CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

20-958

STATISTICAL REVIEW(S)

Statistical Review and Evaluation

NDA 20-958 SEP 3 2000
Name of drug: Pepcid Complete (famotidine-antacid combination)
Applicant: Merck
Indication: heartburn
Documents reviewed: response 4 August 2000 to approvable letter, amendment 10 August 2000
Project manager: Paul Levine
Reviewer: Thomas Permutt

INTRODUCTION

Pepcid (famotidine) is an H_2 antagonist approved for treating heartburn. It might provide slower but longer lasting relief than over-the-counter monograph antacids. The subject NDA concerns a combination of famotidine and antacid, hypothesized to be superior to famotidine alone in speed of relief and to antacid alone in duration of relief.

NDA 20-958 was submitted 20 February 1998. It was found to be not approvable. In the action letter 19 February 1999 the division director indicated that one study (110) provided evidence of the superiority of the combination to each of its components, but that this evidence required replication.

The application was amended 17 December 1999 by a "complete response" to the letter of 19 February 1999. That submission included a report of the results of an additional study (127) of similar design to study 110. The amendment was reviewed by Dr. Michael Elashoff, who has since left the agency. The primary statistical analysis of study 127 was characterized by Dr. Elashoff as "an extremely complex model, with many assumptions and interlocking parts. The study report," he continued, "did not adequately address the validity of these assumptions." I concurred with Dr. Elashoff's conclusions. The action letter 20 June 2000 (approvable) contained similar language.

The present submission addresses these deficiencies in two ways. First, it repairs the original primary analysis by relaxing certain assumptions questioned by Dr. Elashoff, with the aim of showing that the conclusions of the analysis are not much changed even if the assumptions may not be valid. Second, it reports some new, simple analyses suggested by Dr. Elashoff. The reworked original analysis is important because it addresses concerns about validity without raising new concerns about multiplicity in post-hoc analyses. The simpler analysis, however, may give the clearest picture of what the effects of the drug really are. Taken together, the analyses in the present submission permit a fair comparison of the performance of the combination drug to that of its components.

REVISED ORIGINAL ANALYSES

ORIGINAL ANALYSIS

The primary statistical analysis of study 127 was discussed in some detail by Dr. Elashoff. In addition, study 110, which used similar methods, was reviewed by Dr. Mushfiqur Rashid in connection with the original NDA submission. I refer to those reviews for some details. I must nevertheless set forth my understanding of the methods involved to make clear what issues this submission needed to address and how well it does address them.

The measures of outcome involved are, for each of up to four episodes of heartburn in each patient, the time at which relief began and ended. Dr. Elashoff discussed the operational definitions of relief and of the times of beginning and ending. He also noted that the time of ending was inappropriately called "duration" of relief, arguing that duration must be counted from the beginning of relief rather than from the time of dosing. The present submission concedes the semantic point, but the analysis remains unchanged in this regard. The times were then grouped into intervals. Beginning of relief was assigned to one of six categories, with the best being under 15 minutes and the worst over 2 hours. (Ending of relief was handled similarly, with different categories.) This produced, for each patient, up to four scores from 1 to 6. Each such score was then interpreted as a collection of 5 dichotomies. That is, a score could be 1, or it could be more than 1. It could also be 1 or 2, or it could be more than 2, and so on. This collection of 5 binary outcomes per episode times 4 episodes per patient was then related to treatment group, along with covariates for site and baseline severity, by multivariate logistic regression. To be able to test a single parameter of the 5-dimensional outcome for each episode, a constant odds ratio was assumed. That is, if the odds of a response of 2 or more relative to a response of 1 was x times as high in one treatment group than in another, then the odds of a response of 3 or more relative to a response of 1 or 2 must also be x times as high. Additionally, to deal with the four correlated, repeated measures for each patient, a correlation structure had to be assumed. Any two different episodes within a patient were assumed to have the same correlation. It is not clear to me precisely what was meant by correlation in this context because there are, in the logistic analysis, 20 dependent variables per patient (5 dichotomies × 4 episodes) and so 150 correlations between a dichotomy in one episode and a dichotomy in another episode. I do not know if all 150 were assumed to be the same, or if some were assumed to be zero (different episodes and different dichotomies, for example), or if the correlations were related parametrically with more than 1 but fewer than 150 parameters. Dr. Elashoff wrote:

> The study report does not contain a full description of the GEE [generalized estimating equation] models used in the primary analysis. The report contained information about the covariates in the model (treatment, investigator, and episode severity) but no information on the assumed or observed correlation matrix. Discussions with the sponsor (4/26/2000) determined that the correlations between episodes were assumed to be the same ...

but the question of correlations between different dichotomies in different episodes does not appear to have been addressed. In any event, Dr. Elashoff questioned the assumed correlation structure, pointing to differences between treatment groups in the observed correlations.

The present submission addresses these problems in two ways. First, there is an analysis fairly similar to the original analysis. It differs in making slightly less restrictive assumptions about the correlation structure. It also employs a slightly different statistical technique that may improve the reliability of the results even if the new, more general specification of the correlation structure is still not correct. Second, there is a new, more distantly related analysis, which requires neither the assumption of constant odds ratio nor any assumptions about correlation. This dual approach is appropriate in the circumstances, I think. The analysis closely parallel to the original one may repair its defects without raising important new ones about a multiplicity of post-hoc choices. The second, simpler analysis makes many fewer assumptions and so is more likely to be valid, but it was admittedly chosen after the fact. Taken together, the two analyses can shed light on the problem of the robustness of the original analysis without raising serious concerns about multiplicity.

NEW CORRELATION STRUCTURE AND SANDWICH ESTIMATOR

The first new analysis estimates the correlation structure separately in the four treatment groups. It also introduces a sandwich estimator of the variance of the resulting estimates of treatment effects.

The idea of the sandwich estimator is this. In weighted least-squares estimation, of which the generalized estimating equation (GEE) method used in the primary analysis is an extension, the parameter estimates are proportional to $\mathbf{X'WY}$, where $\mathbf{Y}$ is a vector of outcomes, $\mathbf{X}$ a matrix of predictors and $\mathbf{W}$ a matrix of weights. The variance of such an estimator is proportional to $\mathbf{X'WVWX}$, where $\mathbf{V}$ is the covariance matrix of the outcomes $\mathbf{Y}$. The matrix $\mathbf{WVW}$ is the "sandwich": $\mathbf{W}$ is the "bread" and $\mathbf{V}$ is the "meat." The best weights are inversely proportional to the matrix $\mathbf{V}$, in which case the sandwich $\mathbf{WVW} = \mathbf{V}^{-1}\mathbf{V}\mathbf{V}^{-1}$ reduces to $\mathbf{V}^{-1}$. Usually $\mathbf{V}$ is not known and has to be estimated, both for the purpose of choosing the weights $\mathbf{W}$ and for the purpose of estimating the variance of the parameter estimates. Often a model-based estimator $\hat{\mathbf{V}}$ is used for both purposes, so that $\mathbf{WVW}$ is estimated by $\hat{\mathbf{V}}^{-1}$. The sandwich estimator uses two different estimates of $\mathbf{V}$ for these two purposes. The weights are chosen proportional to the inverse of a model-based estimate of $\mathbf{V}$. If the model is wrong, these weights will not be optimal, but valid inference can still be based on them. In estimating the variance of the parameter estimates, however, a different estimate of $\mathbf{V}$ is used for the meat of the sandwich. This estimate uses a more general form for the covariance matrix, with more estimated parameters. It thus produces more variable estimates than the model-based estimator if the model is right, but consistent estimates even if the model is wrong, as long as the more general form is right.

Any use of this technique, with a more specific parametrization of $\mathbf{V}$ in the weight matrix and a more general one in the middle, may be referred to as a sandwich estimator. As I noted earlier, the parametrization of the correlation matrix used for estimation (the bread) is not clear. Neither is the precise form of the meat. An information request to clarify this

point was apparently misunderstood by the sponsor as requesting basic information about the technique, and the response consisted of two pages of a book explaining the theory of the sandwich estimator in general. The most general form of the covariance matrix, which is what is discussed in the book, has in this case 840 parameters (different 20 × 20 symmetric matrix in each of four treatment groups). If this is the form that was used, the robustness of the sandwich estimator, which depends on asymptotic arguments, may be somewhat vitiated by the need to estimate so many parameters.

Table 2c

Protocol 127
Weighted Least Squares Analysis - Onset Data
Mean and Standard Error (SE) for the Overall Log Odds By Treatment Group
All-Patients-Treated Approach (N=1618)

FACT n=406 Tot Eps†=1585		Fam 10-mg FCT n=406 Tot Eps=1598		Antacid 21 mEq n=407 Tot Eps=1565		Placebo n=399 Tot Eps=1533	
Mean	SE	Mean	SE	Mean	SE	Mean	SE
0.79	0.09	0.45	0.08	0.61	0.09	0.35	0.08

† Eps = episodes.

Treatment Comparison	Odds-Ratio (95% CI)	Z-Statistic	p-Value
FACT vs. Fam 10-mg FCT [P]	1.40 (1.09, 1.79)	7.19	0.007
FACT vs. AA 21 mEq	1.19 (0.92, 1.53)	1.75	0.186
FACT vs. Placebo	1.54 (1.21, 1.96)	12.15	<0.001
FAM 10-mg FCT vs. Placebo	1.10 (0.88, 1.38)	0.67	0.412
AA 21 mEq vs. Placebo	1.30 (1.02, 1.64)	4.64	0.031
FAM 10-mg FCT vs. AA 21 mEq	0.85 (0.67, 1.08)	1.81	0.178

[P] = Primary treatment comparison.

Table 2d

Protocol 127
Weighted Least Squares Analysis - Duration Data
Mean and Standard Error (SE) for the Overall Log Odds By Treatment Group
All-Patients-Treated Approach (N=1618)

FACT n=406 Tot Eps†=1585		Fam 10-mg FCT n=406 Tot Eps=1598		Antacid 21 mEq n=407 Tot Eps=1565		Placebo n=399 Tot Eps=1533	
Mean	SE	Mean	SE	Mean	SE	Mean	SE
1.22	0.09	0.73	0.08	0.85	0.08	0.33	0.08

† Eps = episodes.

Treatment Comparison	Odds-Ratio (95% CI)	Z-Statistic	p-Value
FACT vs. Fam 10-mg FCT	1.63 (1.30, 2.05)	17.58	<0.001
FACT vs. AA 21 mEq [P]	1.46 (1.17, 1.83)	10.99	0.001
FACT vs. Placebo	2.45 (1.96, 3.07)	61.83	<0.001
FAM 10-mg FCT vs. Placebo	1.50 (1.21, 1.87)	13.63	<0.001
AA 21 mEq vs. Placebo	1.68 (1.36, 2.08)	23.31	<0.001
FAM 10-mg FCT vs. AA 21 mEq	0.89 (0.72, 1.11)	1.02	0.312

[P] = Primary treatment comparison.

The results of this revised analysis are shown in the two tables left, copied from the submission. For onset of relief, the combination (FACT) was statistically significantly better than famotidine alone (Fam), if any of the confidence interval (1.09 to 1.79), Z-statistic (7.19), or p-value (0.007) is correct. They cannot all be correct, however. A Z-statistic of 7 would produce a p-value with many zeroes and a confidence interval for the odds ratio not so nearly approaching 1. (Possibly the table is mislabeled and what is called the Z-statistic is really its square.) For the "duration" of relief, the relevant comparison is between FACT and antacid alone (AA). Again, the results are statistically significant, if any of them is correct, but they are mutually inconsistent. It is very difficult to give a meaningful interpretation of the magnitude of the effects, which are the average odds ratios for the multiple dichotomies for the categorized times of onset and duration.

All the new analyses in this submission have been applied to the earlier reported study 110 as well as to the new study 127. Neither Dr. Elashoff's review nor this one is intended to revisit in detail the analysis of study 110, reviewed by Dr. Rashid. On the other hand, the original action letter stated that the results of study 110 should be "replicated." Accordingly, in applying new analyses to study 127, the applicant has appropriately considered the question of whether the results of the two studies remain consistent.

The results are indeed consistent between the two studies. Again the critical comparisons are reported to be statistically significant. Again the calculations are erroneous, or perhaps erroneously labeled. Again the numerical results are difficult to interpret.

Table 2a

Protocol 110
Weighted Least Squares Analysis - Onset Data
Mean and Standard Error (SE) for the Overall Log Odds By Treatment Group
All-Patients-Treated Approach (N=1231)

FACT n=305 Tot Eps†=1205		Fam 10-mg FCT n=311 Tot Eps=1229		Antacid 21 mEq n=308 Tot Eps=1212		Placebo n=307 Tot Eps=1217	
Mean	SE	Mean	SE	Mean	SE	Mean	SE
0.53	0.11	0.22	0.11	0.24	0.10	0.13	0.10

† Eps = episodes.

Treatment Comparison	Odds-Ratio (95% CI)	Z-Statistic	p-Value
FACT vs. Fam 10-mg FCT [P]	1.36 (1.02, 1.82)	4.27	0.039
FACT vs. AA 21 mEq	1.34 (1.00, 1.78)	3.90	0.048
FACT vs. Placebo	1.48 (1.11, 1.98)	7.20	0.007
FAM 10-mg FCT vs. Placebo	1.09 (0.82, 1.45)	0.35	0.556
AA 21 mEq vs. Placebo	1.11 (0.84, 1.47)	0.52	0.471
FAM 10-mg FCT vs. AA 21 mEq	0.98 (0.74, 1.31)	0.01	0.903

[P] = Primary treatment comparison.

Table 2b

Protocol 110
Weighted Least Squares Analysis - Duration Data
Mean and Standard Error (SE) for the Overall Log Odds By Treatment Group
All-Patients-Treated Approach (N=1231)

FACT n=305 Tot Eps†=1205		Fam 10-mg FCT n=311 Tot Eps=1229		Antacid 21 mEq n=308 Tot Eps=1212		Placebo n=307 Tot Eps=1217	
Mean	SE	Mean	SE	Mean	SE	Mean	SE
1.26	0.10	1.02	0.09	0.78	0.10	0.66	0.09

† Eps = episodes.

Treatment Comparison	Odds-Ratio (95% CI)	Z-Statistic	p-Value
FACT vs. Fam 10-mg FCT	1.27 (0.98, 1.66)	3.22	0.073
FACT vs. AA 21 mEq [P]	1.62 (1.22, 2.14)	11.48	0.001
FACT vs. Placebo	1.83 (1.41, 2.37)	20.40	<0.001
FAM 10-mg FCT vs. Placebo	1.43 (1.13, 1.83)	8.57	0.003
AA 21 mEq vs. Placebo	1.13 (0.87, 1.46)	0.85	0.356
FAM 10-mg FCT vs. AA 21 mEq	1.27 (0.98, 1.65)	3.27	0.071

[P] = Primary treatment comparison.

BINARY CLASSIFICATION AND FIRST EPISODE

The submission's second approach to questions about the robustness of the primary analysis is a fundamentally similar but much simpler analysis. Instead of the ordinal logistic analysis with its multiple dichotomies, a binary analysis with only two categories is used. The time of onset of relief is classified as more or less than 30 minutes, and the "duration" as more or less than 7 hours. Furthermore, only the first episode of heartburn for each patient is used in the analysis. Thus, there is a single dichotomy for each patient for onset (<30 minutes or not) and similarly for duration, rather than a 20-variate vector. The question of the correlations between variates within patients therefore does not arise.

The results are shown in the tables below, again copied from the submission. For onset of relief, the combination was significantly (p = 0.003) better than famotidine alone. The numerical result is still awkward to interpret: the odds of a patient having relief in less than

Table 1c

Protocol 127
Onset for First Episode - Binary Data
Number (%) of Patients Adequately Relieved
All-Patients-Treated Approach (N=1618)

Episode	Adequate Relief at:	FACT n=406		Fam 10-mg FCT n=406		Antacid 21 mEq n=407		Placebo n=399	
		n	%	n	%	n	%	n	%
1	≤30 Mins	213	52.5	172	42.4	205	50.4	171	42.9
	>30 Mins	193	47.5	234	57.6	202	49.6	228	57.1

Treatment Comparison	Model-Adjusted Odds-Ratio (95% CI)	Chi-Square	p-Value
FACT vs. Fam 10-mg FCT [P]	1.54 (1.15, 2.05)	8.61	0.003
FACT vs. AA 21 mEq	1.12 (0.84, 1.49)	0.61	0.436
FACT vs. Placebo	1.51 (1.13, 2.01)	7.77	0.005
FAM 10-mg FCT vs. Placebo	0.98 (0.73, 1.31)	0.02	0.889
AA 21 mEq vs. Placebo	1.34 (1.01, 1.79)	4.08	0.043
FAM 10-mg FCT vs. AA 21 mEq	0.73 (0.55, 0.97)	4.69	0.030

[P] = Primary treatment comparison.

Table 1d

Protocol 127
Duration for First Episode - Binary Data
Number (%) of Patients Adequately Relieved
All-Patients-Treated Approach (N=1618)

Episode	Adequate Relief for:	FACT n=406		Fam 10-mg FCT n=406		Antacid 21 mEq n=407		Placebo n=399	
		n	%	n	%	n	%	n	%
1	≥7 Hrs	293	72.2	257	63.3	253	62.2	217	54.4
	<7 Hrs	113	27.8	149	36.7	154	37.8	182	45.6

Treatment Comparison	Model-Adjusted Odds-Ratio (95% CI)	Chi-Square	p-Value
FACT vs. Fam 10-mg FCT	1.52 (1.12, 2.05)	7.32	0.007
FACT vs. AA 21 mEq [P]	1.64 (1.21, 2.21)	10.20	0.001
FACT vs. Placebo	2.24 (1.66, 3.02)	27.71	<0.001
FAM 10-mg FCT vs. Placebo	1.47 (1.10, 1.97)	6.93	0.009
AA 21 mEq vs. Placebo	1.37 (1.03, 1.82)	4.57	0.033
FAM 10-mg FCT vs. AA 21 mEq	1.08 (0.81, 1.44)	0.25	0.617

[P] = Primary treatment comparison.

30 minutes are better in the combination than in the famotidine group, by a factor of 1.54. The submission incorrectly asserts that combination patients were "1.54 times more likely ... to experience relief within 30 minutes." In fact they were 52.5/42.4 = 1.24 times more likely. (The odds ratio corresponds closely to the rate ratio when the rates are small, but it is approximately the square of the rate ratio when the rates are near a half.) For "duration" of relief, the combination was significantly better (p = 0.001) than antacid alone. Again the

numerical result is awkward to interpret, and again it is incorrectly interpreted in the submission.

This new analysis was also applied retrospectively to study 110. The results for study 110 are shown right. Again the results are consistent between the two studies.

This simpler analysis is consistent in spirit with the originally planned analysis, but does not require the problematic assumptions of it. The new analysis indicates that the combination is better than famotidine alone with respect to onset of relief, and better than antacid alone with respect to "duration" of relief.

Table 1a

Protocol 110
Onset for First Episode - Binary Data
Number (%) of Patients Adequately Relieved
All-Patients-Treated Approach (N=1231)

Episode	Adequate Relief at:	FACT n=305		Fam 10-mg FCT n=311		Antacid 21 mEq n=308		Placebo - n=307	
		n	%	n	%	n	%	n	%
1	≤30 Mins	146	47.9	112	36.0	129	41.9	92	30.0
	>30 Mins	159	52.1	199	64.0	179	58.1	215	70.0

Treatment Comparison	Model-Adjusted Odds-Ratio (95% CI)	Chi-Square	p-Value
FACT vs. Fam 10-mg FCT [P]	1.66 (1.19, 2.31)	9.11	0.003
FACT vs. AA 21 mEq	1.28 (0.92, 1.77)	2.16	0.142
FACT vs. Placebo	2.17 (1.55, 3.04)	20.18	<0.001
FAM 10-mg FCT vs. Placebo	1.31 (0.93, 1.84)	2.34	0.126
AA 21 mEq vs. Placebo	1.70 (1.21, 2.38)	9.42	0.002
FAM 10-mg FCT vs. AA 21 mEq	0.77 (0.55, 1.07)	2.44	0.119

[P] = Primary treatment comparison.

Table 1b

Protocol 110
Duration for First Episode - Binary Data
Number (%) of Patients Adequately Relieved
All-Patients-Treated Approach (N=1231)

Episode	Adequate Relief for:	FACT n=305		Fam 10-mg FCT n=311		Antacid 21 mEq n=308		Placebo n=307	
		n	%	n	%	n	%	n	%
1	≥7 Hrs	224	73.4	228	73.3	197	64.0	199	64.8
	<7 Hrs	81	26.6	83	26.7	111	36.0	108	35.2

Treatment Comparison	Model-Adjusted Odds-Ratio (95% CI)	Chi-Square	p-Value
FACT vs. Fam 10-mg FCT	1.00 (0.69, 1.45)	0.00	0.998
FACT vs. AA 21 mEq [P]	1.58 (1.11, 2.27)	6.33	0.012
FACT vs. Placebo	1.51 (1.06, 2.16)	5.10	0.024
FAM 10-mg FCT vs. Placebo	1.51 (1.06, 2.16)	5.19	0.023
AA 21 mEq vs. Placebo	0.95 (0.68, 1.35)	0.07	0.791
FAM 10-mg FCT vs. AA 21 mEq	1.59 (1.11, 2.26)	6.43	0.011

[P] = Primary treatment comparison.

In principle, a problem of multiplicity arises whenever a dichotomy is chosen after the fact. That is, the analysis might have focused on some other cut-point than 30 minutes for onset of relief or 7 hours for duration of relief. In this case, that problem is not of great concern. The p-values are small enough that they might be adjusted and still be significant. The two studies both give positive results with the same choices. Furthermore, the new analysis is essentially consistent with the single primary analysis originally proposed. Concerns about the validity of that analysis necessitated some post-hoc alternative analyses, and these cannot therefore be criticized merely for being post-hoc.

ANALYSES BY TIMEPOINT

Another, fundamentally different approach to analysis was suggested by Dr. Elashoff. For each episode, patients gave assessments of whether and how well their heartburn was

relieved at several fixed timepoints after dosing. Rather than reducing these assessments to an "onset" and "duration" of relief for each patient, we might simply consider, for various timepoints, how the treatment groups differed. The applicant conducted such an analysis. In particular, the numbers of patients with "adequate" relief were compared for each pair of treatments at each of several timepoints. The data form a 2×2 contingency table (adequate relief or not, for each of two treatments) for each comparison, and Fisher's exact test of a significant difference between treatments was calculated. Again, the question of how to handle repeated measures arises; and again, the applicant has chosen the simplest if not the most efficient method by only using the first episode for each patient. The results are shown in the sponsor's table below. Again, the earlier study 110 was reanalyzed along with the new study 127.

Fisher's Exact Test by Individual Time Points for First Episode Data
Number (%) of Patients With Adequate Relief by Time Point for First Episode

	Protocol 110								Protocol 127							
Adequate Relief at:	FACT n=305		FAM 10-mg FCT n=311		Antacid 21 mEq n=308		Placebo n=307		FACT n=406		FAM 10-mg FCT n=406		Antacid 21 mEq n=407		Placebo n=399	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
15 min	72	23.6	58	18.6	74	24.0	43	14.1	128	31.5	99	24.4	129	31.7	99	24.8
30 min	142	46.6	104	33.4	124	40.3	85	27.7	204	50.2	159	39.2	198	48.6	165	41.4
45 min	197	64.6	181	58.2	174	56.5	151	49.2	287	70.7	244	60.1	271	66.6	248	62.2
60 min	239	78.4	220	70.7	215	69.8	211	68.7	323	79.6	284	70.0	311	76.4	290	72.7
120 min	252	82.6	244	78.5	229	74.4	229	74.6	328	80.8	296	72.9	312	76.7	284	71.2
180 min	249	81.6	247	79.4	226	73.4	230	74.9	328	80.8	301	74.1	307	75.4	273	68.4
240 min	244	80.0	246	79.1	220	71.4	227	73.9	329	81.0	300	73.9	298	73.2	262	65.7
300 min	247	81.0	241	77.5	211	68.5	223	72.6	323	79.6	289	71.2	291	71.5	253	63.4
360 min	240	78.7	240	77.2	207	67.2	218	71.0	310	76.4	282	69.5	281	69.0	235	58.9
420 min	236	77.4	235	75.6	208	67.5	211	68.7	314	77.3	280	69.0	274	67.3	237	59.4
480 min	237	77.7	235	75.6	207	67.2	208	67.8	313	77.1	278	68.5	272	66.8	235	58.9

Fisher's Exact Test Pairwise Treatment Comparison p-Values by Time Point for First Episode

	Protocol 110					Protocol 127				
Adequate Relief at:	FACT vs. Famotidine 10-mg FCT	FACT vs. Antacid 21 mEq	FACT vs. Placebo	Famotidine 10-mg FCT vs. Placebo	Antacid 21 mEq vs. Placebo	FACT vs. Famotidine 10-mg FCT	FACT vs. Antacid 21 mEq	FACT vs. Placebo	Famotidine 10-mg FCT vs. Placebo	Antacid 21 mEq vs. Placebo
15 min	0.139	0.925	0.003	0.129	0.002	0.028	1.000	0.035	0.935	0.035
30 min	0.001	0.122	<0.001	0.138	0.001	0.002	0.674	0.013	0.565	0.040
45 min	0.116	0.047	<0.001	0.029	0.076	0.002	0.227	0.011	0.564	0.211
60 min	0.033	0.017	0.008	0.600	0.794	0.002	0.310	0.025	0.436	0.226
120 min	0.222	0.014	0.018	0.296	1.000	0.010	0.170	0.002	0.638	0.078
180 min	0.542	0.016	0.050	0.213	0.713	0.029	0.075	<0.001	0.074	0.028
240 min	0.842	0.014	0.084	0.154	0.527	0.019	0.010	<0.001	0.011	0.022
300 min	0.321	<0.001	0.017	0.193	0.288	0.007	0.009	<0.001	0.020	0.016
360 min	0.698	0.001	0.032	0.082	0.337	0.033	0.022	<0.001	0.002	0.003
420 min	0.635	0.007	0.018	0.060	0.795	0.009	0.002	<0.001	0.005	0.023
480 min	0.568	0.004	0.006	0.032	0.931	0.007	0.001	<0.001	0.005	0.024

The applicant expressed several reservations about these analyses:

> Although we have conducted the suggested Fisher's exact test calculations, it should be noted that there are certain limitations and difficulties in interpreting the results. Protocols 110 and 127 were not powered to detect differences among treatment groups for individual time points. Fisher's exact test does not allow for inclusion of covariates in the model. A very large number of tests is being proposed without any adjustment to the alpha level for multiplicity, either for time points or treatment comparisons, It is unclear how the results of this series of p-values across time points should be interpreted in order to evaluate onset and duration for the famotidine/antacid combination, since no definition of "onset" or "duration" is being given a priori.

I agree with these remarks for the most part. This analysis would not stand on its own as a primary analysis because of questions about multiplicity. Furthermore, it would be rather undesirable as a primary analysis because it is inefficient. Besides the debatable inefficiency of leaving out covariates, it wastes all the data from episodes after the first.

"Nevertheless," the sponsor continues, "the results of Fisher's exact test clearly demonstrate the efficacy advantage of the combination versus famotidine or antacid alone in both protocols." I also agree with this assertion. In both studies, the early timepoints favor the combination over famotidine alone, and the late timepoints favor the combination over antacid alone. (Interestingly, in study 127 the combination was better than famotidine alone at all timepoints, early and late.) In addition, although the interpretation of the p-values may be clouded by multiplicity, the interpretation of the numerical results is here for the first time straightforward.

As with the binary analysis discussed earlier, I do not think the problem of multiplicity seriously interferes with the interpretation of the results in this case. There is a consistent pattern of significant results, not a sprinkling of nominally significant tests among many nonsignificant ones. Some of the p-values are small enough to withstand considerable adjustment. The results of the two studies are broadly consistent. Finally, there is "protection" (in the sense of protected multiple comparisons) in the significance of the prespecified analysis. That is, this is not a post-hoc attempt to rescue a nonsignificant primary analysis, but rather to elucidate a statistically significant but very obscure one. In light of these considerations, I think the results of this simple analysis may be taken at face value; and they have a face value that is clearly readable.

CONCLUSIONS AND RECOMMENDATIONS

In study 127 the combination product produced adequate relief in more patients than famotidine alone early in the time-course of an episode of heartburn, and more than antacid alone late in the time-course. The requirement of evidence that each component of a fixed-ratio combination product contributes to the claimed effects of the drug has therefore been met.

I base these conclusions on all the analyses in the subject submission considered together. I do not think the reported primary analysis alone would justify such a conclusion. The analysis has not been described in sufficient detail to permit independent verification. The results as reported are clearly erroneous, and I do not have sufficient information to be able to correct them. Nonetheless, I think it is very likely that a correct analysis would have led to similar findings. The simpler analyses all do so, and the primary analysis might be expected to be even more sensitive. Even supposing correct and positive tests of significance, however, this analysis is so obscure that it would be very difficult to assess in practical terms the benefit of the drug.

The simpler analyses give a clear enough picture of the effects of the combination product. As the applicant points out, these analyses taken by themselves would be subject to some problems of interpretation. The main theoretical problems relate to sensitivity and

multiplicity, although the results turn out to be strong enough to allay both concerns. In conjunction with the reworked original analyses, they are clearly sufficient.

I do not think any claim about duration of relief should be allowed because none of the analyses address the question of duration properly, calculating from the beginning to the end of relief rather than from the time of dosing. This should not, however, preclude the later consideration of other analyses that do address duration, properly defined.

This review is not intended to revise the conclusions of the original review concerning study 110. However, any of the analyses of study 127 discussed in this review may be said to replicate findings of study 110, as they were all applied retrospectively to study 110 with similar results.

/S/ 10/3/00

Thomas Permutt, Ph.D.
Mathematical Statistician (Team Leader)

Concur: S. Edward Nevius, Ph.D. /S/ 10-3-00
Director, Division of Biometrics II

archival: NDA 20-958
cc:
HFD-715/Nevius, Permutt
HFD-180/Levine, Gallo-Torres, Aurecchia, Talarico
HFD-103/Houn, Collier
HFD-700/Anello
HFD-180/division file

Statistical Review and Evaluation

NDA	~~20-968~~ 20-958
Drug Name	Pepcid Complete
Indication	Treatment of heartburn
Review Date	5/24/2000
Medical Reviewer	Dr. Scheldon Kress
Statistical Reviewer	Dr. Michael Elashoff

This NDA consists of one study, designed to demonstrate that the combination of pepcid and antacid is more effective than either pepcid or antacid alone. The NDA was previously submitted with data from three similar clinical trials that showed mixed results. That NDA was not approved, and the sponsor was advised that an additional, positive study would be needed. Study 127 was conducted in order to meet that goal.

Design of Study 127

Prior to receiving study medication, patients participated in a 7 day run in period. In the run-in period, all patient received antacid tablets. Over the 7 day period, the patients recorded how many times they experienced heartburn that required use of the antacid. Patients also filled out diary cards for one hour following each heartburn episode. After the run-in period was over, patient had to meet four entry criteria to be randomized into the main phase of the study:

1. Antacid use on 3 or more days
2. Antacid use twice in a 24 hour period
3. Relief of heartburn within one hour in at least 50% of episodes
4. Completed diary cards

Subjects meeting these criteria were randomized to one of four treatment groups in a blinded fashion. The treatment groups were Pepcid, Antacid, Pepcid + Antacid, and Placebo. Each patient received 4 doses of study medication. Over the next two weeks, patients would take one dose of study medication when they experienced heartburn, up to a maximum of 4 episodes per patient.

For each episode of heartburn, patient recorded the following information in their diary cards:

- Day and time of episode
- Initial severity of episode
- Relief/No Relief at time 15 min, 30 min, 45 min, 1 hour, 2 hours, 3 hours, 4 hours, 5 hours, 6 hours, 7 hours, and 8 hours
- Use of "rescue" medication

Goals of Study 127

Since the trial was designed to assess the efficacy of a combination product, the sponsor needed to show that the combination was better than each of the individual components. And since the individual components have different effects (quick but short lasting effect with antacid, delayed but durable effect with Pepcid), the combination had to show superiority on two endpoints.

1. Onset of relief: combination must be superior to Pepcid. Also need to show that antacid superior to placebo.

2. Duration of relief: combination must be superior to antacid. Also need to show that Pepcid superior to placebo.

While the components are already approved, the comparisons of the components to placebo are necessary to ensure the efficacy of the components in this particular trial setting.

Statistical Analysis of Study 127

1. Onset of Relief: For each episode, the onset of relief was defined as the first time at which the patient recorded relief in the diary card. These times were then grouped into 6 categories: 15 min, 30 min, 45 min, 1 hour, 2 hours, >2 hours. The data were analyzed using logistic regression on the 6 ordered categories. This analysis models a patient's odds of experiencing relief in each category given that the patient had not yet experienced relief. The model then calculates an odds-ratio for each category that compares the odds between two treatment groups. The odds-ratios are then averaged across the categories to produce a final odds-ratio that is the basis for hypothesis testing. Implicit in this averaging procedure is the assumption of constant odds-ratios across the categories.
2. Duration of Relief: For each episode, the duration of relief was defined as the time of the first No Relief diary card entry that followed a Relief entry. For example, if a patient recorded No Relief at 15 min, Relief at 30 and 45 min, and No Relief at 1 hour, the duration of relief would be 1 hour. This can be thought of as a time to loss of relief analysis. These times were then grouped into 6 categories: >=7 hours, 6 hours, 5 hours, 4 hours, <4 hours, and never experienced relief. The data were analyzed using logistic regression on the 6 ordered categories. This analysis models a patient's odds of experiencing loss of relief in each category given that the patient had not yet experienced loss of relief. The model then calculates an odds-ratio for each category that compares the odds between two treatment groups. The odds-ratios are then averaged across the categories to produce a final odds-ratio that is the basis for hypothesis testing. Implicit in this averaging procedure is the assumption of constant odds-ratios across the categories.

Each of the two endpoints were analyzed in the context of a GEE model. The GEE model allows for the inclusion of multiple episodes per patient by accounting for the correlation between episodes. In addition to specification of the terms in the model (eg treatment effect, center effect), a GEE analysis requires specification of the correlation structure between the episodes. The model the sponsor used is the constant correlation model, meaning a) that patients were assumed to have constant correlation (estimated in the analysis) between episodes, and b) the correlation structure was assumed to be the same for all patients. The GEE models used by the sponsor included terms for center and baseline severity of heartburn episodes in addition to the treatment effects.

Patient Accounting

Study 127 included 1651 randomized subjects. Of these subjects, 1618 (98%) took at least one dose of study medication and had data for at least one episode of heartburn. Efficacy was assessed in this subset of patients. Sample size by treatment arm is shown in Table 1. The table also shows that the number of episodes per patient was similar across treatment groups, and most patients had 4 episodes of heartburn.

Table 1

	Placebo	Antacid	Pepcid	Combination
Sample Size	399	407	406	406
4 episodes	357	368	387	377
3 episodes	27	18	13	21
2 episodes	9	18	5	6
1 episode	6	3	1	2
Mean Episodes	3.84	3.85	3.94	3.90

Sponsor Analyses

The sponsor found that demographics and other baseline characteristics were similar across treatment groups. Patient ranged in age from 18 to 90, with a gender breakdown of 60% female and 40% male. Caucasians made up 81% of the efficacy population, with the remainder split between Black and Hispanic.

Onset of relief

The first time a subject recorded relief in their diary cards in shown in Table 2. The table incorporates all episodes for all patients.

Table 2

Time of first relief	Placebo	Antacid	Pepcid	Combination
15 min	25%	32%	27%	34%
30 min	17%	16%	19%	18%
45 min	19%	18%	17%	18%
1 hour	12%	12%	9%	12%
2 hours	5%	5%	5%	5%
>2 hours	21%	16%	20%	13%

Notice that the difference between combination and pepcid is that ~7% of subjects had their onset of relief shifted from the latest category (>2 hours) to the earliest (15 min). Otherwise, the distributions are quite similar. The same pattern was seen for the antacid versus placebo comparison. No difference was seen between combination and antacid or between pepcid and placebo.

The sponsor GEE analysis of the results in Table 2 found that the differences for the two comparisons of interest were significant.

Table 3

Comparison	Odds-Ratio	p-value
Combination vs. Pepcid	1.42	.001
Antacid vs. Placebo	1.35	.003

The sponsor found a treatment by center interaction ($p<.001$) and a treatment by age interaction ($p=.012$), but concluded that the interactions were "quantitative in nature".

Duration of relief

The first time a subject recorded non-relief in their diary cards (that followed a recording of relief) in shown in Table 4. The table incorporates all episodes for all patients.

Table 4

Time of first non-relief	Placebo	Antacid	Pepcid	Combination
7+ hours	51%	59%	60%	70%
6 hours	2%	2%	2%	1%
5 hours	4%	4%	3%	3%
4 hours	4%	4%	3%	2%
<4 hours	22%	21%	16%	14%
No onset	17%	12%	16%	10%

Since the intermediate categories had so few patients, it may be clearer to look at group these data into 0-4 hours of relief and 4-8 hours of relief (4 hours is included in the 4-8 category).

Table 5

Time of first non-relief	Placebo	Antacid	Pepcid	Combination

[4-8) hours	61%	68%	68%	76%
[0-4) hours	39%	32%	32%	24%

Here, the difference between combination and pepcid is ~8%, with a similar difference for the pepcid versus placebo comparison. Interestingly, pepcid and antacid were quite similar (sponsor p-value for Table 4 pepcid vs. antacid was .855).

The sponsor GEE analysis of the results in Table 4 found that the differences for the two comparisons of interest were significant.

Table 6

Comparison	Odds-Ratio	p-value
Combination vs. Antacid	1.60	.001
Pepcid vs. Placebo	1.37	.001

The sponsor found a treatment by center interaction ($p<.001$), but concluded that the interaction was "quantitative in nature".

Reviewer Comments

The comments on the trial design and analysis are organized into five sections. The first two discuss the definitions of the onset of action endpoint and the durability endpoint. The next two sections discuss the two statistical techniques underlying the analysis, since the two endpoints were analyzed similarly these comments will be applicable to both endpoints. The last section summarizes the issues raised.

Onset of Relief

The sponsor assessed the early effects of the combination by an analysis of time to first relief. For example, consider these three hypothetical patients. Each of them would be scored as having a time to relief of 30 minutes.

	15 min	30 min	45 min	1 hour	2 hours
Patient A	No Relief	Relief	No Relief	No Relief	No Relief
Patient B	No Relief	Relief	Relief	Relief	Relief
Patient C	No Relief	Relief	No Relief	Relief	Relief

However, these patients have qualitatively quite different early responses. Time-to-event analyses have a clear interpretation when the event is a distinct, meaningful event (death, tumor recurrence, etc.). Here, however, the event is not really distinct (eg when does Patient C experience "onset of relief" ?). Further, the analysis ignores potentially informative data after the first diary card entry indicating relief. The dichotomous scale may contribute to the problem, as patients with a moderate level of relief must enter either relief or no relief in the diary card.

Duration of Relief

The sponsor assessed the durable effects of the combination by an analysis of time to loss of relief. In this analysis, the duration of relief was the first time of no relief that occurred after a time of relief. For example, Patients A, B, C, and D would all be scored as duration of relief equal to 6 hours. However, this ignores the time it took to achieve relief, which for patient like A and C would exaggerate the true duration of relief. Also, as in the analysis of onset discussed above, data after the endpoint was discarded.

	1 hour	2 hours	3 hours	4 hours	5 hours	6 hours	7 hours	8 hours
Patient A	No	No	Relief	Relief	Relief	No	No	No
Patient B	Relief	Relief	Relief	Relief	Relief	No	No	No
Patient C	No	No	Relief	Relief	Relief	No	Relief	Relief

Patient D	Relief	Relief	Relief	Relief	Relief	No	Relief	Relief

The duration analysis grouped patients with first time of no relief equal to 1 hour, 2 hours and 3 hours together. Patients E, F, and G would all be scored as duration <4 hours.

	30 min	45 min	1 hour	2 hours	3 hours	4 hours	5 hours	6 hours
Patient E	Relief	Relief	Relief	Relief	No	No	No	No
Patient F	No	No	Relief	No	No	No	No	No
Patient G	Relief	No	No	No	No	No	No	No

The duration analysis grouped patients with first time of no relief equal to 7 hours, 8 hours and >8 hours together. Patients E, F, and G would all be scored as duration >=7 hours.

	1 hour	2 hours	3 hours	4 hours	5 hours	6 hours	7 hours	8 hours
Patient H	No	Relief	Relief	Relief	Relief	Relief	Relief	Relief
Patient I	No	Relief	Relief	Relief	Relief	Relief	Relief	No
Patient I	No	Relief	Relief	Relief	Relief	Relief	No	No

No analyses were provided that tested how sensitive the conclusions were to variations in these endpoint definitions.

Ordered Categorical Regression

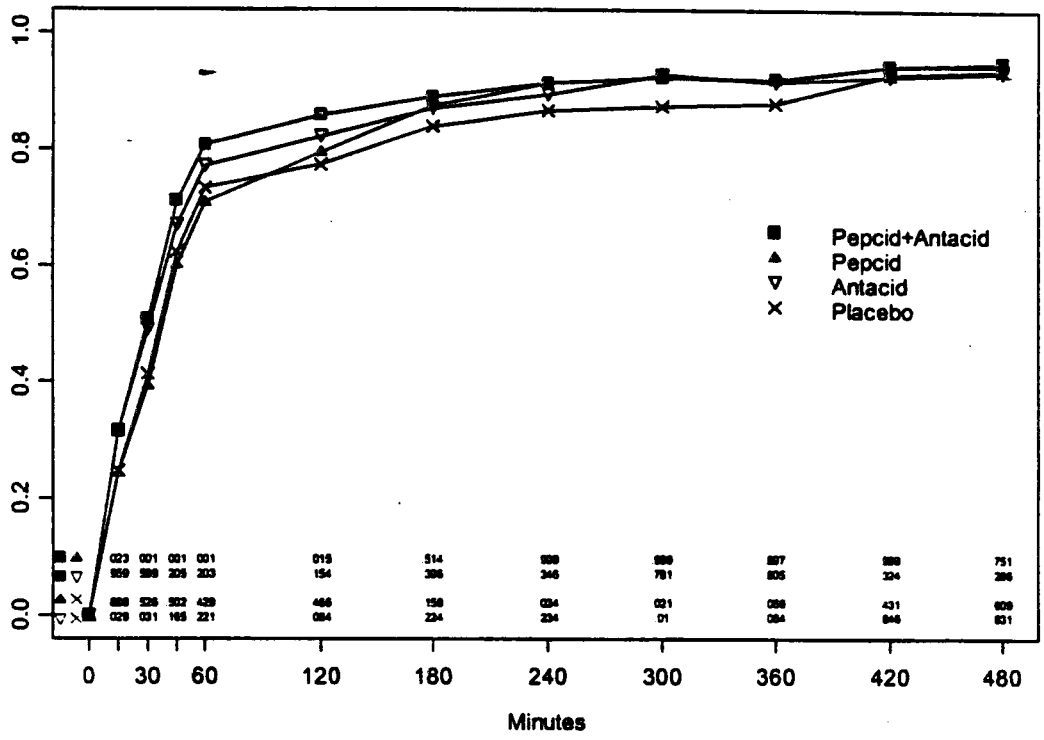
The ordered categorical regression model assumes a constant odds ratio across response categories. As mentioned above, these response categories may not adequately reflect the pattern of response over time for patients. In any case, the study report does not address the validity of the constant odds ratio assumption, how this assumption might be tested, what the power of such a test would be, or how violations of this assumption would impact the treatment effects and p-values for the primary comparisons.

GEE Models

The study report does not contain a full description of the GEE models used in the primary analysis. The report contained information about the covariates in the model (treatment, investigator, and episode severity) but no information on the assumed or observed correlation matrix. Discussions with the sponsor (4/26/2000) determined that the correlations between episodes were assumed to be the same (this is referred to as a constant correlation structure). Further, each patient was assumed to have the same correlation structure.

The study report did not include any discussion of these assumptions. Published articles have typically found that, while the estimated treatment effects may still be valid when the assumptions are not met, the significance level is no longer guaranteed to be .05.

This is particularly relevant since it appears that the significance of the treatment comparisons, for the duration of relief endpoint in particular, was a consequence of the pooling of the episodes. Reviewer analyses found that when episodes were analyzed separately the treatment effects were no longer significant. The proportion recording relief at each time point is shown in Figure 1. This figure represents data from each patient's first episode of heartburn. Running along the x-axis at the bottom of figure 1 are p-values for the various treatment comparisons. Note that at the early time points the combination is superior to pepcid, but at no time point is the combination arm statistically superior to the antacid arm. The other episodes showed similar results.



To assess the validity of the GEE model constant correlation assumptions, correlations between episodes were estimated separately for each treatment group. Table 7 shows the average of the six correlations (correlation between episode 1 and 2, 1 and 3, 1 and 4, 2 and 3, 2 and 4, 3 and 4) that must be estimated, and the range of these correlations.

Table 7

Treatment Group	Average correlation between episodes	Range of correlation between episodes
Placebo	.57	
Antacid	.65	
Pepcid	.52	
Combination	.72	

As the table indicates, the correlation between episodes was different across treatment arms, with patient on the combination arm having the highest within patient correlation and pepcid having the lowest within patient correlation. Correlations were compared statistically using Fisher's Z-transformation, and found to be significantly different. This means that the assumption of constant correlation across all patients is not appropriate. It also appeared that patients with more severe episodes tended to display greater correlations between episodes compared with patient with milder episodes; and patients whose episodes were clustered temporally showed higher correlation than patients whose episodes were more spread out.

Conclusions

The primary analyses for onset and duration were based on logistic regression on ordered categorical data (grouping time points), including investigator (33 binary terms) and severity (linear effect) with a GEE

model (constant correlation between episodes and patients) to incorporate multiple episodes per patient. This is an extremely complex model, with many assumptions and interlocking parts. The study report did not adequately address the validity of these assumptions. Further, the study report did not assess the robustness of the model compared to similar models or to other analytic methods. The review of the initial NDA noted in the conclusion section: "The results across the three studies appear to be **analysis method dependent.**" (emphasis added) We cannot conclude any differently in this case. Therefore, the determination of this review is that efficacy was not conclusively established and thus approval of Pepcid Complete is not recommended.

/S/ for

Michael Elashoff, PhD

5/24/2000

/S/ : 6/2/00

Concur: Dr Permutt

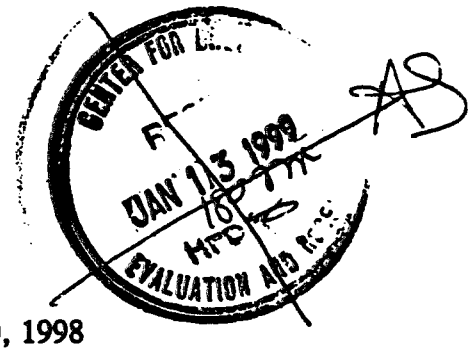
STATISTICAL REVIEW AND EVALUATION
(Addendum)

NDA #: 20-958

JAN 12 1999

Drug: Nonprescription Pepcid —— (Famotidine/Antacid Combination Tablet).

Indication: Treatment of intermittent heartburn symptom relief.



Sponsor: Merck Research Laboratories

Clinical Reviewer: John Senior, M.D.

Statistical Reviewer: Mushfiqur Rashid, Ph.D.

Documents Reviewed: Volumes 1- 2, 18-23, Dated February 20, 1998

User Fee Due Date: February 20, 1999.

This addendum addresses the onset of adequate relief endpoint in study # 110. (After Table 4.4 on page 27)
This reviewer also performed Fisher's exact test for onset of adequate relief up to 30 minutes. The results summarized in the Table 4.4a below indicate there is a statistical FACT advantage over placebo and Famotidine alone for onset of adequate relief by 30 minutes.

Table 4.4a (Reviewer's): Number of Episodes Adequately Relieved Up to and Including 30-Minutes

Adequately Relieved	FACT (Total # of Episodes relieved =1205)	Famot (Total # of Episodes relieved 1229)	Antacid (Total # of Episodes relieve= 1212)	Plac (Total # of Episodes relieved =1217)	p-value FACT Vs Famot	p-value Famot Vs Plac	p-value FACT Vs Plac
Number (%)	544 (45.14)	464 (37.75)	491 (40.51)	401 (32.94)	.002	.0142	<.0001

Note: Famot:Famotidine; Plac: Placebo

Note that the treatment effect size is only 7.4% for the FACT versus Famotidine comparison alone.

This reviewer also performed Fisher's exact test for onset of relief based on first per patient episode. The results are summarized in the following table.

Table 4.4b (Reviewer's): Number of Patients Whose First Episode were Adequately Relieved at 30-Minute Time Point

Adequately Relieved	FACT (n=306)	Famot (n=311)	Antacid (n=308)	Plac (n=307)	p-value	p-value	p-value
					FACT Vs Famot	Famot Vs Plac	FACT Vs Plac
Number (%)	142 (46.14)	103 (33.1)	124 (40.3)	85 (27.7)	.0010	.1618	<.0001

Note: Famot:Famotidine; Plac: Placebo

Note that this analysis is equivalent to a per-patient based analysis. We see from the above table that FACT is statistically significantly superior to placebo and Famotidine alone for onset of adequate relief. However, Famotidine alone is not significantly better than placebo. The treatment effect size for the FACT versus Famotidine comparison is 13.0%

/S/ 4/17/99

M. Mushfiqur Rashid, Ph.D.
Mathematical Statistician

Concur:
Dr. Sankoh /S/
Dr. Welch

cc: Archival NDA # 20958
HFD - 180
HFD - 180/ Dr. Talarico
HFD - 180/ Mr. Folkendt
HFD - 180/ Dr. Senior
HFD - 715/ Dr. Nevius
HFD - 715/ Dr. Welch
HFD - 715/ Dr. Sankoh
HFD - 715/ Dr. Rashid
HFD - 715/ File Copy
Rashid/x73121/MMR/

CSO | Folkendt

STATISTICAL REVIEW AND EVALUATION

NDA #: 20-958

JAN - 8 1999

Drug: Nonprescription Pepcid ——— (Famotidine/Antacid Combination Tablet).

Indication: Treatment of intermittent heartburn symptom relief.

Sponsor: Merck Research Laboratories

Clinical Reviewer: John Senior, M.D.

Statistical Reviewer: Mushfiqur Rashid, Ph.D.

Documents Reviewed: Volumes 1- 2, 18-23, Dated February 20, 1998

User Fee Due Date: February 20, 1999.

1. INTRODUCTION

Heartburn is a common symptom for which a variety of treatments exist. Single doses of Antacid alone and Famotidine 10 mg alone relieve heartburn more effectively than placebo. Although both agents are believed to act by reducing intralumoinal acidity, their mechanisms of action and pharmacodynamic profiles differ substantially. Antacids are believed to work rapidly by neutralizing intralumoinal acid on contact. Their duration of action is limited by physiologic clearing mechanisms. Famotidine reduces gastric acid production via competitive antagonism of the histamine H_2 receptor. Famotidine 10 mg is believed to require a longer time to onset of pharmacodynamic effect than Antacid, but Famotidine has appreciably longer duration of effect than Antacids. These differences suggest that a combination of Famotidine and Antacid in one tablet would potentially offer the benefits of more rapid relief of symptoms than Famotidine alone, and a longer duration of heartburn relief than Antacid alone.

This submission addresses the efficacy and safety of the Famotidine/Antacid Combination Tablet (FACT) for use as an over the counter drug product for the treatment of heartburn, acid indigestion, and sour stomach. FACT is a chewable tablet that contains Famotidine (10 mg), calcium carbonate (800 mg) and magnesium hydroxide (165 mg). The amount of Antacid in each tablet provides 2mEq of acid-neutralizing capacity (ANC), which is within the range of doses typically used on OTC Antacid products for the treatment of intermittent heartburn. This is primarily derived from and supported by a clinical program that consisted of nine clinical studies, including three (Protocols 106, 109 and 110) large phase III clinical trials.

Key Words: Combination therapy, GEE, heartburn, logistic regression, non-prescription,

These three large, double blind, placebo controlled studies, comprising 3645 randomized patients, form the basis for demonstrating the efficacy of Pepcid —— in offering the benefits of more rapid relief of symptoms than Famotidine alone, and a longer duration of heartburn relief than Antacid alone in a single chewable tablet.

1. 1 STUDY PROTOCOLS 106, 109 and 110

The objectives of the three Phase III studies were identical: to determine whether FACT has a faster onset of symptom control than Famotidine, and to determine whether FACT provides a longer duration of relief than Antacid alone. The studies were not designed to determine precisely when onset occurred but rather to show relative differences in onset.

Three different heartburn models were employed (see Table 1.1). As mentioned earlier, all three trials were randomized, double blind, double-dummy, multi-center, factorial, parallel design with four equal-sized treatment groups: FACT, Famotidine, Antacid, and placebo. Study medication was administered with 60 ml (2 oz) of water to facilitate swallowing of the film-coated tablet (FCT). The three studies enrolled similar patient populations: patients aged 18 years or older who reported experiencing heartburn at least three times per week that was generally relieved with Antacids or nonprescription acid reducers.

Table 1.1: Descriptions of the Three Phase III Studies

Protocol Numbers	Reference Numbers	Number of Sites	Planned Sample Size/Group	Description
106	[C-19]	5	300	Single-dose daytime study where patients dosed for spontaneous heartburn (before 3 P.M.) and recorded relief for 8 hours. Standard meal was eaten 4 hours after dosing.
109	[C-20]	10	275	Single-dose evening provocative meal study with onset evaluated in-clinic and duration assessed over night at home.
110	[C-21]	5	300	Multiple (four)-dose study where patients dosed for spontaneous heartburn and recorded relief 8-hours after each dose.

2. Daytime heartburn Study (Protocol 106)

2.1 Description

The Daytime heartburn study is a double-blind, multicenter (five centers) randomized, parallel group study comparing the onset and duration of symptom relief with Famotidine/Antacid combination, Famotidine 10 mg, Antacid 21 mEq, and placebo in patients with frequent heartburn. The primary objective of this study is (1) to determine whether Famotidine/Antacid combination has a faster onset of symptom control than Famotidine alone and (2) to determine whether Famotidine/Antacid combination has a longer duration effect than Antacid alone.

In this five-site study, patients were randomized and given medication, a timer, two diaries, and their choice of a standardized meal (frozen pizza or lasagna, cola, and brownie). Patients were instructed to take study medication to treat heartburn that occurred spontaneously before 3 P.M. They then rated their heartburn every 15 minutes for 2 hours, then at less frequent intervals through 8 hours post-dose. Since uncontrolled food or beverage intake during the 8-hour period could confound analysis of duration of effect, patients were instructed not to eat anything during that period except the meal provided as part of the study. They were instructed to eat the study meal at 4 hours post-dose. Global evaluations were recorded immediately before the meal was eaten and at the end of the 8 hour period.

In the following table we describe patient disposition.

Table 2.1 Patients Dispositions

Population	FACT	Famotidine	Antacid	Placebo	Total
Entered	317	316	315	318	1266
Patients Treated	309	311	306	311	1237
All Patients Treated	306	308	304	307	1225

Of the 1266 patients who were randomized, 26 did not experience spontaneous heartburn within allotted time and did not medicate, 3 withdrew from the study without dosing, 9 were lost to follow up, 2 did not return their diary cards, and one admitted to falsifying her efficacy data. The sponsor did not describe the disposition of these 41 patients among the four treatment groups. The remaining 1225 patients were termed as all patients treated and were assumed to be the ITT population. Twenty-one of the 1225 were described protocol violators by the sponsor. The remaining 1204 non-protocol violators constituted the evaluable patient population.

Demographic characteristics are summarized in Table A.1 in the Appendix. Study patients were predominantly female (61.5%). The treatment groups did not differ significantly in gender and age distribution. No notable differences in other baseline characteristics (e.g. race, baseline heartburn and average number of episodes per week) across treatment groups were evident.

Patient Selection:

A: Inclusion Criteria

Male and female patients who are at least 18 years of age with a history of food-induced heartburn of at least 2 months duration with at least three episodes per week. Patients must have used Antacids or OTC acid reducers for effective relief of their heartburn.

B: Exclusion Criteria:

Patients were excluded from participation in this study if they met any of the following:

a) Had a history of a serious medical condition or evidence of impaired renal function;

b) Had a history of duodenal ulcer, gastric ulcer, atrophic gastritis or divertcultis within the 2 years prior to start of the study; history of upper gastro intestinal tract surgery or vagotomy, esopphageal strictures, Barrett's esophagus, endoscopically identified erosive esophagatis of moderate or greater severity, Zollinger-Ellison syndrome, inflammatory bowel disease, or was known to have gallstones;

c) Were pregnant or lactating. Women of child bearing potential were instructed to use adequate means of contraception;

d) Recent use (within 1 week of entering the study), or continued use of sucralfate, cisapride, metoclopramide, misoprostol, or Rx does of nizatidine, cimitidine, ranitidine, or Famotidine. In addition, chronic use of orally administered corticosteriods, tricyclic antidepressants, anticholinergics, anticoagulants, and antineoplastics were prohibited;

e) Recent use (within 1 week of the treatment session) of OTC H2-receptor antagonist. If the patient used these for the relief of heartburn, patients discontinued the usage for 1 week prior to the treatment session and replaced with Antacid usage to (but not including) the day of study session;

f) Recent use (within 4 weeks of screening visit) of omeprazole or lansoprazole;

g) Had received any form of tetracycline;

h) Had a history of drug or alcohol abuse, psychosis, or other condition that made the patient unlikely to comply with the protocol;

i) Had previously participated in a heartburn study (within 3 months prior to the start of this study);

j) Had used an investigational drug within 30 days prior to start of this study or within five half-lives of the investigational drug, which ever was longer;

k) Had a prior adverse reaction to Antacids, H2 antagonists, or any of the components of the study medication or a prior adverse reaction to any ingredient(s) of the test meals;

l) Other conditions that would interfere with data interpretation or create under risk;

Study Endpoints:

Efficacy:

Primary Endpoints:

There are two primary endpoints in this protocol: (1) heartburn relief and (2) duration of heartburn severity during 4 to 8 hours post-dose.

(1) Heartburn Relief Endpoint:
Proportion of patients with adequate relief after 30 minutes of post-dose.

Adequate relief was assessed by the patients answering the following question each time they evaluated their heartburn severity during the 4-hour evaluation period after taking the study drug: "Do you have adequate relief of your heartburn at this time?"
1=Yes; 2=No.

Note that a 'No' response means that the patient's heartburn severity score is equal to 0.

Patients were instructed to select the grade that most accurately reflected their perception of degree of discomfort since the immediately preceding assessment. Patients used the four-point scale below to assess their perception of heartburn discomfort:

Grade	Severity
0=	None
1=	Mild
2=	Moderate
3 and 4 =	Severe

(2) Duration Endpoint:
Proportion of patients with severe heartburn (peak heartburn severity) during 4 to 8 hours of post-dose. Note that heartburn severity was assessed during 4 to 8 hours post dose.

Secondary Endpoints

(1) Time to 1-grade reduction in heartburn discomfort

(2) Global evaluation of treatment conducted 8 hours after dosing.

Safety:

(1) Proportion of patients with one or more adverse events.

An adverse experience is defined as any unfavorable and unintended change in the structure, function, or chemistry of the body, or worsening of a preexisting condition, temporally associated with any use of a Merck product whether or not considered related to the use of the product.

(2) Proportion of patients who experienced serious clinical adverse events.

A serious adverse experience is one which:

a) Results in death;
b) Is immediately life-threatening;
c) Results in permanent or substantial disability;
d) Results in or prolongs an existing inpatient hospitalization;
e) Is a congenial anomaly;
f) Is cancer; Or
g) Is the result of an overdose.

Screening, Randomization and Sample Size Determination:

Screening:

Patients provided a medical history including a detailed gastrointestinal history. The study coordinator discussed with the patient the nature of the study, its requirements, and restrictions. Patients satisfying all the inclusion criteria signed informed consent and received instructions, a diary card, a 2-oz measuring cup, a timer, study medication and a prepared meal, dessert, and beverage.

Patients were stratified according to choice of meal (pizza or lasanga) to ensure balanced distribution of each among treatment groups. Patients who pre selected pizza were sequentially assigned to the lower numbers in consecutive decreasing order.

Sample Size Determination:

A sample of size 300 patients per treatment groups (total 1200) was planned to

1) detect a 13-percentage point increase in proportion of patients with adequate symptom relief from the FACT group compared to the Antacid 21 mEq group (38% VS 25%) with at least 93% power and a type I error rate (two-sided) of .05;
2) detect a 19-percentage point increase from the FACT compared to the Antacid group (83% vs. 64%) with at least 99% power and a type I error rate (two-sided) of .05.

The method to used to determine sample size is the normal approximation to the binomial distribution.

Diet and Other Activities

Food and drink intake were as specified earlier. A standard meal (test meal) was provided for the patients for the home assessments phase. No other food or drink other than water was to be consumed during the assessment period following study medication and after the test meal. One cup of coffee was permitted immediately after the meal, but at no other time during the 8-hour assessment period.

If the patient was a smoker, smoking was permitted during the assessment period according to the patients' normal habits.

If the patient was a smoker, smoking was permitted during the assessment period according to the patient's normal habits.

The patient was to remain awake and was not to lie down for the full 8-hour assessment period. It was therefore important that the episode of spontaneous heartburn that preceded the home assessment phase occurred sufficiently early in the day. Study medication was therefore to be taken before 3 P.M.

Test Meal:

The test meal consisted of a pre-selected standard provocative meal. A choice of two Supermarket ready prepared meals (pizza and lasanga), a brownie, and 12 oz of cola were provided. Patients could eat for up to 30 minutes at a rate consistent with their usual eating habits. Patients were encouraged to consume at least 8 oz of the meal and 12 oz of drink. If not fully consumed, an estimate of beverage volume consumed was made and recorded on the diary card. No other food, except the test meal was allowed during the 8-hour assessment period. Other than water, patients were not allowed to drink any liquids. One cup of coffee was permitted immediately after the meal, but at no other time during the 8-hour assessment period.

Rescue Medication:

Patients were informed that rescue medication was available for use before and after the test meal if needed, although use within 2 hours of taking study medication was discouraged. Two Mylanta Double strength Antacid tablets were administered as rescue medication for continued or recurrent heartburn symptoms for which the patient felt additional relief was essential. If rescue medication was used, the time it was taken was recorded on the diary cards. Immediately prior to use of rescue medication, the severity of symptoms and the appropriate global evaluation was recorded. The sponsor provided commercially available MYLANTA

double strength Antacid tablets to each study site.

2.2 Sponsor's Statistical Analyses Methods/Reviewer's Comments

All patients who took treatment and provided efficacy data are included in the analyses. These patients were called all treated patients and used in primary the analysis. Per Protocol analyses, excluding patients who were major protocol violators, were also performed for the primary efficacy variables. The proportion of patients in each day who were major protocol violators was small, and per protocol analyses yielded very similar results to those of all patients-treated approach. All efficacy analyses were predefined in the protocols and data analysis plan.

The primary treatment comparisons were FACT versus Famotidine for the onset of adequate relief, and FACT versus Antacid for the duration parameter.

Summary of Sponsor's Analysis Results:

Onset of Relief Results:

As mentioned earlier that the null hypothesis for the onset adequate relief is the same proportion of patients report adequate heartburn relief at 30 minutes post-dose between FACT - FACT and the Famotidine treatment groups.

The protocol specified the use of logistic regression method as the primary analysis method. It should, however, be noted that the sample size for this study was not determined using the logistic model. It was determined by simple difference in proportions of adequate relief. It is therefore likely that the study is somewhat oversized for the more sensitive logistic model. According to the protocol, the sponsor performed an analysis of binary data (e.g. proportion of patients with adequate relief at 30-minute post-dose) using logistic regression models with treatment groups (four) and centers in the model. In the following table we summarize the pair-wise odds ratios. It was found that there were no significant differences among the four treatment groups.

Table 2.2 (sponsor's): Model Adjusted Odds Ratios (n=1215) from Sponsor's Table C-19, from Volume 2.

Minutes Post-dose	Treatment Comparisons	Model Adjusted Odds-Ratios	Chi-square	p-value	% of Patients with Adequate Relief			
					FACT	Famotidine	Antacid	Placebo
30	FACT Vs Famotidine	1.11 (.74,1.65)	.24	.621	.21.4	19.7	23.3	18.4
	Famotidine Vs Placebo	.89(.60,1.30)	.38	.539				
	FACT Vs Placebo	1.62(.93, 2.82)	2.86	.091				
	FACT Vs Antacid	1.30 (.77,2.20)	.94	.333				
	Antacid Vs Placebo	1.24 (.70, 2.22)	.55	.460				

There was no significant difference between FACT and Famotidine in producing relief at the protocol specified primary time-point 30-minute post dose. However, significant differences between FACT treated group and Famotidine treated group were observed only at 45-minute post-dose (secondary time-point). Note that 10 observations were deleted from the logistic regression analysis because of missing values. The sponsor did not impute the missing values. Six patients responded assessments "Not done" at 30–minute post-dose. The sponsor replaced "Not done" by "No" response.

This reviewer performed a sensitivity test via the Fisher's exact test (more appropriate for the way the trial was sized) at 30-minute post-dose by assuming responses corresponding to missing observations were failures. Table 2.3 displays the analyses of proportion of patients with adequate relief at the primary time point (at 30 minutes post-dose). These analysis results indicate no FACT advantage over Famotidine alone. Compared to Antacid alone, FACT is numerically worse than Antacid.

Table 2.3 (reviewer's): Proportion of Patients with Heartburn Relief at 30-Minute Postdose

Onset of Action Parameters	FACT (n=306)	Famotidine (n=308)	Antacid (n=304)	Placebo (n=307)	p-value	p-value	p-value
					FACT Vs Famotidine	Famotidine Vs Placebo	FACT Vs Placebo
Number (%) of Patents with Adequate Relief at	64 (20.92)	60 (19.48)	70(23.03)	56 (18.24)	.688	.757	.417

Furthermore, neither active treatment was significantly better than placebo.

Duration of Relief Results:

The null hypothesis for the "duration" of adequate relief is that the distributions of peak heartburn responses during the four hours (4-8 hours post-dose) following the test meal are the same for the FACT and the Antacid treatment groups.

According to the protocol, the sponsor performed an analysis of ordered categorical data (e.g. peak heartburn severity: 0=none; 1; mild; 2=moderate and 4= severe) using logistic regression models (with treatment groups and centers in the model) for ordinal data. The time range was used as 15-minute after post-dose, 2 hours post-dose (114 minutes to 150 minutes post-dose), 3 hours post-dose (151 minutes to 210 minutes post-dose) and at end of the 8 hours post-dose. The model adjusted odds ratios are summarized in the following table.

Table 2.4 (sponsor's): Model Adjusted Odds-Ratios

Treatment Comparison	Model-Adjusted Estimate (difference)	Model-Adjusted Odds-Ratio (95% CI)	p-value (odds-ratio)
FACT vs. Famotidine	.215	1.24 (0.93, 1.66)	0.145
FACT vs. Antacid	.344	1.41 (1.06, 1.89)	0.020
FACT vs. Placebo	.329	1.39 (1.04, 1.85)	0.027
Famotidine vs. Placebo	.113	1.12 (0.84, 1.49)	0.450
Antacid vs. Placebo	-.020	0.98 (0.74, 1.31)	0.897
Famotidine vs. Antacid	.131	1.14 (0.85, 1.52)	0.377

Patients receiving FACT experienced significantly lower severe symptom severity peak as compared to patients receiving Antacid (p-value .020) and patients receiving placebo (p-value .027). The odds- ratios indicate that FACT patients were 1.41 and 1.39 times more likely to report lower symptom severity peak than Antacid and placebo patients, respectively. Note again that Famotidine 10 mg is not shown superior to placebo.

This reviewer performed a sensitivity analysis by the Fisher's exact test when heartburn severity scores are 0 (none), 1 (mild), 2 (moderate), and 3 and 4 (severe) by replacing the missing values of the response as severe condition. The results are summarized in the following table.

Table 2.5 (reviewer's): Observed (%) Heartburn Severity 4 to 8 hours Post-Dose

	FACT (n=306)	Famotidine (n=308)	Antacid (n=304)	Placebo (n=307)	p-value		
					FACT Vs. Famotidine	FACT Vs. Antacid	FACT Vs. Placebo
None	114 (37.25)	94 (30.52)	84 (27.63)	86 (28.01)	.088	.012	.016
Mild	99 (32.35)	114 (37.01)	115 (37.82)	111 (36.15)	.236	.175	.349
Moderate	46 (15.03)	45 (14.61)	41 (13.48)	54 (17.58)	.910	.644	.444
Severe	46 (15.03)	52 (16.88)	63 (20.72)	54 (17.59)	.582	.073	.444

The proportion of patients with no heartburn 4 to 8 hours post-dose was 9.7 percentage points greater with FACT than Antacid. Numerically, fewer FACT patients reported mild heartburn symptom severity score compared with the rest of the treatment groups. Almost equal proportion of patients in all treatment groups reported moderate heartburn severity symptoms. However, numerically fewer FACT patients reported severe heartburn score compared with Antacid.

This reviewer's analysis results indicate the following:

1) Among patients who reported no heartburn, FACT was statistically significantly superior to both placebo and Antacid (i.e., more FACT patients reported no heartburn).
2) Among those who reported mild heartburn, numerically fewer patients were in FACT than in both placebo and Antacid.

Patients Requiring Rescue Medication:

The following table summarizes the proportion of patients who required rescue medication 4 to 8 hours of post-dose.

Table 2.6 (reviewer's): Proportion of Patients Requiring Rescue Medication 4 to 8 Hours Post-dose (n=1225)

	FACT (n=306)		Famotidine (n=308)		Antacid (n=304)		Placebo (n=307)		p-value (FACT Vs Famotidine)	p-value (FACT Vs Antacid)	p-value (FACT Vs Placebo)
	n	%	n	%	n	%	n	%			
Rescue 4 to 8 hours	41	13.4	49	15.9	55	18.1	48	15.6	.425	.120	.492

This reviewer performed the Fisher's exact test for the proportion of patients who required rescue medications during 4 to 8 hours post-dose. There were no significant treatment differences for the proportion of patients who required rescue medication between the FACT treated group and any other treated group.

Subgroup Analyses:

Gender:

The sponsor's analyses results indicated that there was no evidence of treatment-by-gender interaction for the primary endpoints indicating that the treatment effects were consistent for both males and females.

This reviewer's gender group analysis tables for onset and duration parameters are given in the appendix (onset: Table A.2; duration: Table A.3) and are consistent with those by the sponsor.

Age:

The sponsor's analyses results showed no evidence of a treatment by age interaction when patients were classified as age less than or equal to median or greater than median age (see Table A.1 for median ages in different treatment groups), indicating that the treatment effects were consistent across the age group. There were not enough patients aged 65 or older in each efficacy study to analysis of that demographic subgroup.

This reviewer's age group (<65 and >=65) analysis tables for onset and duration parameters are given in the appendix (onset: Table A.4; duration: Table A.5) and are consistent with those by the sponsor.

Race:

There was no evidence of treatment-by-race (Caucasian or non-Caucasian) interaction suggesting that FACT should be equally effective in all races.

This reviewer's racial origin (Caucasian and non-Caucasian) analysis tables for onset and duration parameters are given in the appendix (onset: Table A.6; duration: Table A.7) and are consistent with those by the sponsor.

2.3 Summary of Safety Analyses

In the following table we summarize the patient's adverse experiences.

Table 2.7 (Reviewer's): Adverse Experiences Summary by Treatment Groups

Clinical Adverse Experiences (AEs)	FACT (n=309)		Famotidine (n=311)		Antacid (n=306)		Placebo (n=311)		p-value	p-value	p-value
									FACT Vs Famotidine	FACT Vs Antacid	FACT Vs Placebo
Number (%) of patients	n	(%)	n	(%)	n	(%)	n	(%)			
With 1 or more AEs	20	(6.5)	21	(6.8)	18	5.9)	12	(3.9)	1.0	.867	.151
With drug related AEs	9	(2.9)	6	(1.9)	4	(1.3)	4	(1.3)	.448	.262	.174
With serious AES	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)	1.0	1.0	1.0
Discontinued due to AEs	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)	1.0	1.0	1.0

It is seen that safety event rates between FACT and Antacid, FACT and Famotidine, and FACT and placebo are similar.

2.4 Conclusions:

Onset:

1) For the primary time point 30-minute, the efficacy data in this study indicate no significant (p-value .621) FACT advantage over Famotidine alone regarding the onset of adequate relief.
2) Furthermore, The data indicate no FACT advantage over placebo.

Duration:

3) The efficacy data in this study indicate a FACT advantage over Antacid alone and placebo regarding duration of adequate relief for 4-8 hours post-dose. FACT has a longer duration of effect than Antacid as measures by peak heartburn severity 4 to 8 hours post-dose.

Safety:

4) FACT treated patients are as safe as Antacid, Famotidine and placebo treated patients. The data suggest no significant safety differential among the treatment groups.

3. Evening Heartburn Study (Protocol 109)

3.1 Description

The Evening heartburn study is a randomized, double blind, parallel group, single dose study comparing the efficacy of FACT, Famotidine, Antacid and placebo on symptoms following an in-clinic evening provocative meal.

In this 10-site study, patients with a history of frequent nocturnal heartburn ate a provocative meal (chili and ice tea) at 7 P.M. in the clinic. They then rated their heartburn severity (none, mild, moderate and severe) at 10-minute intervals through 10 P.M. When patients developed heartburn of at least moderate severity before 8:30 P.M., they were randomized and received study medication. They continued to rate their heartburn intensity every 10 minutes and also recorded whether they had adequate relief (yes or no) at each post-dose time point. Patients were discharged to go home at 10 P.M. They were given a diary in which to record any rescue Antacid use, awakenings with heartburn, and a global evaluation.

Table 3.1 Patients Dispositions

Population	FACT	Famotidine	Antacid	Placebo	Total
Entered/Randomized	283	285	284	287	1139
Patients Treated	283	285	284	286	1138
All Patients Treated	282	285	284	286	1137

A total of 1139 patients were randomized to four treatment groups. One patient discontinued from the study before medicating and provide no efficacy data. Of the 1138 evaluable patients for efficacy, one (randomized to FACT treated group) discontinued because of an adverse experience during the in-clinic period and did not provide data for the over night period. The remaining 1137 patients formed the all treated patient population. Thirteen (3 in FACT group, none in Famotidine group, 6 in Antacid group and 4 in placebo group) of the all treated patient population were classified as serious protocol violators and were excluded from the per-protocol analyses.

Demographic characteristics are summarized in Table A.8 in the Appendix. Study patients were predominantly female (58.6%). The treatment groups did not differ significantly in gender distribution. No notable differences in other baseline characteristics (e.g. age, race, and baseline heartburn and average number of episodes per week) were observed across treatment groups.

Patient Selection:

A. Inclusion Criteria:

Male or female patients who are at least 18 years of age with a history of food-induced heartburn of at least 2 months duration with at least three episodes per week, and who use

Antacids or OTC acid reducers. Patients must have frequently experienced episodes of nocturnal heartburn.

B. Exclusion Criteria:
The exclusion criteria are similar to those for study protocol 106.

Primary Endpoints :

Efficacy:

There are two primary endpoints in this protocol: (1) adequate relief, (2) duration of relief.

(1) Adequate relief: The time to adequate symptom relief (defined as the time to the first of two consecutive adequate relief ratings) for "onset of action."

Adequate relief is assessed by patients answered the following question each time they evaluated their heartburn severity:
"Do you have adequate relief of your heartburn at this time?"
1= Yes; 0 = No.

Note that a 'No' response means that the patient's heartburn discomfort score is equal to 0 (defined as follows).

Patients were instructed to select the grade that most accurately reflected their perception of the degree of discomfort when they received study medication and for every 10 minutes for > = 90 minutes post-dose.

Grade Discomfort

0= None.
1= Mild
2= Moderate
3= Severe.

(2) Duration of relief: evaluated by means of the proportion of patients reporting no awakenings with heartburn.

Secondary Endpoints:

(1) Proportion of patients who reported adequate heartburn relief 30 minutes after dosing,
(2) Proportion of patients who reported at least a 1-grade reduction in heartburn severity 30 minutes after dosing.
(3) Proportion of patients who required rescue medication during the night, and

(4) Proportion of patients who assigned a good or excellent overall global evaluation.

Safety:

(1) Proportion of patients with one or more clinical adverse experiences.
(2) Proportion of patients who experienced a serious clinical adverse event.

Patients were questioned at the return visit regarding any adverse experiences:
Adverse experiences were graded as:

None: No Symptoms

Mild: Awareness of sign or symptom but easily tolerated

Moderate: Discomfort enough to cause interference with usual activity

Severe: Incapacitating with inability to work or do usual activity.

The investigator was asked to evaluate any adverse experiences as to their severity, seriousness, relationship to test drug, action taken, and outcome.

This reviewer could not locate the description/definition of adverse event in the protocol

Screening/Randomization and Sample Size Determination:

Sample Size:

With 275 patients per treatment group, this study was designed to have 91% power to detect a hazard ratio of 1.39, indicating 1.39 times greater likelihood of achieving adequate relief in the FACT group compared to the Famotidine group (type I error rate = .05, two tailed). The study was also designed to have 89% power to detect a 13% point increase for FACT compared to Antacid for the percentage of patients who do not awaken with heartburn (type I error rate = .05, two-tailed).

3.2 Sponsor's Statistical Analyses Methods/Reviewer's Comments

The primary alternative hypotheses are:

1) Compared to Famotidine, FACT will produce a faster time to adequate symptom relief after dosing to treat symptoms induced by a provocative evening meal;

2) Compared to Antacid, FACT will produce a smaller proportion of patients who are awaken with heartburn.

Summary of Sponsor's Analysis Results:

Onset of Relief:

The objective is to determine whether FACT has faster onset of symptom control than Famotidine alone. The sponsor used Cox regression model for survival data to compare the time to adequate relief for the four treatment groups. The model included terms for the treatment groups and the sites in PROC PHREG command of SAS. This reviewer also used Cox regression model to obtain the model-based estimates of the treatment differences and risk ratios. In the following we summarize the results from PROC PHREG of SAS.

Table 3.2 (reviewer's): Model-based estimates (treatment differences), Hazard Ratios with 95% Confidence Intervals and Risk Ratios

Treatment Comparison	Hazard Ratio (sponsor's) (95% CI)	Estimates (Model-based)	Chi-square (Wald)	p-value	Risk-ratio
FACT Vs. Famotidine	1.24 (1.02,1.51)	-.2132	4.51626	.0336	.808
FACT Vs. Placebo	1.18 (.97, 1.44)	-.0536	2.5744	.1086	.948
Famotidine Vs. Placebo	.95 (.78, 1.17)	.0224	.1938	.6597	1.023
FACT Vs Antacid	1.08(.89, 1.31)	.0769	.6100	.4370	.9259
Antacid Vs Placebo	1.09 (.90, 1.33)	.0862	.7400	.3890	.9174

The hazard ratio indicates 1.24 times greater likelihood of achieving adequate relief after treatment with FACT than Famotidine alone. There is only significant difference between the FACT treated group and Famotidine treated group. Although the FACT has numerical advantage over placebo, Famotidine has no numerical advantage over placebo.

The Cox regression results indicate the following:

1) FACT is significantly better than Famotidine in achieving adequate relief 30 minutes after treatment with test drug; the data indicate 1.24 (<1.39 protocol specified) hazard ratio (likelihood of achieving adequate relief compared to Famotidine alone).
2) No FACT advantage over placebo.
3) Also Famotidine alone is numerically worse than placebo. This is indicated by the model-base estimate (see Table 3.2)

This reviewer performed the Fisher's exact test for the proportion of patients with heartburn relief at 30-minute post-dose time point. Note that the proportion of patients with heartburn relief at 30-minute post-dose is the primary onset relief parameter in Protocol 106. Table 3.4 displays the analyses of proportion of patients with adequate relief at a secondary time point (30 minutes).