BMC Pediatrics

RESEARCH

Open Access

Comparing the Effect of Beractant (Beraksurf™) with That of Poractant Alfa (Curosurf®) on the Need for Intermittent Positive Pressure Ventilation in Neonatal Respiratory Distress Syndrome by Adopting a Semi-parametric Approach: Re-Analyzing Data of a Randomized Controlled Trial

Yosra Khazani[1], Sirous Fathi Manesh[2], Elnaz Shaseb[3] and Parvin Sarbakhsh[1*]

Abstract

Introduction Randomized controlled trial (RCT) data are analyzed by two challengeable adjusted and non-adjusted approaches. Performing appropriate adjusted analysis leads to a more interpretable and efficient estimation of treatment effects. Semiparametric adjustment approach modifies the estimating equations solved by the marginal treatment effect estimator by adding an augmentation function, which makes use of the baseline covariates and estimate the unbiased marginal treatment effect with improved precision. The effect of the intervention obtained using the semi-parametric adjustment method, similar to the unadjusted method and contrary to the adjusted parametric method, is marginal, resulting in better interpretability. Moreover, due to leveraging baseline covariates, it is more efficient compared to the unadjusted models. This study aimed to estimate the effect of beractant (Beraksurf™, Tekzima Company), compared with the Poractant alfa (Curosurf®, Chiesi Pharmaceuticals), as surfactant replacement therapy, on the need for Intermittent Positive Pressure Ventilation (IPPV) in Neonatal Respiratory Distress Syndrome (NRDS) more precisely by fitting a semi-parametric efficient model adjusted for appropriate covariates.

Method This study is secondary and we re-analyzing data of a published RCT. This RCT was conducted in the NICU of Alzahra Hospital in Tabriz, Iran for eight months, and 200 infants were assigned to two groups receiving either 100 mg/kg BeraksurfTM ($n=99$) or 200 mg/kg Curosurf® ($n=101$). The effect of the treatments was evaluated regarding the need for IPPV by fitting semi-parametric logistic regression models, adjusted for the best subset of covariates selected by the forward variable selection algorithm and confounders identified by the expert panel. IPPV in our study was administered via an endotracheal tube, as per the protocol followed in the primary trial. The need for IPPV was determined based on the clinical judgment of neonatologists, considering the infants'

*Correspondence:
Parvin Sarbakhsh
sarbakhshp@tbzmed.ac.ir
Full list of author information is available at the end of the article

respiratory distress levels, oxygen saturation, and arterial blood gas measurements. The decision was made within the first 72 h after surfactant administration.

Results The proportions of the required IPPV outcome were 29.3% and 59.4% in the BeraksurfTM group and Curosurf®, respectively. While the unadjusted comparison between the two treatments was significant (OR = .283, 95% CI: (.157, .509), *P*-value < .001), adjusted OR in semi-parametric logistic regression by adjusting for the best subset of covariates selected by forward method including steroids, Apgar score at min1, and initial Fio2 wasn't significant (OR = .751, 95% CI: (.510, 1.111), *P*-value = .151). The efficacy of this semi-parametric model over the unadjusted model was 1.81. The results of the expert-based adjusted model, adjusting for the gestational age, birth weight, and initial FiO2, were consistent with those derived from the R^2-based selection, supporting the robustness of our findings (OR = .685, 95% CI: (.449, 1.529), *P*-value = .078).

Conclusion Fitting a semi-parametric model adjusted for the baseline covariates resulted in a marginal, unbiased, more efficient and interpretable estimation of BeraksurfTM versus Curosurf® effects. Although the crude model showed that BeraksurfTM was more effective than Curosurf®, the results of the efficient semi-parametric model with adjustment for the best subset of covariates revealed no statistically significant difference between the two drugs regarding their effects. We hope that the use of this method and its findings will contribute to a better understanding of covariate adjustment.

Keywords Semi-parametric model, Adjustment, IPPV, Preterm infant, Beractant, Poractant alfa, Respiratory distress syndrome

Introduction

Randomized controlled trials (RCT) are one of the clinical trial types in which a treatment is randomly allocated to the participants. The main goal in these studies is to estimate the effect(s) of an intervention on the intended outcome. The medium and large-sample clinical trials are usually set up so that other auxiliary variables, in addition to the outcome variable information, are also collected for each participant. Covariates include demographic characteristics (e.g., age, gender, etc.), clinical characteristics (e.g., treatment history, mental status, etc.), and the baseline values of the outcomes. Some of these auxiliary variables may be correlated with the outcome variable and may facilitate investigating the difference between two treatments [1]. The main goal in RCT studies is not to investigate the effect of these auxiliary variables; however, their inclusion in RCT studies together with the intervention variable improves the accuracy of estimations and efficiency of the model when there is a correlation between these variables and the outcome variable [2]. The randomized clinical trial studies are statistically analyzed by adopting two adjusted and non-adjusted approaches. In a non-adjusted approach, only the effect of the intervention variable is included in the model whereas in an adjusted analysis, the effect of the intervention variable is examined in the presence of auxiliary variables. Although an appropriate adjustment increases the accuracy and efficiency of the intervention effect estimate, the application of auxiliary variables for adjustment is accompanied by certain subtleties and sensitivities, and an inappropriate adjustment may cause bias in the estimations and even produce misleading results. Furthermore, the complexity of the interpretations of the results produced by adjustment methods has discouraged some researchers to employ these methods. Due to the above-mentioned reasons, the issue of adjustment or non-adjustment in RCT studies has become a challenging issue [2, 3].

To address this challenge and facilitate an optimal application of the information collected in an RCT study, Tsiatis et al [1] introduced a semi-parametric adjustment method for simple linear regression models that separates the estimation of treatment differences from the adjustment process, reducing bias concerns common in regression-based methods. Their method facilitated the estimation of treatment effects that were more powerful than those produced by similar methods as well as permitted the estimators to ensure an asymptotical normal distribution.

In this method, the systematic modeling of the relationship between auxiliary variable and outcome as well as the evaluation of a treatment effect are separated. It enjoys the simplicity of the non-adjustment methods in terms of interpretability. Zhang et al [4] generalized a method developed by Tsiatis et al. based on a linear regression model to a general statistical model (e.g., logistic regression, Cox regression, mixture model, etc.), and presented an innovative semi-parametric adjustment method for estimation and hypothesis testing of the treatment effect. This method leverages the flexibility of non-parametric techniques alongside the structure of parametric models, offering a robust alternative to commonly used methods like traditional multivariable models or propensity score adjustment. It provides

researchers with a powerful tool to adjust for covariates, especially in studies where the relationships between covariates and outcomes are complex. Unlike traditional methods such as propensity score analysis or multivariable regression, the semiparametric approach, separates the estimation of treatment differences from the adjustment process. This separation allows for unbiased and efficient estimation while maintaining interpretability.

The presented semi-parametric adjustment approach modifies the estimating equations solved by the marginal treatment effect estimator by adding an augmentation function, which makes use of the baseline covariates and estimate the unbiased marginal treatment effect with improved precision. The effect of the intervention obtained using the semi-parametric adjustment method, similar to the unadjusted method and contrary to the adjusted parametric method, is marginal, resulting in better interpretability. Moreover, due to leveraging baseline covariates, it is more efficient compared to the unadjusted models.

On the other hand, infant respiratory distress syndrome (RDS) is a type of lung development defect that usually occurs after a premature birth due to the surfactant deficiency. RDS is one of the main causes of death among the premature infants.

The goal of managing RDS is to provide interventions to maximize the survival while minimizing the potential complications. One of the most important treatments for infants suffering from RDS is the administration of exogenous surfactant. Different types of surfactant drugs derived from natural sources (e.g., Curosurf, Infasurf Surfacten, and [Beractant] Surventa) and artificial ones (e.g., Pneumactant, Venticute, Exosurf, and Syrfaxim) have been recognized as successful treatments [5, 6]. Due to the importance of this drug, much research has been devoted to investigating the effectiveness, side effects, and comparison of the various types of it.

One of the available and most common types of exogenous pulmonary surfactants in Iran is Poractant alfa (Curosurf®, Chiesi, Italy), which has FDA approval. The Iranian version of surfactant named Beractant (Beraksurf™), which is the generic form of Survanta®, has been produced by Tekzima Company (Alborz, Iran) since 2018.

Due to the limited resources, an interesting issue in this context is to compare the efficacy and safety of this version with those of the common type of surfactant replacement therapy in RDS. To assess the feasibility of replacing BeraksurfTM with Curosurf®, it is extremely important to compare their effects on RDS and the adverse events of RDS more accurately. One of the most important adverse events of RDS is the need for Intermittent Positive Pressure Ventilation (IPPV). IPPV provides a non-invasive respiratory support for preterm infants who need endotracheal intubation and ventilation [7]. The current study, therefore, aimed to estimate the effect of BeraksurfTM compared with that of Curosurf® on the outcomes of the need for IPPV in the preterm infants with RDS more precisely by fitting a semi-parametric logistic regression model, adjusted for covariates, as well as to compare adjusted semi-parametric results with the findings produced by the parametric adjusted and unadjusted approaches based on the data from a RCT study by Gharebaghi et al. [8].

Methodology

Data

This study is a secondary study and we re-analyzing data of a published RCT conducted by Gharehbaghi et al. [8]. For this primary study, the study protocol was registered at the Iranian Registry of Clinical Trials (IRCT) (IRCT20180404039187N4) prior to subject recruitment.

The data of primary study were 200 preterm infants with RDS and admitted to the intensive care unit of Alzahra Hospital in Tabriz from November, 2018 to June, 2019 in order to evaluate the efficacy of Beraksurf™ from the Tekzima company (Alborz, Iran) as a newly produced drug vs. Curosurf® as an available exogenous pulmonary surfactant for curing infants with RDS [8]. These infants were randomized into two groups receiving Beraksurf™ ($n=99$) and Curosurf® ($n=101$) by the block randomization technique. The Beraksurf™ group received this medicine at a maximum dose of 4 every six hours, while the other group received the drug Curosurf® at a dose of 2.5 every 12 h and, if re-prescription was needed, at a dose of 1.25 or a maximum dose of 3 in 48 h after birth. The report of this clinical trial with more details is available elsewhere [8]. For the purpose of this article, we just assessed the IPPV outcome. IPPV in our study was administered via an endotracheal tube, as per the protocol followed in the primary trial. The need for IPPV was determined based on the clinical judgment of neonatologists, considering the infants' respiratory distress levels, oxygen saturation, and arterial blood gas measurements. The decision was made within the first 72 h after surfactant administration. While our re-analysis focused on the binary outcome of the need for IPPV, the primary trial collected data on secondary outcomes such as FiO2 changes, duration of mechanical ventilation, and duration of respiratory support. These outcomes were not included in the current re-analysis but remain valuable for future studies.

Statistical methods

Data were presented as mean (SD) and number (percent) for continuous and categorical variables, respectively. To assess the treatment's effect on IPPV, an adjusted

semi-parametric approach was employed, and our results were compared with those of the unadjusted and adjusted parametric models.

Covariates for adjustment were selected adopting two approaches. First, we fitted parametric and semi-parametric models using covariates selected by domain experts to incorporate clinical context or pathophysiological insights. For this models, the initial FiO2, gestational age, and birth weight were selected as potential confounders according to the experts' opinion and consistent with original study report [8]. Second, we used a data-driven approach to identify the best subset of covariates based on R^2; So, the best subset of covariates with the highest R^2 was selected using a semi-parametric forward variable selection algorithm. All analyses were carried out using the speff2trial package in R version 4.3.2.

There are various statistical methods for comparing different levels of treatment in RCT studies. In this section, the method developed by Zhang et al [4] based on the efficient semi-parametric approach was described while briefly stating some common adjustment/non-adjustment methods.

Considering the binary nature of our outcome (i.e., the need for IPPV) and treatment (i.e., Beraksurf™ and Curosurf®), these methods were described for the case where the outcome and intervention variables are both binary (0 or 1). The outcome variable was denoted by Y, the intervention by Z, and the auxiliary variables by X.

Unadjusted logistic regression method

In this method and considering $\text{logit}(p) = \ln\frac{p}{1-p}$, the relationship between outcome variable and treatment is determined by:

$$\text{Logit } p(Y = 1|Z) = \beta_0 + \beta_1 Z + \varepsilon \tag{1}$$

where, p is the proportion of success for the outcome variable. In this model, the interpretation of parameter $\beta_1 = \log \frac{\frac{p(Y=1|Z=1)}{p(Y=0|Z=1)}}{\frac{p(Y=1|Z=0)}{p(Y=0|Z=0)}}$ as is log of the ratio of odds under treatment 1 to treatment 0.

Adjusted logistic regression method

In this model, the effect of the intervention variable is adjusted by auxiliary variables $\mathbf{X} = (X_1, \ldots X_p)$ as:

$$\text{Logit } p(Y = 1|\mathbf{X}, Z) = \beta_0 + \beta_1 Z + \alpha_1 X_1 + \cdots + \alpha_p X_p + \varepsilon \tag{2}$$

In this model,β_1 appears as:$\beta_1 = \log \frac{\frac{p(Y=1|Z=1,X)}{p(Y=0|Z=1,X)}}{\frac{p(Y=1|Z=0,X)}{p(Y=0|Z=0,X)}}$

where the parameter β_1 shows the log ratio of odds in two treatments conditionally, and a zero value represents the sameness of the conditional odds. Therefore, the interpretation of parameter β_1 is not the same in two adjusted and non-adjusted models [1].

Many researchers avoid using adjustment in the logistic regression model since they believe that this method reduces the accuracy of the estimator. This belief is mainly formed due to a confusion over the following two issues.

Firstly, a lack of proper understanding of the difference between conditional and marginal treatment effects causes that the conditional treatment effect has been mistakenly used instead of the marginal treatment effect. Secondly, the selection of auxiliary variables is always a challenging issue because an inappropriate selection of these variables distorts the model. In the adjusted model, for instance, the estimate of parameter β_1 may be significantly altered by changing the covariates and, therefore, the result of the trial study may be reversed.

Logistic regression method with efficient semi-parametric approach

The basic idea of semi-parametric adjustment method introduced by Zhang et al. [4] is that we can modify the estimating equations which are solved by the marginal (unconditional) treatment effect estimator by adding an augmentation function, which makes use of the baseline covariates and estimate the marginal treatment effect with improved precision.

The construction of the augmentation function depends on specifying a parametric working model for predicting the outcome using the baseline covariates. Importantly however, estimates remain (asymptotically) unbiased irrespective of whether this so-called working model is correctly specified.

To this end, the unadjusted logistic regression model was considered as $E(Y|Z) = \frac{\exp(\beta_0+\beta_1 Z)}{1+\exp(\beta_0+\beta_1 Z)}$.

It is a function based on a single observation and parameters, which is used to build the estimation equations. Therefore, its estimation function is:

$$m(Y, Z; \beta_0, \beta_1) = \{1, Z\}^T \left\{ Y - \frac{\exp(\beta_0 + \beta_1 Z)}{1 + \exp(\beta_0 + \beta_1 Z)} \right\} \tag{3}$$

As a result, it is deduced as an unadjusted estimate for the model parameters by solving the equations:

$$\sum_{i=1}^{n} m(Y_i, Z_i; \beta_0, \beta_1) = 0 \tag{4}$$

Since these estimators are in the class of M-estimators, they ensure an asymptotic normal distribution. To improve this estimator and apply the auxiliary variables, Zhang et al. suggested the following estimation function:

$$\mathrm{m}^{*}(\mathrm{Y},\mathbf{X},\mathrm{Z};\beta_0,\beta_1)=\mathrm{m}(\mathrm{Y},\mathrm{Z};\beta_0,\beta_1)-\sum_{\mathrm{g}=0}^{1}\left\{\mathrm{I}_{(\mathrm{Z}=\mathrm{g})}-\pi_{\mathrm{g}}\right\}E\left(m(Y,Z,\beta_0,\beta)|\boldsymbol{X},Z=g\right) \tag{5}$$

where π_0 and π_1 are the ratio of observations at Z levels 0 and 1, respectively:

$$E(Y|\boldsymbol{X},Z)=\frac{\exp(\beta_0+\beta_1\mathrm{Z}+\alpha_1X_1+\cdots+\alpha_pX_p)}{1+\exp(\beta_0+\beta_1\mathrm{Z}+\alpha_1X_1+\cdots+\alpha_pX_p)} \tag{6}$$

By solving these estimation equations:

$$\sum_{i=1}^{n}\mathrm{m}^{*}\left(\mathrm{Y_i},\mathbf{X_i},\mathrm{Z_i};\beta_0,\beta_1\right)=0 \tag{7}$$

the efficient semi-parametric adjusted estimators are obtained for the model parameters. These estimators also belong to the family of M-estimators; therefore, they are asymptotically/normally distributed, and their efficiency can be calculated and/or compared to the non-adjusted state.

Two important advantages of these estimators compared to the non-adjusted mode and the conditional mode are as follows:

1. The interpretation of parameters β_0,β_1 in this model is also similar to that in the non-adjustment mode and, as a result, these parameters reflect the value of the marginal treatment effect and not the conditional effect of the treatment.
2. The obtained estimators are more efficient than those obtained from unadjusted logistic regression.

Similarly, Zhang et al [4] developed an efficient semi-parametric method for hypothesis testing

$$\mathrm{H}_0:\beta_1=0$$

$$\mathrm{H}_1:\beta_1\neq 0$$

The test statistic in this method, similar to the estimation problem, is obtained in two steps; that is, first a statistic is found by considering the non-adjusted model and then is improved by including the adjustment. The adjusted test statistic obtained in this normal asymptotic method has a higher power than that obtained in the non-adjusted state.

Results

In the primary study, 101 infants received 200 mg/kg of poractant alfa (Curosurf®) and 99 infants received 100 mg/kg of beractant (Beraksurf™) as the initial doses. The mother/infant-related baseline characteristics are shown in Table 1.

The study groups were not significantly different in terms of the 1-min Apgar score at 1 min score, mode of delivery, and mother-related risk factors (e.g., the mother's underlying diseases, maternal age, the incidence of Premature Rupture of Membranes (PROM), and preeclampsia), but they were significantly different regarding the gestational age, birth weight, and application of antenatal steroids.

Table 2 shows the estimation obtained for the treatment effect on the need for IPPV outcome after implementing unadjusted/adjusted parametric and semi-parametric models.

The proportions of the required IPPV outcome were 29.3% and 59.4% in the Beraksurf™ group and Curosurf®, respectively. The unadjusted comparison between two treatments was significant the unadjusted estimate of the log-odds ratio, is −1.261 with standard error 0.299 (OR = 0.283, 95% CI: (0.157, 0.509), *P*-value < 0.001). The conditional OR became insignificant after adjusting for the expert-based selected confounders including baseline FiO_2, birth weight, and gestational age through the parametric logistic regression (OR = 0.464, 95% CI: (0.199, 1.082), *P*-value = 0.075).

Table 1 Baseline covariates of mothers and infants

variable	Beraksurf™ (*N* = 99)	Curosurf® (*N* = 101)	*P*-value
Gestational age (week), Mean (SD)	32.9 (2.6)	30.3(2.6)	< .001*
Birth weight (g), Mean (SD)	1957.2 (675.2)	1566.2 (672.3)	< .001*
Maternal age (year), Mean (SD)	29.9 (5.7)	31 (7.2)	.21*
Apgar score at 1 min, Mean (SD)	6.6 (2.1)	5.9 (2.3)	.30*
Apgar score at 5 min, Mean (SD)	8.3 (1.4)	7.7 (1.7)	.01*
Initial Fio_2, Mean (SD)	75.93 (9.66)	81.94 (12.47)	< .001*
Preterm Delivery, n(%)	50 (50.5)	68 (67.3)	.01**
using Antenatal steroid, n(%)	73 (74.5)	83 (92.2)	.01**
Preeclampsia, n(%)	11 (11.1)	18 (17.8)	.22**

* *P*-value from independent T test

** *P*-value from chi-square test

Table 2 Estimating effect of Beraksurf™ versus Curosurf® on need to IPPV outcome by the unadjusted and adjusted models

Model	b	SE(b)	95%CI	*p*-value
Unadjusted parametric model	−1.261	.299	(−1.849, -.674)	< .001
Parametric model adjusted for potential confounders[a]	-.769	.432	(−1.61, .080)	.075
Parametric model adjusted for optimal covariates[b]	-.601	.412	(−1.410, .207)	.145
Semi-parametric model adjusted for potential confounders[a]	-.378	.215	(.-.079, .0425)	.078
Semi-parametric model adjusted for optimal covariates[b]	-.289	.201	(-.685, .106)	.151

[a] Potential confounders: initial fio2, Gestational age, birth weight

[b] Optimal covariates: steroids, Apgar score at 1 min, initial fio2

The results produced after adjusting for these variables via a semi-parametric logistic regression showed a non-significant effect for the treatment (OR = 0.685, 95% CI: (0.449, 1.529), *P*-value = 0.078).

The results produced after a selection of the best covariates subset by the forward selection method demonstrated a model with a higher R^2. Selected variables to adjust in this model were steroids, Apgar score at 1 min, and initial Fio_2. Adjusted OR in semi-parametric model was 0.75 (OR = 0.75, 95% CI: (0.51, 1.111), *P*-value = 0.151).

An OR of 0.548 was achieved after considering the selected variables in the parametric logistic regression (OR = 0.548, 95% CI: (0.244, 1.229), *P*-value = 0.145).

The efficacy of the semi-parametric model with the best covariates subset selected by the forward selection method over the unadjusted model was: $\text{efficacy} = \left(\frac{se\left(\hat{\beta}_{unadjusted}\right)}{se\left(\hat{\beta}_{semi-parametric}\right)}\right)^2 = \left(\frac{.299}{.201}\right)^2 = 2.22$. The efficacy of the semi-parametric model with the potential confounders over the unadjusted model was 1.93. Furthermore, the efficacy of the semi-parametric model to the parametric model for the best covariates subset was 4.20 and for potential confounders was 4.03.

Discussion

Performing appropriate adjusted models leads to a more interpretable and efficient estimation of treatment effects. The semi-parametric adjusted estimations, similar to the unadjusted estimations and unlike the parametric adjusted estimations, are unconditional and more interpretable but their estimations are more efficient than the unadjusted model because of leveraging baseline predictors of the outcomes. While we acknowledge the existence of alternative approaches for adjusting imbalanced baseline prognostic factors, such as classical control for confounders using traditional multivariable models [2], propensity score analysis [9] or confounder summary score adjustment [10], in our study, we used the regression approach with the semi-parametric framework due to its better interpretability and more efficiency. This method can complement existing approaches and serve as a valuable option for investigators aiming to enhance the precision and reliability of their analyses.

In our study, the effect of BeraksurfTM versus Curosurf® on the need for IPPV outcome was estimated by using unadjusted/adjusted parametric and semi-parametric models. Adjusted covariates selected by the expert were initial FiO_2, Gestational age, and Birth weight, whereas those selected by semi-parametric forward variable selection algorithm were steroids, Apgar score at 1 min, and initial fio2.

Our results demonstrated that the coefficients of the parametric adjusted models were lower than that of the crude model, which suggested the existence of heterogeneity between two groups in terms of the adjusted variables and, therefore, a significant change in the effect size of the intervention caused by the adjustment for the covariates. This difference may have been attributed to the existence of a significant difference between two groups regarding the baseline values of variables such as initial Fio_2, Apgar score at 1 min, steroids, gestational age, and birth weight.

In each covariate subset, moreover, the absolute values of regression coefficients in the semi-parametric model were smaller than the coefficients of the parametric model with the same covariates. This may have been due to the fact that the values of regression coefficients in the parametric models are conditional effect of the intervention and not its marginal effect, while the obtained coefficient in the semi-parametric model is marginal and indicative of the unconditional effect of the intervention on the outcome like the unadjusted value; seemingly, the contribution of other variables was removed from this value.

In addition, the results generated by parametric analysis with adjustment for potential confounders (borderline *P*-value) were more comparable to those produced by the unadjusted model (significant *P*-value) in term of *P*-value; this is while, parametric model with adjustment for optimal covariates, which were selected based on the forward selection method, did not show a significant effect for the treatment.

This issue indicates that in the parametric adjustment model, the correct determination of covariates for adjustment can greatly affect the results; therefore, in the parametric adjustment approach, only an adjustment is not enough, and the adjusted variables and their functional form should be also selected appropriately to have unbiased estimator of intervention effect.

On the other hand, because the semi-parametric approach separates the estimation of treatment differences from the adjustment process, the use of model selection techniques, such as forward selection, to determine covariates to include in the augmentation term models have no effect asymptotically on the properties of the intervention effect estimator. Hence, misspecification of the working models (covariates section) in semiparametric does not introduce bias in the treatment effect estimator, but instead some efficiency is lost [4].

According to our study results, the SE of coefficients in semi-parametric models was, as expected, lower than that in similar parametric models, which was indicative of the fact that the semi-parametric models were more efficient and generated more precise estimations for the intervention effect [4].

Moreover, the results revealed that the most important variable to adjust in all adjusted models was the initial fio2 variable. The results of the semi-parametric and parametric models became similar to those of the unadjusted model after removing fio2 variable from the models, which suggested that more accurate and realistic results may have been obtained for the intervention effect by adjusting it when there was an important covariate demonstrating a high correlation with the outcome.

Gestational age and birth weight were among the confounders selected based on the expert opinion. Although these variables had effect on the outcome of the need for IPPV and the two groups were different in terms of these variables, they were not observed in the model with optimal covariates, which may have been attributable to the fact that these variables were strongly correlated ($r > 0.7$) with initial fio2 and, therefore, were not selected in the forward method in the process of variables selection due to their collinearity and common information but were represented by the initial fio2 variable.

Although the crude model showed greater effectiveness of BeraksurfTM, after adjusting for the optimal covariates (some of them had a confounding role), the final results of the semi-parametric models indicated no significant difference between two drugs concerning their effects.

Taking into account the results from this study, although the crude model showed that BeraksurfTM has a significantly higher effect than Curosurf®, the results of the efficient semi-parametric models with adjustment for the best subset of covariates and potential confounders were argued that the effect of the Iranian version of surfactant (i.e., BeraksurfTM) was not significantly different from its foreign version (i.e., Curosurf®) in reducing the need to IPPV. This does not imply equivalence, as equivalence testing with appropriate margins and larger sample sizes would be required to confirm this. We recommend future studies with equivalence testing to assess the true comparability of these treatments.

Previous studies have investigated the effect of Bractant and Curosurf on the pulmonary outcome and its complications such as the need for IPPV. FiO_2 is one of the most critical factors in assessing the efficacy of drugs in RDS. Studies comparing the effects of Bractant and Curosurf on RDS have reported contradictory results about the efficacy of Bractant. Gharehbaghi et al. analyzed these data and concluded that the two drugs were the same in terms of side effects, but the amount of Fio_2 in the group of infants aged under 32 weeks and receiving Bractant was significantly reduced compared to that in the group receiving Curosurf® [8].

Fujii et al [11] examined the infants aged under 30 weeks and with RDS symptoms in 2010, and concluded that the amount of Fio_2 in the Curosurf group 72 h after birth was lower than that in Bractant. Malloy et al. (2005) demonstrated that the FiO_2 requirement in the first 48 h in the poractant alfa (Curosurf®) group was significantly lower than that in the beractant (Survanta®) group [12]. Saeedi et al. [13] also showed that the two drugs (i.e., Curosurf and Bractant) were similar in terms of mortality, length of hospitalization, and the need for ventilation. Mirzarahimi et al [14] similarly indicated that these two drugs had similar side effects, but Curosurf required less re-dosing than Bractant, and Bractant caused a greater reduction in ventilation time than Curosurf. Dilli et al [15] also reported that the amount of the required oxygen, mechanical ventilation, and length of hospitalization were similar in two groups receiving Bractant and Curosurf®, which was consistent with our results.

While we used a data-driven approach to identify the best subset of covariates based on R^2 we also fitted a model using covariates selected by domain experts to incorporate clinical context or pathophysiological insights, including clinically relevant variables such as gestational age, birth weight, and initial FiO2. These confounders had been selected by experts based on the DAGs approach in the primary study and we considered them for adjustment in the current study.

The results of this expert-based semi-parametric model were consistent with those derived from the R^2-based selection, supporting the robustness of our findings.

Although we fitted a model by considering experts opinions, alternative methods, such as the Structural Equation Modeling (SEM), could provide more clinically oriented insights. We suggest future studies consider these approaches for better evaluation of causal effects.

Limitations

The primary study has been conducted in one center. It was recommended that a multicenter study with a longer follow-up duration should be carried out to evaluate these drugs and obtain more accurate outcomes. While our re-analysis focused on the binary outcome of the need for IPPV, the primary trial collected data on secondary outcomes such as FiO2 changes, duration of mechanical ventilation, and duration of respiratory support. These outcomes were not included in the current re-analysis but remain valuable for future studies.

Conclusion

Fitting a semi-parametric model adjusted for the best subset of covariates resulted in a more interpretable and efficient estimation of BeraksurfTM versus Curosurf® effects. Although the crude model showed that BeraksurfTM has significantly higher effect than Curosurf®, the results of the efficient semi-parametric models with adjustment for the best subset of covariates and potential confounders revealed no statistically significant difference between the two drugs regarding their effects. We hope that the use of this method and its findings will contribute to a better understanding of covariate adjustment.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12887-025-05477-z.

Supplementary Material 1.

Acknowledgements

The authors would like to appreciate Clinical Research Development Unit of Al-Zahra Educational and Treatment Center for kind supports, and the Vice-Chancellor for Research and Technology of Tabriz University of Medical Sciences for approving and supporting this research (grant number: 68967). This paper was extracted from M.Sc. thesis of Yosra Khazani in Biostatistics.

Human ethics and consent to participate declarations

Current study is a secondary study based on the data of the primary study. In the primary study, written informed consent form signed by the parents or legal guardians of the infants.

Authors' contributions

PS initiated the project. PS and ESH designed the project. YKH and SF led the knowledge elicitation, data analysis, and interpretation. PS and YKH led wrote the manuscript. All authors reviewed and approved the final manuscript.

Funding

This study was financially supported by the Vice-Chancellor for Research and Technology of Tabriz University of Medical Sciences with grant number 68967.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

This study is a secondary study based on the data of the primary study that was a clinical trial. For the primary study, ethical approval had been obtained from the Ethics committee of Tabriz University of Medical Sciences (ethics No. TBZMED.REC.1398.334). Also, the study protocol was registered at the Iranian Registry of Clinical Trials (IRCT) (IRCT20180404039187N4) prior to subject recruitment. Current secondary study was approved by the Ethics Committee of Tabriz University of Medical Sciences as well (ethics No. IR.TBZMED.REC.1401.577).

The study adhered to ethical guidelines, including the principles outlined in the Helsinki Declaration of 1975 and the European Convention of Oviedo dated 4 April 1997, which safeguards human rights and dignity concerning the applications of biology and medicine. Before participating in the study, the parents or legal guardians of all included individuals provided written informed consent. The results obtained were maintained in an anonymous manner to ensure confidentiality.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Statistics and Epidemiology, Faculty of Health, Tabriz University of Medical Sciences, Tabriz, Iran. [2]Department of Statistics, Faculty of Science, University of Kurdistan, Kurdistan, Iran. [3]Faculty of Pharmacy, Tabriz University of Medical Science, Tabriz, Iran.

Received: 15 March 2024 Accepted: 30 January 2025
Published online: 26 February 2025

References

1. Tsiatis AA. Semiparametric theory and missing data. New York: Springer; 2006.
2. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med. 2002;21:2917–30.
3. Wang J. Covariate adjustment for randomized controlled trials revisited. Pharm Stat. 2020;19:255–61.
4. Zhang M, Tsiatis AA, Davidian M. Improving efficiency of inferences in randomized clinical trials using auxiliary covariates. Biometrics. 2008;64:707–15.
5. Choi CW, Hwang JH, Yoo EJ, Kim KA, Koh SY, Lee YK, et al. Comparison of clinical efficacy of Newfactan versus Surfacten for the treatment of respiratory distress syndrome in the newborn infants. J Korean Med Sci. 2005;20:591–7.
6. Ramanathan R, Bhatia JJ, Sekar K, Ernst FR. Mortality in preterm infants with respiratory distress syndrome treated with poractant alfa, calfactant or beractant: a retrospective study. J Perinatol. 2013;33:119–25.
7. Owen LS, Morley CJ, Davis PG. Neonatal nasal intermittent positive pressure ventilation: what do we know in 2007? Arch Dis Child Fetal Neonatal Ed. 2007;92:F414–8.
8. Gharehbaghi MM, Sarbakhsh P, Mohammadbager H, Gharibzadeh S, Shaseb E. A Randomized, Single-blind, Comparison Trial of Beractant (Beraksurf^{TM}) versus Poractant Alfa (Curosurf®) in the Treatment of Respiratory Distress Syndrome in Preterm Infants. Pharm Sci. 2023;29:52–8.

9. Xu Z, Kalbfleisch JD. Propensity score matching in randomized clinical trials. Biometrics. 2010;66:813–23.
10. Nguyen TL, Debray TPA, Youn B, Simoneau G, Collins GS. Confounder Adjustment Using the Disease Risk Score: A Proposal for Weighting Methods. Am J Epidemiol. 2024;193:377–88.
11. Fujii AM, Patel SM, Allen R, Doros G, Guo CY, Testa S. Poractant alfa and beractant treatment of very premature infants with respiratory distress syndrome. J Perinatol. 2010;30:665–70.
12. Malloy CA, Nicoski P, Muraskas JK. A randomized trial comparing beractant and poractant treatment in neonatal respiratory distress syndrome. Acta Paediatr. 2005;94:779–84.
13. Saeedi R, Hamedi A, Javadi A, Gholami Robatsangi M, Dinparvar SK. Comparison of side effect of survanta and curosurf in decreasing mortality due to respiratory distress syndrome (RDS) in premature infants admitted in NICU of Ghaem Hospital On 2006–2008. Iranian Journal of Neonatology. 2013;4:7–12.
14. Mirzarahimi M, Barak M. Comparison efficacy of Curosurf and Survanta in preterm infants with respiratory distress syndrome. Pak J Pharm Sci. 2018;31:469–72.
15. Dilli D, Çakmakçı E, Akduman H, Oktem A, Aydoğan S, Çitli R, et al. Comparison of three natural surfactants according to lung ultrasonography scores in newborns with respiratory distress syndrome. J Matern Fetal Neonatal Med. 2021;34:1634–40.

Publisher's Note