this reviewer treated treatment effects as nested within vein sizes. The analyses recommended previously treated these as crossed effects. Further details of the analyses used are in sections 3.1-3.3 above.

.

On February 26, 2004, prior to any FDA analysis of the data, at an internal clinical/ statistics meeting the Clinical Team determined that the primary analysis would be a superiority comparison of the dichotomized complete disappearance of varicosities. That determination was used to guide the analyses presented here.

As noted earlier, at the April 15, 1996 guidance meeting the Division Director recommended that the comparisons to Sotradecol should be conducted as considered as tests of superiority. However, at several later meetings the Sponsor was encouraged to also analyze these as non-inferiority trials. Such an analysis was done for the primary dichotomous endpoint, and for the original 5-point disappearance of variscosities scale. Details are given in Appendix 1.

Findings

A post hoc adjustment separated the results from the three centers into two nominal studies, one with the California and Michigan centers, the other based on the single Ohio center. The primary endpoint was success on complete disappearance of the varicosities by week 16 after treatment, as determined by a panel of three vascular surgeons. Again, the Medical team specified that Aethoxysklerol should be required to show superiority over Sotradecol (Sodium Tetradecyl Sulfate). For the observed success rates in each center see Table 2, page 5. As noted there, even without an adjustment for the multiplicity of tests, in both nominal studies no overall or within vein size comparisons between treatment were statistically significant (over all studies and vein sizes, $p \geq 0.1758$).

5.2 Conclusions and Recommendations

Complete disappearance of varicosities was specified as the primary endpoint. Although the comparator drug was active, because the dose levels of the comparator were below the labeled doses, the Medical team specified these should be superiority comparisons. Even without an adjustment for the multiplicity of tests, in both nominal studies, no overall or within vein size comparisons of the dichotomized complete disappearance were statistically significant (all $p \geq 0.1758$). Further, in neither study did any of these comparisons of this dichotomous endpoint meet the requirements to show non-inferiority (see Appendix 1). This conclusion was supported by a preliminary Bayesian analysis (see Appendix 6).

Note the Sponsor's actual analysis seemed to be consistent with results above. However, the Sponsor concluded that these results show that Aethoxysklerol is as good or better than Sotradecol. From Appendix 1, in both studies we see that non-inferiority was not shown for the primary endpoint.

APPENDICES

Appendix 1.0 Primary Endpoints: Non-inferiority Analyses

In the March 25, 1997, memo to the Sponsor's statistician from the FDA statistician, the Sponsor was told that non-inferiority could be tested by the following procedure:
1) Calculate the 95% CI for the difference test mean - active control mean, giving bounds (LL,UL).
2) LL should not be more than 20% worse than the mean for the active control drug.
3) The test and active control are determined to be equivalent if the 95% CI includes 0 and the lower limit is not less than -0.2 times the active control mean.

One approach to assess non-inferiority of Complete Disappearance is to compute a 95% confidence interval about the test proportion - active control proportion. Ignoring the stratification variables and using simple binomial approximations, one can compute the intervals given in the "Simple/(LU,UU)" column below. The column labeled 20% bound is .20 x Sotradecol proportion. If the lower bound of the confidence interval is greater than the 20% bound, using the procedure above, one can infer non-inferiority. An alternative to the simple difference in proportions is to use the weighted combination used in the calculation of the Mantel-Haenszel statistic. As an aside, note that this is the same contrast used in the computation of SAS Type 2 sums of squares and adjusts for unequal numbers of observations per stratum. Using a product binomial term with a continuity correction to compute the variance of this weighted difference gives the confidence intervals labeled "Weighted/(LU,UU)".

Table A.1.1 Simple comparisons on Success proportions

Study	Sotradecol # succ/N	Aethoxy. # succ/N	Simple (LU,UU)	Weighted Difference	Weighted (LU,UU)	20% bound
MICA Overall	19/87	23/80	(-0.06,0.20)	0.07	(-0.06,0.19)	-0.04
MICA <1mm	5/32	8/26	(-0.07,0.37)	0.13	(-0.08,0.34)	-0.03
MICA 1-3mm	7/28	5/27	(-0.28,0.15)	-0.05	(-0.26,0.17)	-0.05
MICA 3-6mm	(b)					
Ohio overall	12/69	18/73	(-0.06,0.21)	0.08	(-0.05,0.21)	-0.03
Ohio <1mm	5/21	4/25	(-0.31,0.15)	-0.08	(-0.31,0.15)	-0.05
Ohio 1-3mm	3/23	6/23	(-0.10,0.36)	0.13	(-0.10,0.36)	-0.03
Ohio 3-6mm	(b)					

Since none of the confidence intervals are completely above the lower bound, we cannot conclude that Aethoxysklerol has been shown to be non-inferior to Sotradecol.

A similar analysis can be used for the 5-point Disappearance of Varicosities scale. In the tables below the columns labeled "Diff" are the difference in means scores on this Disappearance of Varicosities scale (computed as Aethoxysklerol - Sotradecol). The corresponding confidence

intervals given in Table A.1.2 below are based on simple normal theory assumptions. Note that these are computed ignoring the other factors in the model.

The column labeled 20% bounds are .20 x Stradecol mean. Recall that each mean disappearance score was the mean of three scores giving the extent of disappearances in the target vein as assessed by a vascular surgeon on a five point scale scored as 1=worse, 2=same, up to 5=complete disappearance. To assess a percent improvement, one should adjust the scale so that 0 corresponds to no improvement. That is, in this particular case one should compare the observed confidence limit to a bound computed as .2 x (Sotradecol mean - 2), or just a .4 reduction in the original definition of the bound. Two bounds are provided below. The column labeled "Adjusted 20% bound" has the .4 reduction, and is the recommended bound. The column labeled "Original 20% bound" does not have the .4 reduction, and seems to correspond to the bound recommended by the Agency statistician in 1997. Note that the choice of bound actually has impact upon final conclusions.

Using the procedure given above, noninferiority is inferred if the two sided 95% confidence interval, denoted (LU,UU), is greater than the appropriate 20% bound.

Table A.1.2 Simple comparisons on Mean Scores

Study	Diff (SE)	Sotr Mean	(LU,UU)	Adjusted 20% bound	Original 20% bound
MICA Overall	0.14 (0.09)	4.32	(-0.04, 0.32)*	-0.46	-0.86
MICA <1mm	0.31 (0.16)	4.20	(-0.01, 0.63)*	-0.44	-0.84
MICA 1-3mm	0.02 (0.19)	4.29	(-0.35, 0.38)*	-0.46	-0.86
MICA 3-6mm	(b) (4)				
Ohio Overall	0.06 (0.12)	4.19	(-0.18, 0.30)*	-0.44	-0.84
Ohio <1mm	-0.34 (0.20)	4.30	(-0.74, 0.05)*	-0.46	-0.86
Ohio 1-3mm	0.28 (0.25)	4.00	(-0.22, 0.77)*	-0.40	-0.80
Ohio 3-6mm	(b) (4)				

* Achieve lower bound, i.e., conclude non-inferiority.

In this case, in each comparison, the Sponsor achieves either the original or the more stringent adjusted non-inferiority bound.

Note the intervals above ignore all other factors in the design. The following intervals are computed as parameter contrasts in more complicated models and are adjusted for other factors:

Table A.1.3 Contrast comparisons on Mean Scores

Study	Diff (SE)	Sotr Mean	(LU,UU)	Adjusted 20% bound	Original 20% bound
MICA Overall	0.24 (0.12)	4.32	(0.01, 0.47)[1]*	-0.46	-0.86
MICA <1mm	0.34 (0.18)	4.20	(-0.01, 0.69)[2]*	-0.44	-0.84
MICA 1-3mm	0.09 (0.19)	4.29	(-0.28, 0.46)[2]*	-0.46	-0.86
MICA 3-6mm			(b)		
Ohio Overall	0.06 (0.12)	4.19	(-0.18, 0.30)[1]*	-0.44	-0.84
Ohio <1mm	-0.34 (0.21)	4.30	(-0.76, 0.07)[2]*	-0.46	-0.86
Ohio 1-3mm	0.28 (0.21)	4.00	(-0.14, 0.69)[2]*	-0.40	-0.80
Ohio 3-6mm			(b)		

* Achieve lower bound, i.e., conclude non-inferiority.
[1] From a full factorial models.
[2] From nested models.

Again, for each comparison, the Sponsor achieves the non-inferiority bounds.

Appendix 2.0 Secondary Endpoints: Conclusions Adjusted for Multiplicity

The secondary endpoints are the investigator's level of improvement and the patient's satisfaction with results. Note that in each study, the original protocol specified an analysis stratified within vein size. To control family-wise Type I error Holm's Step-down method is used. For this test, p-values are sorted by increasing size. For k comparisons at level α, the smallest observed p-value is compared to α/k. If it is significant, compare the next smallest to $\alpha/(k-1)$, then the next p-value to $\alpha/(k-2)$, etc., until the last is compared to α. Stop at the first non-significant comparison and declare all remaining comparisons statistically non-significant. For the six comparisons (three vein sizes for clinical improvement and three for patient satisfaction) at a nominal 0.05 level, one would first compare the smallest p-value to 0.05/6 = 0.0083.

For the Ohio and MICA studies we find the following table of observed significance levels:

Table A.2.1 Mean Clinical Improvement on 10-point Scale: MICA Study

		MICA Study		Ohio Study	
Variable	Vein Size	Holm's Bound	Observed p-value	Holm's Bound	Observed p-value
Clin. Improve.	< 1 mm	0.0083	0.0020	0.0250	0.7622
	1 mm - 3 mm	0.0100	0.2174	0.0100	0.0216
	3 mm - 6 mm	(b) (4)			
Satisfaction	< 1 mm	0.0250	0.3851	0.0500	0.8973
	1 mm - 3 mm	0.0500	0.7193	0.0083	0.0205
	3 mm - 6 mm	(b) (4)			

Thus in the MICA study, the smallest observed p-value corresponds to the test of differences in clinical improvement between treatments in veins < 1mm. Since $0.0020 \leq 0.0083$ this difference is statistically significant. The next smallest observed p-value in the MICA study was for the test of disappearances in clinical improvement in the 1 mm - 3 mm vein group, i.e., 0.2174. Since 0.2174 is greater than its corresponding bound (.05/5 = 0.0100) we conclude that this difference is not statistically significant, and hence no other comparisons in the MICA Study are statistically significant. In the Ohio Study the smallest p-value is 0.0205, and this does not fall below its bound (.05/6 = 0.0083). Hence, after adjusting for multiplicity, no differences in the Ohio Study are statistically significant.

Appendix 3.0 Secondary Endpoints: Level of Improvement

The level of improvement was scored on a 0-10 scale. The grouped data distribution is as follows:

Table A.2.1 Level of Improvement

	California				Ohio				Michigan			
	Sotr		Aeth		Sotr		Aeth		Sotr		Aeth	
Vein Size	n	%	n	%	n	%	n	%	n	%	n	%
0+-1mm 0-<2 Poor	1	4.3	.	.	1	4.8	1	4.0	1	11.1	.	.
2-<4 Fair	2	8.7	.	.	1	4.8	2	8.0	2	22.2	.	.
4-<6 Moderate	2	8.7	2	10.5	3	14.3	3	12.0	1	11.1	1	14.3
6-<8 Good	6	26.1	.	.	9	42.9	13	52.0	3	33.3	2	28.6
8-10 Excellent	12	52.2	17	89.5	7	33.3	6	24.0	2	22.2	4	57.1
1-3mm 0-<2 Poor	1	4.5	1	5.3	1	4.3	1	4.3	.	.	.	.
2-<4 Fair	1	4.5	3	15.8	4	17.4	2	8.7	1	16.7	1	12.5
4-<6 Moderate	4	18.2	2	10.5	7	30.4	3	13.0	1	16.7	.	.
6-<8 Good	6	27.3	3	15.8	8	34.8	6	26.1	3	50.0	1	12.5
8-10 Excellent	10	45.5	10	52.6	3	13.0	11	47.8	1	16.7	6	75.0
3-6mm 2-<4 Fair	.	.	(								.	.
4-<6 Moderate	5	(b)	(b)	(4)						.	.	.
6-<8 Good	5	(b)										
8-10 Excellent	14	(b)										
All 0-<2 Poor	2	2.9	1	1.7	2	2.9	2	2.7	1	5.6	.	.
2-<4 Fair	3	4.3	4	6.7	8	11.6	5	6.8	4	22.2	1	5.0
4-<6 Moderate	11	15.9	8	13.3	18	26.1	11	15.1	2	11.1	1	5.0
6-<8 Good	17	24.6	11	18.3	22	31.9	27	37.0	7	38.9	4	20.0
8-10 Excellent	36	52.2	36	60.0	19	27.5	28	38.4	4	22.2	14	70.0

The mean clinical improvement scores were analyzed as below:

Table A.2.2 Mean Clinical Improvement on 0-10 Scale: MICA Study

Treatment	California		Michigan		
	Sotradecol	Aethoxysklerol	Sotradecol	Aethoxysklerol	p-values
Overall	7.2 (2.1)	7.6 (2.2)	5.8 (2.6)	8.1 (1.6)	0.0181[1]
< 1 mm	7.2 (2.3)	8.7 (1.4)	5.4 (3.2)	7.9 (1.5)	0.0020[2]
1 mm - 3 mm	6.9 (2.3)	6.9 (2.7)	6.3 (2.0)	8.0 (1.9)	0.2174[2]
3 mm - 6 mm					(b) (4)

[1] Overall test of treatment differences (H_0: difference=0).
[2] Contrast on within vein size treatment differences.

Note this used the same GLM Model as for mean reduction in varicosities:
Score = μ + $vein_i$ + $center_j$ + $vein*center_{ij}$ + $treat_k(vein_i)$ + $center*treat_{jk}(vein_i)$, for i=1,2,3, j=1,2, and k=1,2.

Thus, in the overall comparison, there do seem to be statistically significant differences between treatments. This is due primarily to the differences in the smallest vein size. Even adjusting for

multiplicity (see Appendix 2), this difference is statistically significant. However, no other differences are statistically significant.

Table A.2.3 Mean Clinical Improvement on 10-point Scale: Ohio Study

Treatment	California		
	Sotradecol	Aethoxysklerol	p-values
Overall	6.3 (2.1)	7.0 (2.1)	0.0634[1]
< 1 mm	6.9 (2.2)	6.7 (2.1)	0.7622[2]
1 mm - 3 mm	5.6 (1.9)	7.0 (2.4)	0.0216[2]
3 mm - 6 mm	(b)	(b)	(b) (4)

[1] Overall test of treatment differences (H_0: difference=0).
[2] Contrast on within vein size treatment differences.

This used the similar GLM Model:
Score = μ + $vein_i$ + $treat_k(vein_i)$, for i=1,2,3 and k=1,2

In the Ohio Study, as shown in Appendix 2, after adjusting for multiplicity no differences were statistically significant.

Appendix 4.0 Secondary Endpoints: Overall Satisfaction

Table A.3.1 Overall Satisfaction

	California				Ohio				Michigan			
	Sotr		Aeth		Sotr		Aeth		Sotr		Aeth	
Vein Size	n	%	n	%	n	%	n	%	n	%	n	%
0+-1mm 1. Unsatis.	1	4.3	2	10.5	3	14.3	3	12.0	.	.	1	14.3
2. Mod. Satis.	6	26.1	8	42.1	4	19.0	6	24.0	1	12.5	1	14.3
3. Satisfied	6	26.1	3	15.8	5	23.8	4	16.0	2	25.0	2	28.6
4. Very Satis.	10	43.5	6	31.6	9	42.9	12	48.0	5	62.5	3	42.9
1-3mm 1. Unsatis.	1	4.5	.	.	3	13.0	.	.	.	.	.	.
2. Mod. Satis.	4	18.2	4	21.1	5	21.7	2	8.7	.	.	2	25.0
3. Satisfied	9	40.9	6	31.6	8	34.8	4	17.4	1	20.0	4	50.0
4. Very Satis.	8	36.4	9	47.4	7	30.4	17	73.9	4	80.0	2	25.0
3-6mm 1. Unsatis.	5	(b)	(									.
2. Mod. Satis.	5	(b)	(									
3. Satisfied	4	(b)	(b) (4)								.	.
3.5 Sat./Very Sat.	.	.	1	4.8	.	.	.	.	.	.	.	.
4. Very Satis.	10	(b)										
All 1. Unsatisfied	7	10.1	3	5.1	6	8.7	6	8.2	.	.	1	5.0
2. Mod. Satis.	15	21.7	14	23.7	11	15.9	10	13.7	2	12.5	5	25.0
3. Satisfied	19	27.5	12	20.3	23	33.3	14	19.2	4	25.0	6	30.0
3.5 Sat./Very Sat.	.	.	1	1.7	.	.	.	.	.	.	.	.
4. Very Satis.	28	40.6	29	49.2	29	42.0	43	58.9	10	62.5	8	40.0

Table A.3.2 Mean Satisfaction: MICA Study

Treatment	California		Michigan		
	Sotradecol	Aethoxysklerol	Sotradecol	Aethoxysklerol	p-values
Overall	3.0 (1.0)	3.2 (1.0)	3.5 (0.7)	3.1 (1.0)	0.8171
< 1 mm	3.1 (0.9)	2.7 (1.1)	3.5 (0.8)	3.0 (1.2)	0.3851
1 mm - 3 mm	3.1 (0.9)	3.3 (0.8)	3.8 (0.4)	3.0 (0.8)	0.7193
3 mm - 6 mm	(b) (4)				

[1] CMH test stratified on center x vein size.
[2] CMH test stratified on center.

In the MICA Study, even without an adjustment for multiplicity as in Appendix 2, no differences were statistically significant.

Table A.3.3 Mean Satisfaction: Ohio Study

Treatment	Ohio		
	Sotradecol	Aethoxysklerol	p-values
Overall	3.1 (1.0)	3.3 (1.0)	0.1916
< 1 mm	3.0 (1.1)	3.0 (1.1)	0.8973
1 mm - 3 mm	2.8 (1.0)	3.7 (0.6)	0.0205
3 mm - 6 mm	(b) (4)		

[1] MH test stratified on vein size.
[2] Chi-Square test stratified on center.

In the Ohio Study, as shown in Appendix 2, after adjusting for multiplicity no differences were statistically significant.

Appendix 5.0 Demographics

Tables A.5.1 and A.5.2 below summarize the provided patient demographic information in the two studies.

Table A.5.1 MICA Demographics

	California		Michigan		
	Sotradecol	Aethoxysklerol	Sotradecol	Aethoxysklerol	Total
Gender Male	1 (1.4%)	1 (1.6%)	0	1 (5.0%)	3 (1.7%)
Female	69 (98.6%)	62 (98.4%)	21 (100%)	19 (95.0 %)	171 (98.3%)
Age groups					
24 to 35 years	12 (17. 1%)	19 (30.2%)	7 (33.3%)	9 (45.0%)	47 (27. 0%)
36 to 50 years	42 (60.0%)	25 (39.9%)	11 (52.4%)	10 (50.0%)	88 (50.6%)
51 to 65 years	16 (22.9%)	19 (30.2%)	3 (14.3%)	1 (5.0%)	39 (22.4%)
Age Mean (SD)	44.1 (8.7)	43.4 (11.3)	41.0 (9.1)	38.6 (8.7)	42.8 (9.9%)

Table A.5.2 Ohio Demographics

	Sotradecol	Aethoxysklerol	Total
Gender Male	0	1 (1.3%)	1 (0.7%)
Female	75 (100 %)	74 (98.7%)	149 (99.3%)
Age groups			
21 to 35 years	33 (44. 0 %)	16 (21.3 %)	49 (32.7%)
36 to 50 years	29 (38.7 %)	43 (57.3 %)	72 (48.0%)
51 to 65 years	13 (17.3 %)	16 (21.3 %)	29 (19.3%)
Age Mean (SD)	38.8 (10.4)	43.1 (9.5)	40.1 (10.1)

In both studies almost all subjects were female. Age means ranged from 38.6 to 44.1, while the corresponding standard deviations ranged from 8.7 to 11.3, and both seem to be consistent across center by treatment combinations.

Appendix 6.0 Preliminary Bayesian Analysis of Primary Endpoint

The following analysis is post hoc, and hence, at best, can only be considered as supporting.

Recall that the studies were conducted under a single protocol with the same set of raters. So, for analysis it makes sense to pool the data from the two studies. The primary endpoint, success, i.e., a score of 5 on the disappearance of variscosities scale, was modeled as a logit of a linear predictor as follows:

$$\text{logit}(p_i) = \alpha_1 + \alpha_2*\text{Cntr1} + \alpha_3*\text{Cntr2} + \alpha_4*\text{Vein1} + \alpha_5*\text{Vein2} + \alpha_6*\text{Cntr1}*\text{Vein1} + \alpha_7*\text{Cntr1}*\text{Vein2} + \alpha_8*\text{Cntr1}*\text{Vein1} + \alpha_9*\text{Cntr2}*\text{Vein2} + \alpha_{10}*\text{Trt} + \alpha_{11}*\text{Cntr1}*\text{Trt} + \alpha_{12}*\text{Cntr2}*\text{Trt} + \alpha_{13}*\text{Vein1}*\text{Trt} + \alpha_{14}*\text{Vein2}*\text{Trt}$$

The indicators for center were Cntr1 and Cntr2 defined as 1 when the subjects were from centers 2 and 3, Ohio and Michigan, respectively. Otherwise they were coded as 0. The similar indicators for vein size were Vein1 and Vein2 indicating vein categories 2 and 3, 1 mm - 3 mm and 3 mm - 6 mm, respectively. The indicator for treatment was coded as 0 for Sotradecol and 1 for Aethoxysklerol.

Note that all variables in the model are indicators, and thus we would expect that the coefficients are of roughly the same magnitude, and at least as far as prior knowledge is concerned are expected to be exchangeable. So we model the knowledge about the coefficients as having the same distribution, namely all $\alpha_i \sim \text{Normal}(0,\tau)$. The precision parameter is not known, and is specified with a second order distribution Gamma (1,100). Note this will have mean 1/100 and variance 1/10000. Thus the variance of the α_i's roughly will have mean 100 with variance 10000, which should be quite over-dispersed.

Fitting these models with WINBUGS we get the following:

Table A.6.1 Models Analyzed

Model	Cntr	Vein	Trt	Cntr*Vein	Vein*Trt	Cntr*Trt	DIC	Mean Loglike
1	1	1	1	1	1	1	355.171	-170.8
2	1	1	1	1	1	0	351.450	-169.9
3	1	1	1	1	0	1	352.207	-170.2
4	1	1	1	0	1	1	348.293	-169.2
5	1	1	1	1	0	0	348.412	-169.3
6	1	1	1	0	1	0	344.458	-168.3
7	1	1	1	0	0	1	344.998	-168.6
8	1	1	1	0	0	0	341.234	-167.6
9	1	1	0	0	0	0	341.245	-168.2

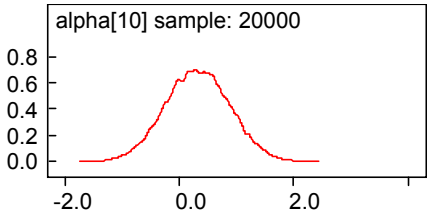
Model 1 is the full model described above. Each of the 9 models is characterized by the terms included in that model. These terms are indicated in the second column in the table above. For example, a 1 under the Cntr*Vein term indicates that all four indicators for such an interaction with their coefficients were included in the model. A 0 indicates their absence. The Deviance Information Criterion is an information measure. Since for a given data set smaller DICs are associated with better models, the DICs for these models suggest main effects model, Model 8, with terms for center, vein size, and treatment is the best among these models. However, the DIC's indicate that Model 9 is indistinguishable from Model 8, suggesting no treatment difference. The models with interaction terms, i.e. models 1-7 all have higher DIC's than Model 8 and 9, although using the usual scaling recommendations for DIC's Models 6 and 7 are close to being indistinguishable from Models 8 and 9. These results do indicate that the various interaction terms are superfluous.

Fitting the full model in WINBUGS 1.4 gives the following summary statistics for the posterior distributions:

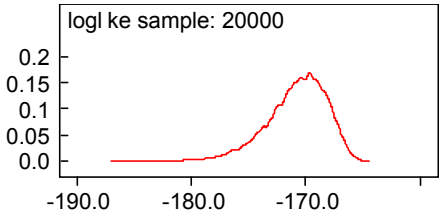
node	mean	sd	MC error	2.5%	median	97.5%	start	sample
alpha[1]	-1.255	0.4294	0.01432	-2.14	-1.241	-0.4436	5001	20000
alpha[2]	-0.4263	0.5812	0.01566	-1.581	-0.4149	0.6879	5001	20000
alpha[3]	-0.9045	0.9208	0.01801	-2.831	-0.8546	0.7915	5001	20000
alpha[4]	-0.1459	0.6014	0.01685	-1.336	-0.139	1.021	5001	20000
alpha[5]	0.1479	0.5676	0.01633	-0.9747	0.1495	1.266	5001	20000
alpha[6]	0.2418	0.7205	0.01696	-1.185	0.2416	1.649	5001	20000
alpha[7]	-0.05222	0.6716	0.01568	-1.362	-0.05348	1.264	5001	20000
alpha[8]	0.8037	1.1	0.01856	-1.355	0.7957	2.987	5001	20000
alpha[9]	0.3313	1.224	0.01662	-2.101	0.3408	2.702	5001	20000
alpha[10]	0.3233	0.5551	0.01691	-0.7619	0.3217	1.412	5001	20000
alpha[11]	0.03618	0.5752	0.01225	-1.084	0.0319	1.17	5001	20000
alpha[12]	-0.1979	0.9767	0.01532	-2.106	-0.2165	1.78	5001	20000
alpha[13]	-0.158	0.6904	0.01648	-1.523	-0.1546	1.183	5001	20000
alpha[14]	0.3807	0.6618	0.01699	-0.9497	0.3872	1.675	5001	20000
loglike	-170.8	2.609	0.03345	-176.7	-170.4	-166.7	5001	20000
pr0	0.7134	0.4522	0.01106	0.0	1.0	1.0	5001	20000

Note that alpha[10] is the differential effect of the Aethoxysklerol over Sotradecol, in the sense that the coefficient is the estimated log odds ratio for Aethoxysklerol over Sotradecol holding all other factors constant. The log likelihood is the estimated data marginal. One caveat is that with only 20000 iterations, even with the moderately low autocorrelations seen here, common wisdom suggests that the mean posterior log likelihood is not well estimated. However limited experience with logit models seems to indicate that the marginal model may be adequately estimated. This needs further exploration and justification.

The estimated posterior distribution of treatment effect alpha[10] is displayed below. Note that it is seems quite symmetric.



The estimated posterior loglikelihood is plotted as:



If $f(y|M_j)$ denotes the marginal posterior for model M_j and $f(y|M_k)$ denotes the marginal posterior for model M_k , the Bayes Factor for model M_j over M_k is defined as the ratio of these marginal posteriors $BF_{jk} = f(y|M_j) / f(y|M_k)$. Thus, in terms of log likelihoods, BF_{jk}= $\exp(\text{loglike}(y|M_j) - \text{loglike}(y|M_k))$. Using the MCMC mean log likelihoods as estimates of these terms we get $BF_{89} = \exp(-167.6 + 168.2) = 1.82$. Values of this magnitude are usually considered to be indicative of extremely weak support for model M_8 over M_9. That is, there is no strong evidence of a treatment difference.

An alternative approach would be to assume all nine models are, a priori, equally likely. That is they all have the same prior, 1/9. Then the estimated posterior probability of model M_8 would be $P(M_8 \mid y) = 1/ (\Sigma\, BF_{j8}) = .3319$, while the estimated posterior probability of M_9 would be $P(M_9 \mid y) = 1/ (\Sigma\, BF_{j8}) = .1812$. However, as noted above, common wisdom suggests that the data marginals, and hence the Bayes factors, are not well estimated. For example, the sum of the 9 computed posterior model probabilities is less than 1.0. One possible improvement in the estimated model probabilities might be to normalize these estimated model posterior probabilities so that they sum to 1. However, the effect of such ad hoc modifications remains to be investigated.

The DIC's suggest that none of the interaction terms are of particular importance, nor are there important treatment differences. Although there may be problems estimating the posterior probability, with equal priors across models the posterior probability of the main effects model with treatment differences is rather underwhelming. Note that as is typical of such analyses, the Bayesian analysis tends to be consistent with the usual frequentist approach, but perhaps gives a better assesment of actual effect size.

WINBUGS 1.4 Programs similar to the following were used to derive these results:

```
model {
   for ( i in 1:N ) {
     succ[i]~dbern(p[i])
     logit(p[i]) <- mu[i]
     mu[i]<- alpha[1] + alpha[2]*nc1[i] + alpha[3]*nc2[i] + alpha[4]*vn1[i] +
          alpha[5]*vn2[i] + alpha[6]*nc1[i]*vn1[i]+ alpha[7]*nc1[i]*vn2[i] +
          alpha[8]*nc2[i]*vn1[i] + alpha[9]*nc2[i]*vn2[i] + alpha[10]*tx[i] +
          alpha[11]*nc1[i]*tx[i] + alpha[12]*nc2[i]*tx[i] +
          alpha[13]*vn1[i]*trt[i] + alpha[14]*vn2[i]*trt[i]
    LL[i]<- succ[i]*mu[i] - log(1+exp(mu[i]))
                    }
    for (m in 1:12 ) {
       alpha[m]~dnorm(0.0,tau)
                    }
   tau~dgamma(1,100)
   sigma<-1/tau
   loglike <- sum(LL[ ])
   like    <- exp(loglike)
   pr0 <- step(alpha[10])
    }

inits
 list(alpha=c(1,0,0,0  ,0,0,0,0  ,0,0,0,0 ),tau=1)
data
  list(N=308)
 tx[ ]  nc1[ ]  nc2[ ]  vn1[  ]  vn2[  ]  succ[  ]
1  0  0  0  0  1
0  0  0  0  0  0
0  0  0  0  0  0
1  0  0  0  0  0
0  0  0  0  0  0
- data -
0  1  0  0  1  0
0  1  0  0  1  0
End
```

SIGNATURES/DISTRIBUTION LIST

Primary Statistical Reviewer: Steve Thomson
Date: 30 April 2004

Concurring Reviewer:

Statistical Team Leader: Mohamed Alosh, Ph.D., HFD-725

cc:
HFD-540/Brenda Vaughan.
HFD-540/Brenda Carr.
HFD-540/Lea Carrington, M.D.
HFD-540/Markham Luke, M.D., Ph.D.
HFD-725/Steve Thomson
HFD-725/ Mohamed Alosh, Ph.D.
HFD-725/Daphe Lin, Ph.D.
HFD-725/Mohammed Huque, Ph.D.,
HFD-725/Charles. Anello, Ph.D.
HFD-725/Robert. O'Neill, Ph.D.
HFD-725/Satya. Dubey, Ph.D.

c:\Data\Profiles\MyDocuments~\N21201AethoxysklerolJobst.doc

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

/s/

Steven Thomson
6/3/04 03:18:43 PM
BIOMETRICS

Mohammad Huque
6/7/04 01:54:23 PM
BIOMETRICS