CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:NDA 20-942

STATISTICAL REVIEW(S)

Statistical Review and Evaluation

NDA 20-942
Drug name: Versed (midazolam HCl), syrup
Applicant: Hoffmann-La Roche Inc.
Indication: Sedation/anxiolysis and amnesia prior to diagnostic/therapeutic or endoscopic procedure or before induction of anesthesia in pediatric patients

Volumes reviewed: 1.1, 1.3, 1.24-61 dated 9 February 1998
(Received HFD-170 10 February 1998)
Electronic data submitted by sponsor
Medical review
Reviewer: Z. Jonathan Ma, Ph.D., HFD-720
Date of Review: 1 June 1998

Project manager: David Morgan
Medical reviewer: Igor Cerny, Ph.D.

Introduction

Midazolam is an approved product in this country for the indication of preoperative sedation/anxiolysis/amnesia through intravenous (IV) or intramuscular (IM) injection. In clinical sites, however, the parenteral midazolam is often prepared as an oral solution mixed with sweeteners or juices for pediatric patients. This "off-label" use is intended to help to reduce the anxiety experienced by pediatric patients before induction of anesthesia due to fear of the medical procedures.

The sponsor developed Versed (midazolam HCl) syrup in order to provide clinicians with a standardized oral formulation of midazolam for the pediatric population. No placebo-controlled trial was conducted for this NDA submission because the sponsor argued that the efficacy of oral midazolam has been demonstrated in many placebo-controlled trials in the literature. In addition, as quoted by FDA medical reviewer Igor Cerny, Ph.D., sponsor's study plan had been agreed by former FDA medical reviewer Barbara Palmisano, M.D. (6/27/96): "*A placebo-controlled study to prove efficacy is not planned and is not needed as long as the efficacy for this formulation can be linked to the efficacy of the approved intravenous formulation. This can be done via clinical pharmacology by establishing the pharmacokinetic/pharmacodynamic relationship between the two formulations. The proposed plan offers a pharmacokinetic study to link the two drugs...*".

Therefore, besides a PK/PD study and an extensive literature review, the only clinical study included in this NDA submission was a dose ranging study, Protocol # NR15345, in which three doses of Versed syrup (0.25 mg/kg, 0.50 mg/kg and 1.0 mg/kg) were studied and compared in efficacy and safety profiles. The main purpose of this trial, as cited by the sponsor, was to provide prescription guidance to clinicians.

This statistical review will, however, only focus on the efficacy results from this multiple-dose clinical trial (Protocol # NR15345) . The safety results have been reviewed by the medical reviewer and no major statistical issue needs to be addressed regarding the safety issue.

Study Design

This was a multicenter (9 sites), randomized, single-dose, double-blind, parallel-group study in pediatric patients. There were three dose groups:

- Group 1: 0.25 mg/kg,
- Group 2: 0.50 mg/kg, and
- Group 3: 1.0 mg/kg midazolam,

up to a maximum dose of 20 mg. These doses were expected to "*produce sub-optimal, optimal, and slightly excessive pharmacologic effects*" by the sponsor.

All patients were first stratified in three age groups: 6 months–2 years, 2–6 years and 6–16 years, and then randomized to one of the three dosage groups. To maintain blinding, a third-party pharmacist at each center dispensed the appropriate volume of study drug in to a dispenser, based on patient randomization number (indicating treatment assignment) and patient weight. The dispenser was then given to a study nurse to administer to the patient. Both the observer and the patient were supposedly blinded to the dosage information.

A total of 405 patients who met the inclusion/exclusion criteria were randomized. Eight (8) patients never received any study medication. Hence, a total of 397 patients received treatment medication: 132 in Group 1, 132 in Group 2 and 133 in Group 3. The sponsor considered these 397 patients as the Intent-to-treat (ITT) population.

The inclusion criteria were as follows:

- Physical status I, II or III based on American Society of Anesthesiologists (ASA) Physical Status Classification criteria.
- Males and females 6 months to 16 years old (up to but not including the 16th birthday).
- Patients scheduled for surgical procedures requiring general anesthesia with mask induction.

The exclusion criteria were as follows:

- Were pregnant or nursing
- Had a known or suspected hypersensitivity to midazolam or any benzodiazepine

- Had a known or suspected hypersensitivity to opioids
- Had a documented history of recurrent seizures or a seizure disorder
- Had a gastrointestinal disorder that might have affected absorption
- Had known or suspected hypersensitivity to cherry or cherry products
- Were treated with any investigational drug or participated in a clinical trial within the 4 weeks preceding the study
- Had used any agent (e.g., benzodiazepine) that might interfere with sedation assessment less than 8 hours before dosing
- Had used a macrolide (i.e., clarithromycin or erythromycin) within 7 days of dosing
- Had any medical condition which, in the judgment of the investigator, rendered them inappropriate for participation in the trial.

The primary objective of this trial was:

- To compare the sedative effect of midazolam syrup prior to induction of anesthesia in pediatric patients who received single oral doses of 0.25 mg/kg, 0.5 mg/kg, and 1.0 mg/kg midazolam, up to a maximum dose of 20 mg.

The secondary objectives were:

- To compare the distribution of time to achieve satisfactory sedation among regimens
- To evaluate the anxiolytic properties of oral midazolam syrup as a premedicant
- To evaluate the safety of three single oral doses of midazolam syrup
- To evaluate the effect of oral midazolam syrup on patient cooperation
- To evaluate the effect of oral midazolam syrup on the recovery parameters of patients following surgical procedures.

One exploratory objective was:

- To evaluate the acceptability of the flavored investigational formulation of oral midazolam syrup.

Primary Efficacy Parameter
The sedation assessment was based on the following 5-point scale:

Sedation Score	Description
1	Alert/Active: Agitated, moving, physical or verbal display (loud or high-pitched crying) of apprehension
2	Upset/Wary: Tearful, may be clinging (to parent or guardian)
3	Relaxed: Calm, silly, responds readily to commands or gentle stimulation
4	Drowsy: Easily arousable, responds to mild shaking or prodding
5	Asleep: Unarousable, does not respond to shaking or prodding

A score of 3, 4 or 5 was considered "satisfactory" sedation and a score of 1 or 2 was considered "unsatisfactory". The primary efficacy parameter was the satisfactory sedation rate (proportion) for each dosing regimen based on each patient's highest sedation score within 30-minute post-treatment.

Baseline sedation level was assessed prior to Versed administration and post-treatment sedation levels were measured at 10-minute intervals up to 45 minutes or the general mask induction as the result of an adequate sedation (score of 3, 4 or 5).

Secondary Efficacy Parameters Includes:

1. Distribution of time to satisfactory sedation;
2. Measure of anxiety, including ease of attempted separation from parent/guardian;
3. Measure of cooperation;
4. Time from discontinuation of last inhaled anesthetic agent until recovery.

Study Populations

The sponsor performed the efficacy analyses based on two populations: the Intent-to-treat (ITT) population and the Standard (STD) population.

The ITT population was defined in the protocol as "all patients who are randomized, receive any amount of test medication, and have at least one post-baseline assessment.". A total of 405 patients met the inclusion/exclusion criteria and were randomized to enter the study. Eight (8) of them never took any study medication and hence had no post-treatment measurement. They were excluded from the ITT population by definition. The intent-to-treat (ITT) analyses were then performed on these 397 ITT patients.

The STD population was a subgroup of the ITT population, excluding those who did not fully comply with the protocol requirements. Exclusion reasons included "Full dose not ingested", "Vomited before mask induction", "Concomitant drug not allowed", and "Rescue medication given", etc. The STD population had a total of 350 patients since 47 out of the 397 ITT patients did not fully comply with the protocol requirements. The major reason for this violation was "Full dose not ingested", accounting for more than two third of the total violations. More investigation on these 47 protocol non-compliers can be found in the Discussion section.

The main advantage of randomization is to produce the comparability between the comparison groups with respect with all potential risk factors. The STD population, or sometimes called per protocol population, however, may damage this comparability by removing some subjects depending on different outcome observations. Therefore, the results based on the STD population would then potentially be subject to bias.

As to sponsor's ITT population, only eight (8) patients were excluded from the original randomization population because they did not take any treatment medication and had no post-treatment measures. Considering that this is less than 2% of the total sample size, the comparability between the comparison groups in sponsor's ITT population should remain valid.

In the rest of this review, only analyses based on the ITT population will be discussed in details. According to sponsor's reports, the STD population in general yielded similar results despite of potential intrinsic bias mentioned earlier.

The ITT Population

The intent-to-treat (ITT) population contained 397 patients who were randomized and dosed with some medication. As shown in Table 4, Module I (Vol. 31), the three treatment groups were comparable with regard to baseline demographic parameters, such as sex, age, race, etc.

Both male and female patients were recruited into the study. However, approximately twice as many males (264, 66%) as females (133, 34%) were enrolled. The sponsor explained that the reason was that several of the procedures were male-specific. Patient age ranged from 0.5 to 15.3 years with a mean of 4.6 years and a standard deviation of 3.74. Most of them were Caucasian (235, 59%); 93 (23%) were Black, 1 (<1%) was Asian, and 68 (17%) were classified as Other.

As indicated in the following two tables, the satisfactory sedation and anxiety rates at baseline were very similar across three treatment groups and were quite different for different age groups. Older patients obviously had larger rates at baseline.

	Baseline Sedation			Baseline Anxiety*		
Regimen	N	Unsatisfactory	Satisfactory	N	Unsatisfactory	Satisfactory
0.25 mg/kg	132	69 (52%)	63 (48%)	132	52 (39%)	80 (61%)
0.50 mg/kg	132	67 (51%)	65 (49%)	131	50 (38%)	81 (62%)
1.0 mg/kg	133	70 (53%)	63 (47%)	131	51 (39%)	80 (39%)
Total	397	206 (52%)	191 (48%)	394	153 (39%)	241 (61%)

*3 patients not assessable for baseline anxiety
Source: Table 4I and Table 5I, Module III (Vol. 34)

	Baseline Sedation			Baseline Anxiety*		
Age Group	N	Unsatisfactory	Satisfactory	N	Unsatisfactory	Satisfactory
0.5-2	147	101 (69%)	46 (31%)	145	86 (59%)	59 (41%)
2-6	123	60 (49%)	63 (51%)	122	39 (32%)	83 (68%)
6-16	127	45 (35%)	82 (65%)	127	28 (22%)	99 (78%)
Total	397	206 (52%)	191 (48%)	394	153 (39%)	241 (61%)

*3 patients not assessable for baseline anxiety
Source: Table 4I and Table 5I, Module III (Vol. 34)

Primary Efficacy Analyses

The primary efficacy parameter was defined as the satisfactory sedation rate, based on the maximum sedation score achieved within 30 minutes post-treatment, i.e., the highest score among the 10-, 20- and 30-minute measurements. The outcome was collapsed into a binary scale: satisfactory (sedation score=3, 4, or 5) or unsatisfactory (sedation score=1 or 2).

As shown in the following table, overall, 385 out of the 397 (97%) ITT patients achieved satisfactory sedation post-treatment and the satisfaction rate seemed positively associated with the dosage regimen (93%, 98% and 100% for the 0.25, 0.50 and 1.0 mg/kg groups, respectively). There appeared to have a similar trend for Anxiety and other secondary endpoints which will be discussed in more details later.

Number and rate of satisfactory response post-treatment by regimen: ITT population

	Treatment Regimen			
	0.25 mg/kg (n=132) N (%)	0.50 mg/kg (n=132) N (%)	1.0 mg/kg (n=133) N (%)	Overall (n=397) N (%)
Sedation*	123 (93%)	129 (98%)	133 (100%)	385 (97%)
Anxiety*	126 (95%)	129 (98%)	132 (99%)	387 (97%)
Ease of Separation	115 (87%)	116 (88%)	119 (89%)	350 (88%)
Cooperation (nitrous oxide)	110 (83%)	112 (85%)	121 (91%)	343 (86%)
Cooperation (halothane)	106 (80%)	106 (80%)	121 (91%)	333 (84%)

*Each patient's highest score within 30 minutes was used to assess response.

To examine the dose-response relationship, a closed-test procedure was applied in order to protect type I error from inflation due to multiple testing. The procedure involves testing a family of hypotheses which is in such a format that one contains another. The procedure starts from testing the smallest hypothesis and move on to test the second smallest only if the first hypothesis is rejected at type I error α. Otherwise, the testing procedure stops and does not reject all the hypotheses that follow.

In the sponsor's primary analysis, the first stage was to test the difference between the 0.25 mg/kg and 1.0 mg/kg groups in sedation rate. It was a one-sided test at a significance level of 0.05. The p-value was 0.002 (Cochran-Mantel-Haenszel, row mean score statistic; Table 12I, Module III, Vol. 34) and hence the null hypothesis that all three sedation rates are equal was rejected. Then the second stage was to test the difference between the 0.50 mg/kg and 1.0 mg/kg groups. The result was not significant (p=0.08). Therefore, a statistically significant difference was detected only between the 0.25 mg/kg and 1.0 mg/kg groups at a one-sided significance level of 0.05.

The Cochran-Mantel-Haenszel test, which was used to test for the association between regimen and response at each stage in the primary analyses, was stratified on both age group and baseline

sedation score in the final statistical analyses. It should be noted that the statistical plan in the original protocol included age group as the only stratification variable. In the final analyses, however, baseline sedation score was also adjusted for in all tests.

As shown in the following table, the post-treatment satisfactory sedation rate was 99% for patients with "satisfactory" sedation at baseline and 95% for those with "unsatisfactory" sedation at baseline. Therefore, the baseline sedation seemed well related to the post-treatment sedation and hence needs to be adjusted for as well.

No. of patients by baseline sedation, post-treatment sedation and regimen

Baseline		Post-Treatment		
Sedation	Regimen (mg/kg)	Satisfactory	Unsatisfactory	Total
Satisfactory	0.25	61	2	
	0.50	65	0	
	1.0	63	0	
	Subtotal	189 (99%)	2 (1%)	191
Unsatisfactory	0.25	62	7	
	0.50	64	3	
	1.0	70	0	
	Subtotal	196 (95%)	10 (5%)	206
Total		385 (97%)	12 (3%)	397

Source: Table 8I, Module III (Vol. 34)

The sponsor also performed a logistic regression to examine the interaction between age group and regimen with the maximum 30-minute post-treatment sedation as the response variable. The interaction term in the regression model came out to be not significant (p=0.75). (Source: Table 68I, Module III, Vol. 35) The "satisfactory" sedation rates by age group and regimen are tabulated in the following table. More discussion regarding the age by regimen interaction can be found in the Discussion section.

Sedation response rate (%): Age Group by Regimen

	Regimen			
Age Group	0.25 mg/kg	0.50 mg/kg	1.0 mg/kg	Overall
0.5 - 2 yrs	94	96	100	97
2 - 6 yrs	88	100	100	96
6 - 16 yrs	98	98	100	98
Overall	93	98	100	97

Source: Table 9I, Module III (Vol. 34)

Secondary Efficacy Analyses

Time to Satisfactory Sedation

This efficacy endpoint was evaluated for all patients with an unsatisfactory baseline sedation score. Overall, out of a total of 206 such patients, 151 (73%) patients reached a satisfactory sedation score within 10 minutes post-treatment, 38 (18%) patients between 10 and 20 minutes, and 7 (3%) patients between 20 and 30 minutes. 5% (10) patients did not reach satisfactory sedation at 30 minutes. Only borderline significant (p=0.05) association between regimen and time to sedation was detected by Cochran-Mantel-Haenszel test stratified by age groups. Sponsor noted that, in general, higher percentages of patients reached satisfactory sedation within 10 minutes in the 0.5 mg/kg and 1.0 mg/kg groups than in the 0.25 mg/kg group. In fact, this trend is more convincingly indicated by the cumulative sedation rate within 20 minutes, as shown in the following table.

Cumulative Sedation Rate (%) by Age and Regimen

Age Group	Regimen (mg/kg)	Time Interval			Unsatisfactory Rate
		0-10 min.	0-20 min.	0-30 min.	
0.5-2 yrs	0.25				9
	0.50				6
	1.0				0
2-6 yrs	0.25				13
	0.50				0
	1.0				0
6-16 yrs	0.25				8
	0.50				7
	1.0				0
Overall					5

Source: Table 17I, Module III (Vol. 34)

Anxiety: based on the highest score achieved within 30 minutes.

Overall, 387 out of 397 patients (97%) achieved a satisfactory anxiety rating post-treatment: 126/132 (95%) in the 0.25 mg/kg group, 129/132 (98%) in the 0.50 mg/kg group, and 132/133 (99%) in the 1.0 mg/kg group. The association between regimen and post-treatment anxiety was borderline significant (p=0.04) by Cochran-Mantel-Haenszel test stratified on baseline anxiety and age group (Table 21I, Module III ,Vol. 34).

Ease of separation from parent/guardian: if the parent or guardian was permitted to be with the child prior to mask induction, ease of separation was assessed at the time of attempted separation. Overall, 350/397 (88%) patients had a satisfactory separation score: 115/132 (87%) in the 0.25 mg/kg group, 116/132 (88%) in the 0.50 mg/kg group, and 119/133 (89%) in the 1.0 mg/kg group. The association between regimen and ease of separation was not significant (p=0.75, Table 32I, Module III, Vol. 35).

The patients' cooperation at the time general mask induction with nitrous oxide and again at introduction of halothane administration was assessed.

Overall, 343/397 (86%) patients had satisfactory *cooperation (nitrous oxide)*: 110/132 (83%), 112/132 (85%), and 121/133 (91%) for the 0.25, 0.50 and 1.0 mg/kg groups, respectively. The test of association between regimen and cooperation (nitrous oxide) was not significant (p=0.07) by the Cochran-Mantel-Haenszel test stratified on baseline anxiety and age group. (Table 39I, Module III, Vol. 34).

Overall, 333/397 (84%) patients had satisfactory *cooperation (halothane)*: 106/132 (80%), 106/132 (80%), and 121/133 (91%) for the 0.25, 0.50 and 1.0 mg/kg groups, respectively. The test of association between regimen and cooperation (halothane) was significant (p=0.03) by the Cochran-Mantel-Haenszel test stratified on baseline anxiety and age group. (Table 46I, Module III, Vol. 34).

Discussion

i) *Protocol Violation and the Taste Test*

As mentioned earlier, 47 out of the 397 ITT patients did not fully comply with the protocol requirements. The main reason was "Full dose not ingested" (see the following table), which accounts for more than 2/3 of total non-compliers (32 out of 47).

Disposition of Study Populations

Study Population	0.25 mg/kg	0.50 mg/kg	1.0 mg/kg	Total
Randomized				405
Intent-to-treat (ITT) -	132	132	133	397
Standard (STD)	123	115	111	350
No. of violators (ITT - STD)	8	17	22	47
Full dose not ingested	2	12	18	32
Others	6	5	4	15

Source: Table 6, Module III (Vol. 34)

Moreover, the "Full dose not ingested" category appeared to have an increasing trend while "Others" remained about the same across the three treatment groups. No explanation was given in sponsor's analyses regarding this issue. The Pearson's Chi-square test on the 2×3 table in the shaded area was significant (p=0.012). Therefore, there might be some evidence that the "Full dose not ingested" category may behave differently from "Others".

The Taste Test performed by the sponsor may be helpful to explore further on this issue. The facial expression of each patient was assessed immediately following the oral dose. Four categories were used to record the outcome: accepted readily, accepted with grimace, accepted with verbal complaint, and rejected entirely or spit out. The first three categories were considered as "satisfactory" and the last one "unsatisfactory". The following table is extracted from sponsor's Table 52i and Table 52s (Module III, Vol. 35).

Taste Test Outcome: the entries represent No. ITT / No. STD

Regimen	Age Group:	6m < 2y	2y < 6y	6y < 16y	Total
0.25 mg/kg	Satisfactory	47 / 45	42 / 40	41 / 39	130 / 124
	Unsatisfactory	1 / 0	1 / 0	0 / 0	2 / 0
	Total	48 / 45	43 / 40	41 / 39	132 / 124
0.50 mg/kg	Satisfactory	42 / 36	37 / 37	44 / 41	123 / 114
	Unsatisfactory	7 / 1	1 / 0	1 / 0	9 / 1
	Total	49 / 37	38 / 37	45 / 41	132 / 115
1.0 mg/kg	Satisfactory	43 / 33	42 / 40	40 / 37	125 / 110
	Unsatisfactory	7 / 1	0 / 0	1 / 0	8 / 1
	Total	50 / 34	42 / 40	41 / 37	133 / 111

Source: Table 52i and Table 52s, Module III (Vol. 35)

As shown in the above table, the most obvious difference between the ITT and STD occurred at the two cells at lower left corner representing patients with age 6m <2y assigned to the 0.50 mg/kg and 1.0 mg/kg groups. They had not only the largest number of non-compliers (12 and 16, respectively) but also the largest difference in the taste "Unsatisfactory" (6 for both cells). This coincidence may suggest that the increased incidence of non-compliance in the youngest patient group could have been associated with the "unsatisfactory" taste of both the 0.50 mg/kg and the 1.0 mg/kg.

ii) *Study Design and the Primary Efficacy Parameter*

It is important to remember that the objective of this trial is to compare the efficacy of the three dosage groups rather than to demonstrate it. The closed-test procedure showed that the low dose (0.25 mg/kg) and high dose (1.0 mg/kg) groups were significantly different in the post-treatment sedation response, while the mid dose (0.50 mg/kg) and high dose (1.0 mg/kg) groups were not.

As the primary efficacy parameter, the "satisfactory" sedation rates achieved by the three treatment groups all looked very impressively high: 93%, 98% and 100% for the low, mid and high dose, respectively. However, since no placebo comparison group presented in the trial, the true attributable improvement due to the drug effect was not available. The following will speculate what the sedation rate might be for a hypothetical placebo group.

According to the definition of sedation score, "satisfactory" meant better than "tearful". Close to half of the patients (48%) were already in a "satisfactory" status at baseline. Suppose that the three sedation measurements (at 10-, 20- and 30-minute post-treatment) are independent of each other, i.e., being "satisfactory" or not at one time point is not related to the statuses at any other points. And also suppose that, at those three post-treatment time points, the patients under placebo would have same probability of being in a certain sedation status as the baseline, i.e., 48% for "satisfactory" and 52% for "unsatisfactory" . Then the probability of having a "satisfactory" sedation for at least one of the three measurements can be approximated as follows:

Prob (at least one "satisfactory" out of three sedation scores)
=1- Prob (all three sedation scores are "unsatisfactory")
=1- Prob ("unsat." at 10 m.)*Prob ("unsat." at 20 m.)*Prob ("unsat." at 30 m.)
=1- (0.52)*(0.52)*(0.52)
=0.86.

In other words, a reasonable approximation of the sedation rate for placebo would be somewhere around 86%.

Remember that this approximation is based on the assumptions that the four observations (baseline + the three post-treatments) were independent of each other and they had same probability of being "satisfactory", or in statistical terms, they were independently and identically distributed. If one believes that correlation may exist between observations, the approximated rate could be affected in either way: a positive correlation (meaning it is more likely being "satisfactory" at one point given being "satisfactory" at the previous point(s)) would reduce the approximation and a negative one would increase it. It is not clear which correlation should dominate in reality. On the other hand, the "same probability" assumption could also be violated in either way. But it seems reasonable to believe that the probability of being "satisfactory" may increase somewhat along the process due to, say, placebo effect and, consequently, the approximated rate should too increase. Therefore, overall, it is quite likely that 86% tends to be a conservative estimate.

The analysis above should be helpful in providing some basis for understanding the range of the attributable efficacy effect of the drug in the absence of a placebo comparison group. With this speculation in mind, the high sedation rates achieved by the three treatment groups might not be as impressive as they looked on the first sight. Furthermore, it was even possible that the high percentage rates might actually have reduced the power for detecting meaningful differences in subgroup analyses, such as the age-by-regimen interaction.

Conclusions

The objective of this trial was to compare the sedative effect of the three doses (0.25 mg/kg, 0.50 mg/kg and 1.0 mg/kg) of Versed syrup. Only the difference in sedation between the high dose (1.0 mg/kg) and low dose (0.25 mg/kg) was statistically significant ($p<0.01$) and the difference between the high dose (1.0 mg/kg) and the mid dose (0.5 mg/kg) was not.

An increased incidence of non-compliance, such as "full dose not ingested", might be associated with the "unsatisfactory" taste of the 0.50 mg/kg and 1.0 mg/kg, especially for the youngest patient group (0.5–2 years of age).

A "satisfactory" sedation was defined as better than "tearful" according to study protocol. The "satisfactory" proportion rate at baseline was 48% overall (31%, 51% and 65% for 0.5-2, 2-6 and 6-16 years, respectively). A post-treatment success was defined as reaching "satisfactory" at any one of the three time points: 10-, 20 and 30-minute post-treatment. It was shown in this review that, given certain conditions, a hypothetical placebo group might obtain an expected sedation rate of more than 86% overall for the same population. Following the exactly same procedure, for age group 6-16 years, the expected placebo success rate would probably reach more than 96%. With a primary efficacy endpoint such defined and expected, the success rates achieved by the three doses (93%, 98% and 100% for the three dose groups, respectively) were not clearly to be "highly" efficacious. Furthermore, this problem might also reduce the sensitivity of detecting meaningful differences in subgroup analyses, such as the age-by-regimen interaction.

Overall, the dose level of 1.0 mg/kg of Versed syrup might offer little advantage in efficacy over the dose level of 0.50 mg/kg and hence may not be needed in clinical practice; the taste of the 0.50 and 1.0 mg/kg Versed may cause an increased incidence of "not full ingestion" in the youngest patient group (0.5–2 years old), but did not seem to seriously damage their efficacy advantage over the 0.25 mg/kg group; further study with a better defined primary efficacy endpoint may be helpful in investigating the age by dosage interaction.

Labeling recommendations:

Z. Jonathan Ma, Ph.D.
Mathematical Statistician

Concur: Thomas Permutt, Ph.D.
Team Leader
Division of Anesthetic, Critical Care
and Addiction Drug Products

Concur: Mike Welch, Ph.D.
Acting Division Director
Division of Biometrics III

Archival: NDA 20-942
cc:
HFD-720/M. Welch
HFD-720 file copy
HFD-344/B. Barton
HFD-170/D. Morgan
HFD-170/I. Cerny
HFD-170/B. Rappaport
HFD-170/C. McCormick
HFD-170/T. Permutt
HFD-170/J. Ma
HFD-170/division file

Statistical Review and Evaluation

NDA 20-942
Drug name: Versed (midazolam HCl), syrup
Applicant: Hoffmann-La Roche Inc.
Drug class: 3P
Indication: Sedation/anxiolysis and amnesia prior to diagnostic/therapeutic or endoscopic procedure or before induction of anesthesia in pediatric patients

Volumes reviewed: 1.1, 1.3, 1.24-61 dated 9 February 1998
(Received HFD-170 10 February 1998)
Electronic data submitted by sponsor
Medical review
Reviewer: Z. Jonathan Ma, Ph.D., HFD-720
User fee date: 9 August 1998

Project manager: David Morgan
Medical reviewer: Igor Cerny, Pharm.D.

Introduction

Midazolam is an approved product in this country for the indication of preoperative sedation/anxiolysis/amnesia through intravenous (IV) or intramuscular (IM) injection in adult and pediatric patients. In clinical sites, however, the parenteral midazolam is often prepared as an oral solution mixed with sweeteners or juices for pediatric patients. This "off-label" use is intended to help to reduce the anxiety experienced by pediatric patients before induction of anesthesia due to fear of the medical procedures.

The sponsor developed Versed (midazolam HCl) syrup in order to provide clinicians with a standardized oral formulation of midazolam for the pediatric population. No placebo-controlled trial was conducted for this NDA submission because the sponsor argued that the efficacy of oral midazolam has been demonstrated in many placebo-controlled trials in the literature. In addition, as quoted by FDA medical reviewer Igor Cerny, Pharm.D., sponsor's study plan had been agreed by former FDA medical reviewer Barbara Palmisano, M.D. (6/27/96): "*A placebo-controlled study to prove efficacy is not planned and is not needed as long as the efficacy for this formulation can be linked to the efficacy of the approved intravenous formulation. This can be done via clinical pharmacology by establishing the pharmacokinetic/pharmacodynamic relationship between the two formulations. The proposed plan offers a pharmacokinetic study to link the two drugs...*".

Therefore, besides a PK/PD study (Protocol #NR15323) and an extensive literature review, a single clinical dose ranging study (Protocol # NR15345) was included in this NDA submission. In this clinical trial, three doses of Versed syrup (0.25 mg/kg, 0.50 mg/kg and 1.0 mg/kg) were studied and compared in efficacy and safety profiles. The main purpose of this trial, as cited by the sponsor, was to provide prescription guidance to clinicians.

The safety results of this study (#NR15345) have been reviewed by the medical reviewer and no major statistical issue needs to be addressed regarding the safety issue. A total of 124 adverse events were reported by 397 patients. A majority of the events were gastrointestinal (49 events) and respiratory (46 events). As shown in the following table, there appeared to have some increasing dose response for "Respiratory" (8, 14 and 24 events for the three regimens, respectively) as well as in the total number (35, 33 and 56 events, respectively). the only two "serious" events identified by sponsor were both respiratory, a 6 year old female in the 0.25 mg/kg group and a 2 year old male in the 0.50 mg/kg group.

Table 1. Summary of adverse events by organ system and regimen

Organ System	0.25 mg/kg N=132	0.50 mg/kg N=132	1.0 mg/kg N=133	Overall N=397
Gastrointestinal	17	10	22	49
Respiratory	8	14	24	46
Psychiatric	1	3	4	8
Cardiovascular	3	3	1	7
Others	6	3	5	14
Total	35	33	56	124

The rest of this statistical review will only focus on the efficacy results from this multiple-dose clinical trial (Protocol # NR15345).

Study Design

This was a multicenter (9 sites), randomized, single-dose, double-blind, parallel-group study in pediatric patients. There were three dose groups:

- Group 1: 0.25 mg/kg,
- Group 2: 0.50 mg/kg, and
- Group 3: 1.0 mg/kg midazolam,

up to a maximum dose of 20 mg. These doses were expected to "*produce sub-optimal, optimal, and slightly excessive pharmacologic effects*" by the sponsor.

All patients were first stratified in three age groups: 6 months–2 years, 2–6 years and 6–16 years, and then randomized to one of the three dosage groups. To maintain blinding, a third-party

pharmacist at each center dispensed the appropriate volume of study drug in to a dispenser, based on patient randomization number (indicating treatment assignment) and patient weight. The dispenser was then given to a study nurse to administer to the patient. Both the observer and the patient were supposedly blinded to the dosage information.

A total of 405 patients who met the inclusion/exclusion criteria were randomized. Eight (8) patients never received any study medication. Hence, a total of 397 patients received treatment medication: 132 in Group 1, 132 in Group 2 and 133 in Group 3. The sponsor considered these 397 patients as the Intent-to-treat (ITT) population.

The inclusion criteria were as follows:

- Physical status I, II or III based on American Society of Anesthesiologists (ASA) Physical Status Classification criteria.
- Males and females 6 months to 16 years old (up to but not including the 16th birthday).
- Patients scheduled for surgical procedures requiring general anesthesia with mask induction.

The exclusion criteria were as follows:

- Were pregnant or nursing
- Had a known or suspected hypersensitivity to midazolam or any benzodiazepine
- Had a known or suspected hypersensitivity to opioids
- Had a documented history of recurrent seizures or a seizure disorder
- Had a gastrointestinal disorder that might have affected absorption
- Had known or suspected hypersensitivity to cherry or cherry products
- Were treated with any investigational drug or participated in a clinical trial within the 4 weeks preceding the study
- Had used any agent (e.g., benzodiazepine) that might interfere with sedation assessment less than 8 hours before dosing
- Had used a macrolide (i.e., clarithromycin or erythromycin) within 7 days of dosing
- Had any medical condition which, in the judgment of the investigator, rendered them inappropriate for participation in the trial.

The primary objective of this trial was:

- To compare the sedative effect of midazolam syrup prior to induction of anesthesia in pediatric patients who received single oral doses of 0.25 mg/kg, 0.5 mg/kg, and 1.0 mg/kg midazolam, up to a maximum dose of 20 mg.

The secondary objectives were:

- To compare the distribution of time to achieve satisfactory sedation among regimens
- To evaluate the anxiolytic properties of oral midazolam syrup as a premedicant
- To evaluate the safety of three single oral doses of midazolam syrup
- To evaluate the effect of oral midazolam syrup on patient cooperation

- To evaluate the effect of oral midazolam syrup on the recovery parameters of patients following surgical procedures.

One exploratory objective was:

- To evaluate the acceptability of the flavored investigational formulation of oral midazolam syrup.

Primary Efficacy Parameter
The sedation assessment was based on the following 5-point scale:

Table 2. Sedation scale

Sedation Score	Description
1	Alert/Active: Agitated, moving, physical or verbal display (loud or high-pitched crying) of apprehension
2	Upset/Wary: Tearful, may be clinging (to parent or guardian)
3	Relaxed: Calm, silly, responds readily to commands or gentle stimulation
4	Drowsy: Easily arousable, responds to mild shaking or prodding
5	Asleep: Unarousable, does not respond to shaking or prodding

A score of 3, 4 or 5 was considered "satisfactory" sedation and a score of 1 or 2 was considered "unsatisfactory". The primary efficacy parameter was the satisfactory sedation rate (proportion) for each dosing regimen based on each patient's highest sedation score within 30-minute post-treatment.

Baseline sedation level was assessed prior to Versed administration and post-treatment sedation levels were measured at 10-minute intervals up to 45 minutes or the general mask induction as the result of an adequate sedation (score of 3, 4 or 5).

Secondary Efficacy Parameters Includes:

1. Distribution of time to satisfactory sedation;
2. Measure of anxiety; including ease of attempted separation from parent/guardian;
3. Measure of cooperation;
4. Time from discontinuation of last inhaled anesthetic agent until recovery.

Anxiety and cooperation were defined as follows:

Table 3. Anxiety scale

Criterion	Description	Score
Poor	Afraid, combative, crying, restrained	1
Fair	Fearful, moderate apprehension	2
Good	Slightly fearful, easily calmed by strangers, non-combative	3
Excellent	No fear or apprehension displayed	4
Not Applicable	Patients asleep	

Table 4. Cooperation scale

Criterion	Description	Score
Poor	Strongly refused intervention	1
Fair	Considerable effort required to achieve compliance with intervention	2
Good	Accepted intervention reluctantly	3
Excellent	Accepted intervention readily	4
Not Applicable	Patient asleep	

In both definitions, a score of 3 or 4 was considered "Satisfactory" and a score of 1 or 2 was considered "Unsatisfactory".

Study Populations

The sponsor performed the efficacy analyses based on two populations: the Intent-to-treat (ITT) population and the Standard (STD) population.

The ITT population was defined in the protocol as "all patients who are randomized, receive any amount of test medication, and have at least one post-baseline assessment.". A total of 405 patients met the inclusion/exclusion criteria and were randomized to enter the study. Eight (8) of them never took any study medication and hence had no post-treatment measurement. They were excluded from the ITT population by definition. The intent-to-treat (ITT) analyses were then performed on these 397 ITT patients.

The STD population was a subgroup of the ITT population, excluding those who did not fully comply with the protocol requirements. Exclusion reasons included "Full dose not ingested", "Vomited before mask induction", "Concomitant drug not allowed", and "Rescue medication given", etc. The STD population had a total of 350 patients since 47 out of the 397 ITT patients did not fully comply with the protocol requirements. The major reason for this violation was "Full dose not ingested", accounting for more than two third of the total violations. More investigation on these 47 protocol non-compliers can be found in the Discussion section.

The main advantage of randomization is to produce the comparability between the comparison

groups with respect with all potential risk factors. The STD population, or sometimes called per protocol population, however, may damage this comparability by removing some subjects depending on different outcome observations. Therefore, the results based on the STD population would then potentially be subject to bias.
As to sponsor's ITT population, only eight (8) patients were excluded from the original randomization population because they did not take any treatment medication and had no post-treatment measures. Considering that this is less than 2% of the total sample size, the comparability between the comparison groups in sponsor's ITT population should remain valid.

In the rest of this review, only analyses based on the ITT population will be discussed in details. According to sponsor's reports, the STD population in general yielded similar results despite potential for bias mentioned earlier.

The ITT Population

The intent-to-treat (ITT) population contained 397 patients who were randomized and dosed with some medication. As shown in Table 4, Module I (Vol. 31), the three treatment groups were comparable with regard to baseline demographic parameters, such as sex, age, and race.

Both male and female patients were recruited into the study. However, approximately twice as many males (264, 66%) as females (133, 34%) were enrolled. The sponsor explained that the reason was that several of the procedures were male-specific. Patient age ranged from 0.5 to 15.3 years with a mean of 4.6 years and a standard deviation of 3.74. Most of them were Caucasian (235, 59%); 93 (23%) were Black, 1 (<1%) was Asian, and 68 (17%) were classified as Other.

As indicated in the following two tables, the satisfactory sedation and anxiety rates at baseline were very similar across three treatment groups and were quite different for different age groups. Older patients typically had larger rates at baseline.

Table 5. Baseline sedation and anxiety by regimen

	Baseline Sedation			Baseline Anxiety*		
Regimen	N	Unsatisfactory	Satisfactory	N	Unsatisfactory	Satisfactory
0.25 mg/kg	132	69 (52%)	63 (48%)	132	52 (39%)	80 (61%)
0.50 mg/kg	132	67 (51%)	65 (49%)	131	50 (38%)	81 (62%)
1.0 mg/kg	133	70 (53%)	63 (47%)	131	51 (39%)	80 (39%)
Total	397	206 (52%)	191 (48%)	394	153 (39%)	241 (61%)

*3 patients not assessable for baseline anxiety
Source: Table 4I and Table 5I, Module III (Vol. 34)

Table 6. Baseline sedation and anxiety by age group

	Baseline Sedation			Baseline Anxiety*		
Age Group	N	Unsatisfactory	Satisfactory	N	Unsatisfactory	Satisfactory
0.5-2	147	101 (69%)	46 (31%)	145	86 (59%)	59 (41%)
2-6	123	60 (49%)	63 (51%)	122	39 (32%)	83 (68%)
6-16	127	45 (35%)	82 (65%)	127	28 (22%)	99 (78%)
Total	397	206 (52%)	191 (48%)	394	153 (39%)	241 (61%)

*3 patients not assessable for baseline anxiety
Source: Table 4I and Table 5I, Module III (Vol. 34)

Primary Efficacy Analyses

The primary efficacy parameter was defined as the satisfactory sedation rate, based on the maximum sedation score achieved within 30 minutes post-treatment, i.e., the highest score among the 10-, 20- and 30-minute measurements. The outcome was collapsed into a binary scale: satisfactory (sedation score=3, 4, or 5) or unsatisfactory (sedation score=1 or 2).

As shown in the following table, overall, 385 out of the 397 (97%) ITT patients achieved satisfactory sedation post-treatment and the satisfaction rate seemed positively associated with the dosage regimen (93%, 98% and 100% for the 0.25, 0.50 and 1.0 mg/kg groups, respectively). There appeared to be a similar trend for Anxiety and other secondary endpoints which will be discussed in more detail later.

Table 7. Number and rate of satisfactory response post-treatment by regimen: ITT population

	Treatment Regimen			
	0.25 mg/kg (n=132) N (%)	0.50 mg/kg (n=132) N (%)	1.0 mg/kg (n=133) N (%)	Overall (n=397) N (%)
Sedation*	123 (93%)	129 (98%)	133 (100%)	385 (97%)
Anxiety*	126 (95%)	129 (98%)	132 (99%)	387 (97%)
Ease of Separation	115 (87%)	116 (88%)	119 (89%)	350 (88%)
Cooperation (nitrous oxide)	110 (83%)	112 (85%)	121 (91%)	343 (86%)
Cooperation (halothane)	106 (80%)	106 (80%)	121 (91%)	333 (84%)

*Each patient's highest score within 30 minutes was used to assess response.

To examine the dose-response relationship, a closed-test procedure was employed by the sponsor. This procedure assumes that the dose effect is monotonic (seems true from Table 7) and sequentially performs two pairwise comparisons: low vs. high and mid vs. high. If the first null hypothesis, i.e., equality between the low and high doses, is rejected, move on to test the second null hypothesis, i.e., equality between the mid and high doses. If the first null hypothesis is not rejected, stop the procedure and conclude that the second null hypothesis is not rejected

either. The main advantage of the closed-test procedure is to protect the type I error from inflation due to multiple testing. (See references 1 and 2.)

In the sponsor's report, the test for the difference between the 0.25 mg/kg and 1.0 mg/kg groups in sedation rate yielded a p-value of 0.002 (Cochran-Mantel-Haenszel, row mean score statistic; Table 12I, Module III, Vol. 34) and hence the first null hypothesis was rejected at 0.05 level. Then the test for the difference between the 0.50 mg/kg and 1.0 mg/kg groups yielded a p-value of 0.08. Therefore, second null hypothesis was not rejected at 0.05 level. It should be noted that both tests performed by sponsor were one-sided tests at a significance level of 0.05. The Agency however conventionally requires a two-sided test at a 0.05 level or a one-sided test at a 0.025 level. This obviously does not alter the above conclusions.

The Cochran-Mantel-Haenszel (CMH) test, which was used to test for a general association between regimen and response at each stage in the primary analyses, was stratified on both age group and baseline sedation score in the final statistical analyses. It should be noted that the statistical plan in the original protocol included age group as the only stratification variable. In the final analyses, however, baseline sedation score was also adjusted for in all tests.

As shown in the following table, the post-treatment "satisfactory" sedation rate was 99% for patients with "satisfactory" sedation at baseline and 95% for those with "unsatisfactory" sedation at baseline. Therefore, the baseline sedation seemed well related to the post-treatment sedation and hence needed to be adjusted for as well.

Table 8. No. of patients by baseline sedation, post-treatment sedation and regimen

Baseline		Post-Treatment		
Sedation	Regimen (mg/kg)	Satisfactory	Unsatisfactory	Total
Satisfactory	0.25	61	2	
	0.50	65	0	
	1.0	63	0	
	Subtotal	189 (99%)	2 (1%)	191
Unsatisfactory	0.25	62	7	
	0.50	64	3	
	1.0	70	0	
	Subtotal	196 (95%)	10 (5%)	206
Total		385 (97%)	12 (3%)	397

Source: Table 8I, Module III (Vol. 34)

To examine whether baseline sedation made a difference, this reviewer repeated the CMH tests with the age group as the only stratification variable. The p-values are tabulated in the following table in comparison with the ones using both stratification variables. The two columns are quite similar.

Table 9. P-values from the CMH tests with different stratification variables

Regimen	Stratified on age group only	Stratified on both age group and baseline sedation score
0.25 vs. 1.0	0.002	0.002
0.50 vs. 1.0	0.075	0.080
0.25 vs. 0.50	0.085	0.082

The sponsor also performed a logistic regression to examine the interaction between age group and regimen with the maximum 30-minute post-treatment sedation as the response variable. The interaction term in the regression model came out to be not significant (p=0.75). (Source: Table 68I, Module III, Vol. 35) The "satisfactory" sedation rates by age group and regimen are tabulated in the following table. More discussion regarding the age by regimen interaction can be found in the Discussion section.

Table 10. Sedation response rate (%): Age Group by Regimen

	Regimen			
Age Group	0.25 mg/kg	0.50 mg/kg	1.0 mg/kg	Overall
0.5 - 2 yrs	94	96	100	97
2 - 6 yrs	88	100	100	96
6 - 16 yrs	98	98	100	98
Overall	93	98	100	97

Source: Table 9I, Module III (Vol. 34)

Secondary Efficacy Analyses

Time to Satisfactory Sedation

This efficacy endpoint was evaluated for all patients with an unsatisfactory baseline sedation score. Overall, out of a total of 206 such patients, 151 (73%) patients reached a satisfactory sedation score within 10 minutes post-treatment, 38 (18%) patients between 10 and 20 minutes, and 7 (3%) patients between 20 and 30 minutes. 5% (10) patients did not reach satisfactory sedation at 30 minutes. To test the association between regimen and time to sedation, Cochran-Mantel-Haenszel test (Row mean score statistic) was performed on the sedation counts (3×4 tables) stratified by age group in Table 11 and only borderline significance (p=0.05) was obtained. In addition, sponsor noted that, in general, higher percentages of patients reached satisfactory sedation within 10 minutes in the 0.5 mg/kg and 1.0 mg/kg groups than in the 0.25 mg/kg group. In fact, this trend is more convincingly indicated by the cumulative sedation rate within 20 minutes, as shown in Table 11 (percentages in parentheses).

Table 11. Sedation Counts and Cumulative Sedation Rate (%) by Age and Regimen

Age Group	Regimen (mg/kg)	Time Interval 0-10 min.	0-20 min.	0-30 min.	Unsatisfactory Rate
0.5-2 yrs	0.25 (N=34)				3 (9%)
	0.50 (N=34)				2 (6%)
	1.0 (N=33)				0 (0%)
2-6 yrs	0.25 (N=23)				3 (13%)
	0.50 (N=18)				0 (0%)
	1.0 (N=19)				0 (0%)
6-16 yrs	0.25 (N=12)				1 (8%)
	0.50 (N=15)				1 (7%)
	1.0 (N=18)				0 (0%)
Overall	N=206				10 (5%)

Source: Table 17I, Module III (Vol. 34)

Anxiety: based on the highest score achieved within 30 minutes.
Overall, 387 out of 397 patients (97%) achieved a satisfactory anxiety rating post-treatment: 126/132 (95%) in the 0.25 mg/kg group, 129/132 (98%) in the 0.50 mg/kg group, and 132/133 (99%) in the 1.0 mg/kg group. The association between regimen and post-treatment anxiety was borderline significant (p=0.04) by Cochran-Mantel-Haenszel test stratified on baseline anxiety and age group (Table 21I, Module III ,Vol. 34).

Ease of separation from parent/guardian: if the parent or guardian was permitted to be with the child prior to mask induction, ease of separation was assessed at the time of attempted separation. Overall, 350/397 (88%) patients had a satisfactory separation score: 115/132 (87%) in the 0.25 mg/kg group, 116/132 (88%) in the 0.50 mg/kg group, and 119/133 (89%) in the 1.0 mg/kg group. The association between regimen and ease of separation was not significant (p=0.75, Table 32I, Module III, Vol. 35).

The patients' cooperation at the time general mask induction with nitrous oxide and again at introduction of halothane administration was assessed.

Overall, 343/397 (86%) patients had satisfactory *cooperation (nitrous oxide)*: 110/132 (83%), 112/132 (85%), and 121/133 (91%) for the 0.25, 0.50 and 1.0 mg/kg groups, respectively. The test of association between regimen and cooperation (nitrous oxide) was not significant (p=0.07)

by the Cochran-Mantel-Haenszel test stratified on baseline anxiety and age group. (Table 39I, Module III, Vol. 34).

Overall, 333/397 (84%) patients had satisfactory *cooperation (halothane)*: 106/132 (80%), 106/132 (80%), and 121/133 (91%) for the 0.25, 0.50 and 1.0 mg/kg groups, respectively. The test of association between regimen and cooperation (halothane) was significant (p=0.03) by the Cochran-Mantel-Haenszel test stratified on baseline anxiety and age group. (Table 46I, Module III, Vol. 34).

Discussion

i) *Protocol Violation and the Taste Test*

As mentioned earlier, 47 out of the 397 ITT patients did not fully comply with the protocol requirements. The main reason was "Full dose not ingested" (see the following table), which accounts for more than 2/3 of total non-compliers (32 out of 47).

Table 12. Disposition of Study Populations

Study Population	0.25 mg/kg	0.50 mg/kg	1.0 mg/kg	Total
Randomized				405
Intent-to-treat (ITT)	132	132	133	397
Standard (STD)	123	115	111	350
No. of violators (ITT - STD)	8	17	22	47
Full dose not ingested	2	12	18	32
Others	6	5	4	15

Source: Table 6, Module III (Vol. 34)

Moreover, the "Full dose not ingested" category appeared to have an increasing trend while "Others" remained about the same across the three treatment groups. No explanation was given in sponsor's analyses regarding this issue. The Pearson's Chi-square test on the 2×3 table in the shaded area was significant (p=0.012). Therefore, there might be some evidence that the "Full dose not ingested" category may behave differently from "Others".

The Taste Test performed by the sponsor may be helpful to explore further on this issue. The facial expression of each patient was assessed immediately following the oral dose. Four categories were used to record the outcome: accepted readily, accepted with grimace, accepted with verbal complaint, and rejected entirely or spit out. The first three categories were considered as "satisfactory" and the last one "unsatisfactory". The following table is extracted from sponsor's Table 52i and Table 52s (Module III, Vol. 35).

Table 13. Taste Test Outcome: the entries represent No. ITT / No. STD

Regimen	Age Group:	6m < 2y	2y < 6y	6y < 16y	Total
0.25 mg/kg	Satisfactory	47 / 45	42 / 40	41 / 39	130 / 124
	Unsatisfactory	1 / 0	1 / 0	0 / 0	2 / 0
	Total	48 / 45	43 / 40	41 / 39	132 / 124
0.50 mg/kg	Satisfactory	42 / 36	37 / 37	44 / 41	123 / 114
	Unsatisfactory	7 / 1	1 / 0	1 / 0	9 / 1
	Total	49 / 37	38 / 37	45 / 41	132 / 115
1.0 mg/kg	Satisfactory	43 / 33	42 / 40	40 / 37	125 / 110
	Unsatisfactory	7 / 1	0 / 0	1 / 0	8 / 1
	Total	50 / 34	42 / 40	41 / 37	133 / 111

Source: Table 52i and Table 52s, Module III (Vol. 35)

As shown in the above table, the most obvious difference between the ITT and STD occurred at the two cells at lower left corner representing patients with age 6m <2y assigned to the 0.50 mg/kg and 1.0 mg/kg groups. They had not only the largest number of non-compliers (12 and 16, respectively) but also the largest difference in the taste "Unsatisfactory" (6 for both cells). This coincidence may suggest that the increased incidence of non-compliance in the youngest patient group could have been associated with the "unsatisfactory" taste of both the 0.50 mg/kg and the 1.0 mg/kg.

ii) *Study Objective and Design*

Conventionally, the primary objective of a phase III clinical trial should be to demonstrate the efficacy of a treatment drug. In this trial, only the high dose Versed (1.0 mg/kg) was shown to be effective in this pediatric population if the low dose (0.25 mg/kg) was considered as an active control and could be assumed no worse than a placebo. According to this reviewer's analysis (Table 9), the mid dose (0.50 mg/kg) was not significantly different from the low dose (p=0.082) and hence failed to demonstrate its efficacy in this trial.

However, it seems to this reviewer that the objective of this trial was more to compare the efficacy of the three dosage groups rather than to demonstrate it. Two facts may be considered. First, this was not a placebo-controlled trial. Second, sponsor planned no pairwise comparison between the low and mid doses and hence might not primarily consider the low dose as an active control. The closed-test procedure was designed to compare the low and mid doses to the high dose rather than to compare the high and mid doses to the low dose.

iii) *The Primary Efficacy Analyses*

As the primary efficacy parameter, the "satisfactory" sedation rates achieved by the three treatment groups all appeared high: 93%, 98% and 100% for the low, mid and high dose, respectively. However, since no placebo comparison group was present in the trial, the true

attributable improvement due to the drug effect was not available. The following speculates what the sedation rate might be for a hypothetical placebo group.

According to the definition of sedation score, "satisfactory" meant better than "tearful". Close to half of the patients (48%) were already in a "satisfactory" status at baseline. Suppose that the three sedation measurements (at 10-, 20- and 30-minute post-treatment) are independent of each other, i.e., being "satisfactory" or not at one time point is not related to the statuses at any other points. And also suppose that, at those three post-treatment time points, the patients under placebo would have same probability of being in a certain sedation status as at baseline, i.e., 48% for "satisfactory" and 52% for "unsatisfactory". Then the probability of having a "satisfactory" sedation for at least one of the three measurements can be approximated as follows:

Prob (at least one "satisfactory" out of three sedation scores)
=1- Prob (all three sedation scores are "unsatisfactory")
=1- Prob ("unsat." at 10 m.)*Prob ("unsat." at 20 m.)*Prob ("unsat." at 30 m.)
=1- (0.52)*(0.52)*(0.52)
=0.86.

In other words, a reasonable approximation of the sedation rate for placebo would be somewhere around 86%.

This approximation is based on the assumptions that the four observations (baseline + the three post-treatments) were independent of each other and they had same probability of being "satisfactory", or in statistical terms, they were independently and identically distributed. If one believes that correlation may exist between observations, the approximated rate could be affected in either way: a positive correlation (meaning it is more likely being "satisfactory" at one point given being "satisfactory" at the previous point(s)) would reduce the approximation and a negative one would increase it. It is not clear which correlation should dominate in reality. On the other hand, the "same probability" assumption could also be violated in either way. But it seems reasonable to believe that the probability of being "satisfactory" may increase somewhat along the process due to, say, placebo effect and, consequently, the approximated rate should too increase. Therefore, overall, 86% may be a conservative estimate.

Following the exactly same procedure, for age group 6-16 years which has a baseline sedation rate of 65%, the expected placebo success rate would be: 1-0.35*0.35*0.35=96%.

The analysis above should be helpful in providing some basis for understanding the range of the attributable efficacy effect of the drug in the absence of a placebo comparison group. With this speculation in mind, the high sedation rates achieved by the three treatment groups might not be as impressive as they looked at first sight. Furthermore, it was even possible that the high percentage rates might actually have reduced the power for detecting meaningful differences in subgroup analyses, such as the age-by-regimen interaction.

Conclusions

An increased incidence of non-compliance, such as "full dose not ingested", might be associated with the "unsatisfactory" taste of the 0.50 mg/kg and 1.0 mg/kg, especially for the youngest patient group (0.5–2 years of age).

A "satisfactory" sedation was defined as better than "tearful" according to study protocol. The "satisfactory" proportion rate at baseline was 48% overall (31%, 51% and 65% for 0.5-2, 2-6 and 6-16 years, respectively). A post-treatment success was defined as reaching "satisfactory" at any one of the three time points: 10-, 20 and 30-minute post-treatment. It was shown in this review that, given certain conditions, a hypothetical placebo group might obtain an expected success sedation rate of more than 86% overall for the same population. This expected success rate could reach even higher for some sub-populations, e.g., 96% for age group 6-16 years. With a primary efficacy endpoint expected to this high for a hypothetical placebo arm, the success rates achieved by the three doses (93%, 98% and 100% for the low, mid and high, respectively) were not clearly shown to be "highly" efficacious. Furthermore, the narrow margins of efficacy differences among dose groups might also reduce the sensitivity of detecting meaningful differences in subgroup analyses, such as the age-by-regimen interaction.

It seems to this reviewer that the primary objective of this clinical trial was to compare the efficacy of the three doses of Versed (0.25, 0.50 and 1.0 mg/kg) rather than to demonstrate it. This is somewhat unusual for a phase III clinical trial. Besides the brief note regarding this issue on Page 1, more detailed rationale argued by the sponsor can be found in the medical reviewer's report.

Overall, the high dose (1.0 mg/kg) of Versed syrup demonstrated its efficacy in this trial if the low dose (0.25 mg/kg) was considered as an active control; however, the high dose (1.0 mg/kg) was not shown to be more effective over the mid dose (0.50 mg/kg) and nor was the mid dose (0.50 mg/kg) over the low dose (0.25 mg/kg); the taste of the 0.50 and 1.0 mg/kg Versed may have caused an increased incidence of "not full ingestion" in the youngest patient group (0.5–2 years old), but did not seem to seriously alter their efficacy when comparing with the 0.25 mg/kg group; further study with a better defined primary efficacy endpoint may be helpful in investigating the age by dosage interaction.

Labeling recommendations:

Reference

1. Bauer, P. Multiple testing in clinical trials. *Statistics in Medicine*, Vol. 10, 871-90 (1991).

2. Rom, DM, Costello, RJ, and Connell LT. On closed test procedures for dose-response. *Statistics in Medicine*, Vol. 13, 1583-96 (1994).

/S/ 6/3-/98

Z. Jonathan Ma, Ph.D.
Mathematical Statistician

/S/ 6/30/98

Concur: Thomas Permutt, Ph.D.
Team Leader
Division of Anesthetic, Critical Care
and Addiction Drug Products

/S/ 7/10/98

Concur: Mike Welch, Ph.D.
Acting Division Director
Division of Biometrics III

Archival: NDA 20-942
cc:
HFD-720/M. Welch
HFD-720 file copy

HFD-344/B. Barton
HFD-170/D. Morgan
HFD-170/I. Cerny
HFD-170/B. Rappaport
HFD-170/C. McCormick
HFD-170/T. Permutt
HFD-170/J. Ma
HFD-170/division file