CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

21-551

STATISTICAL REVIEW(S)

2 Page(s) Withheld

/ § 552(b)(4) Trade Secret / Confidential

___ § 552(b)(5) Deliberative Process

___ § 552(b)(5) Draft Labeling

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH
OFFICE OF BIOSTATISTICS
DIVISION OF BIOMETRICS II

STATISTICAL REVIEW AND EVALUATION

Clinical Studies

NDA:	21-551
Name of drug:	HalfLytely® 20 mg/2L Tablets
Applicant:	Braintree Laboratories, Inc.
Indication:	Bowel cleansing prior to colonscopy
Documents reviewed:	NDA volumes 1, 13 – 16, and 24 dated August 16, 2002; W-000BZ Dated 1/26/2003; N-000-BZ dated 1/26/2003 and 2/13/2003.
Project manager:	Ms. Alice Kacuba
Clinical reviewer:	Robert Prizont, M.D.
PDUFA Dates:	June 16, 2003
Statistical reviewer:	Wen-Jen Chen, Ph.D.
Statistics team leader:	Thomas Permutt, Ph.D.
Biometrics division director:	Edward Nevius, Ph.D.

Keywords: NDA review, Clinical studies, Clinical equivalence.

TABLE OF CONTENTS

1.0 EXECUTIVE SUMMARY OF STATISTICAL FINDINGS

1.1 Overview of Clinical Program and Studies Reviewed

The goal of this NDA supplement submission is to support the use of HalfLytely® 20 mg/2L Tablets in bowel cleansing prior to colonscopy. The studied drug HalfLytely® 20 mg/2L Tablets is a reduced volume preparation consisted of 20 mg bisacodyl (four 5 milligram tablets) followed several hours later by 2 liters of NuLytely.

The efficacy and safety for HalfLytely® 20 mg/2L tablets was assessed in four clinical studies: Studies F38-15, F38-23, F38-13/14, and F38-20. Of the four studies, the sponsor indicated that the two studies, F38-15 and F38-23, were considered as supportive studies while the other two studies F38-13/14, and F38-20 were pivotal.

As indicated by the submission, the two principal Studies F38-13/14, and F38-20 were designed to compare the efficacy of studied drug HalfLytely® 20 mg/2L tablets with that of approved drug 4L NuLytely on bowel cleansing prior to colonscopy.

However, noted by this reviewer, Study F38-15 separately compared the efficacy of the two components (bisacodyl and 2L NuLytely) for the studied drug HalfLytely® 20 mg/2L tablets with that of approved drug 4L NuLytely. In the protocol for Study F38-15, the sponsor stated that the reason for conducting this study was that the components of the combination drug should be tested separately to satisfy the regulatory requirements. However, in order to meet the rule for the combination drug stated in code of federal regulations (§ 300.50), the two components should have included in the trial together with HalfLytely® 20 mg/2L tablets and 4L Nutelytely as Dr. Fredd suggested in the meeting minutes on August 25, 1993. Then, demonstrated that the efficacy of 2L+bis superior to that of each of the two components (bisacodyl and 2L NuLytely). Furthermore, as noted by this reviewer, this study was conducted inappropriately. Without providing an interim analysis plan in the protocol, after knowing the non-superiority results on the treatment efficacy comparisons at end of the study, the sponsor expanded the study by increasing sample size from 150 patients to 300 patients to demonstrate the treatment differences. Accordingly, the results of the efficacy comparisons separately for the two components (bisacodyl and 2L NuLytely) reported by Study F38-15 are not providing valuable information to support the equivalent claim for the two treatments, HalfLytely® 20 mg/2L tablets and 4L Nulytely. Study F38-15 is not further reviewed.

In addition, Study F38-23 compared the efficacy of the studied drug HalfLytely® 20 mg/2L tablets (*abbreviated as 2L+bis*) with that of Visicol Tablets using total 60 patients (30 patients per arm). However, as indicated by the meeting minutes on August 25, 1993 and the two principal Studies F38-13/14 and F38-20, the main objective of this submission was to demonstrate the clinical equivalence between 2L+bis and 4L Nulytely. Furthermore, the success rate of 2L+bis in bowel cleansing quality was not superior to that of Viscol (p= 0.34 by Fisher Exact test performed by this reviewer using Per-Protocol patients). Thus, the result from this

study is not considered providing useful information to support the efficacy of 2L+bis superior to other effective drug or the clinical equivalence claim between 2L+bis and 4L Nulytely. However, the non-significant result of comparing the efficacy of 2L+bis versus approved drug Viscol is revisited in the section of "Overall Conclusions and Recommendations". Study F38-23 is not further reviewed. Accordingly, the two main Studies F38-13/14 and F38-20 are the focus of this review.

1.2 Principal Findings

1.2.1 Study F38-13/14

- The clinical equivalence of 2L+bis and 4L NuLytely claimed by the sponsor was incorrectly based upon the non-significant result of testing hypothesis of superiority. To demonstrate the equivalence of the two drugs 2L+bis and 4L NuLytely, the sponsor should have selected an adequate margin (Δ) and shown that the 95% two-sided confidence interval on the difference of two success rates for 2L+bis and 4L NuLytely was included in the interval ($-\Delta$, Δ). Accordingly, the equivalence claim based upon the non-significant result made by the sponsor is not correct.
- In order to minimize the biased assessments in favor of test drug 2L+bis induced by the single blinded trial and the efficacy equivalent analysis criteria, the sponsor should have added another arm 20 mg bisacodyl plus 2 liter placebo liquid (2LPL+bis) to the trial. Then, in order for drug 2L+bis to be approved equivalent to 4L NuLytely, the sponsor should have also demonstrated that the success rate of bowel preparations for the study drug 2L+bis was superior to that of its component 2LPL+bis.
- Due to highly potentially biased assessments in favor of the test drug 2L+bis as indicated by the above comment, the sensitivity analysis demonstrates that the clinical equivalence for the two treatments 2L+bis and 4L NuLytely is not statistically persuasive when 16% is used as margin. However, more problematically, since the sponsor did not provide the information on equivalence margin, the clinical equivalence on the two treatments 4L NuLytely and 2L+bis is difficult to be assessed.
- The efficacy analysis on 2L+bis shows that lower bounds for the two-sided 95% confidence interval on the success rate of 2L+bis in bowel cleansing quality are 0.79 and 0.78 respectively using Per-Protocol and ITT patient populations. Since the assessments on the bowel preparations were potentially biased in favor of the studied drug 2L+bis, the true success rate of 2L+bis in bowel cleansing quality is expected to be smaller than 0.79 or 0.78. However, based upon the expertise of the reviewing medical officer Dr. Prizont, the success rate of the placebo is approximately from 0% to 15%. Accordingly, the efficacy of drug 2L+bis can be considered superior to that of placebo. Finally, since the effects of the two components, bisacodyl 20 mg and 2 liter NuLytely, for the studied drug 2L+bis were assessed together, the percentage contributed by individual component to the efficacy of the study drug 2L+bis is not known.

1.2.2 Study F38-20

The first two findings related to the comments on the equivalent claim made by the sponsor and the single blinded trial along with the third finding on this reviewer's sensitivity analysis are the same as that of Study F38-13/14 and are not repeated in this section.

The finding on the efficacy claim for the studied drug 2L+bis is stated below:

- The efficacy analysis on 2L+bis shows that lower bound for the two-sided 95% confidence interval on the success rate of 2L+bis in bowel cleansing quality is 0.70 using Per-Protocol population. Since the assessments on the bowel preparations were potentially biased in favor of the studied drug 2L+bis, the true success rate of 2L+bis in bowel cleansing quality is expected to be smaller than 0.70. However, based upon the expertise of the reviewing medical officer Dr. Prizont, the success rate of the placebo is approximately from 0% to 15%. Accordingly, the efficacy of drug 2L+bis can be considered superior to that of placebo. Finally, since the effects of the two components, bisacodyl 20 mg and 2 liter NuLytely, for the studied drug 2L+bis were assessed together, the percentage contributed by individual component to the efficacy of the study drug 2L+bis is not known.

1.3 Overall Conclusions and Recommendations

- ✓ To demonstrate the equivalence of the two drugs 2L+bis and 4L NuLytely, instead of using the non-significant result of superiority hypothesis testing, the sponsor should have selected an adequate margin (Δ) and shown that the 95% two-sided confidence interval on the difference of two success rates for 2L+bis and 4L NuLytely was included in the interval $(-\Delta, \Delta)$. Accordingly, the equivalence for the two treatments 4L NuLytely and 2L+bis claimed by the sponsor is not correct.
- ✓ In order to minimize the biased assessments in favor of test drug 2L+bis induced by the single blinded trial and the efficacy equivalent analysis criteria, the sponsor should have added another arm 20 mg bisacodyl plus 2 litter placebo liquid (2LPL+bis) to the trial. Then, in order for the study drug 2L+bis to be approved equivalent to 4L NuLytely, the sponsor should have also demonstrated that the success rate of bowel preparations for the study drug 2L+bis was superior to that of its component 2LPL+bis.
- ✓ Due to highly potentially biased assessments in favor of the test drug 2L+bis as indicated by the above comment, the sensitivity analysis demonstrates that the clinical equivalence for the two treatments 2L+bis and 4L NuLytely is not statistically persuasive when 16% is used as margin. However, more problematically, since the sponsor did not provide the information on equivalence margin, the clinical equivalence on the two treatments 4L NuLytely and 2L+bis is difficult to be assessed.
- ✓ The efficacy analysis on 2L+bis shows that lower bound of the two-sided 95% confidence interval on the success rate of 2L+bis in bowel cleansing quality is around 70% using Per-Protocol patients for the two Studies F38-13/145 and F38-20. Due to biased assessments in favor of the studied drug 2L+bis, the true success rate of 2L+bis in bowel cleansing quality is expected to be smaller than 0.70. However, based upon the expertise of the reviewing medical officer Dr. Prizont, the success rate of the placebo effect is approximately from 0%

to 15%. Thus, the efficacy of drug 2L+bis can be considered superior to that of placebo. In addition, the superiority of the studied drug 2L+bis to placebo can be further supported by the non-significant result for the comparison on the efficacy of 2L+bis versus that of an approved drug Viscol, reported by Study F38-23. Finally, since the effects of the two components, bisacodyl 20 mg and 2 liter NuLytely, for the study drug 2L+bis were assessed together, the percentage contributed by individual component to the efficacy of the study drug 2L+bis is not known.

2.0 STATISTICAL REVIEW AND EVALUATION OF EVIDENCE

2.1 Introduction and Background

In Volume 1 of this NDA submission, the sponsor made the following observations with regard to HalfLytely®:

> Although years of clinical research have been expended to make early detection of colerectal cancer (CRC) a high clinical priority, the disease remains the second leading cause of cancer-related death in the United States. In order to reduce the rate of death by detecting early cancer or premalignant polyps, a Senate committee voted to require all private health insurance plans in the United States to provide coverage for colonescopies and other tests to detect colon cancer in people who are 50 or older or have a high risk of developing the disease.
>
> To prepare the colon for endoscopic exam, current cleaning procedures include the combination of reduced food intake with laxatives, enemas, suppositories, bowel evacuants, or orthograde colonic lavage. Of these, orthograde colonic lavage with Polyethylene Glycol/Electrolyte Solutions (PEG-ELS, GoLYTELY or SF-ELS, NuLYTELY) is the most frequently prescribed preparation. These "preps", consisting of 4 L of solution, are generally uncomfortable for the patient to complete. They often complain of a sense of fullness, nausea, cramping, and vomiting, sometimes of such magnitude that they do not complete the prescribed regimen.
>
> In order to improve the symptoms, Braintree Laboratories developed a new preparation 2L+bisacodyl (Half Lytely® brand of bowel lavage) to reduce the volume of the PEG-ELS by 50% (to 2L) while retaining successful colonoscopic visualization.

The goal of this NDA supplement submission is to support the use of HalfLytely® 20 mg/2L Tablets in bowel cleansing prior to colonscopy. The studied drug HalfLytely® 20 mg/2L Tablets is a reduced volume preparation consisted of 20 mg bisacodyl (four 5 milligram tablets) followed several hours later by 2 liters of NuLytely. This preparation is denoted as "2L+bis" in this report.

The efficacy and safety for HalfLytely® 20 mg/2L tablets was assessed in four clinical studies: Studies F38-15, F38-23, F38-13/14, and F38-20. Of the four studies, the sponsor indicated that the two studies, F38-15 and F38-23, were considered as supportive studies while the other two studies F38-13/14 and F38-20 were pivotal.

As indicated by the submission, the two studies F38-13/14 and F38-20 were designed to compare the efficacy of studied drug HalfLytely® 20 mg/2L tablets (2L+bis) with that of approved drug 4L

NuLytely on bowel cleansing prior to colonscopy.

However, noted by this reviewer, Study F38-15 separately compared the efficacy of the two components (bisacodyl and 2L NuLytely) for the test drug HalfLytely® 20 mg/2L tablets with that of approved drug 4L NuLytely. In the protocol for Study F38-15, the sponsor stated that the reason for conducting this study was that the components of the combination drug should be tested separately to satisfy the regulatory requirements. However, in order to meet the rule for the combination drug stated in code of federal regulations (§ 300.50), the two components should have included in the trial together with HalfLytely® 20 mg/2L tablets and 4L Nutelytely as Dr. Fredd suggested in the meeting minutes on August 25, 1993. Then, demonstrated that the efficacy of 2L+bis superior to that of each of the two components (bisacodyl and 2L NuLytely). Furthermore, as noted by this reviewer, this study was conducted inappropriately. Without providing an interim analysis plan in the protocol, after knowing the non-superiority results on the treatment efficacy comparisons at end of the study, the sponsor expanded the study by increasing sample size from 150 patients to 300 patients to demonstrate the treatment differences. Accordingly, the results of the efficacy comparisons separately for the two components (bisacodyl and 2L NuLytely) reported by Study F38-15 are not providing valuable information to support the equivalent claim for the two treatments, HalfLytely® 20 mg/2L tablets and 4L Nulytely. Study F38-15 is not further reviewed.

In addition, Study F38-23 compared the efficacy of the test drug HalfLytely® 20 mg/2L tablets (*abbreviated as 2L+bis*) with that of Visicol Tablets using total 60 patients (30 patients per arm). However, as indicated by the meeting minutes on August 25, 1993 and the two principal Studies F38-13/14 and F38-20, the main objective of this submission was to demonstrate the clinical equivalence between 2L+bis and 4L Nulytely. Furthermore, the success rate of 2L+bis in bowel cleansing quality was not superior to that of Viscol ($p= 0.34$ by Fisher Exact test performed by this reviewer using Per-Protocol patients). Thus, the result from this study is not considered providing useful information to support the efficacy of 2L+bis superior to other effective drug or the clinical equivalence claim between 2L+bis and 4L Nulytely. However, the non-significant result of comparing the efficacy of 2L+bis versus approved drug Viscol is revisited in the section of "Overall Conclusions and Recommendations". Study F38-23 is not further reviewed. Accordingly, the two main studies, Studies F38-13/14, and F38-20, are the focus of this review.

The objective for the two Studies, F38-13/14 and F38-20, was to determine the efficacy and safety of a new reduced volume oral gastrointestinal preparation, HalfLytely® 20 mg/2L tablets (2L+bis), as compared to the approved 4-liter dose of NuLytely.

Both clinical trials were multi-center, Phase III, single blind (investigators unaware of preparation method), multi-center (2), randomized, parallel-group, and 2-arm studies with 3-week duration and one day drug exposure. Patients were selected for inclusion if they were 18 years of age or older and required colonoscopy for routine indications. Patients were excluded if they had acute surgical abdomen, appendicitis, bowel obstruction or perforation, gastric outlet: obstruction or paresis, toxic colitis or toxic megacolon. For each study, 200 qualified patients

were enrolled and randomized into two treatment groups (4-liter NuLytely and HalfLytely) by two study centers, (Mobile, Alabama) and (Rochester, Minnesota) for Study F38-13/14 and (San Antonio) and a private practice in Braintree, Massachusetts for Study F38-20.
The course for both trials was pre-study era, 1 day prior to colonoscopy, and examination day. In the pre-study period, patients signed consent form. On the day prior to colonoscopy, enrolled patients reported to the GI lab to be weighed and for blood draw. The blood was analyzed for CBC, hematology, chemistry and osmolarity. Then, subjects were randomly assigned to a treatment group, 4L NuLytely or HalfLytely 20 mg/2L tablets (2L+bis), and issued the appropriate medication and patient questionnaire. Patient questionnaires were for patients to record symptoms and tolerance to the method of preparation. Symptoms of fullness, cramps, nausea, vomiting and overall discomfort were scored on a five point scale: 1 = "none"; 2 = "mild"; 3 = "bothersome"; 4 = "distressing"; and 5 = "severely distressing".

On the day of the examination, after returning NuLytely bottles and questionnaires, patients were weighed and blood was again taken for analysis. Then, the physicians were to perform colonscopy on patients and scored the bowel cleansing consequences. Bowel cleansing was scored by the colonoscopist on a four point scale: 1 - "poor"; 2 - "fair" 3 - "good"; 4 - "excellent". Scores of 3 and 4 were considered "successful" and score of 1 or 2, "unsuccessful". Patients who were unable to tolerate their preparation and led to not being examined (due to lack of bowel cleansing) were also scored as "unsuccessful". The colonoscopist also scored each preparation as to whether in their opinion the bowel preparation was clinically adequate (yes/no).

The sample size was estimated (recorded at page 108 of Volume 1.14) based upon an efficacy criterion of "successful" preparation (which includes good and excellent preparations) rather than "excellent" only preparations. In previous studies, about 30% of the GI lavage preparations were rated as excellent and 50 % as good. Since both ratings were considered successful, the success rate for the 4-liter preparations was 80%. To detect a 20% difference in this success rate (or a difference of 16% from 80%) would require 100 patients per treatment group or 200 enrolled patients, calculated using one-tailed test at a significance level of 0.05 and a power of 0.80 for a one-sided test.

The primary efficacy variable of the study was the bowel preparation categorized as "successful" and "unsuccessful" outcomes converted from the bowel cleansing score rated by physician. For the two Studies F38-13/14 and F38-20, the primary endpoint (success or failure) and investigator ratings of clinical adequacy (yes or no) were analyzed by chi-square tests while the symptoms of fullness, cramps, nausea, vomiting and overall discomfort were analyzed by ridit method. Furthermore, for Study F38-13/14 bowel cleansing examination scores rated by physician (poor, fair, good, and excellent) were evaluated using ridit analysis while, for Study F38-20, this variable was analyzed by chi-square tests.

Based on the sponsor's submission, it is noted that in the ridit analysis, for each variable or measurement in the experimental group (reduced volume preparation), a mean ridit is calculated using the control group (4-liter preparation) as the reference group. The mean ridit of the control

group is conventionally set at 0.5. Then, the standard normal Z score calculated as follows was used to test the null hypothesis that the mean ridit of experimental group is equal to that of control group:

$$Z = |r - .5|/SE(r); \text{ and } SE(r) = 1/(2 *(3N)^{1/2})$$

Where r is the mean ridit of experimental group and SE(r) is the standard error of r.

Finally, in the document (N-000-BZ dated 2/13/2003) responded to this reviewer's information request letter, the sponsor indicated that the intent to treat (ITT) population included all patients that received study medication while the per-protocol (PP) population included only patients who received study medication and completed a colonoscopy. However, the sponsor did not specify which population would be used as a primary efficacy analysis.

2.2 Statistical Evaluation of Evidence on Efficacy/Safety

2.2.1 Detail Review for Individual Studies

2.2.1.1 Study F38-13/14

Demographic Data

As indicated by the sponsor in Volume 1.14, 200 consenting adult subjects were randomized to receive either treatment 4L NuLytely (R) or 2L+bis. However, fourteen patients did not complete the protocol. Accordingly, efficacy analysis was based upon 186 patients except as noted below. Of the 14 incomplete patients, three were excluded for non-compliance with the protocol; seven were for personal reasons; and four were unable to complete their preparations due to an expected adverse event (nausea or-vomiting). The latter 4 patients were included for "Intent-to Treat" analyses where noted. In addition, two hundred (200) available subjects were included in the safety-related analysis and laboratory data analysis where possible.

The study included 95 female and 105 male study subjects with an average age of 58 years. Table 2.2.1.1.1, extracted from the sponsor's Table 4 in Volume 1.14, presents the demographic data by center.

Table 2.2.1.1.1 Demographic data by center

Center	Mean Age	Mean Weight	Male	Female	Race AA	Race C
1 (Mobile)	54.8	180.5	46	54	23	77
2 (Rochester)	61.9	180.2	59	41	0	100
Totals	58.4	180.4	105	95	23	177

AA = African American, C = Caucasian

The study centers were similar in their proportion of female to male patients. However, as indicated in Table 2.2.1.1.1, there was a difference in the average age of the study populations between the two centers: the mean age in center 2 about 7 years older than that in center 1 ($p<0.01$). In addition, study center 1 enrolled 23% African American patients whereas study center 2 enrolled none, reflecting underlying regional differences. There were no between site differences with respect to pre-study body weights.

Efficacy analysis Results and Conclusions

The results for the efficacy comparisons of the two treatment groups, 4L NuLytely versus 2L+bis, on the bowel preparation success (primary endpoint) and the physician rating of colon cleansing quality are presented by Table 2.2.1.1.2 and Table 2.2.1.1.3, respectively. For Table 2.2.1.1.2, The analysis was performed using Intent-to-treat (ITT) patient population while for Table 2.2.1.1.3, the Per-protocol (PP) patient population was used.

Table 2.2.1.1.2 (Sponsor's) Bowel preparation Success by treatment group (ITT)

Rating	4 L NuLytely	2L+bis
Successful	90% (86)	86% (81)
Unsuccessful	10% (10)	14% (13)

Chisq-value=0.52, p=0.47; 90% CI of 4L NuLytely = 90% ± 7.2%.

Table 2.2.1.1.3 (Sponsor's) Physician rating of colon cleansing quality by treatment group (PP)

Rating	4L NuLYTELY	2L+bis
4 (excellent)	51.6% (48)	45.2% (42)
3 (good)	40.9% (38)	41.9% (39)
2 (fair)	5.4% (5)	10.8% (10)
1 (poor)	2.1% (2)	2.2% (2)
Mean Rating	3.4	3.3
Mean Ridit	0.5	0.542
SE	-	0.030
z	-	1.4
p	-	0.16

In Table 2.2.1.1.2, of the total number of 190 patients (ITT patient population), 186 patients (PP patient population) completed their bowel preparations and four patients were not able to complete due to adverse events and were classified as unsuccessful (3 received 4L Nulytely and 1 received 2L+bis). The result from Table 2.2.1.1.2 indicated that there was no statistically significant difference between the two treatments, 4L NuLytely versus 2L+bis, on bowel preparation success ($p=0.47$). The sponsor further indicated that the observed successful rate 86% of 2L+bis fell within the two-sided 90% confidence interval calculated using successful rate 90% of 4L Nulytely. The sponsor intended to use this result to support the equivalence of the two drugs. However, the 86% was only an estimate for the true success rate of 2L+bis. Obviously, the estimated success rate fell within the two-sided 90% confidence interval of 4L NuLytely did not guarantee that the true success rate of 2L+bis would also fall within it. Actually, the clinical equivalent analysis tries to determine if the absolute difference between the two true success rates of 4L NuLytely and 2L+bis is less than a selected margin (for detail, refer to section of

Reviewer's comments and analysis at page 13). As a consequence, it is utterly inappropriate to apply this method to the clinical equivalent analysis.

Similarly, from the result of Table 2.2.1.1.3, the sponsor indicated that there was no statistically significant difference between the two treatment groups in terms of the quality of bowel cleansing (p=0.16).

Finally, mainly based on the non-significant results for the treatment comparisons between 4L NuLytely and 2L+bis on bowel preparation success, the sponsor claimed that the cleansing efficacy of 2L+bis in bowel preparation is equivalent to that of 4L NuLytely bowel preparation.

Adverse Events

All enrolled patients were included in the safety analysis. In this study, patients were exposed to the study medication for about 4 hours. There were three unexpected and 18 expected adverse events occurred during the study. Of the three unexpected adverse events, two events, rectal bleeding during bowel preparation and elevated total iron following preparation, were reported by 2L+bis group and one event, headache, was reported by 4 L NuLytely group. As for the 18 expected adverse events, all involved vomiting: 5 from 2L+bis and 13 from 4L NuLytely.

Finally, for the adverse events scored by the patients on a five-point scale, the sponsor indicated that there were significantly different (using ridit analysis) between the two preparations, 2L+bis and 4L NuLytely, with respect to fullness, nausea, vomiting, and overall discomfort in favor of 2L+bis.

2.2.1.2 Study F-38-20

Demographic Data

As indicated by the sponsor in Volume 1.14, 200 consenting adult subjects were randomized to receive either treatment 4L NuLytely or 2L+bis. However, twelve patients did not complete the protocol. Accordingly, efficacy analysis was based upon 188 patients except as noted below. Of the 12 incomplete patients, five were excluded prior to preparation for non-compliance with the protocol; three patients withdrew from the protocol prior to preparation for personal reason and two withdrew due to illness prior to preparation; and two patients (45 and 173) were unable to complete their preparations due to an expected adverse event (nausea or vomiting). The latter 2 patients were included for "Intent-to-Treat" analyses where noted.

The study included 84 female and 116 male study subjects with an average age of 56 years. Demographic data for each of the study centers is given in the table below.

Table 2.2.1.2.1 Demographic data by center

Center	Mean Age	Mean Weight	Male	Female	Race AA	C	H
3 (SA)	56.7	191.3	85	65	23	60	67
4 (Braintree)	55.1	188.9	31	19	2	47	0
Totals	56.3	190 7	116	84	25	107	67

AA = African American, C = Caucasian, H = Hispanic

The study centers were similar in their proportion of female to male patients. However, as indicated in Table 2.2.1.2.1, there was a difference between study centers with respect to racial makeup where 23 African Americans and 67 Hispanic patients were enrolled in San Antonio versus 2 African Americans and no Hispanic patients enrolled in Braintree reflecting underlying regional differences. There were no site differences with respect to pre-study body weights.

Efficacy analysis Results and Conclusions

The results for the efficacy comparisons of the two treatment groups, 4L NuLytely versus 2L+bis, on the bowel preparation success (primary endpoint) and the physician rating of colon cleansing quality are presented by Table 2.2.1.2.2 and Table 2.2.1.2.3, respectively. For Table 2.2.1.2.2, the analysis was performed using Intent-to-treat (ITT) patient population while for Table 2.2.1.2.3, the Per-protocol (PP) patient population was used. The information presented by these two tables is extracted from Table 5 (Revised) at page 5 of Appendix E in the document (N-000-BZ) dated 1/25/2003. The difference between original Table 5 at page 238 in Volume 1.4.2 and this revised Table 5 is that in the document N-000-BZ, the sponsor indicated that patient 23 having been miscoded receiving 4L NuLytely actually received 2L+bis and had a cleansing score of "Good" in the physician rating.

Table 2.2.1.2.2 (Sponsor's) Bowel preparation Success by treatment group

Rating	4 L NuLytely	2L+bis
Successful	78% (76)	80% (74)
Unsuccessful	22% (21)	20% (19)

Chisq-value=0.52, p=0.58.

Table 2.2.1.2.3 (Sponsor's) Physician rating of colon cleansing quality by treatment group

Rating	4L NuLYTELY	2L+bis
Excellent	36.8% (35)	39.8% (37)
Good	43.2% (41)	39.8% (37)
Fair	18.9% (18)	18.3% (17)
Poor	1.0% (1)	2.2% (2)

χ^2 = 0.60, p = 0.90

Table 2.2.1.2.2 shows the total number of treatment successes and failures in the study. Of the total number of 190 patients (ITT patient population), 188 patients (PP patient population)

completed their bowel preparations and two patients were not able to complete due to adverse events and not examined (both received 4L Nulytely).

The result from Table 2.2.1.2.2 indicated that there was no statistically significant difference between the two treatments, 4L NuLytely versus HalfLytely (2L+bis), on bowel preparation success (p=0.58).

Similarly, from the result of Table 2.2.1.2.3, the sponsor indicated that there was no statistically significant difference between the two treatment groups in terms of the quality of bowel cleansing (p=0.90).

Finally, mainly based on the non-significant result for the treatment comparisons between 4L NuLytely and HalfLytely (2L+bis) on bowel preparation success, the sponsor claimed that the cleansing efficacy of 2L+bis in bowel preparation is equivalent to that of 4L NuLytely bowel preparation.

Adverse Events

All enrolled patients were included in the safety analysis. In this study, patients were exposed to the study medication for about 4 hours. There were two unexpected and 4 expected adverse events occurred during the study. Of the two unexpected adverse events, one event, a neck ache, was reported by 4 L NuLytely group and the other one, headache, was reported by 2L+bis group. As to the four expected adverse events, all involved volume related complaints (nausea, vomiting, etc.): one in the 2L+bis group and 3 in the 4L NuLytely group.

Finally, for the adverse events scored by the patients on a five-point scale, the sponsor indicated that patients prepared with the 2L+bis preparation were noted significantly less (using ridit analysis) fullness and overall discomfort than patients with 4 L NuLytely.

2.2.2 Statistical Reviewer's Findings

2.2.2.1 Study F38-13/14

2.2.2.1.1 Reviewer's comments and analysis

In order to validate the efficacy claim made by the sponsor, this reviewer first comments on the following two issues with regard to the study design: 1) criteria of equivalence analysis, 2) assessments of colon cleansing quality. Then, this reviewer performs the following four analyses on the primary endpoint to assess the equivalence of the two treatments 2L+bis and 4L NuLytely: 1) two-sided 95% confidence interval, 2) sensitivity analysis, 3) efficacy analysis on 2L+bis, and 4) subgroup analysis.

Reviewer's Comments

1.) Issue on the criteria used for equivalence analysis

At page 109 of Volume 1.14 submitted by the sponsor, the criteria used to assess the equivalence between the two treatment groups, 2L+bis and 4L NuLytely, were based on a non-significant result in bowel preparation and the successful rate of the study drug 2L+bis fallen within the 90% confidence interval of the active control treatment 4L NuLytely. In addition, from document W-000BZ (dated 1/26/2003) responded to this reviewer's information request letter, the sample size was determined using 80% power to detect 16% successful rate difference between the two treatment groups (80% for 4L NuLytely versus 56% for 2L+bis) at 0.05 significance level. Thus, according to the above information provided by the applicant, the equivalent conclusion established for the two treatment groups was mainly based on a non-significant result for testing the equality of the successful rates in bowel preparation between the two treatment groups, 4L NuLytely versus 56% for 2L+bis.

From E10, "Guidance for Industry, E10 choice of Control Group and Related Issues in Clinical Trials", one learns that the equivalence trials are designed to demonstrate efficacy of a new drug similar in efficacy to a standard agent. Most of these are actually non-inferiority trials, attempting to show that the new drug is not less effective than the control by more than a defined amount, generally called margin. Furthermore, following the meeting minutes on June 9, 1993, for the equivalence analysis, Dr. Huque commented that the demonstration of equivalence between treatment groups involves rejecting the hypothesis of some difference. Then, the study should be powered accordingly.

Based on the information provided by E10 and commented by Dr. Huque, instead of using non-significant result to claim equivalence of two treatments, the sponsor first should have chosen an adequate delta margin ($\Delta > 0$), following the guidance stated in E10. Then, with the selected margin (Δ), in order to demonstrate clinical equivalence between 4L NuLytely and 2L+bis, the following two null hypotheses formulated by bowel preparation success (primary endpoint) need to be rejected:

$H_{01}: P_{2L+bis} - P_{4L\ NuLytely} \leq -\Delta$ and $H_{02}: P_{2L+bis} - P_{4L\ NuLytely} \geq \Delta$;

where P_{2L+bis} and $P_{4L\ NuLytely}$ are bowel successful rates for 2L+bis and 4L NuLytely, respectively.

Finally, at 0.025 significant level, a 95% two-sided confidence interval on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ can be constructed to test the null hypothesis $H_0: \cup_{i=1,2} H_{0i}$. If the 95% two-sided confidence interval is included in the interval $(-\Delta, \Delta)$, then the sponsor can claim that the efficacy of 2L+bis is equivalent to that of 4L NuLytely.

As a result, to demonstrate the equivalence of the two drugs 2L+bis and 4L NuLytely, the sponsor should have selected an adequate margin (Δ) and shown that the 95% two-sided

confidence interval on the difference of two success rates for 2L+bis and 4L NuLytely was included in the interval (-Δ, Δ). It follows that the equivalence claim based upon the non-significant result made by the sponsor is not correct.

To further assess the clinical equivalence for the two treatments, 2L+bis and 4L NuLytely, this reviewer performs the 95% two-sided confidence intervals on the difference of two success rates for 2L+bis and 4L NuLytely and presents them in the section of "Reviewer's Analysis".

2) Issue on the assessments of colon cleansing quality

Based on the sponsor's study design, the biased assessments on the colon cleansing quality are very possibly induced by i.) nature of single blinded design and ii.) defect of efficacy equivalent analysis criterion.

i.) Issue on the single blinded design: as indicated by the sponsor, this trial was conducted by single blinded study in which investigators were blinded as to the methods of preparation. However, since patients knew which drugs were used for their bowel preparations, the investigators were easy to recognize the bowl preparation drug used by patients. Therefore, in reality, the single blinded trial was highly potential to be an open label trial. Furthermore, noted by this reviewer, the definitions of "fair" (enough feces or fluid to prevent a completely reliable exame) and "good" (small amounts of feces or fluid not interfering with the exame) in bowel cleansing quality are not clearly cut and may be assessed subjectively. Accordingly, as long as the investigator realized which drug was used by the patient, the assessment on the successful bowel preparation (scored as "excellent" and "good" by investigators) was likely to be biased in favor of study drug 2L+bis.

In the section of Fairness of comparisons (1.4.3) in E10, it states that for the comparative trial to be informative concerning relative safety and/or efficacy, the trial needs to be fair; ie., the conditions of the trial should not inappropriately favor one treatment over the other. In order to avoid the potential biased assessments, the sponsor should have included another arm 20mg bisacodyl plus 2 litter placebo liquid (2LPL+bis) in this trial as suggested by Dr. Fredd in the industrial meeting on June 9, 1993. Then, in order for drug 2L+bis to be approved, the sponsor should have also demonstrated that the success rate of bowel preparations for the combination drug 2L+bis was superior to that of its component 2LPL+bis.

In reality, due to different appearances shown by the two treatments, 2L+Bis and 4L NuLytely, it may be difficult for the sponsor to conduct a double blinded trial. However, the concerns on the issues induced by such trials can not be ignored and the biased conditions induced by the nature of single blinded trial may be improved if another arm 2LPL+bis had been included in this trial.

ii.) Issue on the efficacy equivalent analysis criterion: Based on the efficacy equivalent analysis criteria, one notes that if the outcomes of the bowel preparations for the two treatment

groups, 2L+bis and 4L NuLytely, are assessed as similar/comparable as possible then the two drugs will be claimed equivalent. As indicated in the above sub-section i.), due to the ambiguous definition on the scores "good" and "fair" of the bowel cleansing quality, the bowel preparation quality might not be assessed objectively. Therefore, with only two arms 2L+bis and 4L NuLytely in the trial, it was very probably for the investigator to assign similar scores to the bowel preparations for the two treatment groups. As long as investigators assessed the outcomes of the bowel preparations for the two treatment groups as close as possible, the chance of the efficacy equivalence for the two drugs is greatly increased. However, the equivalence of the two treatment groups established by the above assessments is a biased result. To avert the bias, the section (§ 314.126) of adequate and well-controlled studies in code of federal regulations recommends including additional treatment groups such as dose-comparison control. Thus, as commented by this reviewer in the above sub-section i.), in order to prevent the potentially biased assessments, the sponsor should have included another arm 2LPL+bis in this trial.

Reviewer's Ananlysis

1) Two-sided 95% confidence interval

As indicated by Comment 1.), at 0.025 significant level, if the 95% two-sided confidence interval on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ is included in the interval $(-\Delta, \Delta)$, then the efficacy equivalence between 2L+bis and 4L NuLytely is supported. However, since the sponsor did not select a margin for the equivalence analysis, the 95% two-sided confidence interval on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ performed by this reviewer using both ITT and Per-Protocol populations only tries to understand the closeness of the two proportions. Table 2.2.2.1.1 presents the asymptotic results of the two-sided 95% confidence intervals on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$.

Table 2.2.2.1.1 (Reviewer's) 95% two-sided confidence intervals on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$

	4L NyLytely		2L +bis		95% Confidence Interval on
Patient Population	No. Success	Success Rate (n/N)	No.Success	Success Rate (n/N)	$P_{2L+bis} - P_{4L\ NuLytely}$
Per-Protocol Population	86	0.93 (86/93)	81	0.87 (81/93)	(-0.15, 0.004)
Intent-to-Treat Population	86	0.90 (86/96)	81	0.86 (81/94)	(-0.13, 0.006)

Table 2.2.2.1.1 indicates that for Intent-to-Treat population, the two preparations, 2L+bis and 4L NuLytely, shows efficacy equivalence if margin 14% was selected while for Per-Protocol population, the margin needed 16% in order to demonstrate the efficacy equivalence between the two preparations. Since unlike Intent-to-Treat patients, all Per-Protocol patients completed the endoscopy examination, the result from Per-Protocol patient population is more reliable/realistic than that of Intent-to-Treat population. Thus, the sensitivity analysis performed by this reviewer is based on Per-Protocol Population.

2) Sensitivity analysis

As commented on the issue of the assessments on the colon cleansing quality, the assessments of colon cleansing quality were very likely biased in favor of 2L+bis based on the sponsor's design. In order to evaluate the impact of the potentially biased bowel quality assessments on the treatment comparisons, this reviewer performs a sensitivity analysis by sequentially reversing one failed patient in the 4L NuLytely group to success and simultaneously, reversing one successful patient in the 2L+bis group to failure. Then, at each sequential step, the 2-sided 95% confidence interval on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ is calculated using Per-Protocol patient population. The results were presented in Table 2.2.2.1.2.

Table 2.2.2.1.2 (Reviewer's) Sensitivity analysis on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ using Per-Protocol patient population

4L NyLytely		2L +bis		2-sided 95% Confidence Interval
No. Success	Success Rate (n/N)	No. Success	Success Rate (n/N)	on $P_{4L\ NuLytely} - P_{2L+bis}$
86	0.93 (86/93)	81	0.87 (81/93)	(-0.15, 0.04)
87	0.94 (87/93)	80	0.86 (80/93)	(-0.17, 0.01)
88	0.95 (88/93)	79	0.85 (79/93)	(-0.19, -0.01)

Table 2.2.2.1.2 indicates that the lower bound of the 2-sided 95% confidence interval on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ was dropped by .02 if only one more patient in 4L Nutely reversed as success and one more in 2L+bis reversed as failure. Now, if 16% was selected as the margin, the efficacy of 2L+bis shows inferior to that of 4L NuLytely as long as one more patient in 4L Lytely group and one more in 2L+bis corrected as success and failure, respectively. Due to highly potentially biased assessments in favor of the test drug 2L+bis, it may be fair to reverse one patient in 4L NuLytely group from failure to success and one in 2L+bis group from success to failure, when performing the equivalent analysis. However, more problematically, since the sponsor did not use the equivalent margin to perform the equivalent analysis, the information on the margin was not provided. Because of no information on the margin provided, the clinical equivalent assessment on the two treatments 4L NuLytely and 2L+bis is difficult to be performed.

3) Efficacy analysis on 2L+bis

As noted by the above sub-section, the clinical equivalent claim between the two treatments 4L NuLytely and 2L+bis can not be established due to lack of information on equivalent margin. In order to determine if the test drug 2L+bis has efficacy superior to placebo, in this sub-section, this reviewer performs the two-sided 95% confidence interval on the success rate of 2L+bis using both Per-Protocol and ITT patient populations. Table 2.2.2.1.3 presents the results.

Table 2.2.2.1.3 (Reviewer's) 95% two-sided confidence intervals on P_{2L+bis}

	2L +bis		95% Confidence Interval on
Patient Population	No.Success	Success Rate (n/N)	P_{2L+bis}
Per-Protocol Population	81	0.87 (81/93)	(0.79, 0.93)
Intent-to-Treat Population	81	0.86 (81/94)	(0.78, 0.92)

Table 2.2.2.1.3 shows the lower bounds for the two-sided 95% confidence intervals on the success rate of bowel cleansing quality are 0.79 and 0.78 respectively for Per-Protocol and ITT patient populations. Since as noted by this reviewer, the assessments on the bowel preparations were potentially biased in favor of the test drug 2L+bis, the true success rate of 2L+bis may be lower than the one 0.87 or 0.86 estimated using the sponsor' data. Accordingly, the lower bound of the 95% two-sided interval for 2L+bis calculated using the data from more reliable population is expected to be smaller than the one 0.79 or 0.78 presented in Table 2.2.2.1.3. However, based upon the expertise of the reviewing medical officer Dr. Prizont, the success rate of the placebo effect is approximately from 0% to 15%. Accordingly, the efficacy of study drug 2L+bis can be considered superior to that of placebo. Since the effects of the two components, bisacodyl 20mg and 2 liter NuLytely, for the test drug 2L+bis were assessed together, the percentage contributed by individual component to the efficacy of the study drug 2L+bis was not known.

4) Subgroup analysis

In order to assess the consistency of the treatment effect of 2L+bis versus 4L NuLytely across subgroups, this reviewer performed the subgroup analysis on the success rate in bowel cleasning quality using Per-Protocol patient population. The subgroups analyzed are Gender (Male and Female), Race (Caucasian and non-Caucasian), and Age group (age $\leq$ 65 and age > 65).

Gender

Table 2.2.2.1.4 presents the two treatment success rates in bowel cleansing quality along with the two-sided 95% confidence intervals on the success rate difference between 2L+bis and 4L NuLytely, while Table 2.2.2.1.5 presents success rate and the two-sided 95% confidence interval for 2L+bis by gender.

Table 2.2.2.1.4 (Reviewer's) 95% two-sided confidence intervals on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ Using Per-Protocol patients by Gender

	4L NyLytely		2L +bis		95% Confidence Interval on
Gender	No. Success	Success Rate (n/N)	No.Success	Success Rate (n/N)	P_{2L+bis} - $P_{4L\ NuLytely}$
Male	51	0.91 (51/56)	39	0.87 (39/45)	(- 0.17, 0.08)
Female	35	0.95 (35/37)	42	0.88 (42/48)	(-0.19, 0.05)

Table 2.2.2.1.5 (Reviewer's) 95% two-sided confidence intervals on P_{2L+bis} by Gender

	2L +bis		95% Confidence Interval on
Gender	No.Success	Success Rate (n/N)	P_{2L+bis}
Male	39	0.87 (39/45)	(0.73, 0.95)
Female	42	0.88 (42/48)	(0.75, 0.95)

Table 2.2.2.1.4 indicates that the lower bound of the two-sided 95% confidence interval on the success rate difference for female is 2 percents lower than that of male. Intuitively, the success rate differences between the two treatments can be considered stable across male and female subgroups.

Table 2.2.2.1.5 shows that the two lower bounds 0.73 and 0.75 of 2L+bis respectively for male and female subgroups are higher than the effect range of placebo (success rate approximately from 0% to 15%). Therefore, the effect of 2L+bis is superior to that of placebo for both males and females.

Race

Since almost 90% of patients recruited were Caucasian (White), no subgroup analysis on race is performed.

Age group

Table 2.2.2.1.6 presents the two treatment success rates in bowel cleansing quality and the two-sided 95% confidence intervals on the success rate difference between 2L+bis and 4L NuLytely, while Table 2.2.2.1.7 presents success rate and the two-sided 95% confidence interval for 2L+bis by age group.

Table 2.2.2.1.6 (Reviewer's) 95% two-sided confidence intervals on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ Using Per-Protocol patients by age group (age ≤ 65 and age > 65)

	4L NyLytely		2L +bis		95% Confidence Interval on
Age group	No. Success	Success Rate (n/N)	No.Success	Success Rate (n/N)	$P_{2L+bis} - P_{4L\ NuLytely}$
Age ≤ 65	64	0.91 (64/70)	59	0.87 (59/68)	(-0.15, 0.06)
Age > 65	22	0.96 (22/23)	22	0.88 (22/25)	(-0.23, 0.08)

Table 2.2.2.1.7 (Reviewer's) 95% two-sided confidence intervals on P_{2L+bis}

	2L +bis		95% Confidence Interval on
Age group	No.Success	Success Rate (n/N)	P_{2L+bis}
Age ≤ 65	59	0.87 (59/68)	(0.76, 0.94)
Age > 65	22	0.88 (22/25)	(0.69, 0.97)

Table 2.2.2.1.6 indicates that the lower bound of the two-sided 95% confidence interval on the success rate difference for patients with age greater than 65 is 8 percents lower than that of patients with age less than 65. It indicates that the success rate difference between 4L NuLytely and 2L+bis for senior patients may be smaller than that of patients with age less than 65 years.

However, Table 2.2.2.1.7 shows that the two lower bounds 0.76 and 0.69 of 2L+bis respectively for patients with age less than 65 and patients with age greater than 65 are higher than the effect range of placebo (success rate approximately from 0% to 15%). Therefore, the effect of 2L+bis is superior to that of placebo shown by both age groups.

2.2.2.1.2 Remarks

Firstly, as commented by this reviewer, the clinical equivalence of 2L+bis and 4L NuLytely claimed by the sponsor was incorrectly based upon the non-significant result of superiority hypothesis testing. To demonstrate the equivalence of the two drugs 2L+bis and 4L NuLytely, the

sponsor should have selected an adequate margin (Δ) and shown that the 95% two-sided confidence interval on the difference of two success rates for 2L+bis and 4L NuLytely was included in the interval (-Δ, Δ). Accordingly, the equivalence claim based upon the non-significant result made by the sponsor is not correct.

Secondarily, due to the nature of the single blinded trial and the efficacy equivalent analysis criterion, the assessments for colon cleansing quality was very likely biased in favor of the test drug 2L+bis. In order to minimize the biased assessments, the sponsor should have added another arm 20mg bisacodyl plus 2 liter placebo liquid (2LPL+bis) to the trial, as recommended by Dr. Fredd in the industrial meeting on June 9, 1993. Then, in order for 2L+bis to be approved equivalent to 4L NuLytely, the sponsor should have also demonstrated that the success rate of bowel preparations for the study drug 2L+bis was superior to that of its component 2LPL+bis.

Thirdly, due to highly potentially biased assessments in favor of the tested drug 2L+bis, it may be fair to reverse one patient in 4L NuLytely group from failure to success and one in 2L+bis group from success to failure when performing equivalent analysis. After the corrections, the clinical equivalence for the two treatments 2L+bis and 4L NuLytely is not established when 16% was used as margin. However, more problematically, since the sponsor did not provide the information on equivalence margin, the clinical equivalence on the two treatments 4L NuLytely and 2L+bis is difficult to be assessed.

Finally, the efficacy analysis on 2L+bis shows that lower bounds for the two-sided 95% confidence interval on the success rate of 2L+bis in the bowel cleansing quality are 0.79 and 0.78 respectively using Per-Protocol and ITT patient populations. Since the assessments on the bowel preparations were potentially biased in favor of the test drug 2L+bis, the true success rate of 2L+bis in bowel cleansing quality is expected to be smaller than 0.79 or 0.78. However, based upon the expertise of the reviewing medical officer Dr. Prizont, the success rate of the placebo is approximately from 0% to 15%. Accordingly, the efficacy of drug 2L+bis can be considered superior to that of placebo. Finally, since the effects of the two components, bisacodyl 20 mg and 2 liter NuLytely, for the test drug 2L+bis were assessed together, the percentage contributed by individual component to the efficacy of the study drug 2L+bis is not known.

2.2.2.2 Study F38-20

2.2.2.2.1 Reviewer's comments and analysis

It is noted that the trial design and the efficacy analysis method for this study were similar to that of Study F38-13/14. Thus, alike to Study F38-13/14, to validate the efficacy claim made by the sponsor, this reviewer first comments on the following two issues with regard to the study design: 1) criteria of equivalent analysis, 2) assessments of colon cleansing quality. Then, this reviewer performs the following three analyses on the primary endpoint using both ITT and/or Per-Protocol populations to assess the equivalence of the two treatments 2L+bis and 4L NuLytely: 1) two-sided 95% confidence interval, 2) sensitivity analysis, 3) efficacy analysis on

2L+bis, and 4) subgroup analysis.

Reviewer's Comments

As indicated by the sponsor at page 227 of Volume 1.14, Study F38-20 was the secondary efficacy trial recommended by FDA in the second industrial meeting at October 5, 1998 to be performed in support of the proposed indication. It followed that the trial design and the efficacy analysis method employed by this study were the same as that used for Study F38-13/14. Therefore, the comments on the issues of equivalent analysis and assessments of colon cleansing quality made by this reviewer for this study were the same as that made for Study F38-13/14. For detail of these two comments, refer to the sub-section of Reviewer's Comments for Study F38-13/14 at page 13.

Reviewer's Ananlysis

1) Two-sided 95% confidence interval

As the logic stated in the sub-section of two-sided 95% confidence interval for Study F38-13/14, the 95% two-sided confidence interval on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ performed by this reviewer using both ITT and Per-protocol populations tries to understand the closeness of the two proportions. Table 2.2.2.2.1 presents the asymptotic results of the two-sided 95% confidence intervals on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$.

Table 2.2.2.2.1 (Reviewer's) 95% two-sided confidence intervals on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$

	4L NyLytely		2L +bis		95% Confidence Interval on
Patient Population	No. Success	Success Rate (n/N)	No.Success	Success Rate (n/N)	$P_{2L+bis} - P_{4L\ NuLytely}$
Per-Protocol Population	76	0.80 (76/95)	74	0.80 (74/93)	(-0.12, 0.11)
Intent-to-Treat Population	76	0.78 (76/97)	74	0.80 (74/93)	(-0.11, 0.13)

Table 2.2.2.2.1 indicates that for Intent-to-Treat population, the two preparations, 2L+bis and 4L NuLytely, shows efficacy equivalence if margin 12% was selected while for Per-Protocol population, the margin needed 13% in order to demonstrate the efficacy equivalence between the two preparations. Since unlike Intent-to-Treat patients, all Per-Protocol patients completed the endoscopy examination, the result from Per-protocol patient population is more reliable/realistic than that of Intent-to-Treat population. Thus, the sensitivity analysis performed by this reviewer is based on Per-Protocol Population.

2) Sensitivity analysis

As commented on the issue of the assessments on the colon cleansing quality, the assessments of colon cleansing quality were very likely biased in favor of 2L+bis based on the sponsor's design. In order to evaluate the impact of the potentially biased bowel quality assessments on the treatment comparisons, this reviewer performs a sensitivity analysis by sequentially reversing one failed patient in the 4L NuLytely group to success and simultaneously, reversing one successful

patient in the 2L+bis group to failure. Then, at each sequential step, the 2-sided 95% confidence interval on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ is calculated using Per-Protocol patient population. The results were presented in Table 2.2.2.2.2.

Table 2.2.2.2.2 (Reviewer's) Sensitivity analysis on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ using Per-Protocol patient population

4L NyLytely		2L +bis		2-sided 95% Confidence Interval
No. Success	Success Rate (n/N)	No. Success	Success Rate (n/N)	on $P_{4L\ NuLytely} - P_{2L+bis}$
76	0.80 (76/95)	74	0.80 (74/93)	(-0.12, 0.11)
77	0.81 (77/95)	73	0.78 (73/93)	(-0.14, 0.09)
78	0.82 (78/95)	72	0.77 (73/93)	(-0.16, 0.07)

Table 2.2.2.2.2 indicates that the lower bound of the 2-sided 95% confidence interval on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ was dropped by .02 if one more patient in 4L Nutely reversed as success and one more in 2L+bis reversed as failure. Now, if 16% was selected as the margin, the efficacy of 2L+bis shows inferior to that of 4L NuLytely as long as two more patients in 4L NuLytely group and two more in 2L+bis group corrected as success and failure, respectively. Due to highly potentially biased assessments in favor of the tested drug 2L+bis, it may be fair to reverse two patient in 4L NuLytely group from failure to success and two in 2L+bis group from success to failure when performing equivalent analysis. However, more problematically, since the sponsor did not use the equivalent margin to perform the equivalent analysis, the information on the margin was not provided. Because no information on the margin was provided, the clinical equivalent assessment on the two treatments 4L NuLytely and 2L+bis is difficult to be performed.

3) Efficacy analysis on 2L+bis

As noted by the above sub-section, the clinical equivalent claim between the two treatments 4L NuLytely and 2L+bis can not be established due to lack of information on delta margin. In order to determine if the test drug 2L+bis has efficacy superior to placebo, in this sub-section, this reviewer performs the two-sided 95% confidence interval on the success rate of 2L+bis using both Per-Protocol and ITT patient populations. Since the success rate for ITT and Per-Protocol populations are equal, only Per-protocol result is presented. Table 2.2.2.2.3 presents the results.

Table 2.2.2.2.3 (Reviewer's) 95% two-sided confidence intervals on P_{2L+bis}

	2L +bis		95% Confidence Interval on
Patient Population	No.Success	Success Rate (n/N)	P_{2L+bis}
Per-Protocol Population	74	0.80 (74/93)	(0.70, 0.87)

Table 2.2.2.2.3 shows the lower bound for the two-sided 95% confidence intervals on the success rate of bowel cleansing quality is 0.70 calculated using Per-Protocol population. Since as noted by this reviewer, the assessments on the bowel preparations were potentially biased in favor of the test drug 2L+bis, the true success rate of 2L+bis may be lower than 0.80 estimated using the sponsor' data. Accordingly, the lower bound of the 95% two-sided interval for 2L+bis calculated using the data from more reliable population is expected to be smaller than the one 0.70

presented in Table 2.2.2.2.3. However, based upon the expertise of the reviewing medical officer Dr. Prizont, the success rate of the placebo is approximately from 0% to 15%. Accordingly, the efficacy of drug 2L+bis can be considered superior to that of placebo. Finally, since the effects of the two components, bisacodyl 20 mg and 2 liter NuLytely, for the test drug 2L+bis were assessed together, the percentage contributed by individual component to the efficacy of the study drug 2L+bis was not known.

4) Subgroup analysis

In order to assess the consistency of the treatment effect of 2L+bis versus 4L NuLytely across subgroups, this reviewer performed the subgroup analysis on the success rate in bowel cleasning quality using Per-Protocol patient population. The subgroups analyzed are Gender (Male and Female), Race (Caucasian and non-Caucasian), and Age group (age ≤ 65 and age > 65).

Gender

Table 2.2.2.2.4 presents the two treatment success rates in bowel cleansing quality and the two-sided 95% confidence intervals on the success rate difference between 2L+bis and 4L NuLytely, while Table 2.2.2.2.5 presents success rate along with the two-sided 95% confidence interval for 2L+bis by gender.

Table 2.2.2.2.4 (Reviewer's) 95% two-sided confidence intervals on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ Using Per-Protocol patients by Gender

	4L NyLytely		2L +bis		95% Confidence Interval on
Gender	No. Success	Success Rate (n/N)	No.Success	Success Rate (n/N)	$P_{2L+bis} - P_{4L\ NuLytely}$
Male	49	0.82 (49/60)	43	0.90 (43/48)	(0.05, 0.21)
Female	27	0.77 (27/35)	31	0.69 (31/45)	(-0.28, 0.11)

Table 2.2.2.2.5 (Reviewer's) 95% two-sided confidence intervals on P_{2L+bis} by Gender

	2L +bis		95% Confidence Interval on
Gender	No.Success	Success Rate (n/N)	P_{2L+bis}
Male	43	0.90 (43/48)	(0.77, 0.97)
Female	31	0.69 (31/45)	(0.53, 0.82)

Table 2.2.2.2.4 shows that the lower bound of the two-sided 95% confidence interval on the success rate difference for females is negative 28 percents while for males, the lower bound is positive 5 percents. It indicates that the gender may be qualitatively interacted with the success rate differences between the two treatments. Thus, it may not be appropriate to perform the clinical equivalent analysis for the two treatments, 2L+bis versus 4L NuLytely, using data pooled across gender.

However, Table 2.2.2.2.5 shows that the two lower bounds 0.77 and 0.53 of 2L+bis respectively for male and female subgroups are higher than the effect range of placebo (success ratc approximately from 0% to 15%). Therefore, the effect of 2L+bis is superior to that of placebo for both males and females.

Race

Table 2.2.2.2.6 presents the two treatment success rates in bowel cleansing quality and the two-sided 95% confidence intervals on the success rate difference between 2L+bis and 4L NuLytely, while Table 2.2.2.2.7 presents success rate and the two-sided 95% confidence interval for 2L+bis by race (Caucasian and non-Caucasian).

Table 2.2.2.2.6 (Reviewer's) 95% two-sided confidence intervals on the difference of P_{2L+bis} and $P_{4L\ NuLytely}$ Using Per-Protocol patients by age group (age ≤ 65 and age > 65)

Race	4L NyLytely No. Success	4L NyLytely Success Rate (n/N)	2L +bis No.Success	2L +bis Success Rate (n/N)	95% Confidence Interval on $P_{2L+bis} - P_{4L\ NuLytely}$
Caucasian	41	0.87 (41/47)	49	0.91 (49/54)	(-0.09, 0.16)
Non-Caucasian	35	0.73 (35/48)	25	0.64 (25/39)	(-0.28, 0.11)

Table 2.2.2.2.7 (Reviewer's) 95% two-sided confidence intervals on P_{2L+bis}

Race	2L +bis No.Success	2L +bis Success Rate (n/N)	95% Confidence Interval on P_{2L+bis}
Caucasian	49	0.91 (49/54)	(0.79, 0.97)
Non-Caucasian	25	0.64 (25/39)	(0.47, 0.79)

Table 2.2.2.2.6 shows that the lower bound of the two-sided 95% confidence interval on the success rate difference for Non-Caucasian is 19 percents lower than that of Caucasian. It indicates that the success rate differences between the two treatments may not be homogeneous across Caucasian and Non-Caucasian.

However, Table 2.2.2.2.7 shows that the two lower bounds 0.79 and 0.47 of 2L+bis respectively for Caucasian and Non-Caucasian subgroups are higher than the effect range of placebo (success rate approximately from 0% to 15%). Therefore, the effect of 2L+bis is superior to that of placebo for both Caucasian and Non-Caucasian.

Age group

Since almost 80% of patients recruited were younger than 65 years old, no subgroup analysis on age group is performed.

2.2.2.2.2 Remarks

The first two comments on the equivalent claim made by the sponsor and the nature of the single blinded trial are the same as that of Study F38-13/14 stated in the section of "Remarks" at page 19 and are not repeated in this section.

Thirdly, due to potentially biased assessments in favor of the tested drug 2L+bis as commented in the section of "Reviewer's comments", it may be fair to reverse two patient in 4L NuLytely group from failure to success and two in 2L+bis group from success to failure. After the

corrections, the clinical equivalence for the two treatments 2L+bis and 4L NuLytely is not established when 16% was used as margin. However, more problematically, since the sponsor did not provide the information on equivalence margin, the clinical equivalence on the two treatments 4L NuLytely and 2L+bis is difficult to be assessed.

Finally, the efficacy analysis on 2L+bis shows that lower bound for the two-sided 95% confidence interval on the success rate of 2L+bis in bowel cleansing quality is 0.70 using Per-Protocol population. Since the assessments on the bowel preparations were potentially biased in favor of the test drug 2L+bis, the true success rate of 2L+bis in bowel cleansing quality is expected to be smaller than 0.70. However, based upon the expertise of the reviewing medical officer Dr. Prizont, the success rate of the placebo is approximately from 0% to 15%. Accordingly, the efficacy of drug 2L+bis can be considered superior to that of placebo. Finally, since the effects of the two components, bisacodyl 20 mg and 2 liter NuLytely, for the test drug 2L+bis were assessed together, the percentage contributed by individual component to the efficacy of the study drug 2L+bis is not known.

2.2.3 Overall Conclusions and Recommendations

From the finding remarks of the two Studies F38-13/14 and F38-20, the Overall Conclusions and Recommendations on the efficacy of study drug 2L+bis are made as follows:

- To demonstrate the equivalence of the two drugs 2L+bis and 4L NuLytely, instead of using the non-significant result of superiority hypothesis testing, the sponsor should have selected an adequate margin (Δ) and shown that the 95% two-sided confidence interval on the difference of two success rates for 2L+bis and 4L NuLytely was included in the interval $(-\Delta, \Delta)$. Accordingly, the equivalence for the two treatments 4L NuLytely and 2L+bis claimed by the sponsor is not correct.
- In order to minimize the biased assessments in favor of test drug 2L+bis induced by the nature of the single blinded trial and the efficacy equivalent analysis criteria, the sponsor should have added another arm 20 mg bisacodyl plus 2 litter placebo liquid (2LPL+bis) to the trial. Then, in order for drug 2L+bis to be approved equivalent to 4L NuLytely, the sponsor should have also demonstrated that the success rate of bowel preparations for the study drug 2L+bis was superior to that of its component 2LPL+bis.
- Due to highly potentially biased assessments in favor of the test drug 2L+bis as indicated by the above comment, the sensitivity analysis demonstrates that the clinical equivalence for the two treatments 2L+bis and 4L NuLytely is not statistically persuasive when 16% is used as margin. However, more problematically, since the sponsor did not provide the information on equivalence margin, the clinical equivalence on the two treatments 4L NuLytely and 2L+bis is difficult to be assessed.
- The efficacy analysis on 2L+bis shows that lower bound of the two-sided 95% confidence interval on the success rate of 2L+bis in bowel cleansing quality is around 70% using Per-Protocol patients for the two Studies F38-13/145 and F38-20. Due to biased assessments in favor of the studied drug 2L+bis, the true success rate of 2L+bis in bowel cleansing quality is

expected to be smaller than 0.70. However, based upon the expertise of the reviewing medical officer Dr. Prizont, the success rate of the placebo is approximately from 0% to 15%. Thus, the efficacy of drug 2L+bis can be considered superior to that of placebo. . In addition, the superiority of the studied drug 2L+bis to placebo can be further supported by the non-significant results for the comparison on the efficacy of 2L+bis versus that of an approved drug Viscol, reported by Study F38-23. Finally, since the effects of the two components, bisacodyl 20 mg and 2 liter NuLytely, for the study drug 2L+bis were assessed together, the percentage contributed by individual component to the efficacy of the study drug 2L+bis was not known.

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

/s/

Wen-Jen Chen
5/12/03 03:23:32 PM
BIOMETRICS

Thomas Permutt
5/12/03 04:29:50 PM
BIOMETRICS
concur