CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER: 19-627/S-035

STATISTICAL REVIEW

Statistical Review and Evaluation

NDA 19-627/S-035
Name of drug: DIPRIVAN (propofol)
Applicant: Zeneca Pharmaceuticals
Indication: anesthesia/sedation
Documents reviewed: Pediatric supplement, volumes 1-24
Project manager: Laura Governale
Clinical reviewer: Pat Hartwell, Carole Freeland
Dates: submitted 21 May 1999

Reviewer: Stella Grosser

Diprivan (propofol) injectable emulsion is an intravenous sedative-hypnotic agent for use in the induction and maintenance of anesthesia. This pediatric supplement reports on two trials, one to determine the safety and efficacy of 1% Diprivan and 2% Diprivan for sedation of pediatric subjects in the ICU, and the second to determine the safety of 1% Diprivan used as anesthesia in surgery for pediatric subjects less than 36 months of age.

Trial 1 A multicenter, comparative, randomized trial to determine the overall safety and efficacy of 1% Diprivan ™ vs. 2% Diprivan vs. Standard Agents without Disodium Edetate for sedation of trauma, postsurgical or critically ill pediatric subjects (0859IL/0068-Pediatric Trial 1)

The primary objective was "to compare the safety and efficacy of Diprivan 2% versus Diprivan 1% versus standard sedative agents without disodium edetate (SSA) in trauma, postsurgical, and critically ill pediatric patients" [Summary, v. 2. p. 15]. The study population consisted of "approximately 360 trauma, postsurgical, or critically ill patients who were on mechanical ventilation, expected to require sedation for at lest 24 hours, and had comfort scores of 26 points or greater"; patients were stratified by age (0-1 year, 2-11 years, and 12-16 years) and randomized within stratum to one of the three arms. Twenty-four centers participated in this trial.

The information contained in this report pertains mostly to the safety and appropriate dosing rather than the efficacy of Diprivan 1% or 2% compared to standard agents. The clinical review looks at these issues in depth. Furthermore, there are no comparative efficacy claims in the label. In my statistical review of this trial, therefore, I comment on

an issue that the sponsor raised and examined in Appendix H, "Supplementary statistical information" (Vol. 11), in the "Exploratory Analysis for Mortality."

An imbalance in the numbers of deaths was observed among the patients randomized to the Diprivan groups (1% or 2%) compared to SSA. Approximately 50% of the deaths overall were from three centers (1 = University Hospital of Cleveland, 5 = Baylor College/ Texas Children's Hospital, 10 = West Virginia University); in center 5, with 17% of the study subjects, eleven deaths were observed, which constituted 44% of the total number of deaths.

(Table 1, "Mortality by center and treatment, " v. 11, p. 290)

CENTER	Diprivan 2%		Diprivan 1%		Standard Sedative Agents	
	Treated	Died	Treated	Died	Treated	Died
1	16	1	18	0	17	0
5	19	5	20	5	17	1
10	19	3	16	1	17	2
Others	59	3	55	3	54	1
All	113	12	109	9	105	4

The sponsor says (v. 11, p. 290):

> An exploratory analysis was done to attempt to identify whether the unexpected finding was in fact due to Diprivan or imbalances that were predictive of death. It is recognized that the number of deaths is small and having the ability to detect associations between any factor and death is low. Baseline information considered to be influential to the mortality rates include age, center, PRISM [Pediatric Risk of Mortality] score, number of intubation days before trial drug start, lipids and TPN administered, sepsis, triglycerides, abnormal inorganic phosphate, (<2.2 mg/dl), and the number of concomitant underlying diseases.

Having identified, apparently from a clinical point of view, factors possibly associated with mortality, the sponsor then used stepwise logistic regression to choose among them: "to search for the statistically significant baseline factors associated with mortality excluding treatment." Center, TPN administered, and PRISM score were identified as jointly associated with mortality using this method. Finally, controlling for these factors statistically, treatment was found to not be associated with mortality.

This analysis by the sponsor is inconclusive. Stepwise logistic regression is a mechanical, computerized variable selection procedure that may or may not identify important predictive factors; simulation studies have shown that it often identifies variables that were designed to be uncorrelated with the outcome as being significantly associated with it. In addition, stepwise logistic regression is even more likely to give inconclusive or misleading results in situations where the proportion of deaths is relatively small and there is a high

association among possible factors (which is the case here: "number of intubation days before trial drug start, lipids administered, triglycerides and sepsis were all highly associated with TPN administration"). It is also not clear how "center" was used in the final regression model—as 24 distinct factors, in 4 groups (center 1, 5, 10, and other), or dichotomized (5 vs. all others).

A more productive approach would be to examine each death for its relation to treatment or to argue, as the sponsor does, that the patients who died were sicker at baseline (Table 2, v. 11, p. 292).

Center 5 and to a lesser extent center 10 had a disproportionate number of deaths. The difference between center 5 and the rest of the centers, combined, in the number of deaths is statistically significant ($p<0.001$). Although the numbers are small, 20% (11/56) of the patients at center 5 died, compared to 12% (6/52) at center 10, and 4% at all other centers combined. The sponsor compares baseline information for patients at center 5 to that at all others (Table 3, v11, p 292). It finds that the mean and median PRISM at center 5 is slightly lower (6.11 and 5 vs. 7.66 and 7), indicating lower risk of mortality; the difference is not statistically significant. There are statistically significant differences ($p<0.05$) between center 5 and the others in mean number of intubation days (7.96 vs. 3.65), percent receiving lipid administration (23% vs. 7%), percent receiving TPN administration (32% vs. 19%), and percent with sepsis (25% vs. 17%). Mean triglycerides is marginally statistically significantly ($p=0.052$) higher at center 5 (140 vs. 120).

The disproportion in the number of deaths at center 5, however, is more marked in the group treated with Diprivan. The ratio of the odds of death – Diprivan to SSA – is 5.5 to 1 at center 5 and 1.8 to 1 at all other centers combined. Arguing that the patients at center 5, in all treatment arms combined, and those given Diprivan, at all centers overall, happen to be sicker does not answer the question of why there is an excess number of deaths in the Diprivan-treated patients at center 5. The numbers are indeed small and the observed excess may indeed be due to chance; however, this issue can not be resolved with statistical analysis.

Conclusion

There is an excess number of deaths in the Diprivan treatment arms and in center 5 and in particular in the interaction between the two. The statistical analysis by the sponsor attempts to, but does not, put this problem to rest; the deaths, specifically those at center 5, should be examined in detail by the medical officer as a safety issue.

Trial 2 The safety of Diprivan ™ (propofol) anesthesia versus standard anesthetic techniques in pediatric subjects less than 36 months of age (0859US/0046)

The primary objective is "to compare the safety profile of Diprivan versus standard anesthetic technique" (v. 14, p. 7) in surgery. Among the secondary objectives is the goal of comparing the recovery profile of Diprivan versus standard anesthetic technique.

One hundred and five subjects, who were less than 36 months of age and who were admitted for surgical or other procedures expected to last for 15 minutes or more, were randomized to Diprivan or to standard anesthetic. One hundred and three subjects were included in the safety analysis (51 Diprivan and 52 standard anesthetic) and 94 (51 Diprivan, 43 standard) were included in the efficacy analysis. Three centers contributed patients to this trial.

The sponsor used five efficacy measures, most of which are times to various recovery endpoints: "time to extubation, time to spontaneous eye opening, time to response to verbal stimuli, time to complete recovery... and incidence of postoperative vomiting." (v. 14, p. 9)

This was primarily a safety study, and no efficacy claims are made in the label. The sponsor does mention in summary (v. 14, p. 9) that "mean recovery times were longer for the Diprivan group than the standard anesthetic group; however, there was no statistically significant difference between treatment groups in any of the recovery endpoints. The longer recovery times in the Diprivan group may be related to the higher number of Diprivan-treated subjects undergoing cardiopulmonary bypass and/or to a center effect." Tables 8 and 9, v.14, pp. 35 and 36, give descriptive statistics for each treatment arm overall and by age group.

Several problems arise with the sponsor's analysis that I mention but, given that label claims are not made, I did not investigate further as efficacy issues. Not only are the mean and median times longer for the Diprivan-treated subjects than for those treated with standard anesthetic, but also the extreme values tend to be higher in the former. For example, time to spontaneous eye opening ranges from 6 to 118 minutes among the Diprivan subjects, but from 2 to 45 minutes among the comparator group. (The mean, median times are 24.5, 14.5 and 14.9, 13.0 minutes, respectively.) The tendencies toward higher values are attributed by the sponsor to excess numbers of bypass subjects in the Diprivan-treated subjects, although longer times are not limited solely to bypass surgeries, and to an unspecified effect due to Diprivan at Center 1 (West Virginia University), where most of the bypass surgeries took place.

Another technical issue is that recovery times "were compared between groups using analysis of covariance (ANCOVA). The model included treatment and center as factors and age, duration of anesthesia, and surgical procedures as covariates" (v 14, p.22). While this analysis is the one specified in the protocol, there are better approaches: measurements of time-to-an-event often have many relatively small values and a few large ones. That is, the distribution of values is skewed. At the least, a transformation to normality before carrying out an ANCOVA, or a survival analysis, such as fitting a Cox proportional hazards regression model, would have been a more appropriate way to analyze recovery times. The power of the resulting test, and therefore the power to detect a difference between groups, would have been increased, and the observed mean

delay in recovery times for Diprivan patients relative to those given SSA may have been found significant.

Conclusions While there are flaws in the statistical analysis, none lead directly to the conclusion that Diprivan is unsafe, and no claims are made for efficacy.

Recommendations

There are no statistical considerations suggesting that Diprivan should not be used in pediatric populations. I suggest, however, that the medical officer review the deaths, paying special attention to potential problems at Center 5 in Trial 1 (0859IL/0068).

Labeling

There are no claims made in the proposed label for efficacy. None of the changes in the label is based on formal statistical tests from either of these two trials.

/S/

Stella Grosser, Ph.D.
Mathematical Statistician

/S/

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

Memorandum

To: Cynthia McCormick, MD
Division Director, DACCADP
Through: Thomas Permutt, PhD
Team Leader, Biostatistics
From: Stella Grosser, PhD
Mathematical Statistician
Date: 3 October 2000
Re: NDA 19-627, S-035 (Diprivan, Pediatric ICU study)

The following issue was raised during our meeting with the Pediatric Implementation Team on Sept. 13, 2000, in our discussion of Diprivan (propofol), study 0859IL/0068--Pediatric Trial 1.

The primary objective of this study was to compare the safety and efficacy of Diprivan 2% versus Diprivan 1% versus standard sedative agents without disodium edetate in trauma, postsurgical, and critically ill pediatric patients. As can be seen in the table below, an imbalance in the numbers of deaths was observed among the patients randomized to the Diprivan groups (1% or 2%) compared to the standard agents.

Diprivan 2%		Diprivan 1%		Standard Sedative Agents	
Treated	Died	Treated	Died	Treated	Died
113	12	109	9	105	4

The question was raised as to the level of statistical significance attached to the observed association between administration of Diprivan and death. Dianne Murphy requested that this information be sent to the committee.

I carried out chi-square tests of this association, comparing various combinations of the Diprivan-treated groups and the standard sedative agents (SSA). For the record, the results are shown below:

Groups Compared	p-value
Diprivan (both groups, combined) vs. SSA	.073
Diprivan 2% vs. SSA:	.054
Diprivan 1% vs. SSA:	.173
Diprivan 1% vs. Diprivan 2 %:	.548

None of these tests achieve significance at the .05 level, although the test of association of death with Diprivan 2% relative to SSA comes close. Note that lack of evidence here does not definitively show that there is a lack of association, as the study was not designed to investigate this particular outcome.