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

STATISTICAL REVIEW AND EVALUATION
CLINICAL STUDIES

NDA/Serial Number:	22101
Drug Name:	Nexium® (esomeprazole magnesium) Delayed-Release Granules
Indication(s):	Short term treatment of gastroesophageal reflux disease (GERD) and healing of erosive esophagitis for patients with ages from one to eleven.
Applicant:	AstraZeneca LP
Dates:	Received: September 27, 2006 PDUFA: July 27, 2007
Review Priority:	Standard
Biometrics Division:	Division of Biometrics III
Statistical Reviewer:	Wen-Jen Chen, Ph.D.
Concurring Reviewer:	Mike Welch, Ph.D., Team Leader
Medical Division:	Gastroenterology Products
Clinical Team:	Wen-Yi Gao, MD. Hugo Gallo-Torres, MD, Team Leader
Project Manager:	Chantal Phillips, LCDR, PHS
Keywords:	Clinical studies; NDA review; Single Study.

TABLE OF CONTENTS

1.0 EXECTIVE SUMMARY OF STATISTICAL FINDINGS

1.1 Conclusions and Recommendations

From a statistical perspective, the single Study D9614C00097 does not provide substantial evidence of efficacy to support the study drug Nexium in use of short term treatment of gastroesophageal reflux disease (GERD) and healing of erosive esophagitis for patients with ages from one to eleven.

However, this submission does satisfy the Agency's Written Request (WR) for a study of pharmacokinetics, safety and clinical outcome for pediatric patients, one to eleven years of age. There was no intent to provide confirmatory evidence of efficacy. With regard to labeling, [redacted]

1.2 Brief Overview of Clinical Studies

In this NDA submission, the applicant has submitted one Phase III Study D9614C00097 to support the use of esomeprazole for short term treatment of gastroesophageal reflux disease (GERD) and healing of erosive esophagitis for patients with ages from one to eleven. This study was submitted in response to a WR filed under NDA 21-153; refer to "Study #4: Pharmacokinetic, Exposure/Response, and Safety Study in Pediatric Patients 1 to 11 Years of Age." dated 12/20/05.

The primary objective of this multi-center, parallel-group study was to evaluate the safety of once-daily esomeprazole treatment in relieving GERD-associated symptoms in pediatric patients, ages one to eleven years, inclusive. A secondary objective of the study was to evaluate the clinical outcome of once daily treatment with esomeprazole in relieving GERD-associated signs and symptoms in pediatric patients with ages one to eleven years, inclusive. One hundred and nine (109) patients were randomized by the study and were included for the clinical outcomes analysis. However, no primary efficacy endpoint was pre-specified since clinical outcome was a secondary objective.

Patients were enrolled at 24 centers located in Belgium, France, Italy, and the US. Patients were stratified based on their weights at screening. Within each of two strata (<20 kg, ≥ 20 kg), patients were randomized on a 1:1 basis to receive one of the two esomeprazole doses: either 5mg or 10 mg for patients with weights less than 20 kg and either 10 mg or 20 mg for patients with weights more than 20 kg. In addition, patients with EE on endoscopy at baseline underwent follow up endoscopic examinations. Follow-up endoscopy allowed for determination of the proportion of patients showing mucosal healing after completion of esomeprazole therapy. Finally, the study duration consisted of seven assessment procedures: Screen Visit 1 (21 days prior to dosing), Randomization Visit 2 (Day 0; dosing day), Visit 3 (Day 14 after dosing), Visit 4 (Day 28 after dosing), Visit 5 (Day 42 after dosing), Visit 6 (Day 56 after dosing), and Follow-up Contact (14 days after last dose). The total treatment period was eight weeks.

1.3 Statistical Issues and Findings

For Study D9614C00097, the following three issues are commented upon: 1) No control arm in study, 2) Severity of erosive esophagitis, and 3) Questionable analysis method.

1) No control arm study

a) Since no comparator was included, investigator knew that all patients took the study drug. In reality, it is an open label study. ICH E9 indicates that blinding is one of the most important design techniques for avoiding bias assessments in clinical trials.
b) Definitions of "Mild" and "Moderate" in physician's global assessments are not completely distinguished in this open label study. Accordingly, the two shortcomings (open label and ambiguous endpoint definition) could induce biased assessments in favor of the study drug.
c) Since no control arm (placebo or active) was included in the study, the efficacy of study drug Nexium can not be objectively assessed by quantitatively comparing the efficacy between Nexium and a concurrent control.
d) ICH Guidance E10 states in effect that baseline controlled studies are really not controlled at all, but implicitly assume an external control or threshold value for efficacy. The validity of the external control is crucial. One approach would be to show the observed changes from Baseline are comparable to those that can be extrapolated from the adult (well controlled) studies.

2) Severity of erosive esophagitis

For patients with esophagitis disease at baseline, less than 4% (2/53) of enrolled subjects with erosive esophagitis had more severe LA grades C and D and the other patients (96%) had mild erosive esophagitis with LA grades A and B. Therefore, due to the small number of severe esophagitis subjects enrolled, the overall healing rates would not properly reflect efficacy of Nexium for pediatric patients with more severe erosive esophagitis at Baseline.

3) Questionable analysis method

a) For the analysis on the physician's global GERD assessment scores, the applicant compared the ordinal categories (none, mild, moderate, and severe) at Baseline versus that of the Final Visit using Cochran-Mantel-Haenszel method for each of the four treatment groups. Since the same patient population was used at Baseline and Final Visit, patient outcomes at Baseline are not independent with those at Final Visit, violating an independence assumption for efficacy comparison between two treatment groups. Accordingly, the analysis method applied by the applicant to analyze the physician's global assessment scores is not legitimate. The p-values generated by the applicant based on dependence data are not correct p-values to assess the improvement from Baseline.

b) For the analysis on the patient diary assessments of GERD symptoms on Heartburn, Acid Regurgitation, and Epigastric Pain, the applicant performed a paired t-test using average of the 7 data points from the Final Visit and one data point at Baseline. Accordingly, the shapes of the distributions for the two components of the paired data are not identical. The p-values generated by the applicant using the paired t-test based upon the paired distributions of un-equal shapes are inadequate to assess the improvement from Baseline.

In order to explore the efficacy of the study drug used in pediatric patients, the reviewer performed the following two analyses on proportions: 1) based upon physician's GERD global assessment score and 2) based upon patient diary assessment scores on GERD symptoms.

1) Proportion analysis based upon physician's GERD global assessment score

a) For the physician's global assessment scores, this reviewer calculated the following two types of proportions for the four treatment groups: proportions of patients with one score improvement from Baseline at Final visit and proportions of patients with two score improvement from Baseline at Final visit.
b) The lower limits of the two-sided 95% confidence intervals for one score improvement are 48%, 43%, 63%, and 60% respectively for treatments I, II, III, and IV.[1] For the two-score improvement, the lower limits are 12%, 5%, 14%, and 10% respectively for treatments I, II, III, and IV.
c) Since this is an open label study, the efficacy assessments of the physician's global assessments may be biased in favor of the study drug. The true proportions for the patients with improved from Baseline may be lower than the ones presented here.
d) Due to the shortcomings of open label and ambiguous definitions of "Mild" and "Moderate", the proportions for two-score improvement may more objectively assess improvement from Baseline.

2) Proportion analysis based upon patient diary assessment scores on GERD symptoms

a) For the patient's diary assessment scores on GERD symptoms, this reviewer calculated the proportions of patients with one score improvement from Baseline and the corresponding two-sided 95% confidence intervals for the GERD symptoms on heartburn, acid regurgitation, and epigastric pain for each of the four treatment groups.
b) For the three GERD symptoms assessed by the patients, the proportions of patients who improved two scores from Baseline at Final visit for treatment II (Esomeprazole 10 mg & Weight <20 kg) are less than that for the other three treatment groups.
c) For the physician's GERD global assessments, the lower bounds of the 95% two-sided confidence intervals for the proportions of patients who improved from Baseline by one score are much higher than that of the three GERD symptoms assessed by patients: range from 9% (eg., Regurgitation in treatment group I) to 33% (eg., Heartburn in treatment

1 I: Esomeprazole 5 mg and Wt < 20 kg II: Esomeprazole 10 mg and Wt < 20 kg III: Esomeprazole 10 mg and Wt ≥ 20 kg IV: Esomeprazole 20 mg and Wt ≥ 20 kg

group II). This may indicate a bias in the physician global assessments in favor of the study drug.

2.0 INTRODUCTION

2.1 Overview

With regard to Nexium (esomeprazole magnesium), the applicant made the following observations:

> Gastroesophageal reflux (GER) is defined as the retrograde passage of gastric contents into the esophagus or extraesophageal regions. It is presumed to be caused by a transient relaxation of the lower esophageal sphincter and is commonly reported in healthy infants and children as a physiologic event. Simple physiologic reflux evolves into pathologic GER or gastroesophageal reflux disease (GERD) when the reflux produces an adverse symptomatology in patients.
>
> Pharmacologic treatment for GERD is aimed at reducing the amount of gastric acid exposure to the esophageal and supraesophageal mucosa. Successful treatment leads to symptom relief and helps to prevent complications of chronic reflux. The proton-transporting enzyme involved in the production of hydrochloric acid in the stomach is known as gastric parietal cell H+/K+-ATPase, or "proton pump." Compounds that inhibit this enzyme are known as Proton Pump Inhibitors, or PPIs. PPIs bind covalently to the proton pump on the apical surface of the gastric parietal cells, irreversibly inhibiting the inward transport of H+ ions by gastric parietal cell H+/K+-ATPase. Esomeprazole, the S-isomer of omeprazole, has the capacity to selectively inhibit this enzyme thereby inhibiting gastric acid production. In the clinical study report (CSR), esomeprazole magnesium will be referred to as esomeprazole.

In this NDA submission, the applicant has submitted one Phase III Study D9614C00097 to support the use of esomeprazole for short term treatment of gastroesophageal reflux disease (GERD) and healing of erosive esophagitis for patients with ages from one to eleven.

The Applicant's submission is in response to a Written Request (WR) filed under NDA 21-153, Study 4: Pharmacokinetic, Exposure/Response, and Safety Study in Pediatric Patients 1 to 11 Years of Age. Amendment #3 (dated 12/20/05) to this WR states that the objectives of this study would be:

(a) To characterize the pharmacokinetic profile of single and repeated doses of esomeprazole magnesium in patients 1 to 11 years of age.

(b) To compare the safety and clinical outcome of pediatric patients 1 to 11 years of age with endoscopically proven GERD across different dosages of esomeprazole magnesium.

(c) To determine the proportion of patients showing endoscopic evidence of healing after completion of therapy across different dosages of esomeprazole magnesium in those pediatric patients 1 to 11years of age who undergo follow-up endoscopy after treatment.

The WR further stipulated that this would be a randomized, double blind, dose-ranging study and the dosages of esomeprazole magnesium would be selected as dosages likely to be therapeutically effective and safe, based on data from the pharmacokinetic component of this study as well as from other studies in pediatric patients and adults. Eligible patients were to be randomized in approximately equal proportions to one of at least two dose levels. After randomization, the overall duration of the trial was to be at least eight weeks. Outcome measures were to be assessed weekly: at clinic visits that occur at least once every other week, as well as by other appropriate means (e.g., telephone questionnaire) during weeks in which no clinic visits are scheduled. At least 40 patients 1 to 5 years of age and 40 patients 6 to 11 years of age were to complete at least 8 weeks treatment.

Thus the primary objective of this multi-center, parallel-group study was to evaluate the safety of once daily treatment with esomeprazole in relieving GERD-associated symptoms in pediatric patients' with ages one to eleven years, inclusive. The secondary objective of the study was to evaluate the clinical outcome of once daily treatment with esomeprazole in relieving GERD-associated signs and symptoms in pediatric patients with ages one to eleven years, inclusive. One hundred and nine (109) patients who met the inclusion and not the exclusion criteria were randomized by the study and were included for the clinical outcomes analysis. However, no efficacy primary endpoint was pre-specified since efficacy assessment is the secondary objective.

Patients were enrolled at 24 centers located in Belgium, France, Italy, and the United States (US). Patients were stratified based on their weight at Screening (<20 kg, ≥ 20 kg). Within each of these two strata, patients were randomized on a 1:1 basis to receive one of the two esomeprazole doses: either 5mg or 10 mg for patients with weights < 20 kg and either 10 mg or 20 mg for patients with weights ≥ 20 kg. In addition, enrolled patients with EE on endoscopy at baseline underwent follow up endoscopic examination. Follow-up endoscopy allowed for determination of the proportion of patients showing mucosal healing after completion of esomeprazole therapy. Finally, the study duration consisted of seven assessment procedures: Screen Visit 1 (21 days prior to dosing), Randomization Visit 2 (Day 0; dosing day), Visit 3 (Day 14 after dosing), Visit 4 (Day 28 after dosing), Visit 5 (Day 42 after dosing), Visit 6 (Day 56 after dosing), and Follow-up Contact (14 days after last dose). Total treatment period is eight weeks.

2.2 Data Sources

This reviewer reviewed "MODULE 5 Clinical Study Reports" submitted by the applicant through electronic system dated Sep 27, 2006 and located at "\\CDSESUB1\N22101\N_000\2006-9-27\Clinstat". Data used by this reviewer's statistical analysis was submitted by the applicant on December 19, 2006.

3.0 STATISTICAL EVALUATION

3.1 Evaluation of Efficacy for Study D9614C00097

Study Design and Endpoints

The primary objective of this multi-center, parallel-group study was to evaluate the safety of once daily treatment with esomeprazole in relieving GERD-associated symptoms in pediatric patients with ages 1 to 11 years, inclusive. The secondary objective of the study was to evaluate the clinical outcome of once daily treatment with esomeprazole in relieving GERD-associated signs and symptoms in pediatric patients with ages 1 to 11 years, inclusive.

The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guideline E1 (ICH-E1 1994), recommends that 100 patients should be included as part of the safety database for a drug. In addition, the Agency issued a Written Request (WR) for pediatric studies that specified enrollment of a minimum of forty 1 to 5 year old patients and forty 6 to 11 year old patients.

The applicant indicated that based on ICH-E1 and the FDA WR, the current study was designed to randomize at least 100 patients in order to ensure that at least 40 patients in each of 2 age groups (1 to 5 years and 6 to 11 years) would complete the study and receive all 8 weeks of treatment.

Patients were enrolled at 24 centers located in Belgium, France, Italy, and the United States (US). Patients were stratified based on their weight (<20 kg, ≥ 20 kg) at Screening. Within each of these two strata, patients were randomized on a 1:1 basis to receive one of the two esomeprazole doses: either 5mg or 10 mg for patients with weights < 20 kg and either 10 mg or 20 mg for patients with weights ≥ 20 kg. One hundred and nine (109) patients who met the inclusion and not the exclusion criteria were randomized by the study and were included for the clinical outcomes analysis. However, of the 109 patients, 108 patients who completed the study were involved for the safety analysis. In addition, forty five enrolled patients with EE on endoscopy at baseline underwent follow up endoscopic examination. Follow-up endoscopy allowed for determination of the proportion of patients showing mucosal healing after completion of esomeprazole therapy.

The study duration consisted of seven assessment procedures: Screen Visit 1 (21 days prior to dosing), Randomization Visit 2 (Day 0; dosing day), Visit 3 (Day 14 after dosing), Visit 4 (Day 28 after dosing), Visit 5 (Day 42 after dosing), Visit 6 (Day 56 after dosing), and Follow-up Contact (14 days after last dose). Total treatment period is eight weeks.

Figure 3.1.1 displayed the design of the study and the sequence of treatment periods.

Figure 3.1.1 (Sponsor's) Study flow diagram

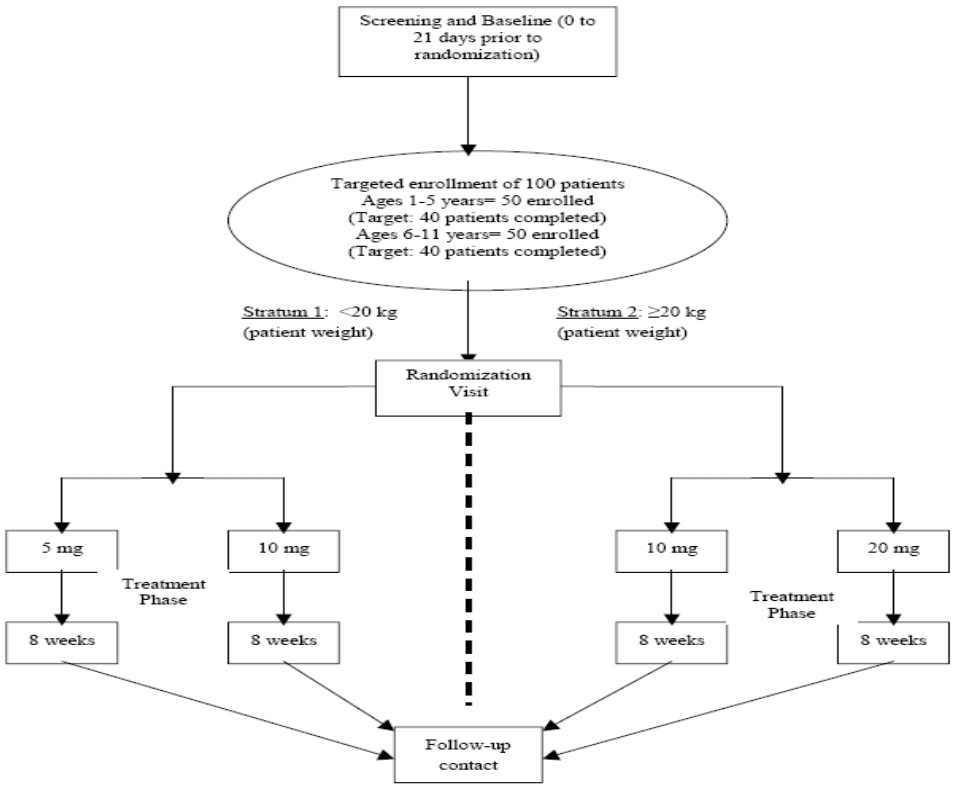


Table 3.1.1 summarized the clinical outcome variables of the study and shows how they relate to the study objective.

Table 3.1.1 (Sponsor's) Efficacy objectives and outcome variables relating to each objective

Objective	Summary outcome variables for analysis
Secondary	Secondary outcome variable
To evaluate the clinical outcome of once daily treatment with esomeprazole in relieving GERD-associated symptoms in pediatric patients ages 1 to 11 years, inclusive.	-changes from baseline in Physician's Global Assessment -changes from baseline in daily patient symptom assessments as reported by parent/guardian -changes from baseline in endoscopic healing of EE

GERD is gastroesophageal reflux disease; EE is erosive esophagitis.

Based upon Table 3.1.1, the applicant indicated that the endpoints assessed in support of this secondary objective were as follows:

• Changes from baseline in the Physician's Global Assessment;
• Mean baseline and weekly patient symptom assessment scores (reported by parent/guardian);

• Proportion of patients with (baseline) EE who exhibited endoscopic evidence of healing after completion of therapy at the follow-up endoscopy. Healing was defined as absence of mucosal breaks in the esophagus.

As indicated by the applicant, the investigators completed the Physician's Global Assessment of patient's symptomatology at Visit 1, Visit 3, Visit 4, Visit 5 and Visit 6. The global assessment instruction was the following.

Please provide overall clinical impression on GERD-related symptoms over the last 7 days as:
• None (no symptoms)
• Mild (symptoms present but not interfering with daily activities)
• Moderate (symptoms present and somewhat interfering with daily activities)
• Severe (symptoms present and greatly interfering or preventing daily activities).

However, no efficacy primary endpoint was pre-specified since efficacy assessment is the secondary objective.

The populations used for analyzing safety and efficacy variables were defined as follows:

The ITT analysis population for the clinical outcomes consisted of patients who had a baseline measurement, at least one post-baseline measurement after randomization, and who took at least one dose of study medication.

Patients were included in the PP population if and only if they completed the study while meeting the conditions described for the ITT population with no major protocol violations or deviations.
All patients who started drug treatment and who had any post-baseline safety information were included in the assessment of AEs (all AEs and treatment-related AEs).

Statistical Methodologies

Measures of clinical outcome included the Physician's Global Assessments, patient symptom assessments (by the parent/guardian), and EE healing rates and other results from endoscopic examinations. Frequency tables of the Physician's Global Assessment at baseline and for each of the on-treatment visits were constructed. Baseline results were then compared to each of the on-treatment results using a chi-square test. Weekly mean scores, from the patient diaries, while on treatment were evaluated against baseline using a paired t-test for each week.

The primary purpose of this clinical trial was to examine the safety of esomeprazole in patients with ages 1 to 11 years, inclusive who had endoscopically proven GERD. Since there are no known safety issues with esomeprazole magnesium, the study was not powered on any specific safety endpoint.

Patient Disposition

In total, 109 patients were randomized in 24 study sites, with no site randomizing more than 14% of all patients. Of these, 101 patients completed the study. The numbers of evaluable patients were 108 patients in the safety population, 109 in the ITT population, and 98 patients in the PP population.
Patients were stratified based on their weight and randomized into 1 of 4 treatment groups: < 20 kg patient weight, 5 mg esomeprazole; < 20 kg patient weight, 10 mg esomeprazole; ≥ 20 kg weight, 10 mg esomeprazole; and ≥ 20 kg weight, 20 mg esomeprazole.

The disposition of patients for the study is summarized in Table 3.1.2.

Table 3.1.2 (Sponsor's) Patient disposition (completion or discontinuation)

	Esomeprazole dose groups									
	5 mg Wt <20 kg (N=26)		10 mg Wt <20 kg (N=23)		10 mg Wt ≥20 kg (N=31)		20 mg Wt ≥20 kg (N=29)		Total (N=109)	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Number randomized	26	(100.0)	23	(100.0)	31	(100.0)	29	(100.0)	109	(100.0)
Evaluable										
Safety population	25[a]	(96.2)	23	(100.0)	31	(100.0)	29	(100.0)	108	(99.1)
ITT population	26	(100.0)	23	(100.0)	31	(100.0)	29	(100.0)	109	(100.0)
PP population	25	(96.2)	22	(95.7)	26	(83.9)	25	(86.2)	98	(89.9)
Completed protocol	24	(92.3)	22	(95.7)	26	(83.9)	29	(100.0)	101	(92.7)
Withdrawals	2	(7.7)	1	(4.3)	5	(16.1)	0		8	(7.3)
Eligibility criteria not met	1	(3.8)	0		0		0		1	(0.9)
Due to AE	0		1	(4.3)	2	(6.5)	0		3	(2.8)
Lack of therapeutic response	0		0		1[b]	(3.2)	0		1	(0.9)
Subject not willing to continue study	0		0		1	(3.2)	0		1	(0.9)
Incorrect randomization	1	(3.8)	0		0		0		1	(0.9)
Other	0		0		1	(3.2)	0		1	(0.9)

a One patient was not evaluable for safety because he did not have any post-baseline safety data.
b Patient E0042009 had "lack of therapeutic response" recorded as reason for withdrawal on termination CRF page.

The applicant indicated that there were 8 early discontinuations from the study; 3 were patients weighing <20 kg and 5 were patients weighing ≥ 20 kg. In the <20 kg weight stratum, 2 patients who withdrew early had received 5 mg and 1 patient had received 10 mg esomeprazole. In the ≥ 20 kg weight stratum, all 5 patients who withdrew early had received the 10 mg dose of esomeprazole.

In this study, there were 3 patients who discontinued due to AEs: one patient in the <20 kg, 10 mg esomeprazole treatment group and 2 patients in the ≥ 20 kg, 10 mg esomeprazole treatment group. The 5 mg and 20 mg esomeprazole dose groups had no discontinuations due to AEs. The applicant clarified that one additional patient had an AE that was considered to have contributed to study discontinuation but was not considered to be the primary reason for withdrawal. Noted by the applicant, this patient had "lack of therapeutic response" recorded as the reason for withdrawal on the termination CRF page. However, on the AE CRF page, he had an AE that was recorded as causing discontinuation. Upon further investigation, the investigator noted that "lack

of therapeutic response" was the correct reason for withdrawal and that the AE was not the reason for discontinuation; therefore, this patient is considered to have discontinued due to "lack of therapeutic response."

Demographics and Baseline Characteristics

The demographic characteristics of study subjects are summarized in Table 3.1.3.

Table 3.1.3 (Applicant's) Demographic for full data set

	Esomeprazole dose groups				
	5 mg Wt <20 kg (N=26)	10 mg Wt <20 kg (N=23)	10 mg Wt ≥20 kg (N=31)	20 mg Wt ≥20 kg (N=29)	Total (N=109)
Demographic characteristics					
Gender [n (%)]					
Male	12 (46.2)	9 (39.1)	17 (54.8)	18 (62.1)	56 (51.4)
Female	14 (53.8)	14 (60.9)	14 (45.2)	11 (37.9)	53 (48.6)
Age in years [n (%)]					
1	12 (46.2)	8 (34.8)	0	0	20 (18.3)
2	6 (23.1)	5 (21.7)	0	0	11 (10.1)
3	4 (15.4)	4 (17.4)	0	0	8 (7.3)
4	2 (7.70)	3 (13.0)	1 (3.2)	2 (6.9)	8 (7.3)
5	1 (3.8)	2 (8.7)	1 (3.2)	1 (3.4)	5 (4.6)
6	1 (3.8)	1 (4.3)	2 (6.5)	0	4 (3.7)
7	0	0	5 (16.1)	3 (10.3)	8 (7.3)
8	0	0	5 (16.1)	9 (31.0)	14 (12.8)
9	0	0	8 (25.8)	6 (20.7)	14 (12.8)
10	0	0	3 (9.7)	6 (20.7)	9 (8.3)
11	0	0	6 (19.4)	2 (6.9)	8 (7.3)
Age category [n (%)]					
1 to 5 years	25 (96.2)	22 (95.7)	2 (6.5)	3 (10.3)	52 (47.7)
6 to 11 years	1 (3.8)	1 (4.3)	29 (93.5)	26 (89.7)	57 (52.3)
Race [n (%)]					
Caucasian	19 (73.1)	19 (82.6)	26 (83.9)	25 (86.2)	89 (81.7)
Black	7 (26.9)	4 (17.4)	5 (16.1)	3 (10.3)	19 (17.4)
Other	0	0	0	1 (3.4)	1 (0.9)
Ethnic group [n (%)]					
Hispanic	2 (7.7)	1 (4.3)	1 (3.2)	0	4 (3.7)
Native Hawaiian/Pacific Islander	0	0	0	1 (3.4)	1 (0.9)
African-American	2 (7.7)	2 (8.7)	3 (9.7)	3 (10.3)	10 (9.2)
Not applicable	22 (84.6)	19 (82.6)	26 (83.9)	25 (86.2)	92 (84.4)
Other	0	1 (4.3)	1 (3.2)	0	2 (1.8)

Based upon Table 3.1.1, the applicant indicated that there was an equitable distribution of males (51.4%) and females (48.6%) and most patients were Caucasian (81.7%). The distributions of these demographic characteristics were similar across the weight/dose groups.

Table 3.1.4 summarizes the key baseline characteristics of study subjects.

Table 3.1.4 (Applicant's) Baseline characteristics for full data set

	Esomeprazole dose groups				
	5 mg Wt <20 kg (N=26)	10 mg Wt <20 kg (N=23)	10 mg Wt ≥20 kg (N=31)	20 mg Wt ≥20 kg (N=29)	Total (N=109)
Baseline characteristics					
Weight (kg)	N=26	N=23	N=31	N=28	N=108
Mean (SD)	12.8 (3.1)	14.1 (2.8)	35.5 (11.7)	34.5 (11.6)	25.2 (13.9)
Median	13	14	37	31	22
Range	8-18	10-18	20-58	21-60	8-60
Height (cm)	N=26	N=23	N=31	N=28	N=108
Mean (SD)	90.0 (11.1)	94.2 (11.7)	134.5 (13.7)	134.5 (11.2)	115.2 (24.4)
Median	89	92	135	133	118
Range	70-109	80-119	108-168	112-159	70-168
BMI (kg/m^2)	N=26	N=23	N=31	N=28	N=108
Mean (SD)	15.7 (2.1)	15.9 (1.7)	19.3 (4.8)	18.6 (3.9)	17.5 (3.8)
Median	15	16	18	17	17
Range	12-20	13-19	14-32	14-29	12-32
Dose/body weight (mg/kg)					
Mean	0.4	0.7	0.3	0.6	0.5
Range	0.3-0.6	0.6-1.0	0.2-0.5	0.3-1.0	0.2-1.0
Biopsy urease test *(H. pylori)*					
Negative	16 (61.5)	12 (52.2)	8 (25.8)	10 (34.5)	46 (42.2)
Positive	0	1 (4.3)	0	0	1 (0.9)
Erosive esophagitis [n (%)][a]	12 (46.2)	12 (52.2)	16 (51.6)	13 (44.8)	53 (48.6)
Nonerosive esophagitis [n (%)]	14 (53.8)	11 (47.8)	15 (48.4)	16 (55.2)	56 (51.4)

Based upon Table 3.1.4, the applicant indicated that within each weight stratum (<20 kg, ≥20 kg), the mean Body Mass Indices (BMIs) of these children were similar between the 2 dose groups. Patients were stratified by weight to approximate the age groups 1 to 5 years of age (weight <20 kg) and 6 to 11 years of age (weight ≥20 kg). In this study, there were some exceptions to this weight/age approximation. Two 6 year old patients weighed <20 kg and three 4 year old and two 5 year old patients weighed ≥20 kg. According to protocol, these patients were randomized according to their weight. In addition, their data are reported according to their weight/dose group. The distribution of patients met the study goal of at least 40 evaluable patients in each age group (1-5 years old, 6-11 years old). In the total study population, 48.6% of patients had erosive esophagitis while 51.4% of patients had non-erosive esophagitis.

Applicant's Efficacy Analysis Results and Conclusions

I. Changes from baseline in Physician's Global Assessment

Table 3.1.5 summarizes the results of the Physician Global Assessments at baseline and the final visit for the ITT population.

Table 3.1.5 (Applicant's) Summary of Physician Global Assessment scores (ITT Population)

Treatment	Timepoint	Assessment	n	(%)	p-value vs Baseline[a]
Esomeprazole 5 mg Weight <20 kg (N=26)	Baseline	None	0		
		Mild	10	39.0	
		Moderate	15	58.0	
		Severe	1	4.0	
		Missing	0		
	Final visit	None	12	48.0	< 0.0001
		Mild	11	44.0	
		Moderate	2	8.0	
		Severe	1		
		Missing	0		
Esomeprazole 10 mg Weight <20 kg (N=23)	Baseline	None	2		
		Mild	6	9.0	
		Moderate	15	26.0	
		Severe	0	65.0	
		Missing	0		
	Final visit	None	9	39.0	0.0004
		Mild	11	48.0	
		Moderate	3	13.0	
		Severe	0		
		Missing	0		
Esomeprazole 10 mg Weight ≥20 kg (N=31)	Baseline	None	1	3.0	
		Mild	14	45.0	
		Moderate	14	45.0	
		Severe	2	7.0	
		Missing	0		
	Final visit	None	18	58.0	< 0.0001
		Mild	12	39.0	
		Moderate	0		
		Severe	1	3.0	
		Missing	0		
Esomeprazole 20 mg Weight ≥20 kg (N=29)	Baseline	None	2	7.0	
		Mild	15	52.0	
		Moderate	11	38.0	
		Severe	1	3.0	
		Missing	0		
	Final visit	None	19	65.0	< 0.0001
		Mild	10	35.0	
		Moderate	0		
		Severe	0		
		Missing	0		

[a]: **Man tel-Haenszel chi-square statistic testing change from baseline.**

Based upon Table 3.1.5, the applicant indicated that a statistically significant reduction in overall GERD-related symptom scores from baseline to the final visit, as assessed by the physician was observed in all treatment groups. In addition, the applicant emphasized that a statistically significant reduction in symptoms from baseline was observed at each study visit (Week 2, Week 4, Week 6, and final visit) for all treatment groups ($p<0.0036$). Similar results were seen in the PP population.

II. Changes from baseline in daily patient symptom assessments as reported by parent/guardian

The applicant indicated that the parent/guardian reported the patient's GERD symptoms on a daily basis throughout the 8 weeks of treatment and during the screening period. The presence or absence of nighttime extra-esophageal symptoms was also recorded within the diary. The patient's daily GERD symptom assessments as reported by the parent/guardian for the ITT population are summarized by symptom and by treatment group in Table 3.1.6.

Table 3.1.6 (Applicant's) Patient diary assessments of GERD symptoms using ITT population

			Baseline (72-hour recall)		Final week in study		Change from baseline		
Treatment	Symptom	N[a]	Mean	SD	Mean	SD	Mean	SD	p-value vs baseline[b]
5 mg Wt <20 kg (N=26)	Heartburn	14	1.50	0.65	0.35	0.61	-1.15	0.93	0.0005
	Acid regurgitation	17	1.71	0.69	0.52	0.70	-1.18	0.81	<0.0001
	Epigastric pain	16	1.69	0.87	0.27	0.52	-1.42	0.61	<0.0001
10 mg Wt <20 kg (N=23)	Heartburn	10	1.70	0.82	0.61	0.97	-1.09	0.86	0.0032
	Acid regurgitation	11	1.64	0.81	0.31	0.41	-1.32	0.93	0.0008
	Epigastric pain	13	1.38	0.65	0.36	0.87	-1.02	0.69	0.0002
10 mg Wt ≥20 kg (N=31)	Heartburn	19	1.42	0.61	0.11	0.20	-1.32	0.67	<0.0001
	Acid regurgitation	20	1.50	0.51	0.26	0.42	-1.24	0.56	<0.0001
	Epigastric pain	15	1.53	0.52	0.23	0.32	-1.30	0.65	<0.0001
20 mg Wt ≥20 kg (N=29)	Heartburn	13	1.46	0.66	0.24	0.40	-1.22	0.85	0.0002
	Acid regurgitation	11	1.55	0.69	0.17	0.35	-1.38	0.81	0.0002
	Epigastric pain	15	1.67	0.72	0.30	0.42	-1.37	0.67	<0.0001

Wt is weight; ITT is intent-to-treat; SD is standard deviation.
[a]: N is the number of patients who had diary data for baseline and their final week in study.
[b]: paired t-test.

Based upon the results shown by Table 3.1.6, the applicant indicated that the GERD symptoms of heartburn, acid regurgitation, and epigastric pain were statistically significantly reduced after treatment with esomeprazole in all treatment groups. The p-values for all of these symptoms were <0.0032 regardless of the weight stratum (<20 kg, ≥20 kg) or esomeprazole dose. Results were similar for the PP population

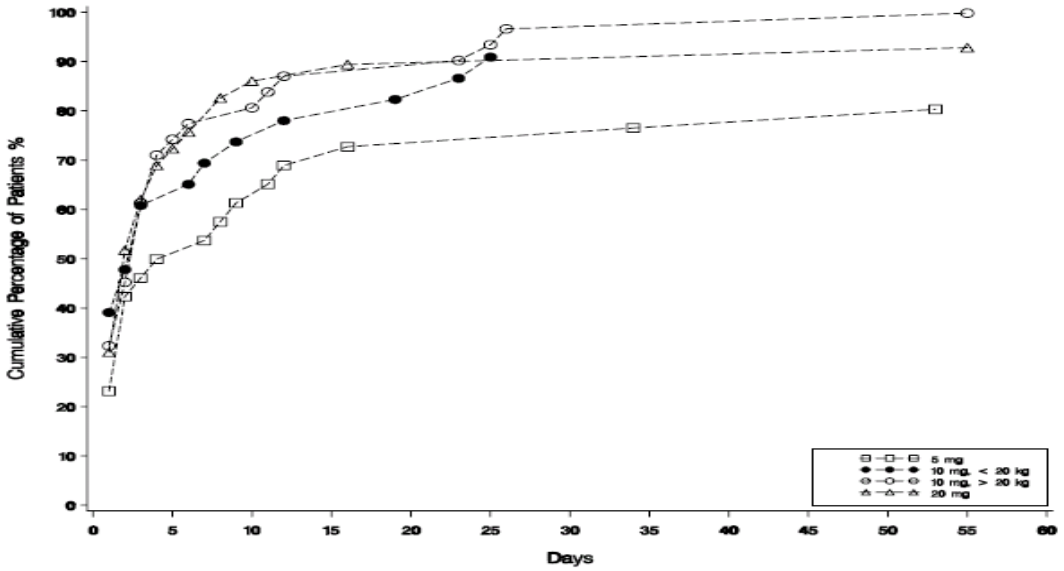
The applicant also indicated that a similar analysis was done for all patients regardless of whether or not they were reported as having these symptoms at baseline. In this analysis, all three symptoms were statistically significantly reduced after esomeprazole treatment with the exception of acid regurgitation in the <20 kg, 10 mg esomeprazole treatment group (p=0.0618, ITT; p=0.0617, PP).

III. Symptom resolution

Time to first resolution and time to first sustained resolution were assessed for the combined GERD symptoms for heartburn, acid regurgitation, and epigastric pain. The applicant indicated that the first resolution was defined as the first day on study drug when the patient indicated "none" for all three symptoms in the diary (IVRS). First sustained resolution was defined as the first day of the first string of 7 consecutive entries of "none" for all three symptoms in the diary (IVRS).

For time to the first resolution, the applicant indicated that the median time to first resolution of symptoms (when at least 50% of the patients achieved first resolution) was 4 days in the esomeprazole 5 mg (<20 kg) treatment group, 3 days in the 10 mg (<20 kg) group, 3 days in the 10 mg (≥20 kg) group, and 2 days in the 20 mg (≥20 kg) group. The cumulative percentage of patients achieving first resolution each day is presented graphically in Figure 3.1.1.

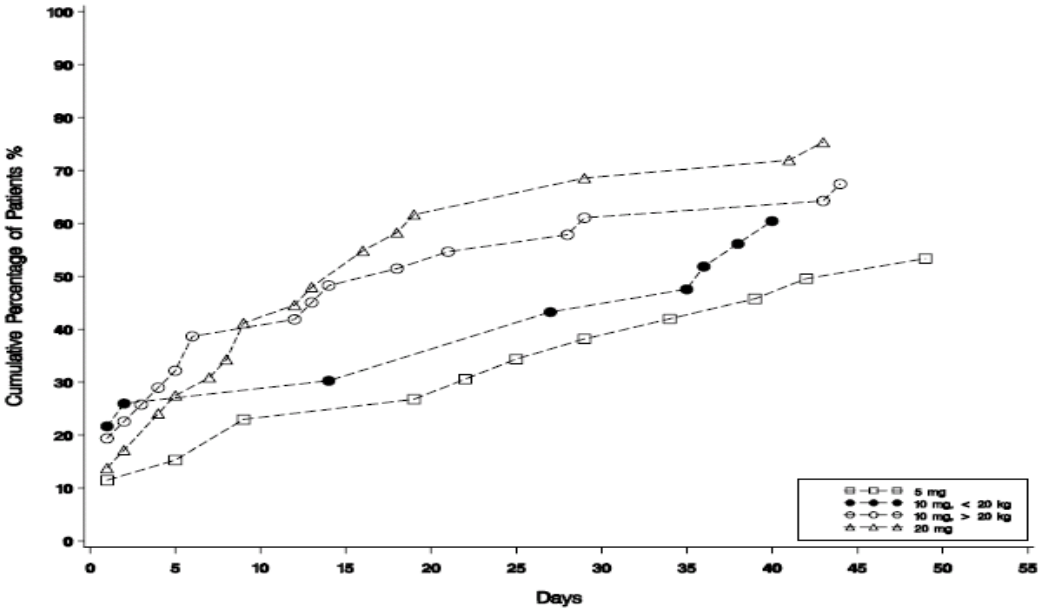
Figure 3.1.1 (Applicant's) Percentage of patients and number of days to first resolution of the combined GERD symptoms heartburn, acid regurgitation, and epigastric pain—ITT population



Based upon this figure, the applicant indicated that it appears that lower percentages of patients in the 5 mg esomeprazole treatment group experienced first resolution of GERD symptoms than in the other treatment groups. Similar results were observed in the PP population.

For the time to the first sustained resolution, the applicant indicated that the median time to reach first sustained resolution was 42 days in the esomeprazole 5 mg (<20 kg) treatment group, 36 days in the 10 mg (<20 kg) group, 18 days in the 10 mg (≥20 kg) group, and 16 days in the 20 mg (≥20 kg) group. The applicant declared that the first sustained resolution was achieved faster in the higher weight children (≥20 kg) than in the lower weight children (<20 kg). The cumulative percentage of patients achieving first sustained resolution each day is presented graphically in Figure 3.1.2.

Figure 3.1.2 (Applicant's) Percentage of patients and number of days to first sustained resolution of the combined GERD symptoms heartburn, acid regurgitation, and epigastric pain—ITT population



Based upon this figure, the applicant indicated that it appears that lower percentages of patients in the 5 mg esomeprazole treatment group experienced first sustained resolution of GERD symptoms than in the other treatment groups. Similar results were observed in the PP population.

IV. Changes from baseline in endoscopic healing of EE

The applicant indicated that eight patients who had EE reported at baseline did not have follow-up endoscopies mainly due to early terminations or revisions of the original EE diagnoses (negating the presence of EE). These patients are not included in the assessment of endoscopic healing, as they were missing post-treatment data.

Patients were considered to be improved if their esophageal erosions at their final endoscopy were one or more LA grades better than they were at baseline. Patients were resolved if their final endoscopy showed no signs of erosions.

Table 3.1.7 summarizes the outcomes for the assessments of endoscopic healing of EE for ITT population. PP population results were similar to those of the ITT population.

Table 3.1.7 Summary of outcome for patients who had EE at baseline and had a follow-up endoscopy—ITT population

	Esomeprazole dose groups									
	5 mg Wt <20 kg (N=11)		10 mg Wt <20 kg (N=11)		10 mg Wt ≥20 kg (N=10)		20 mg Wt ≥20 kg (N=13)		Total (N=45)	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Improved	11	(100.0)	9	(81.8)	9	(90.0)	13	(100.0)	42	(93.3)
Improved but not resolved	0		0		0		2	(15.4)	2	(4.4)
Resolved	11	(100.0)	9	(81.8)	9	(90.0)	11	(84.6)	40	(88.9)
No improvement (same as baseline)	0		2	(18.2)	1	(10.0)	0		3	(6.7)
Worsened	0		0		0		0		0	

Wt is weight

Based upon the results of Table 3.1.7, the applicant declared that overall, 93.3% of patients who had EE at baseline and a follow-up endoscopy were improved at their follow-up endoscopy. In most of these patients (88.9%), the EE was resolved and their erosions had healed. The positive results in improvement and resolution were observed across all treatment groups. An additional note of interest is that all patients whose EE was not healed had received doses in the range of 0.17 to 0.66 mg/kg.

In this study, there were three patients who did not show any improvement. Two of these patients had Grade B erosive esophagitis and one had a single ulcer described in the cardia region. The applicant indicated that it was unclear whether or not this last patient had true EE. These three patients received esomeprazole doses of 0.17, 0.55, and 0.60 mg/kg/day.

Reviewer's Comments and Analysis

In order to explore the applicant's efficacy claim, this reviewer first comments on the following three issues: 1) No control arm in study, 2) Severity of erosive esophagitis, and 3) Questionable analysis method. Then, this reviewer performs analysis to further explore the efficacy of Nexium in use of pediatric patients.

Reviewer's Comments

1) No control arm in study

It is noted that patients were stratified based upon their weight and randomized to receive one of the four treatment groups: < 20 kg patient weight, 5 mg esomeprazole; < 20 kg patient weight, 10 mg esomeprazole; ≥ 20 kg weight, 10 mg esomeprazole; and ≥ 20 kg weight, 20 mg esomeprazole. Basically, the four treatment arms were study drugs; no controlled arm was included. Since no comparator was included in this study, the investigator knew that all patients were administered the study drug. Accordingly, in fact, this was an open label and uncontrolled study. However, ICH E9, "Guidance for Industry, E9 Statistical Principles for Clinical Trials", indicates that blinding is one of the most important design techniques for avoiding bias

assessments in clinical trials. Blinding is intended to limit the occurrence of conscious and unconscious bias in the conduct and interpretation of a clinical trial arising from the influence that the knowledge of treatment may have on the recruitment and allocation of subjects, their subsequent care, the attitudes of subjects to the treatments, the assessment of end-points, the handling of withdrawals, the exclusion of data from analysis, and so on.

Not only was the trial an un-blinded study to the investigators (since they knew that Nexium was the study dose), but the definitions of "Mild" (symptoms present but not interfering with daily activities) and "Moderate" (symptoms present and somewhat interfering with daily activities) in the physician global assessments were not completely distinguished and could be assessed subjectively. It follows that the open label design and endpont subjectivity could lead to biased efficacy assessments in favor of the study drug Nexium.

In addition, since no control arm (placebo or active) was included in the study, there is no basis to assess whether the results of the efficacy analyses demonstrated by the study are quantitatively better than that of placebo or similar to that of an active-controlled drug. In other words, for the drug efficacy assessed by the improvement from baseline, since no results from controlled arm were provided for the efficacy comparisons, different people may have different conclusions on the efficacy assessments of the study drug.

Finally, for the issue on the Baseline controlled study without including a control arm, ICH E10 "Guidance for Industry, E10 choice of Control Group and Related Issues in Clinical Trials", basically states that so-called baseline controlled studies are really not controlled at all, but implicitly assume an external control or threshold. The validity of the external control is crucial. Accordingly, in order to demonstrate the efficacy for Nexium in use of the pediatric patients with ages from one to eleven, one approach would be to show the observed changes from baseline are comparable to those that can be extrapolated from the adult (well controlled) studies. Ideally the applicant would have pre-specified an 'effect' - that is, have stated what changes from baseline were to be expected and justify in terms of the known adult effect. However, extrapolation of efficacy from the adult studies to this population has not been established.

2) Severity of erosive esophagitis for the enrolled patients

Fifty-one percent (56/109) of patients enrolled at baseline had no erosive esophagitis: only 49% (53/109) of enrolled patients had erosive esophagitis at baseline. For the patients with the erosive esophagitis at baseline, Table 3.1.8 presents the LA erosive esophagitis.

Table 3.1.8 (Reviewer's) Erosive esophagitis grade at baseline endoscopy

LA grade at Baseline	Total Patients N = 53 (%)
Grade A	32 (60.4)
Grade B	19 (35.8)
Grade C	1 (1.9)
Grade D	1 (1.9)

For patients with esophagitis disease at baseline, Table 3.1.8 indicates that less than 4 % (2/53)

of enrolled subjects had erosive esophagitis with more severe LA grades C and D and the rest of patients (greater than 96%) had mild erosive esophagitis with LA grades A and B. Therefore, due to lack of sufficient more severe esophagitis subjects enrolled, the assessment on the healing rates performed by the applicant may not be able to properly reflect the efficacy of Nexium for the pediatric patients (ages from 1 to 11) with more severe erosive esophagitis at baseline.

3) Questionable analysis methods

This reviewer comments on the following two statistical analysis methods performed by the applicant: i) physician's global assessment scores and ii) patient diary assessments of GERD symptoms.

i) Analysis on the physician's global assessment scores

Based upon the SAS program submitted by the applicant on December 19, 2006, it is noted that in the analysis for the physician global assessment scores on GERD symptoms, the applicant compared the ordinal categories (none, mild, moderate, and severe) of the Baseline versus that of the Final Visit using Cochran-Mantel-Haensel method for each of the four treatment groups. Since the same patient population was used at Baseline and Final Visit, the patient population in the Baseline is not independent of that in the Final Visit. The asymptotic distribution for Cochran-Mantel-Haensel method may not follow the central Chi-square distribution with one degree of freedom as used in the SAS software program to generate the p-value. Accordingly, the p-values presented by Table 3.1.5 calculated by the applicant using SAS program with the correlated patient populations are not correct p-values to assess the improvement from Baseline based upon the criteria of physician assessment scores. The analysis method (Cochran-Mantel-Haensel test) applied by the applicant to analyze the physician's global assessment scores is not a statistically sound method.

ii) Analysis on the patient diary assessments of GERD symptoms

Based upon the data set and the SAS program submitted by the applicant on December 19, 2006, it is noted that in the analysis for the three GERD symptoms Heartburn, Acid Regurgitation, and Epigastric Pain reported by the parents/guardians, paired t-test was applied to the average of the 7 data points for the Final Visit and one data point at Baseline to test if patients were improved from Baseline. However, since at Final Visit, seven data points were averaged and only one data point was used at Baseline, under the usual assumption of equal standard deviation for each data point, the standard deviation of averaged seven data points from Final Visit is smaller than that of one data point at Baseline by 2.6 ($\sqrt{7}$) times.

It is well known that the paired t-test should be applied to the pairs whose distributions of the two components are identical except for a possible difference in means. However, the paired t-test implemented by the applicant was applied to the distributions of the two components with different shapes. The p-values presented by Table 3.1.6 calculated by the applicant using the distributions of pairs with different shapes are not correct p-values to assess the improvement

from Baseline based upon patient diary assessments of GERD symptoms for each of the four treatment groups. The analysis method (paired-t test) applied by the applicant to analyze the patient diary assessments of GERD symptoms is not legitimate.

Reviewer's Analysis

As commented in the sub-section of "Questionable analysis methods", the applicant's analysis methods for the physician's global assessment scores and the patient diary assessments on GERD symptoms are not legitimate. In order to explore the efficacy of Nexium claimed by the applicant, this reviewer performs the following two proportion analyses: 1) based upon physician's GERD assessment score and 2) based upon patient diary assessment on GERD symptoms.

1) Proportion analysis based upon physician's GERD assessment score

For the physician's assessment scores, this reviewer calculated the following two types of proportions for each of the four treatment groups: proportions of patients with one score lower/better at Final Visit than Baseline and proportions of patients with two scores lower/better at Final Visit than Baseline. The rationale for the calculation for the proportions of two scores lower/better is that the definitions of "Mild" and "Moderate" in physician global assessments are not clearly distinguished and may be assessed subjectively in this open label study. Table 3.1.9 presents the proportions and the associated two-sided confidence intervals on the one and two scores better at Final Visit than Baseline.

Table 3.1.9 (Reviewer's) Proportions of patients improved from Baseline when assessed at Final visit using ITT population

TREATMENT GROUP	ONE SCORE BETTER		TWO SCORES BETTER	
	Proportion (n/N)	95% two-sided CI	Proportion	95% two-sided CI
I. Esomeprazole 5 mg & Weight <20 kg (N=26)	69% (18/26)	(48%, 86%)	27% (7/26)	(12%, 48%)
II. Esomeprazole 10 mg & Weight <20 kg (N=23)	65% (15/23)	(43%, 84%)	17% (4/23)	(5%, 39%)
III. Esomeprazole 10 mg & Weight ≥20 kg (N=31)	81% (25/31)	(63%, 93%)	29% (9/31)	(14%, 48%)
IV. Esomeprazole 20 mg & Weight ≥20 kg (N=29)	79% (23/29)	(60%, 92%)	24% (7/29)	(10%, 44%)

Table 3.1.9 indicates that for the four treatment groups, the lower limits of the two-sided confidence intervals for one score better from Baseline assessed at Final Visit are (48%, 43%, 63%, and 60% respectively for treatments I, II, III, and IV) much higher than that (12%, 5%, 14%, and 10% respectively for treatments I, II, III, and IV) of the two scores better. In addition, since this is an open label study, the efficacy assessments of the physician's global assessments may be biased in favor of the study drug. The true proportions for the patients with improved from Baseline may be lower than the ones presented here. In addition, due to shortcomings of open label and ambiguous definitions of "Mild" and "Moderate", the proportions of two scores

better may be more appropriate to assess the efficacy of Nexium for the pediatric patients with ages from 1 to 11.

2) Proportion analysis based upon patient diary assessment GERD symptom

For the patient's diary assessment GERD symptoms, this reviewer calculated the proportions of patients with one score lower/better at Final Visit than Baseline and the corresponding two-sided 95% confidence intervals for the GERD symptoms on heartburn, acid regurgitation, and epigastric pain for each of the four treatment groups. For the three GERD symptoms, unlike the applicant using the average of the 7 data points of the Final Visit week, the final assessment (one data point) recorded prior to withdrawal is used as the final assessment to analyze the improvement from Baseline at Final Visit.

Data used in this reviewer's analysis is submitted by the applicant on March 1, 2007. Table 3.1.10 presents the proportions and the associated two-sided confidence intervals for the patient's diary assessment GERD symptoms.

Table 3.1.10 (Reviewer's) Proportions of patients improved from Baseline at Final visit by patients' diary assessment GERD symptoms using ITT population

Heartburn

TREATMENT GROUP	ONE SCORE BETTER AT FINAL VISIT	
	Proportion (n/N)	95% two-sided CI
I. Esomeprazole 5 mg & Weight <20 kg (N=26)	48% (12/25)	(28%, 69%)
II. Esomeprazole 10 mg & Weight <20 kg (N=23)	26% (6/23)	(10%, 48%)
III. Esomeprazole 10 mg & Weight ≥20 kg (N=31)	63% (11/30)	(44%, 80%)
IV. Esomeprazole 20 mg & Weight ≥20 kg (N=29)	52% (15/29)	(33%, 71%)

Regurgitation

TREATMENT GROUP	ONE SCORE BETTER AT FINAL VISIT	
	Proportion (n/N)	95% two-sided CI
I. Esomeprazole 5 mg & Weight <20 kg (N=26)	60% (15/25)	(39%, 79%)
II. Esomeprazole 10 mg & Weight <20 kg (N=23)	35% (8/23)	(16%, 57%)
III. Esomeprazole 10 mg & Weight ≥20 kg (N=31)	67% (20/30)	(47%, 83%)
IV. Esomeprazole 20 mg & Weight ≥20 kg (N=29)	48% (14/29)	(29%, 67%)

Epigastric pain

TREATMENT GROUP	ONE SCORE BETTER AT FINAL VISIT	
	Proportion (n/N)	95% two-sided CI
I. Esomeprazole 5 mg & Weight <20 kg (N=26)	60% (15/25)	(39%, 79%)
II. Esomeprazole 10 mg & Weight <20 kg (N=23)	52% (12/23)	(31%, 73%)
III. Esomeprazole 10 mg & Weight ≥20 kg (N=31)	57% (17/30)	(37%, 75%)
IV. Esomeprazole 20 mg & Weight ≥20 kg (N=29)	62% (18/29)	(42%, 79%)

Table 3.1.10 indicates that for each of the three GERD symptoms assessed by the patients, the proportions of patients improved from Baseline as compared to Final Visit for treatment II (Esomeprazole 10 mg & Weight <20 kg) are less than that of the other three treatment groups. In addition, by comparing Table 3.1.9 and Table 3.1.10, it is noted that for the physician's global assessments with respect to each of the four treatment groups, the lower bounds of the 95% two-sided confidence intervals on the proportions of patients improved from Baseline with one score better are much higher than that of the three GERD symptoms assessed by patients: range from 9% (eg., Regurgitation in treatment group I) to 33% (eg., Heartburn in treatment group II). It may provide some clue that physician global assessments might be assessed in favor of the study drug in this open label study.

Since no control arm was included in this study, no objective efficacy comparison is able to be performed regarding the efficacy of the study drug Nexium assessed by the three patient diary GERD symptoms. However, based upon the results presented by the two tables (Table 3.1.9 and

Table 3.1.10) along with the efficacy extrapolated from the adult studies, the Medical division may be able to make decision regarding the efficacy of the study drug in the use of pediatric patients.

3.2 Evaluation of Safety for Study D9614C00097

The applicant indicated that all doses of esomeprazole were generally safe and well tolerated in the studied population of 1 to 11 year old pediatric GERD patients. There were no deaths. There were two serious adverse events (SAEs) that occurred during or after study treatment and one SAE that occurred during the Screening endoscopy, before the patient was randomized. All SAEs were considered to be not treatment related.

There were four patients with discontinuations due to adverse events (DAEs). The DAEs of three of these patients were not considered treatment-related. The one DAE patient whose AEs of asthenia, nausea, and viral infection were considered as possibly treatment related had resolution of all AEs within one day of onset. In addition, there were no clinically important findings and trends in hematology, clinical chemistry, urinalysis, vital signs, or physical examination (including medical history) observed across or within the esomeprazole treatment groups. From the safety assessments made in this study, no new safety signals were identified; therefore the applicant concluded that the 5 mg, 10 mg, and 20 mg doses of esomeprazole were generally safe and well tolerated in the one to eleven year old patients studied.

4.0 SUBGROUP ANALYSIS

Since this is a pediatric trial, no subgroup analyses on gender, race, and age were performed by this reviewer.

5.0 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

For Study D9614C00097, the following three issues are commented upon: 1) No control arm in study, 2) Severity of erosive esophagitis, and 3) Questionable analysis method.

2) No control arm study

a) Since no comparator was included, investigator knew that all patients took the study drug. In reality, it is an open label study. ICH E9 indicates that blinding is one of the most important design techniques for avoiding bias assessments in clinical trials.

b) Definitions of “Mild” and “Moderate” in physician’s global assessments are not completely distinguished in this open label study. Accordingly, the two shortcomings (open label and ambiguous endpoint definition) could induce biased assessments in favor of the study drug.

c) Since no control arm (placebo or active) was included in the study, the efficacy of study drug Nexium can not be objectively assessed by quantitatively comparing the efficacy between Nexium and a concurrent control.

d) ICH Guidance E10 states in effect that baseline controlled studies are really not controlled at all, but implicitly assume an external control or threshold value for efficacy. The validity of the external control is crucial. One approach would be to show the observed changes from Baseline are comparable to those that can be extrapolated from the adult (well controlled) studies.

2) Severity of erosive esophagitis

For patients with esophagitis disease at Baseline, less than 4% (2/53) of enrolled subjects with erosive esophagitis had more severe LA grades C and D and the other patients (96%) had mild erosive esophagitis with LA grades A and B. Therefore, due to the small number of severe esophagitis subjects enrolled, the overall healing rates would not properly reflect efficacy of Nexium for pediatric patients with more severe erosive esophagitis at Baseline.

3) Questionable analysis method

a) For the analysis on the physician's global GERD assessment scores, the applicant compared the ordinal categories (none, mild, moderate, and severe) at Baseline versus that of the Final Visit using Cochran-Mantel-Haenszel method for each of the four treatment groups. Since the same patient population was used at Baseline and Final Visit, patient outcomes at Baseline are not independent with those at Final Visit, violating an independence assumption for efficacy comparison between two treatment groups. Accordingly, the analysis method applied by the applicant to analyze the physician's global assessment scores is not legitimate. The p-values generated by the applicant based on dependence data are not correct p-values to assess the improvement from Baseline.

b) For the analysis on the patient diary assessments of GERD symptoms on Heartburn, Acid Regurgitation, and Epigastric Pain, the applicant performed a paired t-test using average of the 7 data points from the Final Visit and one data point at Baseline. Accordingly, the shapes of the distributions for the two components of the paired data are not identical. The p-values generated by the applicant using the paired t-test based upon the paired distributions of un-equal shapes are inadequate to assess the improvement from Baseline.

In order to explore the efficacy of the study drug used in pediatric patients, the reviewer performed the following two analyses on proportions: 1) based upon physician's GERD global assessment score and 2) based upon patient diary assessment scores on GERD symptoms.

1) Proportion analysis based upon physician's GERD global assessment score

a) For the physician's global assessment scores, this reviewer calculated the following two types of proportions for the four treatment groups: proportions of patients with one score

improvement from baseline at final visit and proportions of patients with two score improvement from Baseline at Final visit.

b) The lower limits of the two-sided 95% confidence intervals for one score improvement are 48%, 43%, 63%, and 60% respectively for treatments I, II, III, and IV. For the two-score improvement, the lower limits are 12%, 5%, 14%, and 10% respectively for treatments I, II, III, and IV.
c) Since this is an open label study, the efficacy assessments of the physician's global assessments may be biased in favor of the study drug. The true proportions for the patients with improved from Baseline may be lower than the ones presented here.
d) Due to the shortcomings of open label and ambiguous definitions of "Mild" and "Moderate", the proportions for two-score improvement may more objectively assess improvement from Baseline.

2) Proportion analysis based upon patient diary assessment scores on GERD symptoms

a) For the patient's diary assessment scores on GERD symptoms, this reviewer calculated the proportions of patients with one score improvement from Baseline and the corresponding two-sided 95% confidence intervals for the GERD symptoms on heartburn, acid regurgitation, and epigastric pain for each of the four treatment groups.
b) For the three GERD symptoms assessed by the patients, the proportions of patients who improved two scores from Baseline at Final visit for treatment II (Esomeprazole 10 mg & Weight <20 kg) are less than that for the other three treatment groups.
c) For the physician's GERD global assessments, the lower bounds of the 95% two-sided confidence intervals for the proportions of patients who improved from Baseline by one score are much higher than that of the three GERD symptoms assessed by patients: range from 9% (eg., Regurgitation in treatment group I) to 33% (eg., Heartburn in treatment group II). This may indicate a bias in the physician global assessments in favor of the study drug.

5.2 Recommendations

From a statistical perspective, the single Study D9614C00097 does not provide substantial evidence of efficacy to support the study drug Nexium in use of short term treatment of gastroesophageal reflux disease (GERD) and healing of erosive esophagitis for patients with ages from one to eleven.

However, this submission does satisfy the Agency's Written Request (WR) for a study of pharmacokinetics, safety and clinical outcome for pediatric patients, one to eleven years of age. There was no intent to provide confirmatory evidence of efficacy. With regard to labeling, [redacted]

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
7/19/2007 05:28:50 PM
BIOMETRICS

Mike Welch
7/20/2007 11:22:23 AM
BIOMETRICS
Concur with review.