

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

021746Orig1s000

STATISTICAL REVIEW(S)



DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

ADDENDUM TO STATISTICAL REVIEW AND EVALUATION CLINICAL STUDIES

NDA #/Serial #: 21-746
DRUG NAME: SURFAXIN (Lucinactant 30 mg/mL) Intratracheal
Suspension
INDICATION: Prevention of Respiratory Distress Syndrome in
Premature Infants
APPLICANT: Discovery Laboratories, Inc.
DATE: October 05, 2005
REVIEW PRIORITY: Standard Review
BIOMETRICS DIVISION: Office of Biostatistics
STATISTICAL REVIEWER: Sue-Jane Wang, Ph.D. (HFD-700)
CONCURRING REVIEWER: Ruthie Davi, M.S., (HFD-715), Pulmonary Team Leader
MEDICAL DIVISION: Division of Pulmonary Drug Product (HFD-570)
CLINICAL TEAM: John Gunkel, M.D., Peter Starke, M.D.
PROJECT MANAGER: Christina Yu (HFD-570)

KEY WORDS: Non-inferiority trial

Review file: DFS

BACKGROUND

Discovery Labs Inc. (the Applicant) submitted its complete response to the deficiencies outlined in the US Food and Drug Administration's (FDA) Approvable Letter dated Feb. 11, 2005 regarding this NDA. The Applicant resubmitted all of the responses including those specific responses that the Agency deemed partial in their August 16, 2005 facsimile.

This addendum is to provide discussion taken place in the process of the drug development and in preparation of this NDA submission. Specifically, this reviewer summarizes Agency's concerns in providing Comment 29b: Studies KL4-IRDS-02 and KL4-IRDS-05 did not "demonstrate efficacy." Efficacy in prevention was demonstrated by only a single study, KL4-IRDS-06. Study KL4-IRDS-02 was of flawed design and the non-inferiority margins are not supportable. (b) (4)

REVIEWER COMMENTS

The non-inferiority efficacy design of study KL4-IRDS-02 seeking that surfaxin would be non-inferior to the comparator Curosurf was not agreed by the Agency. The Agency indicated that the information might be used for safety information.

Below are some of the Agency's concerns with the inappropriate non-inferiority efficacy design of study KL4-IRDS-02. First, there was only one placebo controlled trial available providing limited information on the variability of the Curosurf effect (relative to placebo). The between trial variability cannot be assessed. Secondly, if one assumes that the Curosurf effect could be estimated based on the only trial, the clinical review team discussed a much higher percentage (higher than 50%) preservation of the Curosurf effect that should be used to define the non-inferiority margin. Thirdly, there were concerns on the medical practice changes over time since the approval of Curosurf in 1990s and the ability to correctly quantify the Curosurf effect with limited historical data. Therefore, the convention of using the worst 95% confidence interval limit from the placebo controlled trial to define the non-inferiority margin was recommended. The approach was used in the FDA/CBER considerations on selected aspects of active controlled trial design and analysis for the evaluation of thrombolytics in acute MI discussed in the 1992 advisory committee meeting. In other words, the use of point estimate of Curosurf effect relative to placebo from the placebo controlled trial to define the non-inferiority margin of preserving 50% of the Curosurf effect was not agreed by the Agency. The Agency conveyed to the Applicant and the Applicant agreed that study KL4-IRDS-02 will not be used as a basis for efficacy evaluation, but to provide safety information. Thus, study KL4-IRDS-02 was reviewed on the ground of safety. It is noted that study KL4-IRDS-02 was prematurely terminated. Early termination of a trial often limits the ability to evaluate safety, as safety data of the drug should be collected and evaluated on the longer term basis. According to the Applicant, it was due to economic reasons.

Sue-Jane Wang, Ph.D.
Associate Director, Office of Biostatistics

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

/s/

Sue Jane Wang
2/27/2006 03:27:48 PM
BIOMETRICS

Ruth Davi
2/27/2006 03:30:59 PM
BIOMETRICS



DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE
FOOD AND DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION AND RESEARCH

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA #/Serial #: 21-746/N000
DRUG NAME: SURFAXIN (Lucinactant 30 mg/mL) Intratracheal Suspension
INDICATION: Prevention of Respiratory Distress Syndrome in Premature Infants
APPLICANT: Discovery Laboratories, Inc.
DATE: April 13, 2004
REVIEW PRIORITY: Standard Review
BIOMETRICS DIVISION: Division of Biometrics II
STATISTICAL REVIEWER: Sue-Jane Wang, Ph.D. (HFD-715)
DIVISION DIRECTOR: S. Edward Nevius, Ph.D. (HFD-715)
MEDICAL DIVISION: Division of Pulmonary Drug Product (HFD-570)
CLINICAL TEAM: John Gunkel, M.D., Peter Starke, M.D.
PROJECT MANAGER: Chirstine Yu (HFD-570)

KEY WORDS: co-primary endpoints, adaptive design

Review file: c:\data\my ocuments\hfd570\NDA-surfaxin.doc

TABLE OF CONTENTS

1. EXECUTIVE SUMMARY	3
1.1 Conclusions and Recommendations.....	3
1.2 Brief Overview of Clinical Studies	3
1.3 Statistical Issues and Findings.....	3
2. INTRODUCTION.....	4
2.1 Overview	6
2.2 Data Sources.....	6
3. STATISTICAL EVALUATION.....	6
3.1 Evaluation of Efficacy (KL4-IRDS-06)	6
3.2 Evaluation of Safety	20
4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	20
4.1 Gender, Race and Age.....	20
4.2 Other Special/Subgroup Populations.....	20
4.3 Regional Analysis.....	21
5. SUMMARY AND CONCLUSIONS	22
5.1 Statistical Issues and Collective Evidence	22
5.2 Conclusions and Recommendations.....	24
APPENDICES.....	25
KL4-IRDS-06 Trial	25
1. Retreatment Criteria	25
2. Baseline Characteristics (Sponsor Table 11.2.A)	26
3. Maternal History (Sponsor Table 11.2.B)	27
4. Definition of Incidence of RDS at 24 hours (p.5 of AC SOP)	28
5. RDS related mortality through 14 days of age (p.6-7 of AC SOP)	28
6. Definition of air leak through 7 days of age (p.7 of AC SOP)	29
7. Sample Size Re-Estimation	29
8. Change of co-primary efficacy endpoint	29

1. EXECUTIVE SUMMARY

1.1 Conclusions and Recommendations

Surfaxin[®] yielded a statistically significant reduction in respiratory distress syndrome (RDS) incidence at 24 hours of age and in RDS related mortality by 14 days as compared to Exosurf[®], based on the adjudicated data of study KL4-IRDS-06. However, there was a large discrepancy between the Adjudicated Committee's assessment and CRF reporting in terms of these incidence rates. The case report form (CRF) data did not support the RDS mortality finding. The composite of RDS mortality by 14 days or air leak by day 7 showed a close-to-borderline significant difference in favor of Surfaxin.

All cause mortality by 14 days showed little difference between Surfaxin and Exosurf. Non-RDS related mortality was significantly larger for Surfaxin. Such pattern was observed consistently in both the European region and the Latin American region. The reduction of RDS mortality and the increase of non-RDS mortality with Surfaxin seem to balance each other out. Clinical input is needed to determine whether in light of this situation, the reduction of RDS mortality is a relevant clinical benefit for Surfaxin.

The small numerical advantage for Surfaxin compared to Exosurf for all cause mortality, if any, appears to be driven by the European region. It is not clear whether there is any mortality benefit (relative to Exosurf) in the Latin American region given the slightly worse numerical difference is seen.

1.2 Brief Overview of Clinical Studies

Surfaxin (lucinactant) is a novel peptide-containing synthetic surfactant that mimics the essential characteristics of human surfactant protein B (SP-B). The currently approved, commercially available surfactants are synthetic, non-protein-containing (e.g., Exosurf) or protein-containing, animal-derived (e.g., Survanta, Curosurf, Infarsurf). This new drug application contains two large major clinical studies: KL4-IRDS-06 a major phase 3 trial and KL4-IRDS-02 a supportive phase 3 trial. KL4-IRDS-06 was a multinational (Europe and Latin America), multicenter (50 active centers), randomized, masked, active-controlled (Survanta[®]), prophylaxis superiority trial of the safety and effectiveness of Surfaxin[®] compared to Exosurf[®] in the prevention of RDS in premature neonates. This study consists of two phases: a "short-term" efficacy and safety (36 weeks post conceptual age (PCA)) phase, and a "long-term" outcomes (such as death due to any cause) and safety (6 and 12 months corrected age) phase. KL4-IRDS-02 was a masked, multicenter, randomized, controlled trial comparing the safety and showing Surfaxin is noninferior to Curosurf[®] (poractant alfa) for the prevention and treatment of RDS in premature infants.

1.3 Statistical Issues and Findings

In KL4-IRDS-06, statistical issues pertain to

- (1) possible concern with changing endpoint while monitoring the observed treatment effects,

- (2) possible concern with trial conduct change including sample size reestimation plan as a result of monitoring the observed treatment effects,
- (3) although statistically significant findings were obtained on the co-primary efficacy endpoints, does the cause-specific mortality provide relevant clinical benefit?

This reviewer showed that, based on the three possible co-primary efficacy endpoints using the sponsor's analyses, there appeared little concern with the endpoint change, as the change is mostly due to the sponsor's confusion whether the second co-primary endpoint that should be treated as a composite endpoint or as two separate endpoints. In addition, from the observed data paths on the three possible co-primary efficacy endpoints and the data safety monitoring board (DSMB) meeting minutes report on the course of interim analyses, there appeared to be no clear need for an increase of sample size.

Based on the KL4-IRDS-06 adjudicated data, Surfaxin yielded a statistically significant reduction (17%, 95% Confidence limit: 5%, 28%, $p=0.007$ reviewer's analysis based on the Agency's agreed and protocol specified stratified CMH test, $p=0.005$ sponsor's logistic regression stratifying on center and birthweight) in RDS incidence at 24 hours of age, a statistically significant reduction (52%, 95%CI: 24%, 69%, $p=0.001$ CMH test or logistic method) in RDS related mortality by 14 days. However, these co-primary efficacy endpoints weren't consistently shown to be significant based on the investigator's CRF report (RDS incidence: 15.7% in Surfaxin vs. 20.0% in Exosurf, $p=0.041$ and RDS related mortality by 14 days: 0.2% in Surfaxin vs. 0.6% in Exosurf, $p=0.309$). The originally agreed second co-primary efficacy endpoint, RDS related mortality by day 14 or air leak by day 7, was changed to a secondary efficacy endpoint agreed by the Agency partly due to sponsor's confusion as to whether this endpoint is a composite endpoint or two separate endpoints. The original clinical review team (Dr. Debbie Birenbaum, Dr. Robert Meyer, etc.) proposed this endpoint in lieu of all cause mortality due to the impractical sample size needed to detect a realistic improvement in all cause mortality. For this composite endpoint, the effect of Surfaxin was not robust: the risk reduction was 20% with 95%CI: 2% increase, 38% reduction, $p=0.045$ (logistic analysis), $p=0.068$ (CMH test).

All cause mortality by day 14 was very similar between Surfaxin (15.9%) and Exosurf (16.9%), $p=0.588$ (CMH), $p=0.450$ (logistic). This led the reviewer to also analyze the non-RDS related mortality by day 14. This reviewer identified a statistically significant risk increase in non-RDS related mortality by day 14 (11.2% in Surfaxin, 7.3% in Exosurf, risk increase: 53% (95%CI: 4%, 124%), $p=0.022$ (CMH test). It is noted that the rate in Exosurf is similar to that in Survanta (8.1%), the active control arm studied in the same trial. The increased risk in non-RDS related mortality by 14 days was observed in baby boys (11.0% : 9.5%) and larger in baby girls (11.4% : 5.1%), large in light birthweight (600g-800g) subgroup (26.1% : 16.7%) and in intermediate birthweight (801g-1000g) subgroup (9.5% : 4.9%) but little in heavy birthweight (1001g-1250g) subgroup (4.8% : 4.3%), and also large in both the European region (12.7% in Surfaxin, 8.9% in Exosurf, 8.6% in Survanta) and the Latin American region (9.5% in Surfaxin, 5.4% in Exosurf, 7.6% in Survanta). Although Surfaxin showed a significant risk reduction in RDS incidence at 24 hours of age, the question is whether the clinically relevant benefit of Surfaxin can be supported merely by a decrease in the risk of RDS related mortality at 14 days when the data showed essentially the same mortality of all cause in both Surfaxin and Exosurf treated neonates and yet the significant risk increase in non-RDS related mortality by 14 days.

In light of the similar results for Surfaxin and Exosurf on all-cause mortality, the clinical review team will need to weigh the benefit regarding RDS-related mortality reduction versus increased risk of non-RDS mortality by day 14. The discrepancy between the adjudicated results and the investigators' assessment (e.g., the investigators' assessment indicated that all deaths are essentially non-RDS related) is also of concern to this reviewer.

From this reviewer's evaluation, a lower all cause mortality was observed with Surfaxin than with Exosurf, though not statistically significant different (16% with Surfaxin vs. 17% with Exosurf by 14 days, 19% with Surfaxin vs. 21% with Exosurf by 28 days, and 21% with Surfaxin vs. 24% with Exosurf by 36 weeks PCA, respectively). Furthermore, the analysis by region showed no sufficient evidence for a treatment by region interaction on RDS related mortality or non-RDS related mortality by 14 days. However, numerically, Surfaxin had a higher mortality rate than Exosurf in the Latin American region and a lower mortality rate in the European region.

2. INTRODUCTION

2.1 Overview

This new drug application contains a large major phase III study (KL4-IRDS-06) and a large supportive phase III study (KL4-IRDS-02).

KL4-IRDS-02 was a masked, multicenter, randomized, controlled trial comparing the safety and showing Surfaxin is noninferior to Curosurf for the prevention and treatment of respiratory distress syndrome (RDS) in premature infants. The sample size of approximately 496 neonates was planned. The Agency has concern about the sponsor's noninferiority margin -14.5% (= 50% of the point estimate of the Curosurf effect on the primary efficacy endpoint - percentage of neonates alive and not having bronchopulmonary dysplasia (BPD) at day 28 from study randomization) because limited historical data are available and the endpoint is premature neonate related mortality. An 80% preservation level and the use of a worst confidence limit of Curosurf effect were discussed in addition to several other clinical concerns. The Agency clearly indicated to the sponsor at the pre-NDA meeting and several earlier meetings that the trial as planned, if submitted, will be reviewed mainly for safety consideration. Note that if the noninferiority margin is defined to retain either 50% or 80% of the worst 2-sided 95% confidence limit of the Curosurf effect, this margin is approximately -9.6% and -3.8%, respectively. The sponsor terminated the study after 252 neonates were enrolled (124 Surfaxin neonates and 128 Curosurf neonates) following an extended enrollment period and the sponsor's interest to shift resources to the major clinical study KL4-IRDS-06. According to the sponsor, the treatment difference was 4.75% (37.8% with Surfaxin and 33.1% with Curosurf) with a 95% lower confidence limit of -7.27% using the per-protocol neonates (n=243). Using the ITT neonates, this reviewer computed this bound to be -7.4%. It appeared that the worst limit, either -7.27% or -7.4%, calculated from the noninferiority trial, lies between 50% to 80% preservation level when the noninferiority margin is defined using the 95% worst confidence limit of the Curosurf effect. According to the sponsor, there were no statistically significant differences in all cause mortality at 14 and 28 days of age and through 36 weeks PCA.

This review pertains to the major clinical study KL4-IRDS-06.

2.2 Data Sources

The datasets analyzed are in \N21746\N_000\2004_04_13, \N21746\N_000\2004_09_30, \N21746\N_000\2004-12-01, \N21746\N_000\2004-12-08.

3. STATISTICAL EVALUATION

3.1 Evaluation of Efficacy

KL4-IRDS-06 was a multinational (Europe and Latin America), multicenter (50 active centers), randomized, masked, active-controlled (Survanta[®]), prophylaxis superiority trial of the safety and effectiveness of Surfaxin[®] compared to Exosurf[®] in the prevention of RDS in premature neonates. This study consists of two phases: a "short-term" efficacy and safety (36 weeks PCA) phase and a

“long-term” outcome (such as death due to any cause) and safety (6 and 12 months corrected age) phase. The study was initiated on July 02, 2001 and completed on December 16, 2003 for the short-term phase of the trial. The long-term phase is still ongoing. There were three protocol amendments filed. The major changes that are possibly relevant to statistical validity are: 1) the data safety monitoring board (DSMB) could recommend a sample size re-estimation based on review of the observed treatment differences in order to maintain the study power and all data provided to the DSMB were initially to be masked to treatment groups (Amendment #1), 2) changed the primary efficacy endpoint, restated the sample size estimation based on the revised primary endpoint, and the change in primary efficacy endpoint resulted in a sample size re-estimation (Amendment #3).

Eligible neonates were randomized and stratified by birthweight (stratum 1:600-800g, stratum 2: 801-1000g; stratum 3: 1001-1250g) within each center to receive Surfaxin (n=527), Exosurf (n=509) or Survanta (n=258). Dosing determined by the neonate’s birthweight occurred no sooner than 15 minutes after randomization but no later than 30 minutes of age. Surfaxin and Survanta neonates were eligible to receive up to 3 retreatments, while Exosurf neonates were eligible to receive up to 2 retreatments. Retreatment criteria can be found in the Appendix 1. Three neonates in the Surfaxin group and three in the Exosurf group were not treated, but were included in the all randomized neonates for efficacy analyses. All treated neonates (n=1288) were the basis for safety evaluation.

A “completed” patient defined in the protocol was any patient who is entered into the study and who is evaluated at 28 days of age, 36 weeks of PCA, discharge, or death (whichever came later), and at the 6 and 12-month follow-up. In the short term phase of the evaluation without the 6 and 12-month follow-up, only 5 patients in Surfaxin group (0.9%), 4 patients in Exosurf group (0.8%) and no patient in Survanta group (0%) discontinued the study early. The gestational age of the ITT neonates ranged from 22 to 34 weeks of age with the mean of 28 weeks. The treatment groups were comparable with respect to adverse events: 3 (0.6%) with Surfaxin, 3 (0.6%) with Exosurf, and 1 (0.4%) with Survanta. The demographic, baseline characteristics and material history for all 1294 randomized neonates were comparable between Surfaxin, Exosurf and Survanta groups, except the maternal history on membrane rupture (spontaneous vs. artificial: similar spontaneous rates between Surfaxin [43.4%] and Exosurf [43.8%], but, higher in Survanta treated group [51%]) and Tocolytic therapy (similar between Surfaxin [62%] and Exosurf [60.8%], but, lower in Survanta treated group [50%]). Details can be found in the Appendix 2. The number of neonates in the birthweight subgroup, the randomization stratification factor, was comparable among the three treatment groups, about 22% in 600-800g group, 33% in 801-1000g group and 45% in 1001-1250g group, respectively. Baby boys and baby girls were 50:50, about 78% White, 1% Black and 21% other races. This review focuses on evaluation of the coprimary efficacy variables and some secondary efficacy variables.

Co-Primary Efficacy Endpoints

Change of Co-Primary Efficacy Endpoints

The coprimary efficacy endpoints originally agreed were (1) incidence of RDS at 24 hours, and (2) a composite of either RDS related death through 14 days of age or incidence of air leaks

through 7 days of age, both endpoints each demonstrating the surfaxin effect at a 2-sided 5% level. The sponsor proposed four alternatives in an IND 40, 287 serial #190 submission including alpha allocation for the co-primary endpoints of 0.01 for (1) or 0.04 for (2). An internal clinical/statistical meeting was held on Nov. 07, 2003 to discuss the sponsor's proposed alternatives. It was decided at the internal meeting that the sponsor could either stay with the originally agreed upon co-primary variables or one of the two alternatives from those proposed by the sponsor that are supposedly more stringent than the originally proposed co-variables, see FDA fax dated Nov. 14, 2003.

The co-primary efficacy endpoints were changed to exclude the component "incidence of air leaks through 7 days of age" in the composite endpoint in the protocol amendment #3 dated November 10, 2003, viz., the sponsor's alternative proposal #1. And, the original second co-primary endpoint, a composite endpoint, was considered a secondary endpoint both for the 'or' case and the 'and' case, i.e., "RDS related death at day 14 and incidence of air leaks at day 7", and "RDS related death at day 14 or incidence of air leaks at day 7". In addition, a composite of incidence of RDS at 24 hours, RDS related mortality through 14 days of age, and air leak through 7 days of age is a secondary endpoint. These changes were reflected in the DSMB final statistical analysis plan. This review will compare the original co-primary and the modified co-primary endpoints to assess the robustness of the study finding.

Sample Size Re-Estimation

In the original protocol, the treatment period was to last until 420 events of RDS at 24 hours and 162 total air leaks and/or RDS related mortality in the Surfaxin and Exosurf groups occurred. It was estimated that this would require 1,500 randomized patients (600 patients each for Surfaxin and Exosurf and 300 patients for Survanta). In this estimation, the incidences of RDS were assumed to be 30% and 40% for Surfaxin-treated and Exosurf-treated groups, respectively, and the incidence of air leak or RDS related death were assumed to be 10% and 17% for Surfaxin-treated and Exosurf-treated groups, respectively.

In the protocol amendment #1, the DSMB was allowed to recommend a sample size re-estimation if one or more of the observed rates should be substantially lower than the above assumed rates, viz., based on review of the observed treatment differences, in order to maintain the study power. The details of such a sample size re-estimation based on Cui, Hung, Wang (Biometrics 5, 853-857, 1999), and Chen, DeMetz, Lan (Technical report, University of Wisconsin-Madison, Statistics in Medicine 23, 1023-1038, 2004) were outlined in the DSMB Standard Operating Procedure (SOP). The study called for two formal interim analyses plus the final analysis. At each formal interim efficacy analysis, the two-sided z-critical value and its associated significance level used to assess treatment difference will be based on the O'Brien-Fleming boundaries generated by the Lan-DeMets alpha-spending function, where the overall significance level across all interim analyses is set to 0.05 for each endpoint.

Following the change of the co-primary efficacy endpoints, the sponsor re-estimated the sample size assuming the incidences of RDS related death would be 3.5% and 7.5% for Surfaxin-treated and Exosurf-treated groups. Using an event-driven design, 420 total RDS events and 66 deaths due to RDS events were estimated to be required to detect the assumed differences stated above.

The DSMB monitored the total number of events for each of the two co-primary endpoints for the Surfaxin and Exosurf groups and communicated with the sponsor and study steering committee when the appropriate number of events occurred. If the observed number of events for the co-primary endpoints was lower than what was expected to maintain adequate power based on early information, the recruitment period and/or number of sites may have been increased to obtain the pre-specified total number of events for both endpoints. It is worth pointing out that in the protocol Amendment #1, it was indicated that all data provided to the DSMB were initially to be masked to treatment group.

Reviewer's Comments: According to the chair of the DSMB, Dr. DeMets, the two interim analyses were performed at 17% and 37% information time. The observed effect sizes for the originally planned co-primary efficacy endpoints appeared to be either within the expected range or the trend of the observed effect size wasn't clear, which led the DSMB not to recommend a sample size modification at these two interim analyses. Shortly after the co-primary endpoints were modified, the trial was terminated when the planned events were reached, which resulted in a total of 1294 neonates included in the study.

In the NDA submission, the sponsor's primary efficacy analyses were based on the adjudication committee results. At the review team's request of May 24, 2004, the sponsor submitted the adjudication committee standard operating procedure (AC SOP). It is noted that the number of adjudication committee members were changed from 4-6 to 6-10 in the protocol amendment #2. During the protocol review stage, medical reviewer, Dr. Howard Birenbaum expressed the concern that some adjudication committee members may have conflict of interest. It is noted that in the submitted electronic SAS transport file, the non-adjudicated data of the co-primary efficacy endpoints were not included.

Sponsor's Analysis and Reviewer's Analysis

For the co-primary efficacy endpoints, the specified statistical method was Cochran-Mantel Hanzsel (CMH) adjusting for pooled center, birthweight and gender. The study included an interim analysis rule and involved DSMB for possible sample size reestimation based on the observed effect size. Although the sponsor, in their protocol amendment #3, specified a logistic regression model with independent covariates of birthweight, gender, and region (viz., pooled study centers, those centers with number of neonates less than 10 were pooled into one group within their country), at the pre-NDA meeting held in June 2003, the sponsor was informed that the primary efficacy analysis method is the Cochran-Mantel Hanzsel (CMH) test and not the logistic regression method. Covariates specified at the time for adjustment were pooled center, birthweight strata and gender. In the NDA submission, the sponsor specified in the documentation of statistical methods that "the primary analysis will be performed using specified methods depending on the endpoint adjusted for pooled study center and birthweight stratum. Exploratory analysis will be performed using specified methods depending on the endpoint adjusted for pooled study center, birthweight stratum, gender and race." The sponsor reported the primary analysis results from the model without including the covariate gender. However, at the time of covariate change in the final statistical model, two interim analyses had been performed.

To assess the robustness of the sponsor's results, this reviewer performed both the logistic regression analysis and the CMH test for the co-primary efficacy endpoints adjusting for birthweight and/or gender and/or pooled center. It appeared that p-values were very similar between the unadjusted analysis and adjusted analysis [adjusting for birthweight alone, gender alone, pooled center alone, (birthweight and gender), (birthweight and pooled center), and (birthweight, gender and pooled center)] when the evidence is highly statistically significant. This reviewer reported the analysis result of the final statistical model and provided the relative risk estimate and its 95% confidence limits of Surfaxin vs. Exosurf using the CMH method.

The incidence of RDS at 24 hours following birth (data obtained from independent adjudication)

For the incidence of RDS at 24 hours following birth, this reviewer's analysis results were the same as that reported by the sponsor using the logistic regression stratifying on center and birthweight stratum. All the models showed a nominally significant birthweight effect, the center effect, but not the gender effect. There was no interaction between gender and birthweight stratum. For the relative risk estimate, this reviewer reported the CMH stratifying on weight stratum alone, per the randomization. In addition, the relative risk estimate is generally very similar when the model was stratified by weight stratum alone, by both the weight stratum and the gender, and by all three factors, weight, gender and pooled center.

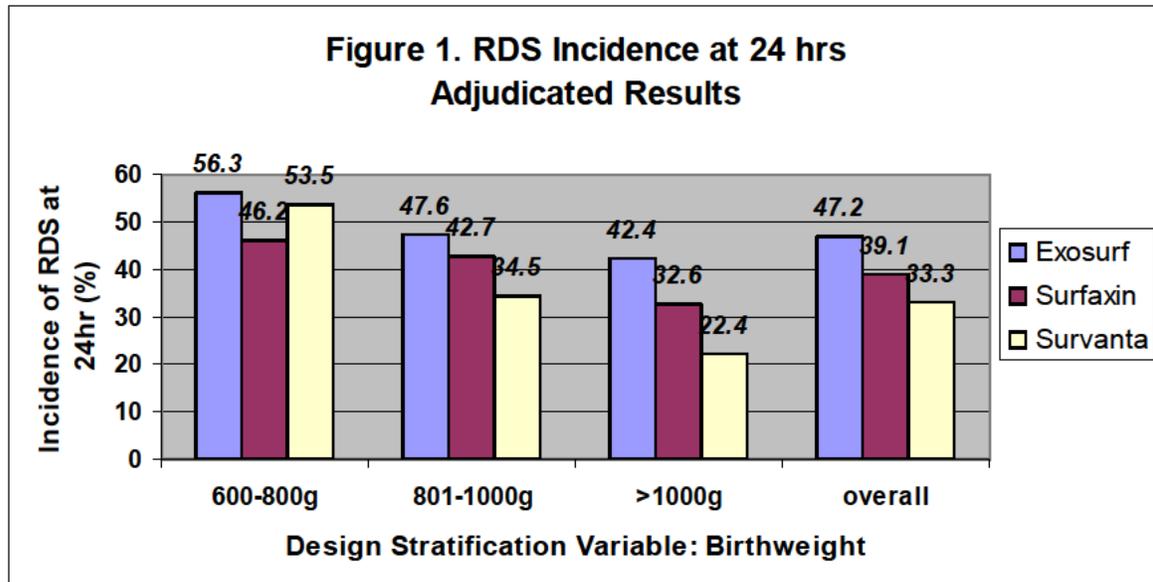
Table 1. Analysis of Incidence of RDS at 24 hrs – 1st co-primary endpoint (the ITT neonates)
[Source: Sponsor's Table 11.3.A. of the Clinical study report]

Incidence of RDS at 24 hours	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)	Relative Risk Surf: Ex (95%CI)*	p-value Surf: Ex
Adjudicated Data	240 (47.2%)	206 (39.1%)	86 (33.3%)	0.83 (0.72, 0.95)	0.005** 0.007*
CRF Data**	102 (20.0%)	83 (15.7%)	31 (12.0%)	NA	0.041**

* reviewer's analysis based on CMH test stratified by birthweight.

** sponsor's report Table 11.3.A, 11.3.B, Volume 5, 1.1, p.51 (logistic model stratifying on center, birthweight).

From Reviewer Table 1, both the sponsor and the reviewer's analysis showed that Surfaxin treated patients have on average a statistically significant lower RDS incidence at 24 hours as compared to Exosurf treated neonates, 39.1% vs. 47.2% ($p \leq 0.007$). The risk reduction for Surfaxin relative to Exosurf in the occurrence of RDS event at 24 hours following birth was 17% with 95% CI: 5%, 28%. Survanta, the reference arm, had the smallest incidence of RDS at 24 hours, 33.3%. Nominally, there was no statistically significant difference between Surfaxin (39.1%) and Survanta (33.3%). Upon a closer look within each of the three birthweight categories, the lower RDS incidence of Surfaxin relative to Exosurf was primarily seen in the 600-800g category (46% vs. 56%) and the >1000g category (33% vs. 42%), and some in the 801g to 1000g group (43% vs. 48%). The consistently lower incidence of RDS at 24 hours in Survanta, the reference arm, over both the experimental treatment Surfaxin and the comparator Exosurf was apparent in the 800g-1000g (only 35%) and >1000g (only 22%) groups, see Reviewer's Figure 1.



- ***Discrepancy Between The Adjudicated Results And The Investigators' Report***

The sponsor also reported the incidence of RDS at 24 hours based on the case report form (CRF) results. This data is not available in the electronic SAS transport file. As shown in Table 1, the trend that the incidence of RDS at 24 hours was highest with Exosurf (20%), followed by Surfaxin (15.7%) and with Survanta (12.0%) being the smallest was maintained, however, the incidence obtained from CRF was consistently much lower than the independent adjudication results. According to the sponsor, the adjudicated committee was given an adjudication data package (ADP), which contains x-ray report, CRF, serious adverse event (SAE) report and autopsy report, but, the investigators may not have all the information. In addition, the adjudication committee was given the definition of RDS incidence at 24 hours, see p.5-6 of AC SOP, Appendix 3. Using the CRF data, the sponsor reported that the comparison between Surfaxin vs. Exosurf was nominally statistically significant with p-value=0.041.

RDS related death through 14 days of age (data obtained from independent adjudication)

For the RDS related death through 14 days of age, the sponsor's results were similar to those analyses performed by this reviewer. From Table 2 of this reviewer's analysis, Surfaxin was shown to have statistically significantly lower RDS related death through 14 days of age compared to Exosurf (4.7% vs. 9.6%, p=0.001). The risk reduction for Surfaxin relative to Exosurf in the occurrence of RDS related death through 14 days of age was 52% with 95% CI: 24%, 69%. In a sponsor's submission during the NDA review clock, one death (patient 092007) in Exosurf found alive after further follow-up. This neonate was a male in the birthweight 801g-1000g group recruited from Chile treated with Exosurf. There was little impact in the result due to this RDS related mortality misclassification, 4.7% vs. 9.4%, p=0.002. The remaining of the review is based on the originally submitted data; i.e., treating this neonate as RDS related death by 14 days.

- ***Discrepancy Between The Adjudicated Results And The Investigators' Report***

The RDS cause-specific death by 14 days based on the CRF data was very few: 0.2% with Surfaxin (1 neonate), 0.6% with Exosurf (3 neonates) and 0% with Survanta (0 neonate). As shown in Table 2, this data did not show a statistically significant decrease in RDS related mortality in Surfaxin treated neonates as compared to Exosurf treated neonates, nominal p-value = 0.309. The large discrepancy of the RDS cause-specific event counts between the adjudicators' versus the investigators' assessment is very troublesome. The criteria for determination of RDS related death stated in the AC SOP can be found in Appendix 4 (p.6-7 of AC SOP). Thus, interpretation of the data heavily relies on adjudication accuracy; so, considering the RDS related mortality accessed using the CRF data, essentially all mortality are non-RDS related. Adjudicated data identified a lot more RDS related death as compared to the CRF data, it is questioned that either the investigators under reported the RDS related death or the adjudicators over reported the RDS related death.

Table 2. Analysis of RDS Related Death by 14 Days – 2nd coprimary endpoint (the ITT patients) [Source: the sponsor's Table 11.3.A. and 11.3.B. of the Clinical study report]

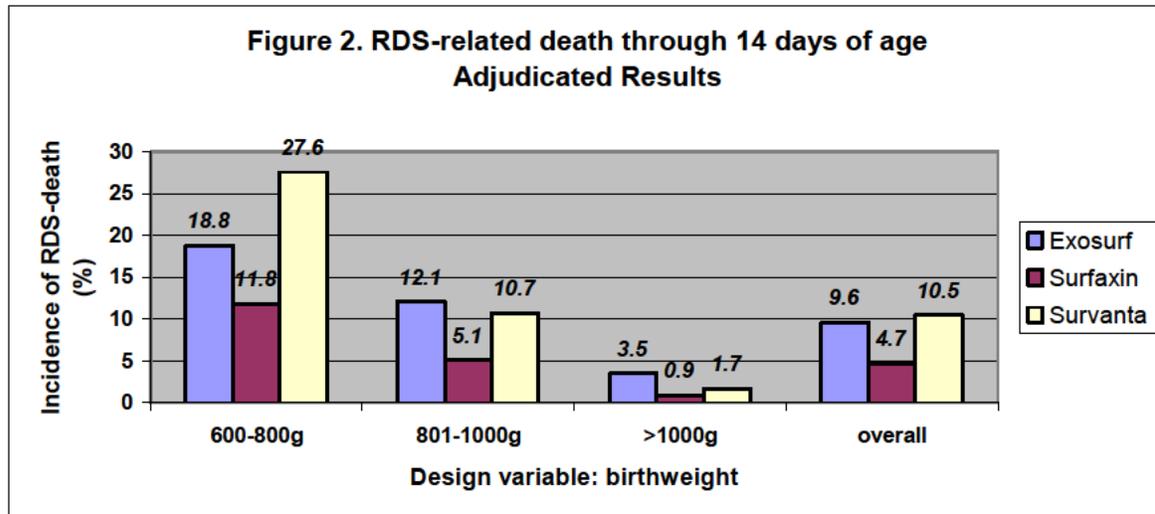
RDS related mortality through 14 days of age	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)	Relative Risk Surf. Ex (95%CI)*	p-value Surf. Ex
Adjudicated Data	49# (9.6%)	25 (4.7%)	27 (10.5%)	0.48 (0.31, 0.76)	0.001**
CRF Data**	3 (0.6%)	1 (0.2%)	0 (0.0%)	NA	0.309**

* reviewer's analysis based on CMH test stratified by birthweight

** sponsor's report Table 11.3.A. and 11.3.B, Volume 5, 1.1, p.51 (logistic stratifying on center, birthweight).

In a sponsor's submission during the NDA review clock, one death (patient 092007) in Exosurf found alive after further follow-up resulted in RDS related mortality 9.4% (48/509), p-value is 0.002.

The birthweight and gender showed significant impact on the differences observed in the RDS related death through 14 days of age using the independent adjudication data. This reviewer presented the Surfaxin effect on RDS related death through 14 days of age by the birthweight, see Figure 2 and by the gender, see Section 4.1. It is clear from Figure 2 that the RDS related mortality decreases as the neonates' birthweight increases for the three treatments Exosurf, Surfaxin, and Survanta studied. However, the RDS cause specific death through 14 days of age with Surfaxin was consistently the lowest for all birthweight categories among the three treatment groups compared.



Reviewer's Analysis and Concerns:

Based on the statistical significance criteria (that both the incidence of RDS at 24 hours of age and the RDS related mortality through 14 days of age must demonstrate statistical significance at 2-sided 5% level) for the co-primary endpoints defined in the final statistical analysis plan, Surfaxin is shown to be superior to Exosurf. This reviewer further performed an analysis on the non-RDS related death after observing the similar result on all cause mortality by day 14 (see Table 3 and Table 4 below).

All cause mortality through 14 days of age

Using data assessed by the independent adjudication committee on death related efficacy endpoint through 14 days of age, this reviewer evaluated all cause mortality and non-RDS related mortality. These results are summarized in Table 3. For all cause mortality endpoint there was little difference between Surfaxin (15.9%) and Exosurf (16.9%); nominal p-value is 0.450 using logistic regression reported by the sponsor and 0.588 using the CMH test performed by this reviewer. The relative risk of Surfaxin versus Exosurf is 0.93 with 95% confidence interval (0.72, 1.21). The all cause mortality with Survanta was 18.6%, numerically higher than both Surfaxin and Exosurf.

Non-RDS related mortality through 14 days of age

Surfaxin appeared to be associated with a higher non-RDS related mortality than Exosurf (11.2% vs. 7.3%, nominal p-value = 0.029 reviewer's analysis), as shown in Table 3. The non-RDS related death rate (11.2%) more than doubled the RDS related death (4.7%) with Surfaxin, but the non-RDS related mortality (7.3%) was smaller than the RDS related mortality (9.6%) with Exosurf.

Table 3. Death related (all-cause or RDS cause-specific) efficacy endpoints (ITT neonates)
 [Source: the sponsor's Table 11.4.1.2.2.A of the Clinical study report]

Data from independent adjudication committee	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)	RR Surf: Exo (95%CI)#	p-value Surf: Ex
RDS related mortality through 14 days of age	49 (9.6%) 48 (9.4%)#	25 (4.7%)	27 (10.5%)	0.48 (0.31, 0.76)	0.001** 0.001*
All cause mortality through 14 days of age	86 (16.9%)	84 (15.9%)	48 (18.6%)	0.93 (0.72, 1.21)	0.450** 0.588*
Non-RDS related mortality by 14 days	37 (7.3%)	59 (11.2%)	21 (8.1%)	1.52 (1.04, 2.23)	Na 0.029*

* reviewer's analysis based on CMH test stratified by birthweight

** sponsor's report using data from independent adjudication committee

In a sponsor's submission during the NDA review clock, one death (patient 092007) in Exosurf found alive after further follow-up resulted in a p-value of 0.002.

Reviewer's Concern on the benefit of RDS related mortality through 14 days

This reviewer is concerned with Surfaxin benefit of reducing RDS mortality. First, with Surfaxin, although there was a significant reduction in RDS related mortality by 14 days [52% with 95% CI: 24%, 69%], in contrast, **there was also a significant increase in non-RDS related mortality by 14 days [52% increase with 95% CI: 4%, 123%]**. Secondly, the non-RDS related death rate (11.2%) more than doubled the RDS related death (4.7%) with Surfaxin, but the non-RDS related mortality (7.3%) was smaller than the RDS related mortality (9.6%) with Exosurf. Thirdly, the discrepancy in reporting/assessment between CRF and the adjudication committee's report clearly needs to be addressed.

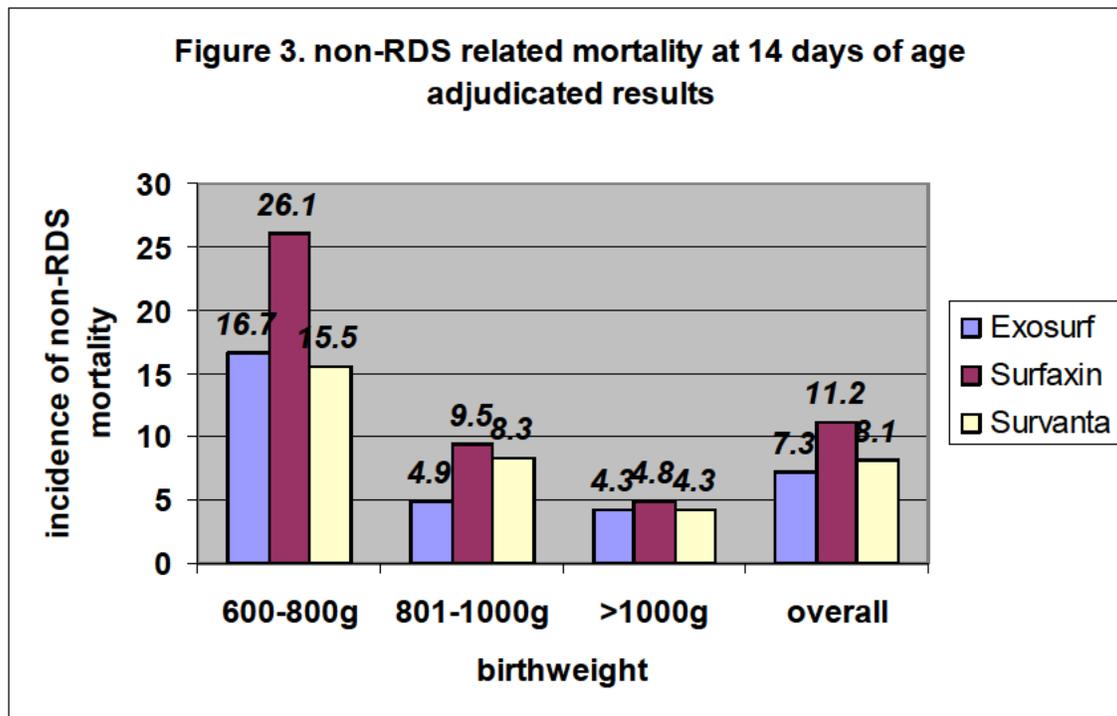
- Analysis of Non-RDS Related Mortality by gender, by birthweight

This reviewer performed additional analyses on non-RDS related mortality by day 14 in addition to the above adjusted analysis. Unadjusted analyses by gender and by birthweight are given in Reviewer Table 4. It appeared that both males and females treated with Surfaxin have a higher non-RDS related mortality than those treated with Exosurf; Surfaxin appears to show much worse risk of non-RDS death in female neonates.

Table 4. Analysis of non-RDS related mortality through 14 days of age

Non-RDS death	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)	Surfaxin:Exosurf	
				RR (95% CI)	Nominal-p
ITT	37 (7.3%)	59(11.2%)	21 (8.1%)	1.53(1.04,2.24)	0.022
Male	9.5%(24/254)	11.0%(29/263)	9.3%(12/129)	1.17(0.70,1.95)	0.555
Female	5.1%(13/255)	11.4%(30/264)	7.0%(9/129)	2.23(1.19,4.18)	0.010
600-800g	16.7%(19/114)	26.1%(31/119)	15.5%(9/58)	1.56(0.94,2.60)	0.082
801-1000g	4.9%(8/163)	9.5%(17/179)	8.3%(7/84)	1.94(0.86,4.36)	0.104
1001-1250g	4.3%(10/232)	4.8%(11/229)	4.3%(5/116)	1.11(0.48,2.57)	0.800

Though Surfaxin had a lower RDS related death (7% lower in 600-800g group, 7% lower in 801-1000g group, and 2.6% lower in 1001-1250g group), it had a higher non-RDS related death (9.4% higher in 600-800g group, 4.6% higher in 801-1000g group, and 0.5% higher in 1001-1250g group) as compared to Exosurf. Except the birthweight 801g-1000g subgroup, Survanta appeared to have similar non-RDS related mortality rate as Exosurf. Here, the data suggest a RDS related mortality reduction with Surfaxin and the data also suggest a non-RDS related mortality increase with Surfaxin. The controversy is the result of numerically very similar all cause mortality between the two treatments yielding essentially no hard evidence of mortality benefit in the premature neonates by 14 days of age. In contrast with Figure 2 for RDS related mortality, Figure 3 seems to suggest that the risk reduction on non-RDS related mortality by 14 days in each birthweight subgroup was balanced out by the increase in non-RDS mortality risk.



- Non-RDS Related Mortality in Prevention Trials for Exosurf Approval

The data available for Exosurf approval in early 1990 did not have RDS related death by day 14 but had all cause mortality by day 28 and RDS related death. This reviewer summarized the data of the three premature infants RDS prevention trials by extracting the information from the statistical review and evaluation report of Dr. James Gebert. Neonates in these trials were mostly Caucasian with birthweight 700g to 1100g. As shown in Table 5, except prevention trial #2 which failed to demonstrate Exosurf effect on both the all cause mortality and the RDS related (and non-RDS related) mortality by 28 days, trial #1 and trial #3 showed consistent mortality benefit with Exosurf as compared to Air treatment. Prevention trial #1 showed a reduction in both the all cause mortality and the RDS related mortality, and no effect on non-RDS related mortality. Prevention trial #3 showed a lower rate in all cause mortality, RDS related mortality, and non-RDS related

mortality with Exosurf. There was no contradictory result on the mortality in the evaluation of Exosurf.

Table 5. Analysis of mortality and RDS related mortality in the approval of Exosurf#

1990	All Cause Mortality by day 28		RDS related Mortality		Non-RDS related Death	
	Air	Exosurf	Air	Exosurf	Air	Exosurf
Prev1	21% (47/222)	15% (34/224)	10.4% (23/222)	5.4% (12/224)	11% (24/222)	10% (22/224)
Prev2	10% (8/80)	10% (8/81)	5.0% (4/80)	6.2% (5/81)	5% (4/80)	4% (3/81)
Prev3	15% (16/110)	7% (8/108)	2.7% (3/109)	0.9% (1/109)	12% (13/110)	6% (7/108)

extracted from Exosurf Statistical Review of Dr. Jim Gebert, 1990.

- Review Team's Assessment on possible explanation of excess of non-RDS with Surfaxin

The medical review team was interested in investigating whether excess of non-RDS related mortality with Surfaxin can be explained by a longer time to death causing events that are indirectly related to RDS as compared to Exosurf. First, Dr. Gunkel assigned the reason of death based on available documents including data listings, case report forms and patient narratives in those deaths whose data seemed unusual or discrepant. The reasons include (1) air leak, (2) Intraventricular haemorrhage (IVH) neonatal, (3) Pulmonary haemorrhage (PE), (4) necrotizing enterocolitis (NEC), (5) renal failure neonatal (RF), (6) Sepsis neonatal, and (7) others. Of these reasons, Dr. Gunkel considered that air leak, IVH and PE are likely the causes of death that are directly related to RDS and the remaining reasons (4-7) are likely indirectly related to RDS. With the medical team's request, Dr. Wang performed a time-to-event analysis stratifying on reasons of death. In this exploratory analysis, the median time to death was shown to differ between Surfaxin and Exosurf across the seven death reason strata. The median times to death among those neonates died of events (air leak, IVH, PH) that might be directly related to RDS were similar or shorter with Surfaxin than with Exosurf and were numerically shorter with Surfaxin than with Exosurf in those events (NEC, RF, Sepsis) that might be indirectly related to RDS, see Table 6. When the directly RDS related reasons are combined, median times to death were 4.5 days (95%CI: 3, 7) with Surfaxin and 5 days (95%CI: 3, 6) with Exosurf, respectively and were 10 days (95%CI: 6, 14) with Surfaxin and 12 days (95%CI: 9, 15) with Exosurf when the indirectly RDS related reasons were combined. It is noted that these median times were calculated for each reason of death alone. The comparison of these median times between Surfaxin and Exosurf in each death reason category is often very difficult to interpret because the comparisons were not based on the ITT neonates, the competing risks among these reasons were not accounted for and number of events is small in most of the strata.

Table 6. Median time (95% CI) to death (in days) due to each specific reason alone*

	Air leak	IVH	PH	NEC	RF	Sepsis	Others
Exosurf time(d) n (%)	4 (2, 12) 11(2.2%)	6 (3, 8) 34(6.7%)	5 (3, 10) 9 (1.8%)	20 (13,47) 6 (1.2%)	21 (10, 56) 5 (1.0%)	12 (8, 14) 32(6.3%)	4.5 (2, 17) 24 (4.7%)
Surfaxin time(d) n (%)	4 (2, 10) 7 (1.3%)	6.5 (4, 9) 22(4.2%)	3 (3, 4) 13(2.5%)	10 (8, 34) 7 (1.3%)	13.5 (10, 22) 8 (1.5%)	8 (5, 13) 38(7.2%)	6.5 (3, 24) 16 (3.0%)

* Dr. Gunkel's assessment based on the data listings, case report forms, and patient narratives.

Secondary Efficacy Endpoints

Reviewer's Comments: RDS related mortality through 14 days of age and/or air leak through 7 days of age, one of the original co-primary efficacy endpoint becomes a secondary endpoint agreed upon by the Agency as stated in the final statistical analysis plan. One might raise the concern with 'changed the component of the primary endpoint while the DSMB was monitoring observed treatment effect during the trial.' This reviewer addressed this issue below.

RDS related mortality through 14 days of age or air leak through 7 days of age

This composite secondary endpoint was originally planned as one of the co-primary variables. Based on the sponsor's analysis, a statistically significant lower RDS related death by 14 days or air leak through 7 days of age with Surfaxin than with Exosurf (17.5% vs. 21.6%) was observed, $p=0.045$ based on sponsor's logistic regression analysis stratifying on center and birthweight, see Table 7. Thus, using the criteria that both co-primary endpoints using the adjudicated data must show statistical significance at a two-sided 5% level, Surfaxin effect (relative to Exosurf) is demonstrated by the sponsor in this major clinical study using the logistic regression model stratifying on center and birthweight (Incidence of RDS at 24 hours, $p=0.005$, RDS related mortality by 14 days, $p=0.001$, a composite of RDS related mortality by 14 days and/or air leak by 7 days, $p=0.045$).

Table 7. RDS related mortality through 14 days of age or air leak through 7 days of age
[Source: sponsor's Table 11.4.1.2.1.A. of clinical Study Report]

RDS death by 14 days and/or air leak by 7 days	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)	Relative Risk Surf: Ex (95%CI)*	p-value Surf: Ex
or	110 (21.6%)	92 (17.5%)	49 (19.0%)	0.80 (0.62, 1.02)	0.045** 0.068*
and	32 (6.3%)	15 (2.8%)	20 (7.8%)	0.44 (0.24, 0.80)	0.006** 0.006*

* reviewer's analysis based on CMH test stratified by birthweight
(RDS death or air leak, nominal $p=0.16$, RDS death and air leak, nominal $p=0.016$ when stratifying on wt, sex, center)

** sponsor's report Table 11.4.1.2.1.A, p.52 (logistic stratifying on center, birthweight).

The evidence of a significant Surfaxin effect on “RDS death by 14 days or air leak by 7 days” might not be a robust finding, as the simple comparison between Surfaxin and Exosurf gave a $p=0.092$, the CMH test adjusting for gender gave a $p=0.092$, adjusting for birthweight gave a $p=0.068$, adjusting for birthweight and gender gave a $p=0.065$, and adjusting for birthweight, gender, center gave a $p=0.160$. In contrast, these nominal p -values are ≤ 0.016 for ‘RDS related mortality by 14 days and air leak by 7 days.’

Survival at day-28 without broncho-pulmonary dysplasia

Table 8. Survival at day-28 without broncho-pulmonary dysplasia (BPD)

[Source: sponsor’s Table 11.4.1.2.43.A. of clinical Study Report]

	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)	Relative Risk Surf:Ex (95%CI)*	p-value Surf:Ex
Alive w/o BPD at day28	190 (37.3%)	221 (41.9%)	106 (41.1%)	0.92 (0.84, 1.01)	0.044** 0.065*
Alive w/o BPD at 36wks PCA	274 (53.8%)	313 (59.4%)	144 (55.8%)	0.87 (0.76, 0.99)	0.022** 0.033*

* reviewer’s analysis based on CMH test stratified by birthweight

(alive w/o BPD at day28, nominal $p=0.152$, at 36wks PCA, nominal $p=0.116$ stratifying on wt, sex, center)

** sponsor’s report Table 11.4.1.2.43.A, p.54 (logistic stratifying on center, birthweight).

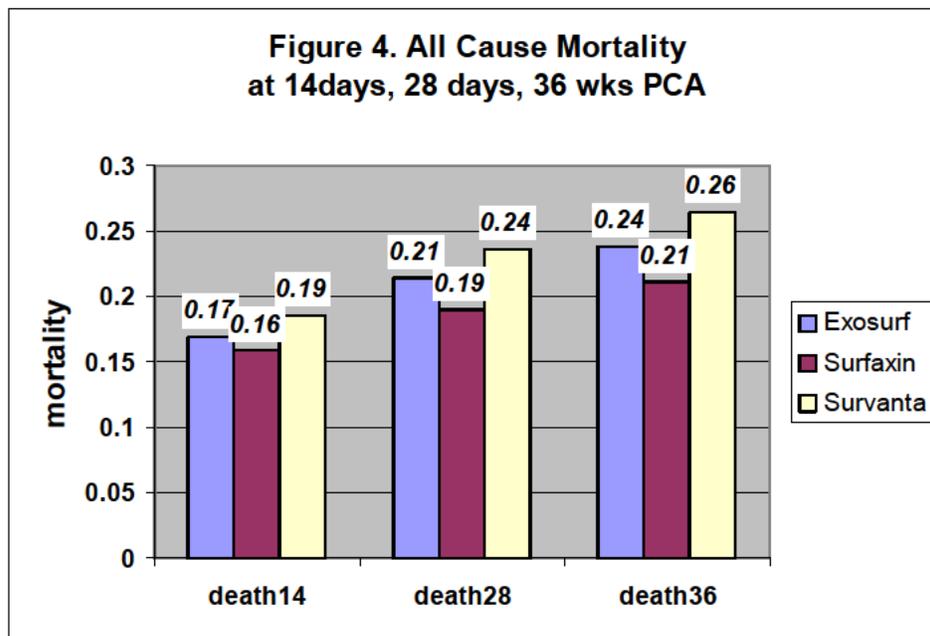
This reviewer also analyzed the secondary endpoint of “being alive and without BPD at day 28” and “being alive and without BPD at 36 weeks PCA.” The former is an efficacy endpoint used as the primary endpoint for the approval for Exosurf and the approval for Survanta. As shown in Table 8 the evidence of a significant Surfaxin effect on “being alive and without BPD at day 28” might not be robust. Although, the sponsor reported a p -value=0.044 using the logistic regression model adjusting for center and birthweight, a nominally significant p -value cannot be found using the CMH test with or without adjusting the covariates, the smallest p -value is 0.063 (adjusting for birthweight and gender), the largest p -value is 0.152 (adjusting for birthweight, gender, and center), and the simple comparison of 41.9% vs. 37.3% gave a nominal $p=0.130$. It is noted that a logistic regression analysis adjusting for center when there are 42 pooled centers, the reported p -value might not be appropriate as it is often the result of a nonconvergence of the maximum likelihood estimate. With the CMH test, the statistical test does not have a convergence issue.

A borderline evidence might be observed with the secondary efficacy endpoint of “being alive and without BPD at 36 weeks PCA.” The simple comparison between 59.4% vs. 53.8% gave a nominal $p=0.071$, a CMH adjusted analysis gave a $p=0.033$ adjusting for weight only, a $p=0.031$ adjusting for weight and gender.

Briefly, these secondary efficacy endpoints showed a numerically lower event rate and a higher response rate with Surfaxin compared to Exosurf supporting the co-primary efficacy endpoint. However, the increased risk in the non-RDS related mortality by 14 days with Surfaxin relative to Exosurf is still of concern, particularly, the unadjusted analysis showed a similar event rate between Survanta and Exosurf, leaving a higher non-RDS related mortality with Surfaxin.

- All cause mortality by 14 day, by 28 days, and by 36 weeks of PCA

This reviewer further investigated all cause mortality. As depicted in Figure 4, all cause mortality was lower with Surfaxin and gradually departed from Exosurf at day 14 (16% vs. 17%, nominal p-value=0.588), at day 28 (19% vs. 21%, nominal p-value=0.252), and at 36 weeks PCA (21% vs. 24%, nominal p-value=0.217). Thus, if this trend of all cause mortality shows a departing trend by the end of the short-term study death by 36 weeks PCA, the concern of the increased risk with non-RDS related mortality at 14 days might be lessen. It is also noted that although numerically Surfaxin had a lower mortality by 36 weeks PCA as compared to Exosurf, the median time to death was not statistically significantly different, nominal p-value 0.302, log-rank test. Apparently, all cause mortality appeared to be the highest with Survanta, the standard of care in US, in this mixed Caucasian and Latin American premature neonates.



- Possible Concerns with Trial Conduct Change due to unblended sample size re-estimation plan

Reviewer's Comments: A role of DSMB is to make a recommendation on modification of the protocol or termination of the study, unblinding may be necessary to appropriately evaluate safety and efficacy including the unblinded sample size re-estimation. The interaction between the sponsor and the DSMB for necessary data is through an independent statistical analysis center. The DSMB, although 'unblinded' to the treatment code in terms of treatment A vs. treatment B, did not appear to involve with the sponsor's request for change of primary endpoint. The sponsor's request was apparently because of the possible confusion on the agency's intended composite endpoint.

The DSMB, while monitoring the data and performing two interim analyses, did not recommend an increase in sample size. According to the entire DSMB meeting minutes report at the interim

analyses, the observed effect sizes on the original co-primary endpoints reach a targeted minimum effect size (assuming an 80% power level) or the sample path did not show a clearly larger or smaller effect size. This reviewer concurred with the DSMB that although the unblinded sample size re-estimation is planned, the need of sample size increase was not apparent. In fact, the study was terminated with less than 600 neonates per treatment arm, as the targeted total number of events was reached.

3.2 Evaluation of Safety

Please read Dr. J. Harry Gunkel's review for safety assessment.

4. FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender

The lower incidence of RDS at 24 hours with Surfaxin than with Exosurf was observed in the baby boys (39.5% vs. 47.2%) and the baby girls (38.6% vs. 47.1%) subgroups, see Table 9. In addition, a lower RDS related mortality by day 14 with Surfaxin than with Exosurf was consistently observed in baby boys (6.1% vs. 11%) and in baby girls (3.4% vs. 8.2%). In general, baby boys had higher RDS related mortality than baby girls. For all cause mortality by day 14, although baby boys showed a lower rate with Surfaxin (17.1%) than with Exosurf (20.5%), baby girls appeared to have a little higher rate with Surfaxin than with Exosurf (14.8% vs. 13.3%). For non-RDS related mortality by day 14, see Table 4.

Table 9. The co-primary endpoints and all cause mortality at 14 days of age by gender*

Incidence of RDS at 24 hours	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)
RDS Incidence – 24hr			
Baby Boy	47.2%	39.5%	33.3%
Baby Girl	47.1%	38.6%	33.3%
RDS related mortality – 14d			
Baby Boy	11%	6.1%	10.9%
Baby Girl	8.2%	3.4%	10.1%
All cause mortality – 14d			
Baby Boy	20.5%	17.1%	20.2%
Baby Girl	13.3%	14.8%	17.1%

* Reviewer's assessment based on the ITT neonates.

4.2 Other Special/Subgroup Populations

Since the major clinical study was stratified by birthweight, special subgroup by birthweight is summarized. For RDS incidence at 24 hours, see Figure 1, for RDS related mortality through 14 days of age, see Figure 2, and for non-RDS related mortality by day 14, see Table 4.

Table 10. The all cause mortality at 14 days of age by birthweight*

	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)
600-800g	35.1%	37.8%	43.1%
801-1000g	17.2%	14.5%	19.1%
1001-1250g	7.8%	5.7%	6.0%

* Reviewer's assessment based on the ITT neonates.

For all cause mortality by day 14, the trend of lower Surfaxin rate than Exosurf was seen in 801g-1000g and 1001-1250g subgroups, but, numerically, Surfaxin rate was higher than Exosurf (37.8% vs. 35.1%) in 600-800g subgroup.

4.3 Regional Analysis

Premature neonates were recruited in 10 countries. A possible explanation of the observed significantly higher non-RDS related mortality by 14 days of life might be that administration of Surfaxin and Exosurf differs among countries. This reviewer performed an analysis on the co-primary endpoints, RDS related, non-RDS related and all cause mortality by day 14 by region. Here, European countries (Russia, Hungary, Poland) are grouped into 'European' region whereas Latin American countries (Ecuador, Chile, Uruguay, Panama, Mexico, Argentina, Brazil) are grouped into 'Latin American' region. In Exosurf treated neonates, 53% were from Europe and 47% were from Latin America. In Surfaxin treated neonates, 54% were from Europe and 46% were from Latin America.

Table 11. The co-primary endpoints and all cause mortality at 14 days of age by region*

	Exosurf (n=509)	Surfaxin (n=527)	Survanta (n=258)
RDS Incidence – 24hr			
European (n=553)	34.9% (n=269)	31.7% (n=284)	30.0%
Latin American (n=483)	60.8% (n=240)	47.7% (n=243)	37.3%
RDS mortality – 14d			
European	11.5%	4.2%	10.0%
Latin American	7.5%	5.4%	11.0%
All cause mortality – 14d			
European	20.5%	16.9%	18.6%
Latin American	12.9%	14.8%	18.6%
Non-RDS related mortality			
European	8.9%	12.7%	8.6%
Latin American	5.4%	9.5%	7.6%
All cause mortality – 28d			
European	26.4%	20.1%	24.3%
Latin American	15.8%	17.7%	22.9%
All cause mortality 36wk			
PCA	29.0%	22.5%	27.1%
European	17.9%	19.3%	25.4%
Latin American			

* Reviewer's assessment based on the ITT neonates.

As summarized in Table 11, a lower incidence of RDS at 24 hours with Surfaxin than with Exosurf was observed in the European and Latin American regions. No region (Europe vs. Latin America) interaction was observed in RDS related mortality and non-RDS related mortality by day 14. Both European and Latin American regions showed a decreased risk in RDS related mortality and an increased risk in non-RDS related mortality and this rate in Surfanta was similar to the European region (8.6% in Surfanta and 8.9% in Exosurf) but higher in Latin American region (7.6% in Surfanta and 5.4% in Exosurf). For all cause mortality, there appeared to be region interaction. The data showed a lower all cause mortality in Latin American region than in European region receiving either the Surfaxin or the Exosurf. In addition, more deaths occurred with Surfaxin than with Exosurf in the Latin American region. In the European region, Surfanta showed a similar all cause mortality as Exosurf, but, Surfaxin had a lower mortality than Surfanta (16.9% vs. 18.6% by day 14, 20.1% vs. 24.3% by day 28, and 22.5% vs. 27.1% by 36 weeks PCA. In the Latin American region, Exosurf had the lowest all cause mortality followed by Surfaxin, and Surfanta had the highest all cause mortality. In terms of all cause mortality risk reduction, there was no statistically significant difference between Surfaxin and Exosurf. It appeared that Surfaxin benefit, if any, was primarily shown in the European treated neonates, but, not in the Latin American treated neonates. It is also noted that the estimated 25th percentile time to death for all causes was 66 days in Surfaxin and was 24 days in Exosurf in the European region (nominal p-value=0.084, log-rank test). These times were not reached in the Latin American region (nominal p-value=0.685, log-rank test).

5. SUMMARY AND CONCLUSIONS

5.1 Statistical Issues and Collective Evidence

In KL4-IRDS-06, statistical issues pertain to

- (1) possible concern with changing endpoint while monitoring the observed treatment effects,
- (2) possible concern with trial conduct change including sample size reestimation plan as a result of monitoring the observed treatment effects,
- (3) although statistically significant findings were obtained on the co-primary efficacy endpoints, does the cause-specific mortality provide relevant clinical benefit?

This reviewer showed that based on the three possible co-primary efficacy endpoints using the sponsor's analyses, there appeared little concern with the endpoint change, as the change is mostly due to the sponsor's confusion whether the second co-primary endpoint that should be treated as a composite endpoint or as two separate endpoints. In addition, from the observed data paths on the three possible co-primary efficacy endpoints and the DSMB meeting minutes report on the course of interim analyses, there appeared to be no clear need for an increase of sample size.

Based on the KL4-IRDS-06 adjudicated data, Surfaxin yielded a statistically significant reduction (17%, 95% CI: 5%, 28%, p=0.007 reviewer's analysis based on the Agency's agreed and protocol specified stratified CMH test, p=0.005 sponsor's logistic regression stratifying on center and birthweight) in RDS incidence at 24 hours of age, a statistically significant reduction (52%, 95%CI: 24%, 69%, p=0.001 CMH test or logistic method) in RDS related mortality by 14 days.

However, these co-primary efficacy endpoints weren't consistently shown to be significant based on the investigator's CRF report (RDS incidence: 15.7% in Surfaxin vs. 20.0% in Exosurf, $p=0.041$ and RDS related mortality by 14 days: 0.2% in Surfaxin vs. 0.6% in Exosurf, $p=0.309$). The originally agreed second co-primary efficacy endpoint, RDS related mortality by day 14 or air leak by day 7, was changed to a secondary efficacy endpoint agreed by the Agency partly due to sponsor's confusion as to whether this endpoint is a composite endpoint or two separate endpoints. The original clinical review team (Dr. Debbie Birenbaum, Dr. Robert Meyer, etc.) proposed this endpoint in lieu of all cause mortality due to the impractical sample size needed to detect a realistic improvement in all cause mortality. For this composite endpoint, the effect of Surfaxin was not robust: the risk reduction was 20% with 95%CI: 2% increase, 38% reduction, $p=0.045$ (logistic analysis), $p=0.068$ (CMH test).

All cause mortality by day 14 was very similar between Surfaxin (15.9%) and Exosurf (16.9%), $p=0.588$ (CMH), $p=0.450$ (logistic). This led the reviewer to also analyze the non-RDS related mortality by day 14. This reviewer identified a statistically significant risk increase in non-RDS related mortality by day 14 (11.2% in Surfaxin, 7.3% in Exosurf, risk increase: 53% (95%CI: 4%, 124%), $p=0.022$ (CMH test). It is noted that the rate in Exosurf is similar to that in Survanta (8.1%), the active control arm studied in the same trial. The increased risk in non-RDS related mortality by 14 days was observed in baby boys (11.0% : 9.5%) and larger in baby girls (11.4% : 5.1%), large in light birthweight (600g-800g) subgroup (26.1% : 16.7%) and in intermediate birthweight (801g-1000g) subgroup (9.5% : 4.9%) but little in heavy birthweight (1001g-1250g) subgroup (4.8% : 4.3%), and also large in both the European region (12.7% in Surfaxin, 8.9% in Exosurf, 8.6% in Survanta) and the Latin American region (9.5% in Surfaxin, 5.4% in Exosurf, 7.6% in Survanta). Although Surfaxin showed a significant risk reduction in RDS incidence at 24 hours of age, the question is whether the clinically relevant benefit of Surfaxin can be supported merely by a decrease in the risk of RDS related mortality at 14 days when the data showed essentially the same mortality of all cause in both Surfaxin and Exosurf treated neonates and yet the significant risk increase in non-RDS related mortality by 14 days.

In light of the similar results for Surfaxin and Exosurf on all-cause mortality, the clinical review team will need to weigh the benefit regarding RDS-related mortality reduction versus increased risk of non-RDS mortality by day 14. The discrepancy between the adjudicated results and the investigators' assessment (e.g., the investigators' assessment indicated that all deaths are essentially non-RDS related) is also of concern to this reviewer.

From this reviewer's evaluation, a lower all cause mortality was observed with Surfaxin than with Exosurf, though not statistically significant different (16% with Surfaxin vs. 17% with Exosurf by 14 days, 19% with Surfaxin vs. 21% with Exosurf by 28 days, and 21% with Surfaxin vs. 24% with Exosurf by 36 weeks PCA, respectively). Furthermore, the analysis by region showed no sufficient evidence for a treatment by region interaction on RDS related mortality or non-RDS related mortality by 14 days. However, numerically, Surfaxin had a higher mortality rate than Exosurf in the Latin American region and a lower mortality rate in the European region.

5.2 Conclusions and Recommendations

Surfaxin yielded a statistically significant reduction in RDS incidence at 24 hours of age and in RDS related mortality by 14 days as compared to Exosurf, based on the adjudicated data of study KL4-IRDS-06. However, there was a large discrepancy between the Adjudicated Committee's assessment and CRF reporting in terms of these incidence rates. The CRF data did not support the RDS mortality finding. The composite of RDS mortality by 14 days or air leak by day 7 showed a close-to-borderline significant difference in favor of Surfaxin.

All cause mortality by 14 days showed little difference between Surfaxin and Exosurf. Non-RDS related mortality was significantly larger for Surfaxin. Such pattern was observed consistently in both the European region and the Latin American region. The reduction of RDS mortality and the increase of non-RDS mortality with Surfaxin seem to balance each other out. Clinical input is needed to determine whether in light of this situation, the reduction of RDS mortality is a relevant clinical benefit for Surfaxin.

The small numerical advantage for Surfaxin compared to Exosurf for all cause mortality, if any, appears to be driven by the European region. It is not clear whether there is any mortality benefit (relative to Exosurf) in the Latin American region given the slightly worse numerical difference is seen.

APPENDICES

KL4-IRDS-06 Trial

1. Retreatment criteria:

- 1). The neonate was still intubated;
- 2). At least 6 hours have elapsed since the previous Surfaxin/Survanta dose or at least 12 hours have elapsed since the previous Exosurf dose;
- 3). The neonate continued to require mechanical ventilation with a MAP of ≥ 6 cm H₂O and an FiO₂ ≥ 0.30 to maintain a PaO₂ between 50 and 80 mm Hg or an oxygen saturation (measured by pulse oximetry) between 88 and 95%; and
- 4). A chest radiograph consistent with RDS.

2. Demographic, baseline characteristics comparability among the three groups

Table 11.2.A. Patient Demographics (all randomized neonates)*

Characteristic	Surfaxin (N=527)	Exosurf (N=509)	Survanta (N=258)	p-value Surfaxin vs. Exosurf
	N (%)			
Birth Status				
Single	426 (80.8)	412 (80.9)	206 (79.8)	0.940
Multiple	101 (19.2)	97 (19.1)	52 (20.2)	
Congenital anomaly	4 (0.8)	9 (1.8)	4 (1.6)	0.144
Mode of delivery				
Vaginal spontaneous	132 (25.0)	122 (24.0)	69 (26.7)	0.600
Vaginal assisted	9 (1.7)	4 (0.8)	1 (0.4)	
Elective C-section	2 (0.4)	2 (0.4)	0	
Emergency C-section	384 (72.9)	381 (74.9)	188 (72.9)	
Race				
White	409 (77.6)	397 (78.0)	204 (79.1)	0.339
Black	3 (0.6)	4 (0.8)	3 (1.2)	
Other	115 (21.8)	108 (21.2)	51 (19.8)	
Gender				
Male	263 (49.9)	254 (49.9)	129 (50.0)	0.892
Female	264 (50.1)	255 (50.1)	129 (50.0)	
Apgar – 1 min				
Mean (S.D.)	5.3 (2.16)	5.3 (2.15)	5.3 (2.12)	0.909
N	526	508	257	
Apgar – 5 min				
Mean (S.D.)	7.1 (1.43)	7.2 (1.42)	7.1 (1.41)	0.971
N	526	508	257	
Apgar – 10 min				
Mean (S.D.)	7.4 (1.32)	7.5 (1.27)	7.4 (1.38)	0.479
N	448	438	221	
Gestational age (weeks)				
Mean (S.D.)	28.2 (1.95)	28.2 (2.03)	28.1 (2.12)	0.976
N	522	507	256	
Weight (g)				
Mean (S.D.)	973.3 (183.41)	970.5 (185.85)	966.6 (187.03)	0.685
N	527	509	258	
Length (cm)				
Mean (S.D.)	36.0 (3.26)	35.8 (3.22)	35.9 (3.26)	0.471
N	525	506	258	
Head circumference (cm)				
Mean (S.D.)	25.3 (2.00)	25.4 (1.97)	25.3 (1.94)	0.894
N	525	504	258	

Source: M5, v 1.1, sec 5.3.5.1, p 48

* extracted from the sponsor Table 11.2.A.

3. Material history comparability among the three groups

Table 11.2.B. Maternal History (all randomized neonates)*

Characteristic	Surfaxin (N=527)	Exosurf (N=509)	Survanta (N=258)	p-value Surfaxin vs. Exosurf
N (%)				
Maternal age (yr)				0.229
Mean (S.D.)	28.4 (6.79)	27.9 (6.42)	28.2 (6.51)	
Gravidity				0.709
Mean (S.D.)	2.5 (2.06)	2.6 (1.89)	2.5 (1.59)	
Parity				0.847
Mean (S.D.)	2.0 (1.36)	2.0 (1.37)	1.9 (1.20)	
Clinical chorioamnionitis	23 (4.4)	19 (3.8)	13 (5.0)	0.613
Gestational diabetes	17 (3.2)	10 (2.0)	5 (2.0)	0.243
Pregnancy-induced hypertension	142 (27.0)	146 (28.7)	62 (24.0)	0.612
Insulin-dependent diabetes	8 (1.5)	3 (0.6)	4 (1.6)	0.162
Labor				0.939
Spontaneous	255 (80.4)	233 (81.5)	133 (83.6)	
Induced	62 (19.6)	53 (18.5)	26 (16.4)	
Missing	210	223	99	
Oligohydramnios > 21 days	7 (1.3)	6 (1.2)	1 (0.4)	0.873
Antenatal steroids (missing)	415 (79.2) 3	394 (78.5) 7	191 (74.3) 1	0.684
Tocolytic therapy	200 (38.0)	198 (39.2)	129 (50.0)	0.571
Diabetes, gestational or insulin-dependent	22 (4.2)	12 (2.4)	9 (3.5)	0.127

Source: M5, v 1.1, sec 5.3.5.1, p 46

* extracted from the sponsor Table 11.2.B.

4. Definition of incidence of RDS at 24 hours (p.5 of AC SOP)

For the purpose of adjudication, RDS will be diagnosed for patients that meet a time-specific operational definition of RDS based on chest x-ray radiography (CXR) and fraction of inspired oxygen (FiO₂) data. RDS will also be diagnosed for patients who die prior to and including 28 hours of age due to RDS, or whose RDS related data are missing prior to 28 hours of age or are lost to follow-up. Specifically, RDS is defined as: *Infant requiring a FiO₂ ≥ 0.30 combined with the demonstration of a reticulogranular pattern consistent with RDS on a chest radiograph obtained between 20 and 28 hours of age.*

Patients who have a CXR positive for RDS between 16 and 20 hours of age and a repeat CXR positive for RDS between 28 and 32 hours of age and a FiO₂ ≥ 0.30 at the times these CXRs were obtained will be counted as having RDS. All other patients whose data are outside the time windows described in the above table will be counted as not having RDS. Patients who die prior to and including 28 hours of age and whose death is due to RDS will be counted as having RDS. Those patients who die prior to and including 28 hours of age due to other causes but have evidence of RDS will be counted as having RDS. Those patients who die prior to and including 28 hours of age due to other causes and have no clinical or radiological evidence of RDS will not be counted as having RDS. Patients who are lost to follow-up prior to 28 hours of age will be counted as having RDS. Patients whose RDS diagnosis is missing will be counted as having RDS.

Diagnosis	CXR at 24 (± 4) hrs post time 0	FiO ₂ at 24 (±4) hrs post time 0
RDS	Positive changes	≥ 30% at 24 (±4) hrs post time 0
No RDS	Positive or indeterminate	< 30%
	If CXR was not taken at 24 (± 4) hrs post time 0	FiO ₂ < 30% prior to or after 24 (±4) hrs post time 0

5. RDS related mortality through 14 days of age (p.6-7 of AC SOP)

For the purpose of adjudication, RDS related mortality will include patients whose death is considered to be due to RDS and its complications through the first 14 days of life based on an evaluation of provided or requested information (e.g., CRF data, autopsy report, any supportive CXR through the first 14 days of life that is consistent with RDS and not associated with sepsis/pneumonia or with pulmonary hypoplasia). Excluded are patients diagnosed with other causes of respiratory failure leading to death or other causes of death. Any infant who dies as a result of pulmonary hemorrhage will be classified as an RDS related death (per the Adjudication SOP definition), if such an infant has RDS that has not resolved prior to the pulmonary hemorrhage. Note: The diagnosis of RDS can occur at any time prior to death of the infant. (The infant does not have to meet the 24 +/- hour RDS definition). In the case of severe intracranial hemorrhage, death will be classified as RDS related if the RDS is clinically significant enough that it is likely to have contributed to this complication. In the less frequent situation where an infant had evidence of RDS that subsequently resolved or significantly improved, then had an intracranial hemorrhage, (e.g., the hemorrhage is not proximately associated with the RDS); those infants will be classified as NON-RDS related death. Note: The diagnosis of RDS can occur at any time prior to death of the infant. (The infant does not have to meet the 24 +/- 4 hours RDS definition). Sepsis can be diagnosed based on substantial clinical evidence of infection including

elevated white count, increased IT ratio, elevated CRPs, metabolic acidosis, hypotension etc., even in the setting of absent or negative blood cultures, especially if pretreatment of the mother or infant with antibiotics occurred. Positive blood cultures may be considered contaminated if the organism was one not commonly associated with early onset sepsis, and there is no laboratory evidence or clinical symptoms consistent with infection. Patients lost to follow-up prior to and including day 14 of life will be counted as having died due to RDS. Patients whose data are missing will be counted as having died due to RDS.

6. Definition of air leak through 7 days of age (AC SOP p.7)

Air leak will be defined as chest radiographic evidence of air leak (e.g., pneumothorax, pulmonary interstitial emphysema, pneumomediastinum, pneumopericardium, subcutaneous emphysema) resulting from lung parenchymal disease. Patients who die due to RDS or other respiratory causes prior to and including 7 days of age will be counted as having air leaks. Patients who die prior to and including 7 days of age when there is no evidence of air leaks due to other causes will be counted as not having air leaks. Patients who are lost to follow-up prior to and including 7 days of age will be counted as having air leaks. Patients whose data (e.g., chest x-rays) are missing will be counted as having air leaks. Adjudication Committee members will determine and specify the etiology of any positively adjudicated events of air leak as caused by barotraumas/volutrauma or by another cause.

7. Sample size re-estimation

A possible sample size re-estimation was planned and is to be recommended by the DSMB if the observed treatment effect is

8. Change of co-primary efficacy endpoint

Due to possible confusion from the sponsor regarding the Agency's view of the co-primary efficacy endpoints, (1) incidence of RDS at 24 hours of age and (2) RDS related mortality through 14 days of age and/or air leak by day 7, where the second component is a composite endpoint, the sponsor proposed four options of the second component as the co-primary efficacy endpoint. Further discussions with the Agency, the co-primary efficacy endpoints used in the final statistical analysis plan are (1) incidence of RDS at 24 hours of age and (2) RDS related mortality through 14 days of age. Statistical criteria for a superior Surfaxin effect as compared to Exosurf are that both components each needs to show statistical significance at a 2-sided 5% level.

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

/s/

Sue Jane Wang
1/13/05 10:14:12 PM
BIOMETRICS

S. Edward Nevius
1/14/05 02:53:29 PM
BIOMETRICS
Concur with review.