RESEARCH

Check for
updates

Comparison of two modeling approaches for the identification of predictors of complications in children with cerebral palsy following spine surgery

Rachel L. Difazio[1*], Tania D. Strout[2], Judith A. Vessey[3], Jay G. Berry[4] and Daniel G. Whitney[5]

Abstract

Background Children with non-ambulatory cerebral palsy (CP) frequently develop progressive neuromuscular scoliosis and require surgical intervention. Due to their comorbidities, they are at high risk for developing peri- and post-operative complications. The objectives of this study were to compare stepwise and LASSO variable selection techniques for consistency in identifying predictors when modelling these post-operative complications and to identify potential predictors of respiratory complications and infections following spine surgery among children with CP.

Methods In this retrospective cohort study, a large administrative claims database was queried to identify children who met the following criteria: 1) ≤ 25 years old, 2) diagnosis of CP, 3) underwent surgery during the study period, 4) had ≥ 12-months pre-operative, and 5) ≥ 3-months post-operative continuous health plan enrollment. Outcome measures included the development of a post-operative respiratory complication (e.g., pneumonia, aspiration pneumonia, atelectasis, pleural effusion, pneumothorax, pulmonary edema) or an infection (e.g., surgical site infection, urinary tract infection, meningitis, peritonitis, sepsis, or septicemia) within 3 months of surgery. Codes were used to identify CP, surgical procedures, medical comorbidities and the development of post-operative respiratory complications and infections. Two approaches to variable selection, stepwise and LASSO, were compared to determine which potential predictors of respiratory complications and infection development would be identified using each approach.

Results The sample included 220 children. During the 3-month follow-up, 21.8% (*n* = 48) developed a respiratory complication and 12.7% (*n* = 28) developed an infection. The prevalence of 11 variables including age, sex and 9 comorbidities were initially considered to be potential predictors based on the intended outcome of interest. Model discrimination utilizing LASSO for variable selection was slightly improved over the stepwise regression approach. LASSO resulted in retention of additional comorbidities that may have meaningful associations to consider for future studies, including gastrointestinal issues, bladder dysfunction, epilepsy, anemia and coagulation deficiency.

*Correspondence:
Rachel L. Difazio
rachel.difazio@childrens.harvard.edu
Full list of author information is available at the end of the article

Conclusions Potential predictors of the development of post-operative complications were identified in this study and while identified predictors were similar using stepwise and LASSO regression approaches, model discrimination was slightly improved with LASSO. Findings will be used to inform future research processes determining which variables to consider for developing risk prediction models.

Keywords Cerebral palsy, Pediatric, Spinal fusion, Post-operative complications, Surgical site infection, Comorbidities, Comorbidity, Respiratory

Introduction

Cerebral palsy (CP) is a neuro-developmental condition that begins in early childhood and persists throughout life. It is the leading cause of physical disability in childhood and international prevalence estimates range from 1 to nearly 4 per 1,000 live births [1]; children with severe, non-ambulatory CP represent approximately 30% of those with CP [2]. The neurologic lesion associated with CP is non-progressive however the co-occurring conditions can worsen over time. Although physical disability is the hallmark of CP, many children experience a vast array of medical comorbidities including neurologic (e.g., epilepsy, visual and hearing abnormalities, cognitive deficits, sleep disorders, pain), gastroenterologic (e.g., esophageal dysmotility, gastroesophageal reflux, delayed gastric emptying, constipation), feeding and growth (e.g., oropharyngeal dysphagia, growth issues, dental problems) and musculoskeletal (e.g., neuromuscular scoliosis, spastic hip instability, joint contractures) [3–5]. Surgeries to correct orthopedic deformities are often indicated but these numerous comorbidities increase surgical risk including the development of post-operative complications [6–9].

Neuromuscular scoliosis affects 50-75% of children with non-ambulatory CP [10–12] and a spine fusion is recommended to prevent worsening of the curvature and diminishing health related quality of life [13–15]. However, studies have demonstrated that children with non-ambulatory CP will experience higher rates of post-operative respiratory complications and infections resulting in poorer clinical outcomes, longer lengths of stay and higher costs than children with idiopathic scoliosis [16–19]. The utility of spine surgery in this population continues to be debated [20, 21].

Rationale and significance

To date the primary focus of clinical outcomes research in this population has been on identifying risk factors for complication development such as the type of surgical intervention, preoperative radiographic measurements, and patient demographics. One critical limitation of the current literature is that most studies have not accounted for the effects that multiple comorbidities have on these outcomes [20, 22]. Without addressing comorbidities, it is uncertain when and for whom surgical intervention would be beneficial as it remains unclear which comorbidities affect the development of specific complications.

Risk prediction models can be useful clinical tools to identify at-risk patients, modify care, and engage in shared decision-making. To develop these models, extensive research must be undertaken to identify potential predictors of the outcome of interest, to determine how to construct each predictor, and to establish how to quantify each predictor's individual or synergistic contribution to the overall risk. Ideally, variable selection is accomplished in sequential phases using different data sources to modify methods as the research process unfolds, with the ultimate goal of enhancing generalizability and the utility of the final model. Bypassing developmental phases can lead to predictive models with insufficient specificity or generalizability, thus reducing the model's validity and usefulness for the intended end-user.

Approaches to variable selection

Variable selection methods are controversial and the superiority of one method over another often depends on the data and context [23]. Traditionally, in exploratory modeling, predictors are either selected based upon clinical experience [24, 25] or selected using a stepwise variable selection strategy from a limited pool of variables [26, 27]. Models that include variables selected based on clinical experience typically focus more on understanding, rather than predicting outcomes and therefore have low prediction accuracy [25, 28]. Stepwise regression uses data-driven and at times arbitrary definitions of thresholds (p-values or F-tests) that are used to decide which variables to include or exclude, which creates an inherent problem that has been identified in previous studies [29–31]. Stepwise strategies can also fail to identify true predictors and overstate predictor-outcome relationships when the sample size is not large [32], which is not uncommon in pediatric research studying post-operative risks, especially for clinical populations like CP.

Another variable selection technique, Least Absolute Shrinkage and Selection Operator (LASSO) [33], can potentially offer an alternative providing improved predictive ability [33, 34]. LASSO is a regularization technique that applies a penalty to non-zero coefficients that "shrinks" the parameter estimate towards zero, optimizing the bias-variance tradeoff to enhance the model's predictive ability [35]. The LASSO approach was developed

to overcome the limitations that occur when there are many predictors within the model. By shrinking variables estimates towards zero, the LASSO model can effectively exclude some irrelevant variables and produce sparse estimations that are more simple and interpretable than models developed using other approaches. The LASSO shrinkage regression model has been increasingly used to adjust various confounders and investigate the associations between several predictors and a health outcome [36, 37].

Consider the following comparison between equations for linear regression and LASSO regression. In the typical set up for linear regression, let Y be the dependent variable, X the independent variables (predictors), n the number of subjects (sample size), and p the number of predictors. The linear regression model assumes $E(Y|X=x) = \beta_0 + x^T \beta$ and the estimated parameter $\hat{\beta}$ (a vector with length p) is the one that minimizes the sum of deviation squares $\sum_{i=1}^{n} (y_i - \beta_0 - x_i^T \beta)^2$ over the space of β. However, the LASSO penalty is to minimize $\sum_{i=1}^{n} (y_i - \beta_0 - x_i^T \beta)^2 /(2n) + \lambda \sum_{j=1}^{p} |\beta_j|$, where λ is a tuning parameter that can be determined using cross validation. Lasso regression will automatically select variables that are useful, discarding the useless or redundant variables [38].

Objectives

In this paper, we constructed models to predict respiratory complications and infection following spine surgery in children with CP. Currently there is little evidence on which variables are contributing factors despite their high prevalence following spine surgery among children with CP [7, 8, 39, 40]. Variables that may contribute to risks include diverse demographic, clinical (e.g., comorbidities), biological, behavioral, and socio-ecological factors [41–48]. Data exploration is ideal in this context as it helps to identify potential predictors for future testing starting from a wide, unknown set of variables. It remains challenging to develop robust variable selection methods in order to enhance predictability.

The primary objective of this exploratory study was to examine the utility of an administrative database for predicting post-operative outcomes in children with CP through examination of the performance of stepwise and LASSO regression techniques in variable selection and the development of clinically useful prediction models. We further sought to identify potential predictors of respiratory complications and infection following spine surgery among children with CP, with the goal of using this exemplar for informing future research processes determining which variables to consider for developing risk prediction models.

Sequestering large sample sizes can be challenging for the pediatric CP population undergoing spine surgery, limiting the number of variables for modelling. We therefore focused on high priority potential predictors, including age, sex, and comorbidities, and used clinical data to address optimally the study objective. In order to meet our objective, we asked the following questions: (1) Which variable selection approach, stepwise or LASSO, is best used to determine potential predictors of respiratory complications and infection development in children with CP following spine surgery? (2) Which comorbidities are associated with the development of respiratory complications and infection following spine surgery in children with CP?

Patients/methods

Design, database, and representation

This retrospective cohort study accessed patient-level medical claims from 01/01/2001-12/31/2018 from Optum's de-identified Clinformatics® Data Mart Database [49]. This database was selected because it represents a large, geographically diverse population from across the United States and allows for tracking of patient claims longitudinally in both the outpatient and inpatient settings. Claims, although primarily used for billing reimbursement of healthcare services, can readily be linked to medical conditions within the database by searching for unique codes attached to patient-level data. The codes used to identify CP, the surgery type, medical comorbidities and complications including respiratory complications and infection are presented in Supplementary Table 1.

Ethical approval

Data were de-identified. All data management protocols were approved and a waiver of informed consent was granted by the University of Michigan's Institutional Review Board (HUM00174549).

Cohort selection

A flow chart of the inclusion and exclusion criterion is presented in Fig. 1. Children were included if they were ≤25 years old by the date of their index spine surgery, underwent surgery between 01/01/2002-09/30/2018, had ≥12-months pre-operative (baseline information) and ≥3-months post-operative (outcomes) continuous health plan enrollment. If children had <3-months of post-operative continuous health plan enrollment but experienced the outcome prior to loss to follow-up, they were included in the study (excluded, *n*=5 [2.2%]). To optimize sensitivity and specificity of the cohort, CP was identified by ≥2 claims with a pertinent code for CP, where each claim for CP was on a separate day within 12-months of one another [50].
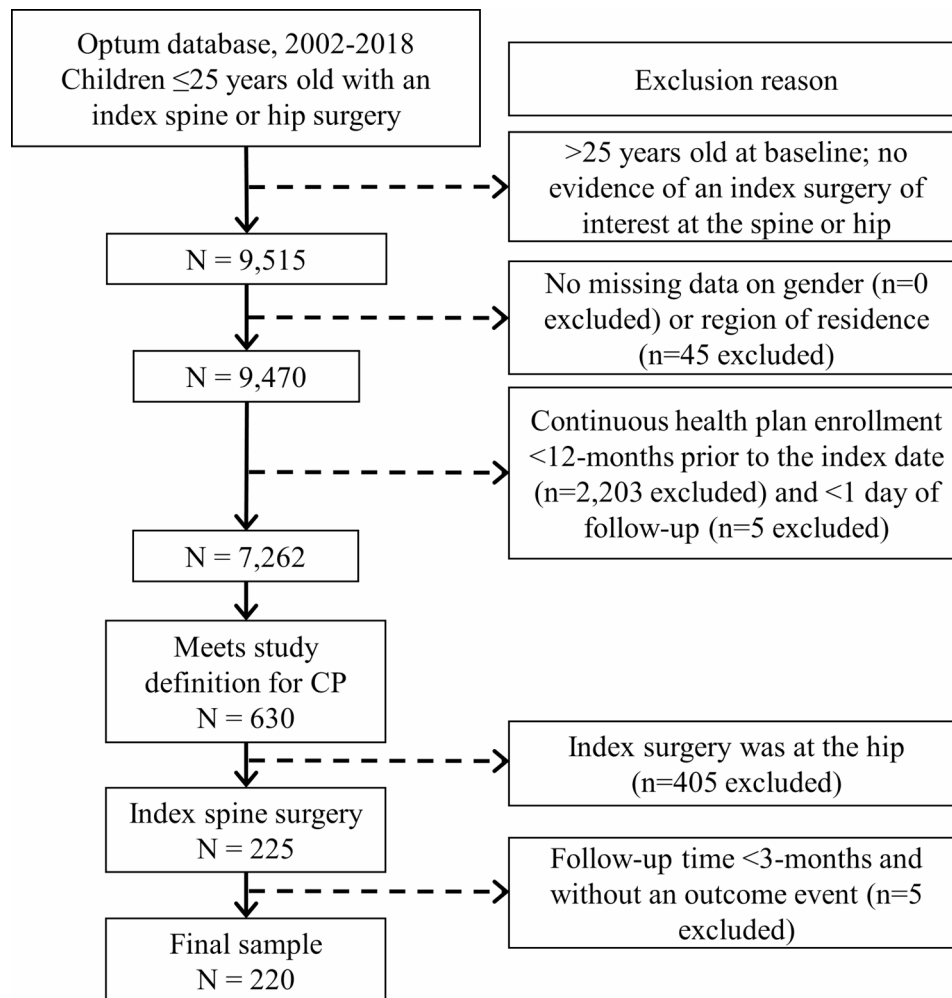
```
┌─────────────────────────────────┐
│     Optum database, 2002-2018    │
│    Children ≤25 years old with an│            ┌─────────────────────────────┐
│     index spine or hip surgery   │            │      Exclusion reason        │
└─────────────────────────────────┘            └─────────────────────────────┘
              │                                  ┌─────────────────────────────┐
              │ - - - - - - - - - - - - - - - → │   >25 years old at baseline; no│
              ▼                                  │  evidence of an index surgery of│
         ┌──────────────┐                        │   interest at the spine or hip │
         │  N = 9,515   │                        └─────────────────────────────┘
         └──────────────┘                        ┌─────────────────────────────┐
              │ - - - - - - - - - - - - - - - → │  No missing data on gender (n=0│
              ▼                                  │   excluded) or region of residence│
         ┌──────────────┐                        │      (n=45 excluded)         │
         │  N = 9,470   │                        └─────────────────────────────┘
         └──────────────┘                        ┌─────────────────────────────┐
              │                                  │ Continuous health plan enrollment│
              │ - - - - - - - - - - - - - - - → │ <12-months prior to the index date│
              ▼                                  │  (n=2,203 excluded) and <1 day of│
         ┌──────────────┐                        │    follow-up (n=5 excluded)  │
         │  N = 7,262   │                        └─────────────────────────────┘
         └──────────────┘
              │
              ▼
         ┌──────────────┐
         │ Meets study  │
         │definition for CP│
         │   N = 630    │                        ┌─────────────────────────────┐
         └──────────────┘     - - - - - - - → │   Index surgery was at the hip│
              │ - - - - - - - - - - - - - - - │      (n=405 excluded)         │
              ▼                                  └─────────────────────────────┘
         ┌──────────────┐
         │Index spine surgery│                   ┌─────────────────────────────┐
         │   N = 225    │                        │  Follow-up time <3-months and│
         └──────────────┘     - - - - - - - → │ without an outcome event (n=5│
              │ - - - - - - - - - - - - - - - │        excluded)             │
              ▼                                  └─────────────────────────────┘
         ┌──────────────┐
         │ Final sample │
         │   N = 220    │
         └──────────────┘
```

Fig. 1 Flow chart of the inclusion and exclusion criterion

Outcomes

The two outcome variables of interest were post-operative respiratory complications and infections. These outcomes were selected because they were identified as two of the leading complications following spine surgery in children with neuromuscular scoliosis with 22.7% of the children experiencing respiratory issues and 10.9% developing an infection [51]. Respiratory complications were defined as the incidence of occurrence respiratory issues within 3-months post-operatively and including the first indication of any of the following conditions: pneumonia, aspiration pneumonia, atelectasis, pleural effusion, pneumothorax, pulmonary edema, or other respiratory complications. The infection outcome included the incidence of infection within 3-months of surgery, including the first indication of the following conditions: surgical site infection, urinary tract infection, meningitis, peritonitis, sepsis, or septicemia. We selected the 3-month timeframe to allow for sufficient time for complications

to develop following spine surgery, consistent with prior work [44].

It was not possible to determine with high confidence if occurrence of some of the specific conditions (e.g., atelectasis, pneumonia) were truly an incident event if that child had an occurrence of the same condition in the baseline period. For example, if a child had evidence of pneumonia pre-operatively, a post-operative claim for pneumonia may have been a follow-up healthcare service for the pre-operative condition. To obtain incident outcome events, specific conditions occurring in the baseline period were not counted as incident events during the follow-up period.

Selection of possible predictors

Using clinical knowledge and informed by the literature [43–48], we initially considered 26 variables as potential prognostic or causal predictors of the outcomes, including age, sex, and 24 comorbidities derived from diagnoses or relevant medications. Given the limitations of our

available sample size, we reduced the number of potential predictors prior to the exploratory modelling: the first reduction phase considered logistical factors and clinical theory and the second reduction phase used data-driven techniques.

For the first reduction phase, we considered the logistics of the data source, such as the sensitivity and specificity, either using the literature or relying on our experience with claims data. This process led to the exclusion of two variables; (1) dysphagia and (2) non-ambulatory status. Variables with $n<5$ were either combined with other physiologically relevant variables or were excluded. This led to the construction of three variables: "cardiovascular disease" combining five variables (congenital heart disease; cardiac conduction disorders and arrythmias; heart failure; hypertension; cerebrovascular disease); "gastrointestinal issues" combining four variables (constipation; gastrointestinal bleeding or obstruction; pancreatitis; evidence of a gastronomy tube); and "anemia/coagulation deficiency" combining two variables (anemia deficiency; coagulation deficiency). This step also led to the exclusion of three variables (chronic kidney disease; liver disease; metabolic disease). At the end of this phase, the number of potential predictors was reduced to 12 variables.

For the second phase, a correlation matrix was developed for the 12 remaining variables to assess for collinearity, as interpretations from the exploratory modelling technique described below can be biased if collinearity is present [52]. Evidence of collinearity among variables was based on a medium effect size (e.g., |0.30|) of the bivariate relationship between each variable. A larger effect size (e.g., |0.40| to |0.70|) has been suggested, but we opted for |0.30| given the relatively small sample size [53]. There was evidence of collinearity between "gastrointestinal issues" and "gastroesophageal reflux" (phi coefficient, 0.32). We combined the latter with the former given the physiological relevance, and there was no longer evidence of collinearity with any variables. At the end of this phase, there were 11 potential predictors for exploratory modelling.

Statistical analysis

Baseline descriptive characteristics (age, sex, race, U.S. region of residence, surgery year, type of CP), prevalence of potential predictors, and outcome events were summarized for the cohort.

Logistic regression models were developed for each outcome using stepwise regression for variable selection. Stepwise regression is ideal for data-driven exploratory screening of potential predictors when there is limited evidence of variable contribution to the outcome [54]. The 11 potential predictors were entered into each model. In separate models, age was treated as continuous, narrow categorical (<9, 9–11, 12–14, 15–18, and

19–25 years), and broad categorical (<12, 12–18, and 19–25 years) to examine for effects on interpretations. Following recommendations [55], $P \leq 0.25$ was used to allow a variable to enter the model and $P \leq 0.20$ was used to retain variables in the final model. We opted for a more lenient threshold for retaining variables, as compared to $P \leq 0.15$ for example [55], given the small sample size and to avoid preemptively excluding possibly important variables for future investigations.

The intended use of this statistical approach was for data exploration. Interpretations from this exploratory modelling approach are analytically and conceptually different than inference-based modelling [23]. This statistical approach does not account for data-driven modelling decisions that give rise to the final model. Thus, the statistical parameters often used to interpret inference-based modelling are biased in this exploratory modelling approach, such as underestimating standard errors creating narrower confidence intervals and lower P-values, overstating the true association [23, 55]. Therefore, the primary interpretation of this study was to identify which variables were retained in the final model, consistent with the goal of the exploratory phase of this work. Other statistical parameters were provided for comparability with future studies, including the effect size as the odds ratio (OR), model discrimination via the c-statistic (≥ 0.70 indicates a "good" predictive model), and model fit using the Hosmer-Lemeshow (HL) goodness-of-fit test ($P \leq 0.05$ indicates poor model fit). Confidence intervals and P-values for the variables are not presented to avoid misinterpretation of the exploratory findings.

Sensitivity analysis

We performed two sets of logistic regression with LASSO using the 11 variables. The first analysis applied no method for choosing an optimal model. This was done recognizing that some variables may be important to include in future studies regardless of the statistical parameters observed in this cohort. The second analysis used a traditional approach for choosing an optimal model using Akaike's Information Criterion (AIC), a measure of model fit that helps to balance the bias-variance tradeoff. The variables retained and their effect size (i.e., OR) from the first analysis were presented. The variables retained from the second analysis are noted, but the effect size is not presented, as this second analysis provides a subset of variables with similar effect sizes as the first analysis.

For patient de-identification purposes, variables with <11 cases were either not reported or suppressed to comply with the Data Use Agreement. Analyses were performed using SAS version 9.4.

Table 1 Demographic and clinical characteristics of the cohort (*N*=220)

	% (*n*)
Age, mean (SD)	14.5 (4)
<9 years	5.5 (12)
9–11 years	15.0 (33)
12–14 years	30.9 (68)
15–18 years	34.1 (75)
19–25 years	14.6 (32)
Sex	
Female	48.6 (107)
Male	51.4 (113)
Race	
Black	7.3 (16)
Hispanic	8.6 (19)
White	74.6 (164)
Asian/missing/other	9.5 (21)
U.S. region of residence	
West	17.3 (38)
Midwest	33.2 (73)
South	38.2 (84)
Northeast	11.4 (25)
Type of CP	
Spastic quadriplegia	43.2 (95)
Spastic diplegia or hemiplegia	8.2 (18)
Other/unspecified	48.6 (107)
Year of index surgery	
2002–2007	16.8 (37)
2008–2013	43.6 (96)
2014–2018	39.6 (87)
Baseline occurrence of outcome	
Respiratory complication	8.2 (18)
Infection	< 4%*

SD, standard deviation. *N<11 cases and not presented for patient de-identification purposes

Table 2 Prevalence of the 11 potential predictors for the development of post-operative respiratory complication and infection (*N*=220)

	% (*n*)
Age, mean (SD)	14.5 (4.0)
<9 years	5.5 (12)
9–11 years	15.0 (33)
12–14 years	30.9 (68)
15–18 years	34.1 (75)
19–25 years	14.6 (32)
Sex	
Female	48.6 (107)
Male	51.4 (113)
Epilepsy	53.2 (117)
Gastrointestinal issues	17.3 (38)
Intellectual disabilities, autism spectrum disorders, or global developmental delay	14.1 (31)
Sleep apnea	11.8 (26)
Asthma/Chronic obstructive pulmonary disease	8.6 (19)
Cardiovascular disease	6.4 (14)
Hypothyroidism	< 5%*
Anemia or coagulation deficiency	< 5%*
Bladder dysfunction	< 5%*

SD, standard deviation. *N<11 cases and not presented for patient de-identification purposes

issues; bladder dysfunction; cardiovascular disease; ID, ASD or global developmental delay; and epilepsy, regardless of how age was treated (Table 3). When treated as continuous or broad categorical, but not narrow categorical, age was also retained in the final model. The c-statistic for each model ranged from 0.76 to 0.78, indicating good model discrimination.

Results

The baseline descriptive characteristics of the 220 children with CP who underwent spine surgery is presented in Table 1.

Approximately 80% (*n*=175) of the children were 12 years or older with almost equal numbers of females (*n*=107, 48.6%) and males (*n*=113, 51.4%). Close to 75% (*n*=164) were white and the vast majority were non-Hispanic (*n*=201, 91.4%). The children resided throughout the United States. The prevalence of the 11 variables considered to be potential predictors of the development of a respiratory complication or an infection is presented in Table 2. During the 3-month follow-up, 21.8% (*n*=48) developed a respiratory complication and 12.7% (*n*=28) developed an infection.

3-Month incidence of respiratory complication

The variables retained in the final model using stepwise regression included five comorbidities: gastrointestinal

Sensitivity analysis for respiratory complication

The results using LASSO for variable selection are presented in Table 4 when predicting 3-month incidence of developing a respiratory complication.

The model discrimination was slightly improved using LASSO over the primary analysis using stepwise logistic regression (c-statistic ranged from 0.78 to 0.80 vs. 0.76–0.78) and the effect size was attenuated as expected due to penalization of the regression coefficients. The first analysis (no method for choosing an optimal model) resulted in all 11 variables retained in the final model regardless of how age was treated, but the post-penalized OR of some variables was close to 1.00, indicating little influence on the outcome. For the second analysis (lowest AIC to choose the optimal model), the final variables retained were largely consistent with the primary analysis, except for the additional retainment of asthma/chronic obstructive pulmonary disease for all models and

Table 3 Results of stepwise regression for variable selection predicting 3-month incidence of post-operative respiratory complications (*N*=220)

Predictors retained in final model	Odds ratio
Age as continuous	
Age (continuous)	1.10
Gastrointestinal issues	6.77
Bladder dysfunction	3.48
Cardiovascular disease	2.56
ID, ASD, or global developmental delay	2.35
Epilepsy	0.49
Model discrimination, c-statistic	0.77
Model fit, HL goodness-of-fit test	0.98
Age, narrow categorical*	
Gastrointestinal issues	5.93
Bladder dysfunction	3.39
Cardiovascular disease	2.36
ID, ASD, or global developmental delay	2.23
Epilepsy	0.47
Model discrimination, c-statistic	0.76
Model fit, HL goodness-of-fit test	0.09
Age, broad categorical**	
Age group, 19–25 years (reference: <12 years)	4.18
Age group, 12–18 years (reference: <12 years)	1.52
Gastrointestinal issues	7.17
Bladder dysfunction	3.77
Cardiovascular disease	2.64
ID, ASD, or global developmental delay	2.37
Epilepsy	0.48
Model discrimination, c-statistic	0.78
Model fit, HL goodness-of-fit test	0.18

ID, intellectual disabilities; ASD, autism spectrum disorders; HL, Hosmer-Lemeshow. *Age categorized as <9, 9–11, 12–14, 15–18, and 19–25 years old. **Age categorized as <12, 12–18, and 19–25 years old

Table 4 Results of LASSO regression for variable selection predicting 3-month incidence of post-operative respiratory complications (*N*=220)

Predictors retained in final model	Odds ratio
Age as continuous	
Age (continuous)	*1.07*
Sex, males (reference: females)	0.98
Gastrointestinal issues	*5.63*
Bladder dysfunction	*2.06*
ID, ASD, or global developmental delay	*1.97*
Chronic obstructive pulmonary disease	*1.77*
Cardiovascular disease	*1.75*
Hypothyroidism	1.01
Sleep apnea	0.96
Anemia or coagulation deficiency	0.87
Epilepsy	*0.54*
Model discrimination, c-statistic	0.78
Model fit, HL goodness-of-fit test	0.64
Age, narrow categorical*	
Age group, 19–25 years	*3.53*
Age group, 15–18 years	*1.49*
Age group, 9–11 years	0.95
Sex, males (reference: females)	0.91
Gastrointestinal issues	*6.36*
Bladder dysfunction	*4.51*
Cardiovascular disease	*2.32*
ID, ASD, or global developmental delay	*2.09*
Asthma/chronic obstructive pulmonary disease	*2.06*
Hypothyroidism	1.21
Sleep apnea	0.70
Epilepsy	*0.53*
Anemia or coagulation deficiency	*0.36*
Model discrimination, c-statistic	0.80
Model fit, HL goodness-of-fit test	0.16
Age, broad categorical**	
Age group, 19–25 years	*3.73*
Age group, 12–18 years	1.33
Sex, males (reference: females)	0.94
Gastrointestinal issues	*6.34*
Bladder dysfunction	*4.36*
Cardiovascular disease	*2.35*
ID, ASD, or global developmental delay	*2.25*
Asthma/Chronic obstructive pulmonary disease	*2.00*
Hypothyroidism	1.09
Sleep apnea	0.70
Epilepsy	*0.53*
Anemia or coagulation deficiency	0.40
Model discrimination, c-statistic	0.80
Model fit, HL goodness-of-fit test	0.32

ID, intellectual disabilities; ASD, autism spectrum disorders; HL, Hosmer-Lemeshow. *Age categorized as <9, 9–11, 12–14, 15–18, and 19–25 years old. **Age categorized as <12, 12–18, and 19–25 years old. Italicized variables indicate the final variables retained in the model when selecting based on the lowest Akaike's Information Criterion

anemia or coagulation deficiency when age was treated as narrow categorical.

3-Month incidence of infection

The variables retained in the final models using stepwise regression included age, sex, cardiovascular disease, and asthma/chronic obstructive pulmonary disease when treating age as continuous or narrow categorical (Table 5). When age was treated as broad categorical, cardiovascular disease was not retained in the final model, but bladder dysfunction was. The c-statistic for each model ranged from 0.64 to 0.71, indicating poor to good model discrimination.

The results using LASSO for variable selection are presented in Table 6 when predicting 3-month incidence of infection.

The model discrimination was slightly improved over the primary analysis (c-statistic ranged from 0.68 to 0.72 vs. 0.64–0.71). The first analysis resulted in 9 to 11 variables retained in the final model depending on how

Table 5 Results of stepwise regression for variable selection predicting 3-month incidence of post-operative infection (*N* = 220)

Predictors retained in final model	Odds ratio
Age as continuous	
Age (continuous)	1.08
Sex, females (reference: males)	2.13
Cardiovascular disease	2.59
Asthma/Chronic obstructive pulmonary disease	2.49
Model discrimination, c-statistic	0.64
Model fit, HL goodness-of-fit test	0.02
Age, narrow categorical*	
Age group, 19–25 years (reference: <9 years)	7.73
Age group, 9–11 years (reference: <9 years)	5.73
Age group, 15–18 years (reference: <9 years)	2.77
Age group, 12–14 years (reference: <9 years)	2.01
Sex, females (reference: males)	2.32
Asthma/Chronic obstructive pulmonary disease	3.12
Cardiovascular disease	2.62
Model discrimination, c-statistic	0.71
Model fit, HL goodness-of-fit test	0.42
Age, broad categorical**	
Age group, 19–25 years (reference: <12 years)	1.80
Age group, 12–18 years (reference: <12 years)	0.58
Sex, females (reference: males)	2.00
Asthma/Chronic obstructive pulmonary disease	2.77
Bladder dysfunction	2.77
Model discrimination, c-statistic	0.70
Model fit, HL goodness-of-fit test	0.67

*Age categorized as <9, 9–11, 12–14, 15–18, and 19–25 years old; HL, Hosmer-Lemeshow. **Age categorized as <12, 12–18, and 19–25 years old

Sensitivity Analysis for Infection

Table 6 Results of LASSO regression for variable selection predicting 3-month incidence of post-operative infection (*N* = 220)

Predictors retained in final model	Odds ratio
Age as continuous	
Age (continuous)	1.06
Sex, males (reference: females)	0.53
Cardiovascular disease	1.73
Asthma/Chronic obstructive pulmonary disease	1.65
Gastrointestinal issues	1.54
Bladder dysfunction	1.48
Hypothyroidism	1.02
Anemia or coagulation deficiency	0.99
Epilepsy	0.78
Model discrimination, c-statistic	0.68
Model fit, HL goodness-of-fit test	0.69
Age, narrow categorical*	
Age group, 19–25 years	5.20
Age group, 9–11 years	3.75
Age group, 15–18 years	1.81
Age group, 12–14 years	1.28
Sex, males (reference: females)	0.45
Asthma/Chronic obstructive pulmonary disease	2.43
Cardiovascular disease	2.23
Bladder dysfunction	2.20
Hypothyroidism	1.46
Gastrointestinal issues	1.26
ID, ASD, or global developmental delay	1.08
Sleep apnea	0.93
Epilepsy	0.75
Anemia or coagulation deficiency	0.54
Model discrimination, c-statistic	0.71
Model fit, HL goodness-of-fit test	0.39
Age, broad categorical**	
Age group, 19–25 years	1.80
Age group, 12–18 years	0.56
Sex, males (reference: females)	0.50
Bladder dysfunction	2.52
Asthma/Chronic obstructive pulmonary disease	2.15
Cardiovascular disease	2.00
Hypothyroidism	1.63
Gastrointestinal issues	1.38
ID, ASD, or global developmental delay	0.96
Sleep apnea	0.79
Epilepsy	0.73
Anemia or coagulation deficiency	0.63
Model discrimination, c-statistic	0.72
Model fit, HL goodness-of-fit test	0.45

ID, intellectual disabilities; ASD, autism spectrum disorders; HL, Hosmer-Lemeshow. *Age categorized as <9, 9–11, 12–14, 15–18, and 19–25 years old. **Age categorized as <12, 12–18, and 19–25 years old. When selecting the final model based on the lowest Akaike's Information Criterion, no variables were retained in any model

age was treated. The relative contribution of age and sex was consistent with the primary analysis, and the retainment of the comorbidities with the highest effect size was also consistent with the primary analysis. However, the LASSO technique resulted in additional retainment of comorbidities that may have meaningful associations to consider for future studies, such as gastrointestinal issues, bladder dysfunction, epilepsy, and anemia or coagulation deficiency. The second analysis resulted in no variables retained regardless of how age was treated.

Discussion

This study demonstrated relative consistency between the two approaches, stepwise and LASSO regression, for identifying potential predictors of respiratory complications and infection in children with CP following spinal surgery. LASSO better modified the effects that fit the predictors, whereas stepwise may not have been as flexible in this regard. This difference in model performance may be related to the small sample size in this study.

These findings can be used to inform the development of clinical risk prediction models by considering the use

of age, sex, and certain comorbidities depending on the risk of interest as explicated in Tables 3, 4, 5 and 6. Five prior studies have begun to examine the risk predictors of respiratory complications [43, 45–48] and infections [47, 48] following spine surgery in children with CP. Significant limitations exist, however, across this body of work. Data for these studies were drawn from retrospective chart reviews (*n*=4; 80%) [43, 45–47] or a pre-established database (*n*=1, 20%) [48], both with their inherent limitations. Relatively small sample sizes (i.e., *n*=74–127), a significant issue in research conducted on this population, was also noted in two-thirds of the studies [43, 45, 46, 48]; thus, impeding the ability to assess for predictors simultaneously in a systematic manner (e.g., limited to bivariate associations). Also, of note, three (60%) of the studies are over a decade old, not capturing current surgical techniques and standards of post-operative care [43, 45, 46]. Future methodologic and clinical studies are needed to test and confirm the observed associations and identify other variables not examined in this study that may be potential predictors (e.g., anthropometrics). Taken together, model findings should be interpreted as hypothesis-generating (exploratory) as opposed to hypothesis-testing (inference).

The primary interpretations for 3-month risk of respiratory complications were largely consistent when variable selection was performed using stepwise selection or LASSO. However, LASSO using AIC to choose the optimal model additionally retained asthma/chronic obstructive pulmonary disease and anemia/coagulation deficiency. These additionally retained variables seem appropriate to consider in future testing to explore the potential underlying mechanisms linking these comorbidities with risk of respiratory complications although, inferences on the observed associations are beyond the scope of this exploratory study. Some of the primary interpretations for 3-month risk of infection were also consistent across variable selection techniques, such that age, sex, cardiovascular disease, asthma/chronic obstructive pulmonary disease, and bladder dysfunction were retained with the highest effect size. However, LASSO using AIC to choose the optimal model retained none of the 11 variables, suggesting that introduction of these variables added to model complexity beyond the regression's intercept alone. The number of post-operative infections was small (*n*=28); limiting model interpretations.

The varied results for infection risk reflect the limitations of modelling the number of predictors for the number of outcome events. One common way to identify how many predictors should maximally fit a given data set is using the ratio of outcome events per independent variable (EPV). A rule-of-thumb is 1:10, such that 1 predictor can be considered per 10 outcome events [56], but stricter (e.g., 1:4) and more lenient (e.g., 1:20) EPVs have been suggested [44], as the bias-variance tradeoff can depend on other factors, such as the magnitude of effect sizes and collinearity among predictors [57, 58]. This study had an EPV of ~4–5 when modelling respiratory complications and ~2–3 when modelling infection. We intentionally aimed to reduce bias from a low EPV by mitigating collinearity among potential predictors prior to entering the statistical model, which decreases bias for a given EPV [55].

It has been recommended that variable selection not be performed when EPV<10 due in part to issues of false discoveries when using variable selection methods (e.g., overstating true associations) [55]. However, this recommendation may have been made under the context of inference and may not fully apply to exploration. On the other hand, it has been suggested that articulating that the modelling goal is exploratory and with appropriate interpretations (e.g., variable retainment rather than *P*-values), the issues of false discoveries are minimized [23]. This relates to the notion that there are no claims of confirmation of the observed associations, and that associations require future testing with independent data. In light of the low EPV, this study is necessary given the importance of the topic and challenges in sequestering large sample sizes with sufficient EPV from children with CP undergoing spine surgery. Further, to augment impact, we performed two commonly used variable selection methods, each with their own set of strengths and weaknesses, to assist interpretations informing future study designs.

Another challenge encountered was how to model age, especially in children with complex chronic conditions. Continuous variables can be complex if the relationship is non-linear. A common data-driven method to transform non-linear variables is using restricted cubic splines [59]. However, this method makes linear assumptions at the ends of the association. In the context of this study, the ends of the age-risk association may behave non-linearly in a way that should be captured due to the clinical relevance. For example, the LASSO method modelling respiratory complications found that 19–25 year olds had the highest adjusted effect, which reduced considerably with the next youngest age group before tapering out to no-to-minimal effect. We therefore opted for a more basic initial assessment by treating age as continuous and in clinically relevant categories. In general, older age was associated with greater risk of respiratory complications, but the association may not be linear across the full age span and especially among the younger half of the cohort. The association appeared more variable for infection.

Limitations

The limitations of this study must be discussed. Information on severity of CP is not provided in claims data. However, studies have demonstrated that children with more severe CP subtypes have a higher proportion and number of comorbidities [4, 60] and this was likely accounted for, at least in part, from the comorbidity variables. Further, spine surgeries are typically performed in children with more severe CP as the incidence and severity of scoliosis is directly proportional to the extent of the child's neurological impairment and inversely proportional to the child's functional abilities [61, 62]. This study was unable to assess the type of surgery due to the relatively small dataset and few outcome events, which may contribute to prediction of the outcome. Future work may incorporate the type of surgery as a potential predictor. The generalizability of findings is not known. It has been suggested that privately insured children with CP represent mild to severe CP, but a slightly less medically complex segment of the broader pediatric population with CP who are eligible for federal insurance, with potential insufficient racial representation of non-white children [13, 50, 63]. In this study, some patient characteristics (e.g., age, sex) and prevalence of comorbidities (e.g., hypothyroidism) were similar to other studies examining risk of complications following spine surgery among children with CP [44, 45]. Moreover, in the study with ∼2,800 children with CP undergoing spine surgery, 41.4% had private insurance which was not strongly associated with outcomes [49], suggesting reasonable use of this private insurance database to meet the study's exploratory goals. There may be other relevant comorbidities not examined in this study to consider in future studies. Suboptimal sensitivity and specificity of comorbidities could underestimate or distort the associations observed in this study. Recording of comorbidities is often accurate but may be incomplete in both numbers and severity. Undercoding is a limitation of all administrative databases [64]. We attempted to mitigate this bias by using comorbidities with reasonable detection in claims, which was based on the literature or our own experience with claims data.

Conclusion

While the model performance was similar between approaches, LASSO had a slight improvement in the c-statistic. Further, LASSO penalizes regression coefficients, which enhances the potential for generalizability of the developed algorithm to other datasets. Further, this exploratory study identified potential age, sex, and comorbidity predictors for risk of respiratory complications and infection following spine surgery among children with CP. These associations will need to be tested in independent datasets for confirmation. The study findings provide novel information to inform the design of future inference-based studies and development of clinical risk prediction models, ultimately to improve post-op monitoring and secondary preventions.

Declarations

Ethical approval
Data were de-identified. All data management protocols were approved and a waiver of informed consent was granted by the University of Michigan's Institutional Review Board (HUM00174549). All methods were carried out in accordance with relevant guidelines and regulations that were approved the University of Michigan's Institutional Review Board (HUM00174549).

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Boston Children's Hospital, Department of Orthopedic Surgery and Sports Medicine, 300 Longwood Avenue, Boston, MA 02115, USA
[2]Maine Medical Center, Department of Emergency Medicine, 22 Bramhall Street, Portland, ME 04102, USA
[3]Boston Children's Hospital, Medicine Patient Services, 300 Longwood Avenue, Boston, MA 02115, USA
[4]Boston Children's Hospital, Complex Care Service, Division of General Pediatrics, Boston Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA

[5]Department of Physical Medicine and Rehabilitation, University of Michigan, Institute for Healthcare Policy and Innovation, Department of Physical Medicine and Rehabilitation, 315 East Eisenhower Parkway, Ann Arbor, MI 48108, USA

References

1. Centers for Disease Control and Prevention. National Center on Birth Defects and Developmental Disabilities. Data and statistics for cerebral palsy: Prevalence and Characteristics, October 18, 2022. https://www.cdc.gov/ncbddd/cp/data.html

2. Reid SM, Carlin JB, Reddihough DS. Using the Gross Motor function classification system to describe patterns of motor severity in cerebral palsy. Dev Med Child Neurol. 2011;53(11):1007–12. https://doi.org/10.1111/j.1469-8749.2011.04044.x.

3. Pruitt DW, Tsai T. (2009). Common medical comorbidities associated with cerebral palsy. Phys Med Rehabil Clin. 2009;3:453–467. https://doi.org/10.1016/j.pmr.2009.06.002

4. Shevell MI, Dagenais L, Hall N, REPACQ Consortium. Comorbidities in cerebral palsy and their relationship to neurologic subtype and GMFCS level. Neurology. 2009;72:2090–6. https://doi.org/10.1212/WNL.0b013e3181aa537b.

5. Venkateswaran S, Shevell MI. (2008). Comorbidities and clinical determinants of outcome in children with spastic quadriplegic cerebral palsy. Dev Med Child Neurol. 2008;3:216–222. https://doi.org/10.1111/j.1469-8749.2008.02033.x

6. Borkhuu B, et al. Prevalence and risk factors in postoperative pancreatitis after spine fusion in patients with cerebral palsy. J Pediatr Orthop. 2009;29(3):256–62. https://doi.org/10.1097/BPO.0b013e31819bcf0a.

7. Chidambaran V, Gentry C, Ajuba-Iwuji C, Sponsellar PD, Ain M, Lin E, Zhang X, Klaus SA, Njoku DB. A retrospective identification of gastroesophageal reflux disease as a new risk factor for surgical site infection in cerebral palsy patients after spine surgery. Anesth Analgesia. 2013;117(1):162–8. https://doi.org/10.1213/ANE.0b013e318290c542.

8. Sponseller PD, Jain A, Shah SA, Samdani A, Yaszay B, Newton PO, Thaxton LM, Bastrom TP, Marks MC. Deep wound infections after spinal fusion in children with cerebral palsy: a prospective cohort study. Spine. 2013;38(23):2023–7. https://doi.org/10.1097/BRS.0b013e3182a83e59.

9. Weissmann KA, Lafage V, Pitaque CB, Lafage R, Huaiquilaf CM, Ang B, Schulz RG. Neuromuscular scoliosis: comorbidities and complications. Asian Spine J. 2021;15(6):778. https://doi.org/10.31616/asj.2020.0263.

10. Hägglund G, Pettersson K, Czuba T, Persson-Bunke M, Rodby-Bousquet E. Incidence of scoliosis in cerebral palsy: a population-based study of 962 young individuals. Acta Orthop. 2018;89(4):443–7. https://doi.org/10.1080/17453674.2018.1450091.

11. Persson-Bunke M, Hagglund G, Lauge-Pedersen H, et al. Scoliosis in a total population of children with cerebral palsy. Spine. 2012;37:E708–13. https://doi.org/10.1097/BRS.0b013e318246a962.

12. Sandstrom K, Alinder J, Oberg B. Description of functioning and health and relations to a gross motor classification in adults with cerebral palsy. Disabil Rehabil. 2004;26:1023–31. https://doi.org/10.1080/09638280410001703503.

13. DiFazio RL, Miller PE, Vessey JA, Snyder BD. Health-related quality of life and care giver burden following spinal fusion in children with cerebral palsy. Spine. 2017;42:E733–9. https://doi.org/10.1097/BRS.0000000000001940.

14. Jones KB, Sponseller PD, Shindle MK, McCarthy ML. Longitudinal parental perceptions of spinal fusion for neuromuscular spine deformity in patients with totally involved cerebral palsy. J Pediatr Orthop. 2003;23:143–9. PMID: 12604940.

15. Tsirikos AI, Lipton G, Chang WN, et al. Surgical correction of scoliosis in pediatric patients with cerebral palsy using the single rod instrumentation. Spine. 2008;33:1133–40. https://doi.org/10.1097/BRS.0b013e31816f63cf.

16. Abousamra O, Nishnianidze T, Rogers KJ, Er MS, Sees JP, Dabney KW, Miller F. Risk factors for pancreatitis after posterior spinal fusion in children with cerebral palsy. J Pediatr Orthop B. 2018;27:163–7. https://doi.org/10.1097/BPB.0000000000000376.

17. Berry JG, Glotzbecker M, Rodean J, Leahy I, Cox J, Singer SJ, O'Neill M, Hall M, Ferrari L. Perioperative spending on spinal fusion for scoliosis for children with medical complexity. Pediatrics. 2017;140(4). https://doi.org/10.1542/peds.2017-1233.

18. Jain A, Modhia UM, Njoku DB, Shah SA, Newton PO, Marks MC, Bastrom TP, Miyanji F, Sponseller PD. Recurrence of Deep Surgical Site infection in cerebral Palsy after spinal Fusion is Rare. Spine Deform. 2017;5(3):208–12. https://doi.org/10.1016/j.jspd.2016.12.004.

19. Berry JG, et al. Comorbidities and complications of spinal fusion for scoliosis. Pediatrics. 2017;139(3):e20162574. https://doi.org/10.1542/peds.2016-2574.

20. Watanabe K, Lenke LG, Daubs MD, Watanabe K, Bridwell KH, Stobbs G, Hensley M. Is spine deformity surgery in patients with spastic cerebral palsy truly beneficial? Spine. 2009;34:2222–32. https://doi.org/10.1097/BRS.0b013e3181948c8f.

21. Whitaker A, Sharkey M, Diab M. Spinal fusion for scoliosis in patients with globally involved cerebral palsy. J Bone Joint Surg Am. 2015;97:82–7. https://doi.org/10.2106/JBJS.N.00468.

22. Berry JG, Glotzbecker M, Rodean J, Leahy I, Hall M, Ferrari L. Comorbidities and complications of spinal fusion for scoliosis. Pediatrics. 2017;139(3). https://doi.org/10.1542/peds.2016-2574.

23. Tredennick AT, Hooker G, Ellner SP, Adler PB. A practical guide to selecting models for exploration, inference, and prediction in ecology. Ecology. 2021;102:e03336. https://doi.org/10.1002/ecy.3336.

24. Donzé J, Aujesky D, Williams D, Schnipper JL. Potentially avoidable 30-day hospital readmissions in medical patients: derivation and validation of a prediction model. JAMA Intern Med. 2013;173(8):632–8. https://doi.org/10.1001/jamainternmed.2013.3023.

25. Shih SL, Gerrard P, Goldstein R, Mix J, Ryan CM, Niewczyk P, Kazis L, Hefner J, Ackerly DC, Zafonte R, Schneider JC. Functional status outperforms comorbidities in predicting acute care readmissions in medically complex patients. Gen Intern Med. 2015;30:1688–95. https://doi.org/10.1007/s11606-015-3350-2.

26. Kroch E, Duan M, Martin J, Bankowitz RA. Patient factors predictive of hospital readmissions within 30 days. J Healthc Qual. 2016;38(2):106–15. https://doi.org/10.1097/JHQ.0000000000000003.

27. Rico F, Liu Y, Martinez DA, Huang S, Zayas-Castro JL, Fabri PJ. Preventable readmission risk factors for patients with chronic conditions. J Healthc Qual. 2016;38(3):127–42. https://doi.org/10.1097/01.JHQ.0000462674.09641.72.

28. Kind AJ, Jencks S, Brock J, Yu M, Bartels C, Ehlenbach W, Greenberg C, Smith M. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med. 2014;161(11):765–74. https://doi.org/10.7326/M13-2946.

29. Hesterberg T, Choi NH, Meier L, Fraley C. Least angle and ℓ 1 penalized regression: a review. Statist Surv. 2008;2:61–93.

30. Greenland S. Invited commentary: variable selection versus shrinkage in the control of multiple confounders. Am J Epidemiol. 2008;167:523–9. https://doi.org/10.1093/aje/kwm355.

31. Hernán MA, Hernández-Díaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol. 2002;155:176–84. https://doi.org/10.1093/aje/155.2.176.

32. Flom PL, Cassell DL. Stopping stepwise: Why stepwise and similar selection methods are bad, and what you should use. InNorthEast SAS Users Group Inc 20th Annual Conference 2007 Nov 11 (Vol. 11).

33. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Ser B. 1996;58:267–88.

34. Steyerberg EW, Eijkemans MJC, Habbema JDF. Application of shrinkage techniques in logistic regression analysis: a case study. Stat Neerl. 2001;55:76–88. https://doi.org/10.1186/s12874-016-0209-0.

35. Ranstam J, Cook JA, LASSO regression. BJS. 2018;105(10):1338. https://doi.org/10.1002/bjs.10895.

36. Avalos M, Adroher ND, Lagarde E, Thiessard F, Grandvalet Y, Contrand B, et al. Prescription-drug-related risk in driving: comparing conventional and lasso shrinkage logistic regressions. Epidemiology. 2012;23:706–12. https://doi.org/10.1097/EDE.0b013e31825fa528.

37. Chen Q, Wang S. Variable selection for multiply-imputed data with application to dioxin exposure study. Stat Med. 2013;32:3646–59. https://doi.org/10.1002/sim.5783.

38. Tong L, Erdmann C, Daldalian M, Li J, Esposito T. Comparison of predictive modeling approaches for 30-day all-cause non-elective readmission risk. BMC Med Res Methodol. 2016;16(1):1–8. https://doi.org/10.1186/s12874-016-0128-0.

39. Furdock R, Luhmann SJ. Preoperative Variables Associated with Respiratory complications after Pediatric Neuromuscular spine deformity surgery. Spine Deform. 2018;332–332. https://doi.org/10.1016/j.jspd.2018.05.005.

40. Luhmann SJ, Furdock R. Preoperative variables associated with respiratory complications after pediatric neuromuscular spine deformity surgery. Spine Deform. 2019;7(1):107–11. https://doi.org/10.1016/j.jspd.2018.05.005.

41. Chambers HG, Weinstein CH, Mubarak SJ, Wenger DR, Silva PD. The effect of valproic acid on blood loss in patients with cerebral palsy. J Pediatr Orthop. 1999;19:792–5. PMID: 10573351.

42. Cloake T, Gardner A. The management of scoliosis in children with cerebral palsy: a review. J Spine Surg. 2016;2:299–309. https://doi.org/10.21037/jss.2016.09.05.

43. Kang GR, Suh SW, Lee IO. Preoperative predictors of postoperative pulmonary complications in neuromuscular scoliosis. J Orthop Sci. 2011;16:139–47. https://doi.org/10.1007/s00776-011-0028-4.

44. Lee NJ, Fields M, Boddapati V, et al. Spinal deformity surgery in Pediatric patients with cerebral palsy: a National-Level analysis of Inpatient and Postdischarge outcomes. Global Spine J. 2022;12:610–9. https://doi.org/10.1177/2192568220960075.

45. Lipton GE, Miller F, Dabney KW, Altiok H, Bachrach SJ. Factors predicting postoperative complications following spinal fusions in children with cerebral palsy. J Spinal Disord. 1999;12:197–205. PMID: 10382772.

46. Master DL, Son-Hing JP, Poe-Kochert C, Armstrong DG, Thompson GH. Risk factors for major complications after surgery for neuromuscular scoliosis. Spine. 2011;36:564–71. https://doi.org/10.1097/BRS.0b013e3181e193e9.

47. Nishnianidze T, Bayhan IA, Abousamra O, Sees J, Rogers KJ, Dabney KW, Miller F. Factors predicting postoperative complications following spinal fusions in children with cerebral palsy scoliosis. Eur Spine J. 2016;25:627–34. https://doi.org/10.1007/s00586-015-4243-0.

48. Samdani AF, Belin EJ, Bennett JT, Miyanji F, Pahys JM, Shah SA, Newton PO, Betz RR, Cahill PJ, Sponseller PD. Major perioperative complications after spine surgery in patients with cerebral palsy: assessment of risk factors. Eur Spine J. 2016;25:795–800. https://doi.org/10.1007/s00586-015-4054-3.

49. Whitney D, Kamdar N, Hirth RA, Hurvitz EA, Peterson MD. Economic burden of paediatric-onset disabilities among young and middle-aged adults in the USA: a cohort study of privately insured beneficiaries. BMJ Open. 2019;9(9):e030490. https://doi.org/10.1136/bmjopen-2019-030490.

50. Whitney DG. 5-year fracture risk among children with cerebral palsy. Pediatr Res. 2022. https://doi.org/10.1038/s41390-022-02207-4.

51. Sharma S, Wu C, Andersen T, Wang Y, Hansen ES, Bünger CE. Prevalence of complications in neuromuscular scoliosis surgery: a literature meta-analysis from the past 15 years. Eur Spine J. 2013;22:1230–49. https://doi.org/10.1007/s00586-012-2542-2.

52. Heinze G, Dunkler D. Five myths about variable selection. Transpl Int. 2017;30:6–10. https://doi.org/10.1111/tri.12895.

53. Stevens JP. Applied Multivariate statistics for the Social Sciences. 2nd ed. Hillsdate, NJ: Erbaum; 1992.

54. Hosmer DW, Lemeshow SL. Applied Logistic Regression. 2nd ed. Hoboken, NJ. 2000.

55. Heinze G, Wallisch C, Dunkler D. Variable selection - A review and recommendations for the practicing statistician. Biom J. 2018;60:431–49. https://doi.org/10.1002/bimj.201700067.

56. Harrell FE Jr., Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med. 1984;3:143–52. https://doi.org/10.1002/sim.4780030207.

57. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165(6):710–8. https://doi.org/10.1093/aje/kwk052.

58. Courvoisier DS, Combescure C, Agoritsas T, Gayet-Ageron A, Perneger TV. Performance of logistic regression modeling: beyond the number of events per variable, the role of data structure. J Clin Epidemiol. 2011;64:993–1000. https://doi.org/10.1016/j.jclinepi.2010.11.012.

59. Harrell FE Jr., Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. J Natl Cancer Inst. 1988;80:1198–202. https://doi.org/10.1093/jnci/80.15.1198.

60. Hollung SJ, Bakken IJ, Vik T, Lydersen S, Wiik R, Aaberg KM, Andersen GL. Comorbidities in cerebral palsy: a patient registry study. Dev Med Child Neurol. 2020;62:97–103. https://doi.org/10.1111/dmcn.14307.

61. Bell DF, Moseley CF, Koreska J. Unit rod segmentation spinal instrumentation in the management of patients with progressive neuromuscular spinal deformity. Spine. 1989;14:1301–7. https://doi.org/10.1097/00007632-198912000-00006.

62. Rinsky L. Surgery of spinal deformity in cerebral palsy. Twelve years in the evolution of scoliosis management. Clin Orthop Relat Res. 1990;253:100–9. PMID: 2317962.

63. Whitney DG, Hurvitz EA, Caird MS. Critical periods of bone health across the lifespan for individuals with cerebral palsy: informing clinical guidelines for fracture prevention and monitoring. Bone. 2021;150:116009. https://doi.org/10.1016/j.bone.2021.116009.

64. Tai D, Dick P, To T, Wright JG. Development of pediatric comorbidity prediction model. Arch Pediatr Adolesc Med. 2006;160:293–9. https://doi.org/10.1001/archpedi.160.3.293.

Publisher's note